TESTING THE MEASUREMENT INVARIANCE OF DATA FROM THE UTRECHT WORK ENGAGEMENT SCALE BY GENERATIONAL COHORT FOR EMPLOYEES IN THE LEISURE AND HOSPITALITY INDUSTRY

Greggory Keiffer
University of Texas at Tyler

Follow this and additional works at: https://scholarworks.uttyler.edu/hrd_grad

Part of the Hospitality Administration and Management Commons, Human Resources Management Commons, and the Management Sciences and Quantitative Methods Commons

Recommended Citation
http://hdl.handle.net/10950/624

This Dissertation is brought to you for free and open access by the Human Resource Development at Scholar Works at UT Tyler. It has been accepted for inclusion in Human Resource Development Theses and Dissertations by an authorized administrator of Scholar Works at UT Tyler. For more information, please contact tbianchi@uttyler.edu.
TESTING THE MEASUREMENT INVARIANCE OF DATA FROM THE UTRECHT
WORK ENGAGEMENT SCALE BY GENERATIONAL COHORT FOR
EMPLOYEES IN THE LEISURE AND HOSPITALITY INDUSTRY

by

GREGGORY LEE KEIFFER

A dissertation proposal submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Human Resource Development

Kim Nimon, Ph.D., Committee Chair

College of Business and Technology

The University of Texas at Tyler
November 2017
The University of Texas at Tyler  
Tyler, Texas  

This is to certify that the Doctoral Dissertation of  

GREGGORY LEE KEIFFER  

has been approved for the dissertation requirement on  
November 16, 2017  
for the Doctor of Philosophy degree

Approvals:  
Dissertation Chair: Kim Nimon, Ph.D.  
Member: Jerry Gilley, Ed.D.  
Member: Paul Roberts, Ph.D.  
Member: Greg Wang, Ph.D.  
Chair, Department of Human Resource Development  
Dean, College of Business and Technology
Acknowledgements

A warm and heartfelt thank you to my dissertation committee members: Drs. Kim Nimon, Jerry Gilley, Paul Roberts, and Greg Wang. Your collaboration and positive approach made the project achievable. I would be remised if I did not specifically recognize Dr. Kim Nimon for her exceptional commitment, coaching, counsel, and patience she demonstrated throughout the process. Thank you for continuously pushing me to elevate my thinking and work product to a high level.

The full faculty and staff of the Human Resource Development program at The University of Texas at Tyler should be commended on creating and maintaining a phenomenal doctoral program. I am grateful to all of those I had the privilege of meeting and learning from throughout the recent years. To my cohort peers and peers from neighboring cohorts – you all made the program enjoyable in how we could all share ideas, create knowledge, and learn together. I look forward the continuing friendships and academic relationships. Special recognition goes to the Magnificent 7 – thank you for the hours of collaboration and discussion.

I could not have completed this program without tremendous support from my family. To my parents, Lee and Carol Keiffer, thank you for instilling a desire for lifelong learning and always supporting my ambitions. To my wonderful children, Grant and Claire, thank you for being positive and patient. I pray that you both will answer your callings in life with a Christ-centered heart and with strong courage. To my wife and my rock, Stacie, a simple thank you cannot begin to express my gratitude to you. Your
selfless generosity over these years of devotion to this project has been nothing short of amazing. Thank you for the countless hours of listening to me drone on about the research and for the encouragement to keep pursuing the calling!
# Table of Contents

List of Tables ........................................................................................................ iv
List of Figures ........................................................................................................ viii
Abstract ................................................................................................................... ix
Chapter 1 - Introduction ....................................................................................... 1
  Background to the Problem .................................................................................. 1
  Statement of the Problem .................................................................................... 9
  Purpose of the Study ........................................................................................... 11
  Theoretical and Conceptual Underpinnings of the Study ................................... 11
  Research Hypotheses .......................................................................................... 14
  Overview of the Design of the Study ................................................................ 17
  Significance of the Study ..................................................................................... 18
  Assumptions ......................................................................................................... 20
  Delimitations ....................................................................................................... 20
  Definitions of Terms ............................................................................................ 21
Chapter Summary and Organization of the Dissertation ....................................... 23
Chapter 2 - Literature Review .............................................................................. 24
  Introduction ......................................................................................................... 24
  The Leisure and Hospitality Supersector ........................................................... 25
    Supersector Description .................................................................................... 25
    Workforce Dynamics ....................................................................................... 26
  Generational Cohort ............................................................................................ 27
    History of Generations ..................................................................................... 28
    Generational Cohorts – An American Concept? .............................................. 29
    Generations in the Workforce ......................................................................... 30
    The Generations Described ............................................................................. 32
    Importance of Generational Cohort Research in Leisure and Hospitality ...... 34
    Generational Cohort and Engagement Research Outside of Leisure and Hospitality ......................................................................................... 44
  Engagement Articles with Measurement Invariance Assessment between
    Generational Cohorts ....................................................................................... 50
Employee Engagement .......................................................................................... 53
  Domains of Employee Engagement .................................................................. 54
  Measurement of Employee Engagement .......................................................... 58
  Engagement Summary ......................................................................................... 60
Psychometrics of the Utrecht Work Engagement Scale ......................................... 61
  Development of UWES ...................................................................................... 61
  Post-Development of the UWES ....................................................................... 62
  UWES Considerations ......................................................................................... 64
List of Tables

Table 1 *Presentation of Generational Cohort Timespans by Author* .......................... 32
Table 2 *Generational Cohort Empirical Research within Leisure and Hospitality* ........ 38
Table 3 *Generational Cohort Research Outside of Leisure and Hospitality* ............... 46
Table 4 *Measurement Invariance Literature for Engagement Related Constructs* .... 68
Table 5 *Distribution of Industries for the Initial Pooled Sample of Boomer and Millennial Groups* ............................................................................................................. 119
Table 6 *Group Comparison Chi-Square Results by Industry for the Initial Pooled Sample of Boomer and Millennial Groups* ............................................................................................................. 119
Table 7 *Group Comparison Chi-Square Results for the Initial Pooled Sample of Boomer and Millennial Groups* ............................................................................................................. 121
Table 8 *Distribution of Demographics for the Initial Pooled Sample of Boomer and Millennial Groups* ............................................................................................................. 122
Table 9 *Group Comparison Chi-Square Results by Boomers and Millennials after Nearest Neighbor Match Level Propensity Score Matching* ........................................... 123
Table 10 *Group Comparison Chi-Square Results by Boomers and Millennials after Genetic Match Level Propensity Score Matching* ................................................................. 124
Table 11 *Distribution of Demographics for the Genetically Matched Pooled Sample of Boomer and Millennial Groups* ............................................................................................................. 125
Table 12 *Sample Representativeness Comparison for BLS and the Initial Sample* ....... 126
Table 13 Sample Representativeness Comparison for BLS and the Final Sample........ 127
Table 14 Distribution of Industries for the Final Sample of Boomer and Millennial
Groups.................................................................................................................................................. 129
Table 15 Group Comparison Chi-Square Results by Industry for the Final Sample of
Boomer and Millennial Groups ............................................................................................................ 129
Table 16 Fit Indices for Original and Respecified Measurement Models ................. 133
Table 17 Pattern (P) and Structure (S) Coefficients for the Pooled Sample Measurement
Model .................................................................................................................................................. 134
Table 18 Implied Correlations, Average Variance Extracted (AVE), and Composite
Reliability (CR) for the Pooled Sample .............................................................................................. 135
Table 19 Fit Indices for the Three-Factor Measurement Model for the Boomer Sample
......................................................................................................................................................... 136
Table 20 Pattern (P) and Structure (S) Coefficients for the Boomer Sample Measurement
Model .................................................................................................................................................. 137
Table 21 Implied Correlations, Average Variance Extracted (AVE), and Composite
Reliability (CR) for the Boomer Sample .............................................................................................. 138
Table 22 Fit Indices for the Three-Factor Measurement Model for the Millennial Sample
......................................................................................................................................................... 138
Table 23 Pattern (P) and Structure (S) Coefficients for the Millennial Sample
Measurement Model.............................................................................................................................. 139
Table 24 Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Millennial Sample ................................................................. 140

Table 25 Fit Indices for Boomer and Millennial Three-Factor Measurement Models .. 140

Table 26 Tests of Measurement Invariance for the Three-Factor Correlated Model .... 144

Table 27 Latent Mean Differences between Boomers and Millennials for the Three-Factor Scalar Invariance Model .......................................................... 145

Table 28 Fit Indices for Original and Respecified Single-Factor Measurement Models 147

Table 29 Pattern (P) and Structure (S) Coefficients for the Pooled Sample Single-Factor Measurement Model ........................................................................ 148

Table 30 Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Single-Factor Pooled Sample Measurement Model ........... 148

Table 31 Fit Indices for the Single-Factor Measurement Model for the Boomer Sample ........................................................................................................ 150

Table 32 Pattern (P) and Structure (S) Coefficients for the Boomer Sample Single-Factor Measurement Model ........................................................................ 151

Table 33 Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Single-Factor Boomer Sample Measurement Model .......... 151

Table 34 Fit Indices for the Single-Factor Measurement Model for the Millennial Sample ........................................................................................................ 152

Table 35 Pattern (P) and Structure (S) Coefficients for the Millennial Sample Single-Factor Measurement Model ........................................................................ 153
Table 36 Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Single-Factor Millennial Sample Measurement Model ........ 153

Table 37 Fit Indices for Boomer and Millennial Single-Factor Measurement Models . 154

Table 38 Tests of Measurement Invariance for the Single-Factor Correlated Model ... 156

Table 39 Latent Mean Difference between Boomers and Millennials for the Single-Factor Scalar Invariance Model................................................................. 157

Table 40 Descriptive Statistics for the Pooled Sample ................................................ 158

Table 41 Descriptive Statistics for the Boomer Sample................................................. 158

Table 42 Descriptive Statistics for the Millennial Sample............................................. 159

Table 43 Common Method Variance Testing with the Pooled Sample ....................... 161

Table 44 Common Method Variance Testing with the Boomer Sample ..................... 162

Table 45 Common Method Variance Testing with the Millennial Sample .................. 162

Table 46 Results of Predicted Hypotheses .................................................................. 164
List of Figures

Figure 1. Configural Invariance for the Three-Factor Model of UWES-9 ..................... 110
Figure 2. Metric Invariance for the Three-Factor Model of UWES-9 ............................ 111
Figure 3. Scalar Invariance for the Three-Factor Model of UWES-9 ............................ 112
Figure 4. Three-Factor Measurement Model (MF) .................................................... 135
Figure 5. Single-Factor Measurement Model (MFS) ................................................... 149
Abstract

TESTING THE MEASUREMENT INVARiance OF DATA FROM THE UTRECHT WORK ENGAGEMENT SCALE BY GENERATIONAL COHORT FOR EMPLOYEES IN THE LEISURE AND HOSPITALITY INDUSTRY

Greggory Lee Keiffer

Dissertation Chair: Kim Nimon, Ph.D.

The University of Texas at Tyler

November 2017

Group comparison in social science research is a common and informative practice. The establishment of measurement invariance between groups is a statistical prerequisite before making group mean comparisons; however, researchers often do not include measurement invariance assessments before making group comparisons. The current study assessed the measurement invariance for data from the short version of the Utrecht Work Engagement Scale (UWES-9) for the generational cohort groups of Boomers and Millennials within the leisure and hospitality industry. Group equivalency was determined by utilizing propensity score matching before conducting measurement invariance assessments through the confirmatory factor analysis technique. Measurement invariance results between groups were discussed for both the three-factor and the single-factor models of work engagement. Latent mean analysis was conducted for each model, and latent mean differences are reported respective to each analysis. The study included an assessment of common method variance using the comprehensive confirmatory factor analysis latent marker technique. The study’s results confirmed measurement invariance
for data from UWES-9 between Boomers and Millennials for the equivalent samples, and the study suggested that Boomers are not more engaged at work than Millennials. This finding was contradictory to much of the leisure and hospitality literature where Boomers are often cited as the more engaged generational cohort. Implications to theory, research, and practice were discussed.

*Key words:* measurement invariance, work engagement, generational cohort, propensity score matching
Chapter 1 - Introduction

Background to the Problem

The workforce of today is dynamic and broad when considering the expanse of generational cohorts. The force of the Millennial generation (Millennials) surge is being felt across industries as more Millennials enter the workforce. The Millennials represented 33% of the workforce as of 2015, and they have surpassed Generation X (GenXers) as the largest labor force (Fry, 2015). Similarly, the Boomer generation (Boomers) is tending to work beyond traditional retirement age (Fry, 2015). This trend has been in place for several years and is continuing as Boomers remain in the workforce due to the economic factors created by the Great Recession and continued need for organizationally provided health benefits and income (Cetron, DeMicco, & Davies, 2006; Fry, 2015).

The United States is moving further away from a manufacturing centered economy and increasingly towards a service based economy (Seaborn & Fels, 2015). Employees must adapt and learn new skills to maintain pace with the contemporary nature of jobs (Torraco, 2005), and the dynamic shift away from a manufacturing centered economy in the United States has, in part, allowed workers to age and remain productive in the workforce for longer (Fry, 2015). Because Boomers are able to remain productive for longer, the workforce employed within the modern service economy is now more generationally diverse compared to previous workforces employed during the manufacturing centered economy era (Fry, 2015). The age distance between the youngest Millennial and oldest Boomer is growing due to these factors.
Research indicates that employees in the different generational cohorts vary in their views, values, and level of employee engagement (Chen & Choi, 2008; Lyons & Kuron, 2013; Park & Gursoy, 2012; Twenge & Campbell, 2008). As work is more complex and more reliant on generationally diverse workforces, organizations are continuously challenged to deal with the confluence of factors that affect organizational outcomes, productivity, and employee engagement (Bakker & Demerouti, 2014; Kim, Shin, & Swanger, 2009; Torraco, 2005). Organizations must address and adapt to workers’ needs that impact employee engagement and account for their generational differences (Dale, 2014; Hoole & Bonnema, 2015; Lyons & Kuron, 2012; Park & Gursoy, 2012).

Meta-analytic research has reported that while there are many empirical claims of generational cohort differences, the “generally small effect sizes contradict such assertions” (Costanza, Badger, Fraser, Severt, & Gade, 2012, p. 387). When considering these various studies and the findings by Costanza and colleagues, the small effect sizes may be a result of low reliability, and as Kline suggested, “poor reliability reduces statistical power ... and … it also generally reduces effect sizes” (2016, p. 92). Chapter 2 will explore this relationship of low score reliability to low effect sizes. Considering the factor of Boomers working later into their lives and the volume of Millennials currently in and those who are still entering the workforce, more recent meta-analytic research focusing on work-attitudes reported that “differences may be growing as generations proceed through their respective life cycles” (Lyons & Kuron, 2013, p. 145). Common in the meta-analytic literature is the call for more methodologically sound and rigorous research on generational cohort differences, especially “time-lag designs, along with
well-constructed cross-sectional studies and the use of validated measures” (Lyons & Kuron, 2013, p. 145). Organizations may “assume that the Millennials will be just like those who have gone before them”, but the reality is different (Lancaster & Stillman, 2002, p. 207). It is therefore imperative for organizational “leaders to understand generational differences and the potential bases for conflict they create” (Lyons & Kuron, 2013, p. 149-150).

As of 2017, the leisure and hospitality supersector in the United States employed approximately 15.9 million employees (United States Department of Labor, Bureau of Labor Statistics [BLS], 2017). The BLS groups several sub-sectors into the leisure and hospitality supersector including lodging, travel, tourism, food and beverage service, gaming, and entertainment. This core industry to the U.S. economy is dealing with many of the same challenges common to other industries. Yet, a unique factor for the leisure and hospitality industry is its dependence on employees to fulfill key functions and interact directly and indirectly with customers (Solnet & Hood, 2008). A leisure and hospitality organization’s employees and human resource functions are paramount to operational success and competitive advantage (Voola, Carlson, & West, 2004), and issues for human resources are often cited as the central focus of concern for leisure and hospitality organizations (Enz, 2004). Engaged employees are of high importance and a strong challenge for an industry that is centrally dependent on its workforce (Schneider, Macey, Barbera, & Martin, 2009; Slatten & Mehmetoglu, 2011; Solnet, Kralj, & Kandampully, 2012). The industry is also experiencing a change in its workforce where it is becoming more dependent on the Boomer generation to fulfill workforce needs (Cetron, DeMicco, & Davies, 2006; Solnet & Hood, 2008). Empirical research in the
leisure and hospitality industry has reported that generational cohort differences between Boomers and Millennials can lead to increased turnover and lower profitability (Chi, Maier, & Gursoy, 2013; Gursoy, Chi, & Karadag, 2013).

Customers of leisure and hospitality organizations are demanding new types of services, such as mobile applications, as part of their customer experience (Deloitte, 2017). For example, Virgin’s mobile application “Lucy” allows guests to control much of their hotel experience from their mobile device, even to the degree of electronically interacting and texting hotel staff and other guests (Virgin, 2017). These customer demands are forcing leisure and hospitality organizations to respond with new technologies and subsequently add new work demands for all generations of employees. Employees from each generation must respond and adapt as their organizations address the “next wave of mobile transformation” (Deloitte, 2017, p. 8). Even so, a paradox remains that customers often still value face to face interactions with hotel employees (Deloitte, 2017). How are leisure and hospitality organizations to address this bevy of demands and keep employees engaged all while controlling for a population of workers that are more generationally divergent and diverse than ever? The leisure and hospitality industry will “benefit from an improved understanding of, and capability to plan for, underlying tensions of intergenerational conflict” caused by the evolution of work (Solnet et al., 2012, p. 37).

In an effort to address the organizational dilemmas and question outlined above, organizations often invest into consultation services, research, intervention programs, and treatments geared to positively influence employee attitudes, such as engagement (Bakker & Demerouti, 2014; Costanza et al., 2012). Engaged employees are linked to
numerous positive organizational outcomes including organizational citizenship behaviors (Dalal, Baysinger, Brummel, & LeBreton, 2012), performance (Crawford, LePine, & Rich, 2010), retention (Schaufeli, Bakker, & Van Rhenen, 2009), and profits for the organization (Harter, Schmidt, & Hayes, 2002). Human resource development (HRD) academic literature on employee engagement is broad and contains several competing definitions and conceptualizations of the phenomenon (Shuck & Wollard, 2010). Regardless of the definition or conceptualization, common among the research is that employee engagement is positive for the organization and employee (Harter et al., 2002; Shuck & Wollard, 2010).

Researchers and organizations often utilize the extensive burnout antithesis domain of engagement research to address organizational HRD questions (Bakker & Demerouti, 2014; Shuck & Wollard, 2010). Rooted in seminal research by Maslach (1982), burnout was originally characterized as a “syndrome of emotional exhaustion, depersonalization, and reduced personal accomplishment” among workers (Demerouti, Bakker, Nachreiner, & Schaufeli, 2001, p. 499). To assess burnout, the Maslach Burnout Inventory (MBI; Maslach & Jackson, 1986; Maslach, Jackson, & Leiter, 1996) incorporated Maslach’s (1982) identified components of burnout (i.e., emotional exhaustion, depersonalization, and reduced personal accomplishment). In order to generalize MBI beyond the original design that was to be used solely with human services populations (i.e., care-givers and health care providers), the Maslach Burnout Inventory-General Survey (MBI-GS) was created and modified the burnout dimensions to exhaustion, cynicism, and reduced professional efficacy (Schaufeli, Leiter, Maslach, & Jackson, 1996).
To address the psychometric limitations of the MBI-GS where each subscale is phrased in the same direction, the Job Demands-Resources (JDR) model was proposed and contained two underpinning processes (Demerouti et al., 2001). First, a stress process with the presence of job demands leads to exhaustion. Second, a motivational process with the lack of job resources leads to disengagement from work (Demerouti et al., 2001). The Oldenburg Burnout Inventory (OLBI) was used within the JDR model to alternatively measure burnout through the dimensions of exhaustion and disengagement from work (Demerouti et al., 2001; Demerouti & Nachreiner, 1998). The OLBI was considered psychometrically more sound compared to MBI as it contained both negatively and positively worded items in the subscales (Demerouti et al., 2001).

As an extension to the original JDR conceptualization (Demerouti et al., 2001), Schaufeli and Bakker (2004) proposed an updated model that positioned the availability of job resources as predictors of work engagement. The Utrecht Work Engagement Scale (UWES; Schaufeli, Salanova, Gonzalez-Roma, & Bakker, 2002) was used to measure the engagement construct within the extended JDR model (Schaufeli & Bakker, 2004). The research represented a pivotal point for the burnout domain to that of the burnout antithesis domain (Schaufeli & Bakker, 2004; Shuck & Wollard, 2010). Schaufeli and Bakker (2004) cited the shift in traditional psychology (i.e., negative psychology) to positive psychology and the shift in burnout research to the positive state of engagement as impetus for their research. Burnout and engagement were seen as independent constructs, not “each other’s complements”, and required independent instruments of measurement (Schaufeli & Bakker, 2004, p. 294). Burnout and engagement were
independent and yet negatively related “because of their antithetical nature” (Schaufeli & Bakker, 2004, p. 294).

The original UWES instrument (UWES-17; Schaufeli et al., 2002) contained 17 questions, and a shortened version (UWES-9) was soon developed to contain nine questions that were a subset of the original 17 (Schaufeli, Bakker, & Salanova, 2006). Both conceptualizations of UWES contain vigor, dedication, and absorption as the three dimensions of engagement. Vigor is the opposite of burnout’s exhaustion, and dedication is the opposite of burnout’s cynicism. Absorption is not the direct opposite of professional efficacy, “rather they are conceptually distinct aspects that are not the end points of some underlying continuum” (Schaufeli et al., 2002, p. 74). Given the pervasiveness of JDR research, positioning of UWES within the JDR model to measure work engagement, and the free availability of the UWES scales through Wilmer Schaufeli’s website, the UWES-17 and UWES-9 scales are the most frequently utilized scales within the body of employee engagement research literature (Saks & Gruman, 2014).

Park and Gursoy (2012), for example, utilized the UWES-9 scale in their study on U.S. hotel employees to investigate the differences in work engagement across Boomers, GenXers, and Millennials. The study reported that Boomers were “likely to be more dedicated to, engrossed in, and even vigorous at work” compared to Millennials (Park & Gursoy, 2012, p. 1198). This study and its generational findings have influenced subsequent work engagement research in the leisure and hospitality industry. The Park and Gursoy (2012) study preceded and influenced the following studies related to generational differences (Dimitriou & Blum, 2015; Gursoy, Chi, & Karadag, 2013; Kim,
Kim, Hand, & Holland, 2016), psychological capital (Paek, Schuckert, Kim, & Lee, 2016), and turnover (Brown, Thomas, & Bosselman, 2015; Chi, Maier, & Gursoy, 2013; Lu & Gursoy, 2016).

As a second example, Hoole and Bonnema (2015) utilized the UWES-9 scale in a work engagement and employee well-being study for professional employees in a South African setting. Reported results indicated a statistically significant difference for work engagement between Boomers and the younger generations (i.e., GenXers and Millennials), but there was not a statistically significant difference between GenXers and Millennials (Hoole & Bonnema, 2015).

Within the limitations section of the second example study, the authors discussed a central caveat of social science research; that is, “one of the downfalls of using well-established instruments is that one sometimes assumes that they remain relevant over time” (Hoole & Bonnema, 2015, p. 9). Respondents from different generational cohorts may “simply have a different understanding or interpretation of the constructs being measured” (Hoole & Bonnema, 2015, p. 9). The UWES instrument is well-established in research (Hoole & Bonnema, 2015; Saks & Gruman, 2014), and the pervasive use of UWES has emboldened many social science researchers. Many researchers may assume that well-established instruments will measure the same construct across their groups of interest (Thompson, 2004).

Human resource development researchers must apply appropriate and rigorous statistical methodology (Reio, 2010), including measurement invariance assessment, since “assessing measurement invariance is vital for meaningful between-group comparisons in social science research” (Nimon & Reio, 2011, p. 210). An extensive
measurement invariance literature review by Vandenberg and Lance synthesized the essential nature of measurement invariance and its importance to empirical research as “(a) knowingly or unknowingly, researchers invoke assumptions about measurement equivalence in conducting tests of substantive hypotheses; (b) although rarely tested, these assumptions are routinely and straightforwardly testable as extensions to the basic CFA framework; (c) if not tested, violations of measurement equivalence assumptions are as threatening to substantive interpretations as is an inability to demonstrate reliability and validity” (2000, p. 6).

**Statement of the Problem**

Employees are an organization’s critical path for achieving organizational outcomes, and engaged employees tend to increase the likelihood of obtaining these positive outcomes (Gilley & Gilley, 2000). Practical application of work engagement research often includes making group comparisons of employees by generational cohort (Costanza et al., 2012; Lyons & Kuron, 2013; Twenge, 2010). Comparing differences across groups and the associated impact to employees and organizations by those group differences is an essential component of social science research (Chen, 2007). In the meta-analysis performed by Costanza et al., the meta-analytic findings suggest that while “generational differences do exist in work-related outcomes” (2012, p. 391), the empirical findings are limited and inconsistent in statistically and practically significant differences by generational cohort. Conversely, statistically significant differences between Boomer and Millennial generational cohorts for work engagement are reported in the body of literature for individual studies within the leisure and hospitality industry, supporting the need to further study these two cohorts in the present study (e.g., Gursoy
et al., 2013; Hoole & Bonnema, 2015; Kralj & Solnet, 2011, Park & Gursoy, 2012; Solnet & Hood, 2008; Solnet, Kralj, & Kandampully, 2012). Empirical research to date is inconclusive and inconsistent as to whether differences exist by generational cohort for work-related outcomes.

While research has been completed to date within the leisure and hospitality industry that explores engagement by generational cohort, prerequisite statistical methodology is often lacking within these generational group comparisons studies (Lyons & Kuron, 2013). Few studies in the body of engagement literature consider measurement invariance by generational cohort before proceeding to make group comparisons. Hoole and Bonnema (2015) and Martins (2015) are notable exceptions. Since “the establishment of measurement invariance across groups is a logical prerequisite to conducting substantive cross-group comparisons” (Vandenberg & Lance, 2000, p. 4), it is essential to understand whether generational cohorts interpret and respond to engagement measures, such as the pervasive UWES-9 measure, the same or not. As part of the methodological design of their studies, “HRD researchers should include measurement invariance assessment” (Nimon & Reio, 2011, p. 210). Only once measurement invariance is established can the differences by group be accurately interpreted; otherwise, the potential for a “loss of statistical conclusion validity” is a substantial concern for researchers (Nimon & Reio, 2011, p. 209). Similarly, impacts to theory and practice are noteworthy concerns as conclusions may be drawn by the researcher based upon an inconclusive statistical foundation (Meredith, 1993; Vandenberg & Lance, 2000).
**Purpose of the Study**

The purpose of the study was to empirically assess measurement invariance by generational cohort using data from the short version of the Utrecht Work Engagement Scale (UWES-9). The population under study was currently employed leisure and hospitality professionals in the United States. Generational cohorts included in the study were Boomers and Millennials. Using confirmatory factor analysis, configural, metric and scalar invariance assessments were tested and reported by generational cohort. Upon assessment of measurement invariance between the Boomer and Millennial groups, latent mean analysis was conducted to determine the latent mean difference of work engagement factors between groups.

**Theoretical and Conceptual Underpinnings of the Study**

Three major theories underpinned and supported the study for three major phenomena: generational cohort, employee engagement, and measurement quality. Strauss and Howe (1991) articulated generational cohort theory. The general tenet of their theory is that social cycles are repetitive and complete a cycle with every four generations (Strauss & Howe, 1991). Strauss and Howe also offered a definition of a generation as “a cohort-group whose length approximates the span of a phase of life and whose boundaries are fixed by peer personality” (1991, p. 429). According to Coomes and DeBard (2004), understanding the differences of each generational cohort’s peer personalities allows for organizations and managers within those organizations to create more effective practices and policies. Generational cohort theory in the workforce “proves particularly useful when people attempt to understand and convey perceived differences between older and younger contemporaries and the social, cultural, and
especially technological changes affecting their lives” (Foster, 2013, p. 212). Three major generations are currently active in the overall workforce today: Baby Boomers, Generation X, and Millennials (Kaifi, Nafei, Khanfar, & Kaifi, 2012). Similarly, the leisure and hospitality industry mirrors the general workforce and houses Baby Boomers, GenXers, and Millennials (Park & Gursoy, 2012; Solnet & Hood, 2008). The newest generation, the Homeland generation or Generation Z, is only beginning to precipitate into the workforce in small numbers as they become of working age (Drago & Cunningham, 2006; Taylor, 2017). The present study used generational cohort theory as the foundation to substantiate that generational cohorts are definable and distinct groups of workers and that those defined groups may have differences related to work life perceptions.

The theory supporting the engagement portion of this study was the multidimensional theory of burnout (Maslach, 1998). Burnout within the theory is defined as “an individual stress experience embedded in a context of complex social relationships” and contains the three core components of “emotional exhaustion, depersonalization, and reduced personal accomplishment” (Maslach, 1998, p. 69). Prior burnout research (Maslach & Jackson, 1981; Maslach, 1982) gave seed to the multidimensional theory of burnout (Maslach, 1998). The Maslach Burnout Inventory (MBI) was developed through psychometric research to measure the three dimensions of burnout (Maslach & Jackson, 1981; Maslach, 1998). A version of the MBI was developed for generalized use as the Maslach Burnout Inventory-General Survey (Schaufeli et al., 1996) and contained the burnout dimensions reclassified to exhaustion, cynicism, and reduced professional efficacy.
Two paths of related burnout antithesis research developed to address the phenomenon of engagement (Maslach, Schaufeli, & Leiter, 2001). The first path viewed engagement as characterized by “energy, involvement, and efficacy – the direct opposites of the three burnout dimensions” and was “assessed by the opposite pattern of scores on the three MBI dimensions” (Maslach et al., 2001, p. 416). The second path led to the development of UWES-17 (Schaufeli et al., 2002). Engagement was considered the positive antithesis of burnout, but engagement was distinct from burnout, requiring its own operationalization (Maslach et al., 2001). In this distinct antithesis conceptualization, engagement was defined as “a persistent and pervasive affective-cognitive state that is not focused on any particular object, event, individual, or behavior” and is characterized by vigor, dedication, and absorption (Schaufeli & Bakker, 2004, p. 295). It is through this definition of engagement, its operationalization through the UWES-17 instrument, and the foundation on the multidimensional theory of burnout that a large volume of academic literature and research has been developed (Saks & Gruman, 2014; Shuck & Wollard, 2010).

The final theory supporting the study was classical test theory (Lord & Novick, 1968). Measurement quality (i.e., the reliability and validity) of observed variables is entrenched in classical test theory (Vandenberg & Lance, 2000). Classical test theory defines that observed scores are composed of both true and error scores where the true and error scores are uncorrelated (Lord & Novick, 1968). The assumption of uncorrelation between the true and error scores allows for the “decomposition of observed score variance into true and error score components”, and the subsequent assumption that “variance in individuals’ observed scores reflects variance contributed by true individual
differences in the measured trait and variance contributed by nonsystematic measurement error” (Vandenberg & Lance, 2000, p. 6). As an extension of classical test theory, measurement invariance is concerned with measurement properties across populations or groups. Research that compares groups is common within the social sciences (e.g., generational cohorts), and it is imperative that invariance between the groups of interest is established before the evaluation of research hypotheses (Meredith, 1993; Nimon & Reio, 2011; Vandenberg & Lance, 2000).

**Research Hypotheses**

A total of four research hypotheses were tested in this study. Within the context of previous work engagement research that considers differences between generational cohorts, studies consistently report mean differences between Boomer and Millennial cohorts by work engagement, albeit with variation by the cohorts and context under consideration (e.g., Bano, Vyas, & Rohini, 2015; Hoole & Bonnema, 2015; Martins & Ledimo, 2016; Park & Gursoy, 2012).

These prior studies failed to evaluate the measurement invariance between generational cohorts before making comparisons between groups, a prerequisite before making such comparisons (Meredith, 1993; Vandenberg & Lance, 2000). Traditional measurement invariance assessment is classified through a hierarchy of four levels of increasingly stronger measurement invariance: configural invariance, metric (weak) invariance, scalar (strong) invariance, and strict invariance (Meredith, 1993; Vandenberg & Lance, 2000). Each assessment level is “considered a necessary but insufficient condition for the next higher level” (Nimon & Reio, 2011, p. 205). Vandenberg and Lance (2000) through an extensive measurement invariance literature review recommend
establishing configural, metric, and scalar invariance before making group comparisons. The Vandenberg and Lance (2000) recommendations were supported in research by Cheung and Rensvold that evaluated the goodness of fit indexes for measurement invariance testing (2002). Strict invariance was determined to be excessively rigorous and related to construct level variance (Byrne, 2001; Cheung & Rensvold, 2002; Teo, Lee, Chai, & Wong, 2009). When utilizing the CFA statistical analysis approach, measurement invariance can be determined by evaluating model fit indices, change in chi-squared tests ($\chi^2$), and change in CFI (CFI) tests for the hierarchical models (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000).

Stemming from the multidimensional theory and three dimensions of burnout, the original development of UWES-17 (Schaufeli et al., 2002) and subsequent shortened UWES-9 (Schaufeli et al., 2006) instruments defined work engagement as containing the three dimensions of vigor, dedication, and absorption. The literature holds that researchers have used UWES-9 as both a three-factor model (vigor, dedication, and absorption) and as a total nine-item scored single indicator to avoid issues with multicollinearity (Schaufeli et al., 2006). The literature review conducted by the researcher confirms the varied use (i.e., three-factor and single-factor) of UWES-9 scores to measure work engagement. Given the varied use of UWES-9 in the literature, the researcher evaluated the measurement invariance between Boomers and Millennials in both the three-factor and single-factor models of work engagement.

In order to assess the measurement invariance between Boomer and Millennial generational cohorts using a three-factor and single-factor UWES-9 models, three hypotheses were proposed. Configural invariance means respondents use the same
conceptual framework to respond to scale items, metric invariance means respondents attribute the same meaning to the latent construct, and scalar invariance means that the meaning of the construct and level of underlying items are equal (Cheung & Rensvold, 2002; Van de Schoot et al., 2012). Support for the hypotheses was provided in Chapter 2.

H1: Evaluation of model fit indices will result in good model fit (i.e., TLI ≥ .95, CFI ≥ .95, RMSEA ≤ .10, and SRMR ≤ .08, Kline, 2016) and support for configural invariance for (a) the three-factor model and (b) the single-factor model.

H2: Evaluation of model fit indices, $\chi^2$, and CFI will result in good model fit and support for metric invariance for (a) the three-factor model and (b) the single-factor model.

H3: Evaluation of model fit indices, $\chi^2$, and CFI will result in good model fit and support for scalar invariance for (a) the three-factor model and (b) the single-factor model.

In order to assess the latent means difference of work engagement between Boomer and Millennial generational cohorts, the following hypothesis was proposed:

H4: There will be statistically and practically significant latent mean differences for vigor, dedication, and absorption between the Boomer and Millennial groups for (a) the three-factor model and a statistically and practically significant latent mean difference for work engagement for (b) the single-factor model. The Boomers will be more engaged compared to the Millennials.
Overview of the Design of the Study

A quantitative research design approach was used for this study. The data were collected at a single time point for a cross-sectional view of the population. Response data were collected through online survey methodology utilizing Qualtrics® survey functionality for the survey design, deployment, and data collection. Participants were sourced from a pre-recruited Qualtrics panel. Qualtrics® e-mail services were utilized for communication to the panel. The targeted sample was those individuals living in the United States who work full-time in the leisure and hospitality industry and whose only difference was by their generational cohort association (i.e., Boomers or Millennials).

The participants were asked to complete the UWES-9 measure, a latent marker variable, the MBI-GS, demographics, and job characteristics. Demographics and job characteristics were used to determine the match of the sample to the leisure and hospitality population demographic profile. Respondents categorized themselves into generational cohort, and statistical analysis confirmed if the cohort groups were comparable by analyzing their associated demographics and job characteristics. Since the groups were found to not be comparable by demographic and job characteristic reviews, propensity score matching (Rubin, 1997) was employed to equate the cohort groups by their associated covariates. The confirmed and comparable generational cohort groups were used for hypotheses testing by Boomer and Millennial groups. Confirmatory factor analysis was used to test for measurement invariance. Configural, metric, and scalar invariance assessments were tested and reported between the generational cohorts of Boomers and Millennials, and latent mean differences for work engagement factors were assessed by generational cohort since scalar measurement invariance held.
Significance of the Study

The study had implications and significance for theory, research, and practice. When considering the theoretical implications of the study, generational cohort theory is commonly cited to provide foundational support for generational comparisons (Lyons & Kuron, 2013). Primarily, the theory was reinforced by elucidating a priori statistical assessments required before making group comparisons (i.e., measurement invariance). If researchers ignore the statistical prerequisites of measurement invariance assessment for group comparisons, invalid conclusions may result from subsequent analysis with negative impact to generational cohort theory (Nimon & Reio, 2011). Before the influx of the Homeland generation into the workforce, stronger and clearer theoretical understanding is needed now before the newest generation begins to impact the organizational landscape and current generations (Lyons & Kuron, 2013). The present study added to the theoretical body of literature for generational cohort theory within the contemporary work environment for the leisure and hospitality industry.

The study was significant to the field of HRD as it answered the calls for additional contextual generational cohort research (Costanza et al., 2012; Park & Gursoy, 2012; Lyons & Kuron, 2013), and it attempted to clarify the methodological procedures required for group comparison research. As noted by in the meta-analysis by Costanza et al., “there is a need for additional, scientifically sound, primary research on generational differences” and there “is a need for improved methodological approaches for studying generational differences” (2012, p. 390). In their meta-analytic discussion, Lyons and Kuron noted “the study of generational differences must progress toward a more mature stage” with “more methodological rigor and a greater consideration of context” (2013, p.
The present study addressed the contextual aspect by evaluating a sample of Boomers and Millennials from the omnipresent leisure and hospitality industry. The study also addressed the calls for stronger methodological rigor in HRD research (Reio, 2010) by systematically demonstrating the CFA technique for measurement invariance assessment. The two major facets of the present study independently addressed calls from the literature for more contextual generational cohort research with increased methodological rigor.

Practice recognized valuable insights from the study. Organizations generally focus investments on projects, programs, and employee interventions that will yield the greatest return or largest perceived outcome. Erroneous investment of those funds can be highly negative for organizations and employees. For those organizations whose working environments and workforce are evolving due to technological advances (Fry, 2015; Torraco, 2005) and generational workforce differences (Park & Gursoy, 2012), investments may be erroneously allocated if they are relying upon “organizational interventions designed to address such differences” when the supporting research “found mixed results … of generational differences” (Costanza et al., 2012, p. 387-388). Meta-analytic research (Costanza et al., 2012; Lyons & Kuron, 2013; Twenge, 2010) has reported that generational differences are not always empirically supported. The inconsistency of meta-analytical findings may cause concern for organizational leaders since empirical research has alternately reported that generational differences do exist in the workforce (Foster, 2013; Gursoy et al., 2013; Lester, Standifer, Schultz, & Windsor, 2012).
The inconclusive empirical landscape and reported perceptions of generational differences are roots for confusion for organizational leaders. Even with the confusion, “managers must recognize that generational differences are valid and important” (Lyons & Kuron, 2013, p. 149). The present study provided a rigorous empirical analysis of the generational differences between Boomers and Millennials for the construct of work engagement within the leisure and hospitality industry context. Organizations within the leisure and hospitality industry now have a solid grounding upon which to make contextual business decisions between Boomer and Millennials generational differences.

**Assumptions**

The researcher held two primary assumptions for the present study. First, the researcher assumed the respondents accurately reported their generational cohort within the demographic section of the survey. Since this demographic variable was used to categorize the respondents into generational cohort for subsequent measurement invariance assessment, it was critical respondents reported their generational cohort accurately. Second, an assumption was made that respondents to the surveys were truthful and diligent in their responses to the UWES-9 instrument. The researcher included survey design considerations that mitigated much of these concerns by insuring anonymity, short survey duration, and ease of use for the survey.

**Delimitations**

Five delimitations existed for the present study. First, two of the four generational cohorts who are of working age were included in the study. Only Boomers and Millennials were included in the data analyses and hypotheses testing; thus, the findings were not generalizable to the GenXer and Homeland generational cohorts.
Second, the study focused on a population of leisure and hospitality workers from the United States. Similar to the delimitation of the generational cohorts, the findings were not be generalizable to other countries beyond the United States. Third, only the UWES-9 (Schaufeli et al., 2006) instrument was used for the study. The UWES-17 (Schaufeli et al., 2002) instrument was not included, and the findings of the study were limited to the scope of the UWES-9 instrument. Fourth, the UWES-9 items were presented in the order as they appear in the UWES test manual and online test form. Contained on Wilmar Schaufeli’s webpage, a statement specifically outlines the UWES data should “adhere to the original answering format and sequential order of items” as part of the usage terms (Notice for potential users of the UWES and the DUWAS section, para. 4). Due to this statement on the webpage, the UWES items were not randomly presented in the survey of the current study. Last, the present study only considered measurement invariance assessments in the confirmatory factor analysis (CFA) framework. Other frameworks such as item response theory (Maurer, Raju, & Collins, 1998) or generalizability theory (Brennan, 1983) were not considered.

Definitions of Terms

The following terms and definitions were relevant to this proposal:

- **Work Engagement:** Work engagement is the “positive, fulfilling work-related state of mind that is characterized by vigor, dedication, and absorption” (Schaufeli et al., 2002, p. 74).

- **Generation:** A generation is defined as “a cohort-group whose length approximates the span of a phase of life and whose boundaries are fixed by peer personality” (Strauss & Howe, 1991, p. 429).
• Baby Boomers: Baby Boomers are individuals “born between 1946 and 1964” (Wieck et al., 2009, p. 170).

• Generation Xers: Generation Xers are individuals “born between 1965 and 1980” (Wieck et al., 2009, p. 170).

• Millennials: Millennials are individuals “born between 1981 and 2000” (Wieck et al., 2009, p. 170).

• Homelanders: Homelanders are individuals “born after 2001” (Codrington, 2008, p. 2).

• Measurement: Measurement in the context of this study can be defined as “the systematic assignment of numbers on variables to represent characteristics of persons, objects, or events” (Vandenberg & Lance, 2000, p. 4).

• Measurement Invariance: Measurement invariance “concerns the invariance of regression intercepts, factor loadings (regression slopes), and error/uniqueness variances” for groups of the same sample (Byrne, Shavelson, & Muthen, 1989, p. 456).

The positionality of employee engagement as an outcome, psychological state, or as a process has gained attention and discussion in HRD literature. The present study utilized the UWES-9 measurement instrument and provided a definitional grounding for work engagement (Schaufeli et al., 2002; Schaufeli et al., 2006). According to Shuck, Osam, Zigarmi, and Nimon (2017), the UWES-9 instrument and work engagement construct has been largely categorized as an outcome. Therefore, the present study positioned UWES-9 as an outcome, rather than as a psychological state or as a process.
Chapter Summary and Organization of the Dissertation

This dissertation was organized into five logical chapters. Chapter 1 of this dissertation included the introduction and background to the problem, statement of the problem, purpose of the study, theoretical underpinning, research hypotheses, overview of the design, significance of the study, assumptions, delimitations, definitions, and concludes with the organization of the proposal.

Chapter 2 included the main literature review for generational cohort, the leisure and hospitality industry, employee engagement, psychometrics of UWES, and hypotheses support. Chapter 3 contained the methodology the study will employ, including the purpose of the study, measurement invariance overview, design of the study, research hypotheses, population and sample, instrumentation for the survey, survey design, data collection procedures, data analysis procedures (i.e., data cleaning, group cohort comparisons and sample representativeness, and statistical assumptions), hypotheses testing, descriptive statistics, common method variance, and limitations. The third chapter concluded with a summary. Chapter 4 reported the results of the study, and Chapter 5 provided the discussion of results, implications, limitations, and paths for future research.
Chapter 2 - Literature Review

Introduction

This chapter reviewed the literature domains relevant to generational cohort and employee engagement in the context of the leisure and hospitality industry. The review was organized into six sections. The first section provided an overview of the leisure and hospitality supersector. The second section reviewed literature relevant to generational cohorts, and the third section reviewed employee engagement, including the domains and measurements of engagement. The fourth section reviewed the psychometrics of the UWES scale, and the fifth section provided support for the four hypotheses. The final section presented a summary of the chapter.

To conduct this literature review, The University of Texas at Tyler Robert R. Muntz Library computer system was utilized as the primary resource. The following databases were searched: Business Source Complete, Emerald, Sage, PsycINFO, Ebscohost, and Science Direct. Google® Scholar was utilized as a secondary resource. The following search terms were used: employee engagement, work engagement, engagement, job engagement, personal engagement, engagement measures, engagement scales, engagement instruments, generations, generational, generational cohort, cross-generation, cross-generational, Millennial, Baby Boomer, Generation Y, leisure and hospitality, hospitality, tourism, hotel management, measurement invariance, measurement equivalence, Utrecht Work Engagement Scale, UWES, UWES-17, and UWES-9. Combinations of the terms were also used. A manual search was conducted within the following HRD and leisure and hospitality field relevant journals: Human

The Leisure and Hospitality Supersector

The leisure and hospitality industry represents a large component of the U.S. economy. Deemed the leisure and hospitality supersector by the Bureau of Labor Statistics, it is part of the service providing industries and contains two major subsectors: the arts, entertainment, and recreation subsector and the accommodation and food services subsector (BLS, 2017). Numerous types of businesses are found within the supersector including accommodations, hotel, resort, lodging, travel, tourism, cruise line, entertainment venues, sporting arenas, recreation facilities, museums, restaurants, catering services, and other variations of food and beverage establishments.

Supersector Description

As of July 2017, the industry employed 15.998 million total workers. Production and nonsupervisory positions accounted for 14.058 million workers or 87.9%. The weekly hours worked for all employees were 26.1 hours and 24.8 hours for production and nonsupervisory positions. Demographically, women represented the majority of the workforce at 50.6%, and men accounted for 49.4%. Race was reported as 12.6% Black or African American, 6.5% Asian, 22.9% Hispanic or Latino, and 58% to a bucket
category to include Caucasian and other non-reported races (BLS, 2017). Notable variations do exist within the subsectors where the accommodation and food services subsector contained more women (52.3%) and more minorities (46.2%) and the arts, entertainment, and recreation subsector contained fewer women (45.2%) and fewer minorities (27.9%) compared to the general leisure and hospitality supersector. The accommodation category of the subsector contained the highest percentage of women (55.5%) and minorities (51.1%) (BLS, 2017).

Leisure and hospitality industry employees must often interact directly with the end customer to fulfill operational functions (Solnet & Hood, 2008). The heavy dependence of employee to customer interactions and associated work stressors of those interactions lends to the frequent study of burnout and engagement within the industry (Solnet & Hood, 2008). Engaged employees are of special importance to leisure and hospitality organizations since operational functions and success are so critically dependent on the employees (Schneider et al., 2009; Slatten & Mehmetoglu, 2011).

There is a wide diversity of literature on engagement within the leisure and hospitality industry. Empirical research on employee engagement is prevalent within the academic literature (e.g., Gursoy et al., 2013; Kralj & Solnet, 2011, Park & Gursoy, 2012; Solnet & Hood, 2008, Solnet et al., 2012), is found in practitioner reports (e.g., Deloitte, 2017; Quantum, 2017), and is found within organization specific reports (e.g., Hilton Hotels, 2016).

**Workforce Dynamics**

The industry is not immune to the generational factors influencing the general U.S. economy and businesses. Many employees find their first employment
opportunities in the leisure and hospitality industry due to the volume of positions and the availability of entry level jobs, and the industry employs a large number workers in the Millennial generation. In tandem with the changing U.S. economic and employment landscape, the industry is now more dependent on the Baby Boomer generation to fill positions (Cetron et al., 2006; Solnet & Hood, 2008).

The hospitality industry frequently changes branding, technologies, and operational concepts to maintain pace with demands of customers (Cetron et al., 2006; Deloitte, 2017). Hilton Hotels is a prime example of an organization that is creating new brands to answer market segment demands, such as the new Tru by Hilton™ brand that is focused on vibrancy and social engagement (Hilton Hotels, 2017). Hilton Hotels currently controls 14 brands and a demographically diverse workforce. According to the 2016 employee survey published on Hilton’s homepage, 53% are women, 69% are minorities, and employees are roughly split into thirds across the generational cohorts of Baby Boomers, Generation X, and Millennials. The Silent generation represents less than 2% of the Hilton employee population. Most are full-time workers (89%), but almost 35% have less than 2 years of tenure and another 20% have between 2 and 5 years of tenure (Hilton Hotels, 2016). It is this workforce diversity within organizations and the changing economic reality that has driven the need for academic generational cohort research to inform practitioners within the leisure and hospitality industry (Gursoy et al., 2013; Park & Gursoy, 2012; Solnet & Hood, 2008).

**Generational Cohort**

Generational research and theory has arisen from two distinct conceptualizations, the social perspective and the cohort perspective (Lyons & Kuron, 2013). The social
perspective purports that social groups form into generations through the course of historical events (Mannheim, 1952). Conversely, the cohort perspective registers groups of people to a generation based upon a defined time period (Laufer & Bengtson, 1974; Strauss & Howe, 1992; Ryder, 1965). The present study presented a background on the social perspective and focused on the cohort perspective. The cohort perspective is prevalent in empirical literature as researchers have concentrated on examining the mean differences among cohorts (Foster, 2013).

**History of Generations**

The social perspective of generations was pioneered by sociologist Karl Mannheim. Mannheim’s social perspective posited that generations form around historical events and contexts with the generation experiencing an “inborn way of experiencing the life and the world” (1952, p. 283). Thus, defining a generation was loose and without specific boundaries. Generations evolve and switch when there is a change in social patterns, lifestyles, and values (Eyerman & Turner, 1998; Laufer & Bengtson, 1974). The social perspective is less concerned about attitudes and behaviors of individuals and is more concerned with the interaction of generations as a social change mechanism, and new generations are the drivers of social change (Lyons & Kuron, 2013; Mannheim, 1952).

The cohort perspective offers a benefit to empirical research in that it specifically defines the generations by birth year ranges (Laufer & Bengtson, 1974; Ryder, 1965). A generational cohort is defined as a group of individuals who “experience the same events within the same time interval” (Ryder, 1965, p. 845). Generational cohorts are considered to be similar enough to carry the same “observable commonalities that are
relatively fixed and measureable by mean scores on attitudinal and behavioral variables” (Lyons & Kuron, 2013, p. 141).

An individual’s association with a generation defines the basis by which they view and understand life experiences, and the generation’s specific historical context is constant and shapes how that generation’s course of life unfolds (Eyerman & Turner, 1998; Mannheim, 1952; Ryder, 1965). The aging process is defined within the generational context and influences how the individuals perceive and respond to stimuli. Individuals in different generations will respond differently to historical events based upon their life cycle stage and generational cohort association (Lyons & Kuron, 2013; Ryder, 1965).

**Generational Cohorts – An American Concept?**

The theory posited by Mannheim (1952) suggests that generations are defined by their location and context in the timeline for social and historic events. Therefore, it would be incorrect to superimpose an American generational cohort context into a non-American context such as another country without considering the similar and dissimilar natures of the social and historic events between the countries (Lyons & Kuron, 2013). The history of generations provided by Strauss and Howe (1991), the most common typology of generations (Costanza et al., 2012), is through an American lens. Dimitriou and Blum suggested that non-American generational cohort research, especially for Millennials, is currently lacking and is needed “in other countries for comparative purposes” (Dimitriou & Blum, 2015, p. 63).

Development of UWES occurred near the turn of the century in Europe and did not consider generational cohorts as part of the measurement invariance assessments
during the developmental phases of the instrument (Schaufeli et al., 2002). Invariance was later tested across countries (Schaufeli, Martinez, et al., 2002; Schaufeli et al., 2006), racial groups (Storm & Rothmann, 2003), and occupational groups (Seppala et al., 2009). Invariance by generational cohort group may not have been a focus of European researchers since the generational cohort concept leans toward the American viewpoint (Strauss & Howe, 1991) and is only recently becoming more important outside of the United States (Dimitriou & Blum, 2015).

Generations in the Workforce

Strauss and Howe (1991) reviewed generations in the United States dating back over 400 years and helped to define the common cohort typology. Strauss and Howe’s definition of a generation was “a special cohort-group whose length approximately matches that of a basic phase of life, or about twenty-two years” (1991, p. 34). Their concept was that generations repeat in cycles of four, which approximates the current workforce of the Silent Generation, Baby Boomers, Generation Xers, and Millennials. The aforementioned labels were not necessarily coined by Strauss and Howe (1991).

The Silent Generation label appeared in *TIME* magazine in the early 1950s (TIME, 1951, November 5). The Baby Boomer label first appeared in the literature describing fertility rates by Westoff (1954). The label of Generation X was first found in Hamblett and Deverson’s (1965) book. The Millennial label was common in various literature sources near the turn of the century and Strauss and Howe (2000) discussed the generation in detail (Costanza et al., 2012).

Each generation has multiple labels depending on the source, author, and citation, including the new generation just now entering the workforce, the Homeland
(Homelanders) generation (Codrington, 2008; Drago & Cunningham, 2006). For the purpose of the present study, generational cohort labels will be Boomers, GenXers, and Millennials. The Silent generation and Homelanders will not be discussed as each represents a very small percentage of the working population as of 2017. For example, Hilton Hotels reported that the Silent generation represented just 1.31% of their workforce in a 2016 employee survey (Hilton Hotels, 2016).

Just as the labels for the generations lack consensus, so do the time period definitions for the generations. Generational cohort timespans are inconsistent and little cohesion is found among authors (Laufer & Bengtson, 1974; Cordeniz, 2002; Gursoy et al, 2012; and Yahr & Schimmel, 2013). Table 1 presents a summary of the various timespans applied to the generational definitions within selected empirical studies.

As can be observed in Table 1, the lack of consensus in the generational timespan definition is large, but a general theme is present. Boomers range from 1940-1964, GenXers range from 1961-1981, and Millennials range from 1977 to roughly 2000. Several authors (Parry & Urwin, 2011; Twenge, 2010) did not prescribe an end date to the Millennials. For the present study, the researcher followed the timespan definition by Wieck et al. (2009) that defines Boomers as 1946-1964, GenXers as 1965-1980, and Millennials as 1981-2000. This decision was supported in that one of the more influential empirical studies (Park & Gursoy, 2012) in the leisure and hospitality industry also used this same generational cohort timespan, and this study influenced later literature on generational research (Brown et al., 2015; Chi et al., 2013; Dimitriou & Blum, 2015; Gursoy et al., 2013; Kim et al., 2016; Lu & Gursoy, 2016; Paek et al., 2016).
Table 1

*Presentation of Generational Cohort Timespans by Author*

<table>
<thead>
<tr>
<th>Author</th>
<th>Boomers</th>
<th>GenXers</th>
<th>Millennials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bristow, Amyx, Castleberry, &amp; Cochran (2011)</td>
<td></td>
<td>1965-1980</td>
<td></td>
</tr>
<tr>
<td>Weston (2001)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* All authors did not report every generation.

**The Generations Described**

Baby Boomers came up through their early years in the post-World War II era and experienced a period of economic prosperity. They lived during a historic period in American history that included the assassinations of John F. Kennedy and Martin Luther King, the Vietnam War, Watergate, and the civil rights movement (Twenge et al., 2010).

Research has suggested the Boomers find work to be more central in their lives and
overall value work more compared to younger generations (Meriac et al., 2010; Twenge et al., 2010). Because Boomers value work and believe it pays off, they are loyal and organizationally committed, expecting commensurate rewards for their work (Smola & Sutton, 2002). A study by the Families and Work Institute (2006) suggested that Boomers are driven by rewards and workplace goals, a defined organizational structure, and value obtaining positions with higher levels of responsibilities.

The GenXers grew up during a period rife with a series of economic recessions and significant political events (e.g., end of the Cold War). Family relocations were common during their formative years due to high unemployment (Twenge et al., 2010). Due to these events, GenXers tend to be more individualistic and have a higher regard for their own career, rather than loyalty toward an organization (Kupperschmidt, 2000). Extrinsic rewards are motivators and the pursuit of new challenges and new careers are facilitators to achieve the rewards (Kupperschmidt, 2000). Even though GenXers are more focused on work advancement than Boomers, they have a lower work centrality than Boomers and value work-life balance (Smola & Sutton, 2002) and leisure time and activities (Twenge et al., 2010). GenXers also value less supervision in the workplace and require more autonomy compared to Boomers (Jurkiewicz, 2000).

Millennials have experienced a time of globalization, communication proliferation through the Internet, social networking, and economic prosperity (Park & Gursoy, 2012). They have been deemed the always connected generation (Burke, 2014). Akin to GenXers, Millennials value work-life balance and workplace autonomy more than Boomers (Smola & Sutton, 2002). Millennials have low work centrality and higher leisure centrality compared to the older generations (Twenge et al., 2010), but they
conversely have higher workplace pay and promotion expectations compared to Boomers and GenXers (Ng et al., 2010). Hill reported a large gap between Millennials’ workplace expectations and what they can actually achieve in the workplace (2002). Millennials value meaningful and intrinsically fulfilling work compared to older generations (Lancaster & Stillman, 2002). Compared to Boomers and GenXers, Millennials value team-based work and an unstructured workplace (Yahr & Schimmel, 2013), and they are receptive to sharing rewards with peers (Byrne, 2007).

Importance of Generational Cohort Research in Leisure and Hospitality

The leisure and hospitality industry is critically dependent on employees “to fulfill the industry’s basic functions, and the predominance of transactions in hospitality involve either direct or indirect employee-customer interaction” (Solnet & Hood, 2008, p. 59). There is a growing workforce demand in the leisure and hospitality industry, but the industry is faced with supply imbalances and changing demographics within the workforce (Solnet & Hood, 2008). With each progressive year, the workforce dynamics are evolving as more Millennials work at higher levels for an organization while Boomers often fulfill lower level positions (Solnet et al., 2012). As a further complication to the industry dynamics, it has been suggested that the work values of Millennials do not align with the “prevalent employment conditions in the hospitality industry” (Solnet & Kralj, 2010, p. 3). Solnet and Hood (2008) defined a research agenda to explore these topics and additional generational differences within the leisure and hospitality industry and offered the following six propositions to guide future research: (a) many Millennial descriptions are merely myths, stereotypes, or transitory states, (b) Millennial employees’ organizational commitment will be directly related to
the level of commitment they perceive the organization has in them personally, (c) for Millennials, organizational commitment and retention will have a less significant relationship in comparison to previous generational groups, (d) Millennials’ job satisfaction will be derived from intrinsic factors, where the opportunity to take ownership and responsibility for a variety of work tasks and meaningful projects with proper support, training, and development opportunities will be of high importance, (e) Millennials’ perception of supervisor support will directly influence their job satisfaction, and (f) leisure and hospitality organizations can improve their appeal to potential Millennial employees and communicate better with existing employees by harnessing the Millennials’ innate habit of social networking.

Several empirical studies are present in the literature base for the leisure and hospitality industry that evaluate the differences in various constructs by generational cohort. Examples in the section below and in Table 2 show the reported differences between Boomers and Millennials across the varied constructs of study, including work engagement. Chi et al. (2013) surveyed a U.S. hotel chain and collected a sample of 905 respondents representing $n = 677$ line-level employees and $n = 228$ manager level employees. The researchers investigated the employee perceptions of older and younger managers while controlling for tenure (Chi et al., 2013). Chi et al. reported “no significant differences across generational cohorts of line level employees’ perceptions of older and younger managers while controlling for employees’ tenure” (2013, p. 47). When reporting mean differences between Boomers and Millennials for the manager level sample and while controlling for tenure, Boomers were reported to have positive
perceptions of their older managers with a mean difference of $M = .48$ with a $p < .05$, and effect size of $d = .23$ (Chi et al., 2013).

Researchers have studied the differences in work values by generational cohort as well. Chen and Choi (2008) collected a sample of 398 managers and supervisors from U.S. hospitality organizations. A total of 15 work values were assessed and through principal component analysis, four dimensions were found: (a) comfort and security ($\alpha = .79$), (b) professional growth ($\alpha = .80$), (c) personal growth ($\alpha = .75$), and (d) work environment ($\alpha = .68$) (Chen & Choi, 2008). A one-way ANOVA followed and findings indicated that Boomers mean score for the personal growth dimension was significantly higher than Millennials ($F = 6.24$, $df = 2,345$) with a $p < .01$ and effect size of $d = .35$ (Chen & Choi, 2008). The findings also indicated that Millennials mean score for the work environment dimension was significantly higher than Boomers ($F = 5.87$, $df = 2,345$) with a $p < .01$ and effect size of $d = .34$ (Chen & Choi, 2008). No statistically significant differences were found for the comfort and security or professional growth dimensions (Chen & Choi, 2008).

A qualitative study of graduates from U.S. hospitality programs surveyed participants regarding their turnover intentions (Brown, Thomas, & Bosselman, 2015). The researchers estimated a 10% response rate and collected data from 107 graduates who were working in the hospitality industry and 39 who were not in the hospitality industry at the time of the survey. All of the respondents were Millennials, even though the study targeted Millennials and GenXers (Brown et al., 2015). Overall qualitative findings revealed that for those graduates who stayed in the hospitality industry after graduation did so because they “enjoyed working with people, enjoyed serving others,
and enjoyed the excitement of hospitality” (Brown et al., 2015, p. 136). For those who left the hospitality industry, they did so “because of the long hours and compensation” (Brown et al., 2015, p. 136).

A different qualitative study utilizing focus group methodology was conducted with employees of a U.S. based hotel chain (Gursoy et al., 2008). A total of 10 focus sessions were held in the northwestern region of the U.S. (Gursoy et al., 2008). Qualitative findings for Boomers were reported as follows: “Boomers live to work … and are willing to wait their turn for promotions and rewards, and are very loyal” (Gursoy et al., 2008, p. 448). Millennials were reported to “believe in collective action, with optimism for the future, and trust in central authority … like team work, showing a strong will to get things done with great spirit” (Gursoy et al., 2008, p. 448).

A quantitative study (Gursoy et al., 2013) was conducted as a continued research effort after the qualitative study (Gursoy et al., 2008) and utilized the same U.S. hotel chain, collecting a total of 717 survey responses from frontline and service contact employees. The survey assessed seven work values: (a) work centrality ($\alpha = .77$), (b) non-compliance ($\alpha = .71$), (c) technology challenge ($\alpha = .73$), (d) work-life balance ($\alpha = .63$), (e) leadership ($\alpha = .74$), (f) power ($\alpha = .70$), and (g) recognition ($\alpha = .81$) (Gursoy et al., 2013). A one-way ANOVA was conducted using Tukey’s HSD for post hoc pair-wise comparisons by generational cohort for each of the work value dimensions. Pair-wise comparisons for Boomers and Millennials resulted in statistically significant differences in all of the work value dimensions except for power at the $p < .05$ level (Gursoy et al., 2013).
Table 2
*Generational Cohort Empirical Research Within the Leisure and Hospitality*

<table>
<thead>
<tr>
<th>Authors</th>
<th>Study Type</th>
<th>Context</th>
<th>Generations (ns by cohort provided if available)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown, Thomas, &amp; Bosselman (2015)</td>
<td>Qualitative</td>
<td>L &amp; H</td>
<td>M(n=107)</td>
<td>Employees stayed in L&amp;H due to working with people, service to others, and excitement of the industry.</td>
</tr>
<tr>
<td>Chen &amp; Choi (2008)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>B(n=92), X(n=144), M(n=112)</td>
<td>Boomers higher than Millennials for Personal Growth with a mean difference of $M = .55$, $p &lt; .01$, $d = .44$. Millennials higher than Boomers for Work Environment with a mean difference of $M = .78p &lt; .01$, $d = .47$.</td>
</tr>
<tr>
<td>Chi, Maier, &amp; Gursoy (2013)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>B(n=237), X(n=248), M(n=192)</td>
<td>Line level employees were found to have no differences by generational cohort in perceptions of older or younger managers. For the manager sample, Boomers had more positive perceptions of their older managers ($p &lt; .05$, $d = .23$).</td>
</tr>
<tr>
<td>Gursoy, Maier, &amp; Chi (2008)</td>
<td>Qualitative</td>
<td>L &amp; H</td>
<td>B, X, M(varied participants from 10 focus sessions)</td>
<td>Focus groups. Boomers live to work, Millennials like teamwork.</td>
</tr>
<tr>
<td>Gursoy &amp; Chi (2013)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>B(n=257), X(n=260), M(n=200)</td>
<td>Boomers to Millennials pair-wise comparison ($M =$ mean differences): (a) work centrality $M = .27$, $d = .44$, (b) non-</td>
</tr>
<tr>
<td>Study</td>
<td>Type</td>
<td>Generations</td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Kralj &amp; Solnet (2011)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>Boomers/GenXers group more engaged than Millennials ($p &lt; .001, d = .45$).</td>
<td></td>
</tr>
<tr>
<td>Lu &amp; Gursoy (2013)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>In the moderating effect of generation in the MBI burnout context, Millennials have lower JS and higher turnover ($p &lt; .05, R^2 = .34$) when emotionally exhausted than Boomers.</td>
<td></td>
</tr>
<tr>
<td>Lub, Bijvank, Bal, Blomme, &amp; Schalk (2011)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>Boomers more affectively committed ($p &lt; .01, d = .53$) and continuously committed ($p &lt; .001, d = .80$) than Millennials. Millennials have higher turnover intention ($p &lt; .05, d = .45$) than Boomers.</td>
<td></td>
</tr>
<tr>
<td>Park &amp; Gursoy (2012)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>Boomers scored higher than GenXers and Millennials for VI ($p &lt; .05, d = .27$), DE ($p &lt; .01, d = .62$), and AB ($p &lt; .01, d = .35$).</td>
<td></td>
</tr>
<tr>
<td>Solnet, Kralj, &amp; Kandampully (2012)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>Boomers/GenXers group more engaged than Millennials ($p &lt; .001, d = .45$).</td>
<td></td>
</tr>
<tr>
<td>Solnet &amp; Kralj (2010)</td>
<td>Quantitative</td>
<td>L &amp; H</td>
<td>Boomers/GenXers group more engaged than Millennials ($p &lt; .005, d = .36$). Boomers/GenXers group has more job satisfaction than Millennials ($p &lt; .011, d = .33$).</td>
<td></td>
</tr>
</tbody>
</table>

*Notes: MI = Measurement Invariance, B = Baby Boomers, X = GenXers, M = Millennials, L & H = Leisure and Hospitality, PEB = Pro-Environmental Behaviors, EC = Environmental Concerns, AM = Autonomous Motivation, SE = Self-Efficacy, MBI = Maslach Burnout Inventory, VI = Vigor, DE = Dedication, and AB = Absorption.*
The mean differences were reported for the remaining six dimensions; with Boomers to Millennials pair-wise comparison: (a) work centrality \( M = .27, d = .44 \); (b) non-compliance \( M = -.27, d = -.33 \), (c) technology challenge \( M = .28, d = .40 \), (d) work-life balance \( M = -.28, d = -.44 \); (e) leadership \( M = -.19, d = -.31 \), and (f) recognition \( M = -.45, d = -.47 \) (Gursoy et al., 2013).

In a study of 358 respondents from a Dutch hotel chain, empirical findings were reported for affective commitment \( (\alpha = .92) \), continuance commitment \( (\alpha = .76) \), and turnover intention \( (\alpha = .79) \) by generational cohort (Lub et al., 2011). Results of a Bonferroni post-hoc test indicated that Boomers were more affectively committed with a mean difference of \( M = .48, p < .01 \), and \( d = .53 \) (Lub et al., 2011). As well, Boomers scored more positively on continuance commitment with a mean difference of \( M = .71, p < .001 \), and \( d = .80 \) (Lub et al., 2011). Results for turnover intention were opposite of the commitment results. Millennials had a higher intention to leave compared to Boomers with a mean difference of \( M = .41, p < .05 \) and \( d = .45 \) (Lub et al., 2011).

Employees from 29 mid and upscale level North American hotel locations were the focus of the next study by Lu and Gursoy (2013). The study included Boomers \((n = 236)\), GenXers \((n = 248)\), and Millennials \((n = 193)\), with an average organizational tenure of just 4.2 years (Lu & Gursoy, 2013). Overall, 70% of the employees reported an organizational tenure of less than 5 years. The sample was skewed towards females, accounting for 65% (Lu & Gursoy, 2013). The study included the constructs of job satisfaction \( (\alpha = .80) \), turnover intention \( (\alpha = .77) \), and the three factors of burnout, exhaustion \( (\alpha = .87) \), cynicism \( (\alpha = .80) \), and personal efficacy \( (\alpha = .78) \) (Lu & Gursoy, 2013). The study “found a significant moderating effect of generation on the relationship
between emotional exhaustion and job satisfaction and turnover intention” (Lu & Gursoy, 2013, p. 15). Within the study, turnover intention was used as the dependent variable, and Millennials were coded as 1 and Boomers as 0. Hierarchical regression analysis showed a standardized regression coefficient of $r = -.54$ with a $p$-value of $< .05$ and $R^2 = .34$. Millennials were statistically ($p < .05$) and practically ($R^2 = .34$) significantly lower in their levels of job satisfaction and higher in their levels of turnover intention when compared to Boomers (Lu & Gursoy, 2013).

An empirical study conducted in five hotels in Brisbane Australia collected data from 264 employees categorized into Gen Y (Millennials) and Non-Gen Y (Boomers and GenXers) (Solnet & Kralj, 2010). The Millennial respondents represented 62.5% of the sample, and 71.6% of the overall sample were employees of the hotels in non-supervisory roles (Solnet & Kralj, 2010). The researchers utilized independent samples $t$-testing to analyze the differences between the cohort by job satisfaction ($\alpha = .70$) and engagement ($\alpha = .74$). The job satisfaction and engagement of Millennial employees were found to be lower than the cohort of older employees (Boomers and GenXers). For engagement, statistics were reported as $t(257) = -2.82$, $p = .005$, $d = -.36$ ($M_{\text{Millennials}} = 5.43$, $M_{\text{Boomers and GenXers}} = 5.73$), and for job satisfaction, statistics were reported as $t(262) = -2.56$, $p = .011$, $d = -.33$ ($M_{\text{Millennials}} = 5.13$, $M_{\text{Boomers and GenXers}} = 5.56$) (Solnet & Kralj, 2010).

The same Australian researchers as in the prior example (Solnet & Kralj, 2010) conducted a second empirical study the following year (Kralj & Solnet, 2011). Data were collected via survey methodology from 914 employees across 24 Queensland area hospitality organizations (Kralj & Solnet, 2011). Reliability for the engagement measure was reported as $\alpha = .70$. Again, the researchers categorized the generations by Gen Y
(Millennials) and Non-Gen Y (Boomers and GenXers). Consistent with the prior study, the researchers again reported that the Millennial cohort was statistically significantly lower on engagement compared to the older generational cohort category (i.e., Boomers and GenXers). Statistics were derived from conducting an independent samples t-test and reported as $t(912) = -6.53$, $p < .001$, $d = .45$ (Kralj & Solnet, 2011).

The consistency of the prior two engagement examples (Solnet & Kralj, 2010; Kralj & Solnet, 2011) is echoed in Park and Gursoy (2012) and their findings regarding work engagement (vigor, dedication, and absorption) differences between Millennials, GenXers, and Boomers. Park and Gursoy (2012) collected data from 677 customer contact employees from 29 mid and upscale hotel properties in the United States. The sample consisted of 65% females, was split almost evenly by generational cohort (Boomers, GenXers, and Millennials), and nearly 70% of the employees had a tenure of less than 5 years with their organization (Park & Gursoy, 2012). Cronbach’s alpha was reported for vigor ($\alpha = .80$), dedication ($\alpha = .83$), and absorption ($\alpha = .74$). The study used a cross-sectional design and evaluated the respondents’ levels of engagement both by age and by generational cohort. Age was regressed against the three dimensions of work engagement and reported “that age was significantly related to vigor ($\beta = 0.10$, $p < .01$), dedication ($\beta = 0.23$, $p < .01$), and absorption ($\beta = 0.13$, $p < .01$), suggesting that older employees tend to be more engaged” than younger employees (Park & Gursoy, 2012, p. 1198).

The study (Park & Gursoy, 2012) next reported on the generational differences by way of ANOVA with post hoc tests by generation for the three dimensions of work engagement. For vigor, Boomers scored higher than GenXers and Millennials ($F (2, 674)$
= 3.59, \( p < .05, d = .27 \)). Also, Boomers scored higher than GenXers and Millennials on dedication (\( F (2, 674) = 19.50, p < .01, d = .62 \)). And Boomers scored higher once again for the absorption dimension compared to GenXers and Millennials (\( F (2, 674) = 5.46, p < .01, d = .35 \)) (Park & Gursoy, 2012). These reported findings were followed by a statement from the researchers that “employees in the older generations are likely to be more dedicated to, engrossed in, and even vigorous at work” (Park & Gursoy, 2012, p. 1198).

Park and Gursoy conducted regression analyses and found “significant main effects of three dimensions of work engagement and generation variables on turnover intention” (2012, p. 1199). Vigor had a significant negative effect on turnover intention for Millennials versus Boomers (\( \beta = -.43, p < .01 \)), generation variables had a significant effect on turnover intention for Millennials versus Boomers (\( \beta = .18, p < .01 \)), and the interaction between vigor-generation added 0.9% in explained variance, with an overall \( R^2 = .25 \) (Park & Gursoy, 2012). Dedication had a significant negative effect on turnover intention for Millennials versus Boomers (\( \beta = -.47, p < .01 \)), generation variables had a significant effect on turnover intention for Millennials versus Boomers (\( \beta = .10, p < .05 \)), and the interaction between dedication-generation added 1.1% in explained variance, with an overall \( R^2 = .27 \) (Park & Gursoy, 2012). Absorption had a significant negative effect on turnover intention for Millennials versus Boomers (\( \beta = -.24, p < .01 \)), generation variables had a significant effect on turnover intention for Millennials versus Boomers (\( \beta = .19, p < .01 \)), and the interaction between absorption-generation added 1.9% in explained variance, with an overall \( R^2 = .13 \) (Park & Gursoy, 2012).
Differences by generational cohort have been found through empirical research in the leisure and hospitality industry (Chen & Choi, 2008; Chi et al., 2013; Gursoy et al., 2008; Gursoy et al., 2013; Kralj & Solnet, 2011; Lub et al., 2011; Park & Gursoy, 2012; Solnet & Hood, 2008; Solnet & Kralj, 2010). Findings by Chen and Choi (2008) confirmed the generational differences and recommended new industry practices related to retention activities and selection. Park and Gursoy focused on generational cohort differences by work engagement and suggested that “low work centrality, varying work values, preferences, and psychological characteristics may act as drivers for generational differential impacts to work engagement” (2012, p. 1197). Researchers have reported statistically significant mean differences for engagement between generational cohorts within the leisure and hospitality industry, with Boomers consistently being reported as more engaged (e.g., Kralj & Solnet, 2011; Solnet & Kralj, 2010; Solnet et al., 2008; Solnet et al., 2012). The findings suggest “that engaging employees is critically important to Millennial employees in order to retain them relative to older employees” (Park & Gursoy, 2012, p. 1201).

**Generational Cohort and Engagement Research Outside of Leisure and Hospitality**

The preceding section reviewed generational cohort literature within the leisure and hospitality industry. As generational cohort research is not limited to the leisure and hospitality industry, the following section reviewed relevant generational cohort and engagement literature outside of the leisure and hospitality industry. Examples in the section below and in Table 3 show the reported differences between generational cohorts across the varied engagement constructs, including work engagement being operationalized through the use of UWES.
The study by Lester et al. (2012) utilized a sample of Boomers, GenXers, and Millennials from a professional Midwestern organization. The sample was reported to have been pulled from all job levels of the organization, including the CEO, and data were collected by way of online survey methodology. The survey contained 15 assessments for the work contexts of (a) teamwork, (b) autonomy, (c) security, (d) professionalism, (e) flexibility, (f) formal authority, (g) technology, (h) face-to-face communication, (i) e-mail communication, (j) social media, (k) structure at work, (l) involvement, (m) continuous learning, (n) fun-at-work, and (o) recognition (Lester et al., 2012). Principal component analysis with Promax rotation was conducted, and the evaluation of eigenvalues revealed four factors including engagement (teamwork, face-to-face communication, participation, continuous learning, fun-at-work, and recognition) (Lester et al., 2012). Using generational cohort theory, the study aimed to untangle the perceived and actual differences of the 15 work contexts by generational cohort. Of the actual reported differences for the engagement classified items, Millennials had more fun-at-work ($M = .66, p \leq .01, d = .72$) than Boomers, and Millennials also valued continuous learning ($M = .51, p \leq .05, d = .56$) more than Boomers (Lester et al., 2012).
Table 3

*Generational Cohort Research Outside of Leisure and Hospitality*

<table>
<thead>
<tr>
<th>Authors</th>
<th>Study Type</th>
<th>Context</th>
<th>Generations (ns by cohort provided if available)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bano, Vyas, &amp; Rohini (2015)</td>
<td>Quantitative</td>
<td>Western Indian Professionals</td>
<td>X(n=50), M(n=50)</td>
<td>GenXers reported higher for POS ($M_X = 34.86, SD_X = 8.50, M_Y = 25.38, SD_Y = 8.40, $p &lt; .01, d = .12$) and WE ($M_X = 75.66, SD_X = 11.55, M_Y = 55.70, SD_Y = 17.34, p &lt; .01, d = 1.36$).</td>
</tr>
<tr>
<td>Hoole &amp; Bonnema (2015)</td>
<td>Quantitative</td>
<td>Financial Organizations</td>
<td>B(n=64), X(n=93), M(n=104)</td>
<td>Boomers more engaged than GenXers ($p = .017, d = .50$). Boomers more engaged than Millennials ($p &lt; .001, d = .64$). Age was found to not be statistically significantly related to MBI-GS or UWES (VI, DE, AB).</td>
</tr>
<tr>
<td>Kim, Shin, &amp; Swanger (2009)</td>
<td>Quantitative</td>
<td>Quick Service Restaurants</td>
<td>Age Only</td>
<td>15 different work contexts. Millennials had more fun-at-work ($p &lt; .01, d = .72$) and valued continuous learning ($p &lt; .05, d = .56$) more than Boomers. For UWES-9, 31-50 group more engaged than Under 30 group ($M = .20, p &lt; .01, d = .21$), Over 50 group more engaged than Under 30 group ($M = .49, p &lt; .01, d = .47$), and Over 50 group more engaged than 31-50 group ($M = .29, p &lt; .01, d = .30$). Support for configural, metric, and scalar</td>
</tr>
</tbody>
</table>
Martins (2015)  MI  Multiple Industries  B(n=1150),
               X(n=1690),
               M(n=1285)
Using the EEI, support for configural and metric invariance in a narrowed (5 industries instead of 7) industry sample.

Martins & Ledimo (2016)  Quantitative  Government  B(n=88),
                         X(n=174),
                         M(n=165)
Using the EEI, found Millennials to be more engaged than Boomers ($p = .005, d = .31$).

Review of generational differences for various work attitudes. Numerous effect sizes reported. Boomers higher in work centrality and work ethic. Millennials higher in leisure values, extrinsic values, and individualistic traits and attitudes.

The long version of the UWES instrument was used in a 2009 study of quick service restaurant employees (Kim et al., 2009). Since the present study excluded restaurant and food service oriented organizations from the sample, the Kim et al (2009) study provided insight into a different sector of the leisure and hospitality industry. The study included the MBI-GS (MBI-GS; Schaufeli et al., 1996) along with the UWES-17 instrument. The researchers reported the correlation of professional efficacy as being more correlated to vigor \( (r = .60) \), dedication \( (r = .55) \), and absorption \( (r = .50) \) than compared to exhaustion \( (r = -.12) \) and cynicism \( (r = -.37) \) and “regarded (professional efficacy) as an element of engagement (Kim et al., 2009, p. 100). Thus, work engagement was operationalized as a four factor construct containing vigor \( (\alpha = .73) \), dedication \( (\alpha = .87) \), absorption \( (\alpha = .74) \), and professional efficacy \( (\alpha = .69) \) (Kim et al., 2009). Demographic variables were collected, including age, gender, marital status, education, position, and work experience (Kim et al., 2009). The authors cited prior work engagement literature (Schaufeli & Bakker, 2003) as support to include age as one of the key individual level control variables in the study (Kim et al., 2009). Yet, when a series of ANOVAs was performed, age was not statistically significantly related to engagement or burnout and was not included as a control variable in the subsequent analyses (Kim et al., 2009). This finding was counter to other empirical findings (e.g., Schaufeli et al., 2002; Schaufeli et al., 2006; Schaufeli & Bakker, 2003; Solnet & Hood, 2008), and the authors attributed their finding to a small and homogenous sample (Kim et al., 2009).

A study sampling GenXers and Millennials from Western India provided insight to overseas generational cohort research (Bano et al., 2015). Using the short form of
UWES (i.e., UWES-9) and operationalizing work engagement as a single factor, the reported reliability was $\alpha = .83$ (Bano et al., 2015). Perceived organizational support (POS) was operationalized using the Perceived Organizational Support Scale (Rhoades & Eisenberger, 2002), and the reported reliability was $\alpha = .72$ (Bano et al., 2015). Overall, the GenXer group scored statistically significantly ($p < .01$) and practically significantly ($r = .49$) higher for POS and statistically significantly ($p < .01$) and practically significantly ($r = .56$) higher for work engagement compared to the Millennial group (Bano et al., 2015).

The recently developed Employee Engagement Instrument (EEI: Nienaber & Martins, 2014) was used as the engagement instrument for a study of $n = 427$ government employees (Martins & Ledimo, 2016). The researchers collected demographics to allocate and subsequently analyze the sample by the generational cohorts of Baby Boomers ($n = 88$), GenXers ($n = 174$), and Millennials ($n = 165$) (Martins & Ledimo, 2016). The EEI contained the six factors of team ($\alpha = .94$), organizational satisfaction ($\alpha = .95$), immediate manager ($\alpha = .94$), organizational commitment ($\alpha = .93$), strategy and implementation ($\alpha = .91$), and customer service ($\alpha = .82$) (Martins & Ledimo, 2016). Comparatively to leisure and hospitality empirical literature that reports Boomers to be more engaged than Millennials (e.g., Park & Gursoy, 2012; Solnet & Hood, 2008; Solnet & Kralj, 2010), Martins and Ledimo reported that overall “Millennials are significantly more positive than the Baby Boomers,” with $p = .005$ and $d = .31$ (2016, p. 23). Recommendations were made for government organizations: “networks should be created to pool experiences and to develop new ideas to enhance levels of engagement”
since “millennials … appear to experience the work environment differently than the
other age groups” (Martins & Ledimo, 2016, p. 24).

The empirical evidence review conducted by Twenge (2010) considered a range
of work attitudes and the generational differences therein. The conclusion of the review
was that “overall, generational differences are important where they appear” (Twenge,
2010, p. 201). Compared to Millennials, Boomers showed higher levels of work
centrality and work ethics. Alternately, Millennials showed higher levels of leisure
values, extrinsic values, and individualistic traits and attitudes compared to Boomers. No
differences by generation were found for altruistic values or intrinsic values (Twenge,
2010). Selected effect sizes were reported by Twenge (2010) as support for the
aforementioned comparisons. The conclusions and recommendations included the
following: recruiting Millennials by focusing on work-life balance and flexible schedules;
corporate programs supporting volunteering, altruistic values, and meaning in work made
no difference by generation; and Millennials were no more likely to leave their work
compared to other generations if they are satisfied with work (Twenge, 2010).

**Engagement Articles with Measurement Invariance Assessment between
Generational Cohorts**

The empirical literature search for engagement studies that conducted
measurement invariance analyses between generational cohorts returned scant few
articles. This result was anticipated as it signaled one of the main issues of generational
cohort research. That is, research that makes mean comparisons between groups (e.g.,
generational cohorts) often lacks measurement invariance testing. The sourced
engagement articles (Table 3) were reviewed in this section. Due to the overall lack of
specific engagement studies that included both measurement invariance and generational cohort, a supplemental view for engagement related constructs (e.g., job satisfaction and organizational commitment) was provided later (Table 4) in this paper in the employee engagement section.

In a study of financial organizations, the constructs of work engagement and meaningful work were analyzed for a sample of Baby Boomers \((n = 64)\), GenXers \((n = 93)\), and Millennials \((n = 104)\) (Hoole & Bonnema, 2015). To measure meaningful work, the Psychological Meaningfulness Scale (PMS: Tymon, 1988) was used, and UWES-9 was used to measure work engagement as a single factor (Schaufeli et al., 2006). Reported reliabilities were \(\alpha = .72\) for PMS and \(\alpha = .93\) for UWES (Hoole & Bonnema, 2015). Reported results for work engagement indicated that Boomers were statistically \((p = .017)\) and practically \((d = .50)\) significantly more engaged than GenXers, Boomers were statistically \((p < .001)\) and practically \((d = .64)\) significantly more engaged than Millennials, but GenXers were not statistically \((p = .148)\) or practically \((d = .14)\) significantly more engaged than Millennials (Hoole & Bonnema, 2015). As part of the limitations section of the article, the authors commented on the need to conduct measurement invariance testing as the cohorts could “simply have a different understanding or interpretation of the constructs being measured” (Hoole & Bonnema, 2015, p. 9). While the researchers did not conduct measurement invariance to confirm invariance between groups, identifying the need is an important addition to the limitations of the article.

Data from the Employee Engagement Instrument (EEI: Nienaber & Martins, 2014) was used to assess measurement invariance across industry sectors in a South
African setting (Martins, 2015). Agriculture, manufacturing, electricity, wholesale and retail, transportation, financial, and community sectors were sampled with a total of 4,125 questionnaires received (Martins, 2015). Six factors were reported for EEI consisting of team (\(\alpha = .93\)), organizational satisfaction (\(\alpha = .94\)), immediate manager (\(\alpha = .93\)), organizational commitment (\(\alpha = .93\)), strategy and implementation (\(\alpha = .90\)), and customer service (\(\alpha = .81\)) (Martins, 2015). The reported demographics included Boomers \((n = 1,150)\), GenXers \((n = 1,690)\), and Millennials \((n = 1,285)\), but the study did not control for demographic variables (Martins, 2015). Configural invariance was supported across the seven industrial sectors, but metric invariance was not supported. The researcher removed the manufacturing and community sectors and conducted invariance testing again, finding support for configural and metric invariance with the narrower industrial sample (Martins, 2015).

In a 2017 article, the Russian version of the UWES-9 instrument was tested for measurement invariance between gender and age groups (Lovakov et al., 2017). The researchers defined the age groups as follows: a) Under 30, b) 31-50), and c) Over 50. The participants were all from a single Russian energy corporation, and a total of 1,783 completed the survey that contained the UWES-9 instrument. After removing cases for incomplete responses, the sample was reduced to Under 30 \((n = 554)\), 31-50 \((n = 999)\), and Over 50 \((n = 173)\). The researchers established the factorial structure of UWES-9 for both the three-factor and single-factor structures before continuing to measurement invariance testing. Measurement invariance testing was conducted using the three-factor model and revealed support for configural, metric, and scalar invariances across the three age groups. Mean scores were compared across the age groups since support for scalar
invariance was found. For UWES-9, the 31-50 group was more engaged than the Under 30 group ($M = .20$, $p < .01$, $d = .21$), the Over 50 group was more engaged than the Under 30 group ($M = .49$, $p < .01$, $d = .47$), and the Over 50 group was more engaged than the 31-50 group ($M = .29$, $p < .01$, $d = .30$). Even though the age groups were not directly comparable to the generational cohort definitions, the study does provide directional support that the younger generation (Millennials) is more engaged (UWES-9) compared to the older generations (Boomers and GenXers) (Lovakov et al., 2017).

**Employee Engagement**

The phenomenon of employee engagement is a preeminent focus of research for academics and practitioners; yet, employee engagement literature was “practically non-existent just 10 years ago” (Saks & Gruman, 2014, p. 156). The field of human resource development has specifically witnessed an upward trend of literature in the field’s leading journals, and it remains an active topic in practitioner circles.

The swell of academic literature attention to employee engagement begs the questions: why and why now? To answer the former question, researchers are interested in employee engagement as it purportedly addresses many questions regarding organizational performance and success (Gilley & Gilley, 2000; Macey, Schneider, Barbera, & Young, 2009; Rich et al., 2010; Saks, 2006). Organizations with engaged employees have higher customer satisfaction ratings, productivity, profitability, and returns for shareholders (Crawford, LePine, & Rich, 2010; Harter et al., 2002).

Regarding the timing of the interest and to answer the latter question, research has pointed to a decline in engaged employees (Bates, 2004; Richman, 2006). With nearly
half of the U.S. workforce reportedly disengaged or not fully engaged (Bates, 2004; Saks & Gruman, 2014), researchers see the need for organizations to engage their workforce.

**Domains of Employee Engagement**

The breadth of literature attention for employee engagement has not yet remedied a basic problem with the phenomenon. Due to the conceptual overlap of employee engagement and related organizational concepts (e.g., organizational commitment and job satisfaction), multiple interpretations exist of the phenomenon’s meaning and definition (Saks & Gruman, 2014; Shuck & Wollard, 2010). Employee engagement is represented by four major domains (Shuck and Wollard, 2010): (a) needs satisfying (Kahn, 1990), (b) satisfaction engagement (Harter et al., 2002), (c) multidimensional (Saks, 2006), and (d) burnout antithesis (Maslach et al., 2001). Of the four domains identified by Shuck and Wollard (2010), a single domain does not dominate the literature (Christian, Garza, & Slaughter, 2011). The four domains are more similar than disparate, only diverging on approach, and agreeing on relating employee engagement to positive organizational consequences (Harter et al., 2002; Maslach et al., 2001; Saks, 2006; Shuck and Wollard, 2010; Shuck, 2011).

The seminal article for employee engagement appeared in the *Academy of Management Journal* in 1990 (Kahn). The seminal conceptualization was influenced by the sociology and psychology fields and extrapolated employee (individual) behaviors into an organizational context. When explaining an individual’s engagement, Kahn (1990) divided the concept into personal engagement and personal disengagement. This binary concept finds grounding in Goffman’s (1961) point in time attachment and detachment role performance theory. Conversely to Goffman (1961), Kahn (1990)
viewed an employee’s work perspective and engagement as continuous rather than point in time or momentary. Kahn offered the following definition: “personal engagement is the simultaneous employment and expression of a person’s ‘preferred self’ in task behaviors that promote connections to work and to others, personal presence (physical, cognitive, and emotional), and active, full role performances” (1990, p. 700). In addition to offering the definition, Kahn (1990) noted three psychological factors that either enable or inhibit personal engagement and personal disengagement: meaningfulness, safety, and availability. May, Gilson, and Harter (2004) found statistically significant empirical support for each of the three psychological factors. Kahn’s (1990) research guided later work in the employee engagement space, including the following three other major domains.

The second academic domain of satisfaction engagement was formulated in 2002 through a meta-analysis utilizing a database of 7,939 business units collected across multiple industries by the Gallup Organization (Harter et al., 2002). Using this database, Harter et al. (2002) included the business unit level and the employee level in their research. The definition offered is: an “individual’s involvement and satisfaction with as well as enthusiasm for work” (Harter et al., 2002, p. 269). The instruments used to survey the respondents were the Gallup Organization’s proprietary property. While this makes the instrument validation difficult, the Gallup and other practitioner instruments are widely used and generally accepted as measures of both employee engagement and job satisfaction.

approach including the antecedents of job characteristics (e.g., work conditions, autonomy, and feedback), leadership, and personality traits. The job characteristics identified by Macey and Schneider (2008) overlap with the job resources identified by Demerouti et al. (2001), demonstrating the interconnectedness of the domains positioned by Shuck (2011). Saks offered the definition for engagement of “a distinct and unique construct consisting of cognitive, emotional, and behavioral components … associated with individual role performance” (2006, p. 602). Saks (2006) empirically tested and confirmed the multi-dimensional concept using numerous antecedents and outcomes. Saks (2006) delineated employee engagement into job engagement (as related to one’s work-related role) and organizational engagement (as related to one’s role within the organization).

Burnout antithesis is the fourth major domain of engagement research and is rooted in the multidimensional theory of burnout (Maslach, 1982; Maslach, 1999). Burnout itself preceded the burnout antithesis domain and was defined by Maslach (1982; Maslach, 1999) as a “syndrome of emotional exhaustion, depersonalization, and reduced personal accomplishment” (Demerouti et al., 2001, p. 499). Burnout can be measured through the MBI and contains the dimensions of emotional exhaustion, depersonalization, and reduced personal accomplishment (Maslach, 1982; Maslach, 1999; Maslach & Jackson, 1986; Maslach et al., 1996). A generalized scale, the MBI-GS (Schaufeli et al., 1996), was later created to be used with a wider, general industry footprint and contained the burnout dimensions of exhaustion, cynicism, and reduced professional efficacy. Near the time period of the MBI-GS creation, two major antithesis burnout branches of research developed.
The first branch directs burnout and engagement as polar opposites of the same construct (Cole et al., 2012). Three dimensions of burnout (i.e., exhaustion, cynicism, and reduced professional efficacy) are countered by a corresponding three dimensions (i.e., energy, involvement, and efficacy) of engagement (Maslach & Leiter, 1997; Maslach et al., 2001; Saks, 2006). Defined as “a persistent positive affective-motivational state”, the concept originated as positive scores on the Maslach Burnout Inventory (Maslach & Leiter, 1997, p. 417), which supports the burnout antithesis concept by linking burnout and engagement on the same continuum (Maslach, 1999).

The second research branch of burnout antithesis research forms the center of the present study. The grounding tenet and differentiator of the second branch is that employee engagement and job burnout are not the same construct or on the same continuum; rather, they are related but independent states, should be measured with independent instruments, and are negatively correlated (Russell & Carroll, 1999; Schaufeli et al., 2002). The following definition was offered: “work engagement is the positive, fulfilling work-related state of mind that is characterized by vigor, dedication, and absorption” (Schaufeli et al., 2002, p. 74). Three major components of the definition are notable. First, engagement is called work engagement and is specific to working. Second, engagement is a state of mind, rather than a trait of the individual. Last, three major characterizations are called out: vigor, dedication, and absorption. Within the Utrecht Work Engagement Scales (UWES-17: Schaufeli et al., 2002; UWES-9: Schaufeli et al., 2006), engagement was positioned specifically as work engagement and includes measurements for the factors of vigor, dedication, and absorption. Within the definition and these notable components, engagement was differentiated as an independent
construct and still as the opposite of burnout (Schaufeli et al., 2002; Schaufeli et al., 2006). Further support and clarification was provided in the definitions of vigor, dedication, and absorption. Defined as “high levels of energy, mental resilience while working, the willingness to invest one’s work and persistence in the face of difficulties”, vigor is the “direct opposite of the core burnout dimension of exhaustion” (Schaufeli et al., 2006, p. 702). Similarly positioned opposite of the corresponding burnout dimension of cynicism, dedication represents “enthusiasm, inspiration, pride, and challenge” (Schaufeli et al., 2006, p. 702). Absorption is treated as a unique characteristic of work engagement, falling outside of the burnout dimension comparisons (Schaufeli et al., 2006). The antithesis domain does present some challenge to researches arguing whether engagement is truly a unique construct or simply the opposite of burnout (Saks & Gruman, 2014).

This section of the present study’s literature review synthesized the four academic domains of engagement proposed by Shuck and Wollard (2010). In addition to structuring the four domains, Shuck and Wollard offered the engagement definition of “an individual employee’s cognitive, emotional, and behavioral state directed toward desired organizational outcomes” (2010, p. 103). The definition includes the multi-dimensional components similar to Saks (2006) and the concept of the phenomenon being a state is similar to Kahn (1990) and Schaufeli et al. (2002).

**Measurement of Employee Engagement**

Given the numerous definitions and conceptualizations of engagement, numerous measurement instruments are also available. At least 12 relevant scales are available to measure engagement. The needs satisfying (Shuck & Wollard, 2010) domain from Kahn
(1990) contains prominently used scales. May et al. (2004) leaned heavily on Kahn’s (1990) conceptualization of engagement and developed a 13-item scale measuring cognitive, emotional, and physical engagement. Again supporting Kahn’s (1990) conceptualization, Rich et al. (2010) developed an 18-item scale to measure cognitive, emotional, and physical engagement. The Rich et al. (2010) scale is split evenly with 6-items each measuring the three dimensions of Kahn’s (1990) conceptualization. The satisfaction engagement domain (Shuck & Wollard, 2010) is dominated by proprietary practitioner measures. Chiefly, the Gallup instrument is frequently used and was the backbone of the research by Harter and colleagues (Harter et al., 2002). The multidimensional domain (Shuck & Wollard, 2010) is measured by two factors of engagement. To measure job engagement and organization engagement, Saks (2006) developed a 12-item scale, using only 11 items after dropping one of the job engagement items after testing.

The burnout antithesis domain (Maslach, 1982; Schaufeli & Bakker, 2004) contains several measurement instruments. Both the Maslach Burnout Inventory and Oldenburg Burnout Inventory have been used to measure engagement through positive scores on the burnout construct (Demerouti et al., 2001; Maslach & Leiter, 1997). The UWES-17 scale was developed specifically for the burnout antithesis domain and measures three dimensions of engagement: vigor, dedication, and absorption (Schaufeli et al., 2002). A shortened UWES-9 version was developed and validated four years later (Schaufeli et al., 2006). The UWES scales are the most widely used instruments within the academic employee engagement literature (Saks & Gruman, 2014). Several engagement scales fall outside of the domains outlined by Shuck and Wollard (2010). A
9-item scale to measure attention and absorption was created in 2001 (Rothbard, 2001). To measure intellectual engagement, affective engagement, and social engagement, the Soane, Truss, Alfes, Shantz, Rees, and Gatenby (2012) 9-item scale was developed with three items measuring each of the factors. Nienaber and Martins (2014, 2015) developed a six factor engagement scale that has been tested and validated in the African continent. And last, to measure felt engagement and behavioral engagement, Stumpf, Tymon, and van Dam (2013) created a 14-item scale that focused on technology work groups.

**Engagement Summary**

It is understandable why organizations and researchers are interested in employee engagement due to its purported influence on organizational performance and success (Cole et al., 2010; Saks, 2006; Shuck, 2011). Even though the construct suffers from a confounding number of definitions, numerous measurement instruments, and no single accepted theory (Saks & Gruman, 2014; Shuck, 2011), the empirical research has suggested it to be a unique and distinct “higher-order construct” (Christian et al., 2011, p. 94). Determining the one appropriate engagement construct is imperative for empirical research. Shuck argued that “no research has suggested one approach is academically more accepted than another” (2011, p. 17). Shuck (2011) continued and compared that determining the appropriate engagement construct is similar to picking the most appropriate research method. The engagement construct used should align across the research spectrum of design, question, definition, and measurement tool (Shuck, 2011). Conceptualization of the construct as a state, rather than a trait, that changes over time or situationally for the individual is consistent with the burnout antitheses domain.
Psychometrics of the Utrecht Work Engagement Scale

The UWES scale was developed near the turn of the century in response to the positive psychology movement (Seligman & Csikszentmihalyi, 2000) and the burnout antithesis domain of engagement research (Maslach & Leiter, 1997; Maslach et al., 2001). Containing the three dimensions of vigor, dedication, and absorption, work engagement was deemed as an independent construct from burnout and required its own instrumentation of measurement (Schaufeli et al., 2002). Instrumentation development began through empirical research assessing the relationships between burnout and engagement using the MBI-GS (Schaufeli et al., 2001) and newly constructed UWES instruments (Schaufeli et al., 2002).

Development of UWES

The development of UWES began with a 24-item self-constructed instrument that was written in both English and Spanish. The instrument was written to reflect the three underlying dimensions of work engagement. Vigor contained nine items, dedication contained eight items, and absorption contained seven items. All items were scored on a 7-point Likert scale and positively worded, requiring no reverse coding. The Likert scale and directional wording was done in this fashion to align with the MBI-GS instrument (Schaufeli et al., 2001). The 24 UWES items were randomly merged with the MBI-GS items to create a 40-item questionnaire. The full questionnaires were administered to two samples. Since one of the samples consisted of students, a student version of the full 40-item questionnaire was developed. Sample one was administered in the spring of 1999 to 314 undergraduate students using the student version. Sample two was administered in
1999 to 619 employees of 12 Spanish companies using the employee version (Schaufeli et al., 2002).

The statistical analysis began with a reliability analysis. Cronbach’s α was calculated across the three engagement scales: vigor α = .69$_{S1}$, .80$_{S2}$; dedication α = .91$_{S1, S2}$; and absorption α = .73$_{S1}$, .75$_{S2}$. A subsequent analysis was conducted to remove items “that either negatively affected values of α or that did not make a positive contribution to the level of α. Vigor was reduced from nine items to six with α = .78$_{S1}$, .79$_{S2}$ for the final reliability. Dedication was reduced from eight items to five with α = .84$_{S1}$, .89$_{S2}$ for the final reliability. Absorption was reduced from seven items to six with α = .73$_{S1}$, .72$_{S2}$ for the final reliability (Schaufeli et al., 2002).

Multi-group SEM analysis was conducted fitting several different models across samples using both three-factor and single-factor models. Analysis results were as expected that burnout and engagement were negatively correlated. The engagement scales were interrelated at $r = .63_{S1}$ and $r = .70_{S2}$. The three-factor hypothesized model of engagement was fitted and compared to a single-factor model of engagement. Comparison of fit indices revealed that the three-factor model fit the data better RMSEA = .05 and CFI = .90 compared to the single-factor model RMSEA = .07 and CFI = .85. It was observed that the correlations between vigor and absorption were considered “quite high”, and a two-factor model was fit with vigor and absorption “collapsed into one factor”; the fit was inferior to the three-factor model (Schaufeli et al., 2002, p. 82).

**Post-Development of the UWES**

The 17-item version of UWES was used in the 2004 extension of the JDR model (Schaufeli & Bakker, 2004). In this study, the JDR model was reconditioned from using
the OLBI to measure burnout and disengagement to using the MBI-GS to measure burnout and UWES-17 to measure engagement (Schaufeli & Bakker, 2004). In 2006, a shortened version of UWES was developed for “basically pragmatic” purposes as “researchers strive to include as few items as possible in measuring a particular construct” (Schaufeli et al., 2006, p. 702).

Using a 10 country sample consisting of 14,251 respondents, the 17 UWES items were iteratively reduced to the final 9-item questionnaire, consisting of three items for each of the three dimensions. Face validity was used by the researchers to pick the most characteristic item for each scale. Next, the chosen item was regressed on the remaining items within the given scale. The item with the highest $\beta$ value was retained, and this process continued until no additional variance was added by subsequent items. Cronbach’s $\alpha$ was evaluated across each of the 10 countries. Vigor carried a median $\alpha = .77$, dedication carried a median $\alpha = .85$, and absorption carried a median $\alpha = .78$. The total nine-item scale, for all 10 countries, carried a median $\alpha = .92$. Three-factor and single-factor models were tested using the shortened version, and the three-factor model was found to have a better fit than the single-factor (Schaufeli et al., 2006).

The UWES-17 and UWES-9 usage has continued and the sister instruments are the most widely used engagement instruments within the literature (Saks & Gruman, 2014), and they have been translated into at least 23 languages (Schaufeli, 2017; Taras, Rowney, & Steel, 2009). The instrument is freely available on Wilmar Schaufeli’s website, including a fully downloadable instructional manual. The free availability of the instrument and the prevalence of usage within the JDR model has contributed to Saks and Gruman’s (2014) findings. Within the leisure and hospitality industry, the instrument has
been used for studies that include work engagement. Reported reliabilities within empirical studies for food and beverage (Kim et al., 2009) and hotel (Park & Gursoy) samples range from .73 to .80 for vigor, from .83 to .87 for dedication, and are consistent at .74 for both samples.

**UWES Considerations**

As part of the 2006 study, Schaufeli and colleagues conducted measurement invariance assessments across 10 countries. Configural invariance was confirmed across the countries, but metric invariance was not found as “the one-factor and three-factor models were not invariant … across countries” (Schaufeli et al., 2006, p. 709). Prior research had established invariance across countries for the UWES-17 instrument (Schaufeli, Martinez, Marques-Pinto, Salanova, & Bakker, 2002) and for racial groups (Storm & Rothmann, 2003). It was concluded that the lack of invariance in the 2006 study sample was attributable to “different occupational groups from different countries” (Schaufeli et al., 2006, p. 713). It was recommended that future research on UWES measurement invariance should contain respondents from “similar occupational groups” (Schaufeli et al., 2006, p. 713). A 2009 empirical study conducted a multi-sample analysis and found modest support for UWES-9 to be invariant across the occupational samples when the paired sample comparisons were those of similarly small sized samples (Seppala, Mauno, Feldt, Hakanen, Kinnunen, Tolvanen, & Schaufeli, 2008). Furthermore, the same study conducted time invariance analysis using the same samples and data from UWES-9 and found support that vigor, dedication, and absorption remained highly stable over a three year measurement period (Seppala et al., 2009).
Schaufeli and colleagues included analyses by age, gender, and occupation group as part of the development of the UWES-9 (Schaufeli et al., 2006). Converse to burnout, engagement was purported to slightly increase with age, but the relationships were very weak for each of the three factors. Correlations for Vigor ranged from .00 to .28 (median = .08), Dedication ranged from .02 to .28 (median = .09), and Absorption ranged from .00 to .27 (median = .12). Gender did not yield practically significant differences between males and females as all effect sizes were small (i.e., $d < .20$). The gender differences varied by country. Australian, Canadian, and French samples were reported to have no gender differences. Men scored slightly higher in Belgian, German, Finnish, and Norwegian samples. Women scored slightly higher in South African, Spanish, and Dutch samples. Occupational group was evaluated and it was found that blue-collar workers ($M_{\text{Vigor}} = 3.47$, $M_{\text{Dedication}} = 3.40$, and $M_{\text{Absorption}} = 2.74$) were less engaged than police officers ($M_{\text{Vigor}} = 4.14$, $M_{\text{Dedication}} = 4.55$, and $M_{\text{Absorption}} = 4.05$), teachers ($M_{\text{Vigor}} = 4.41$, $M_{\text{Dedication}} = 4.40$, and $M_{\text{Absorption}} = 3.70$), and managers ($M_{\text{Vigor}} = 4.40$, $M_{\text{Dedication}} = 4.48$, and $M_{\text{Absorption}} = 3.78$). Caution was mentioned in the interpretation and usefulness of these findings in the article as convenience samples were used rather than random sampling. No measurement invariance testing was conducted by age, gender, occupation, or generational cohort groups within the study to determine if the instrument carried the same interpretation between groupings (Schaufeli et al., 2006).

The dimensionality of UWES has been studied and support has been found for both single factor and three factor models (e.g., Coetzer & Rothmann, 2007; Salanova, Agut, and Peior, 2005; Schaufeli et al., 2002; Schaufeli et al., 2006). The samples and contexts for the studies were different and contributed to the lack of consensus on
dimensionality (Seppala et al., 2009). A multi-sample analysis study reported the structure of UWES-17 was not stable over time due to a “significant loss of fit” and “did not measure work engagement the same way at the two time points” (Seppala et al., 2009, p. 459). This lack of stability was reported by comparing a baseline UWES-17 model to a constrained stability UWES-17 model and evaluating changes in reported chi-squared ($\chi^2$) statistics (Seppala et al., 2009). Conversely, the UWES-9 structure “remained relatively unchanged” over the two time periods according to $\chi^2$ testing and factor loading evaluation (Seppala et al., 2009, p. 459). Effectively, UWES-9 “measured work engagement rather similarly over time” (Seppala et al., 2009, p. 459).

In addition to the lack of consensus on factor structure, high intercorrelations have been reported among the three factors (De Bruin & Henn, 2013). Meta-analytics findings reported mean correlations of .88 between vigor and dedication, .95 between vigor and absorption, and .90 between dedication and absorption (Christian & Slaughter, 2007). Noting the high intercorrelation of the factors, Schaufeli et al. (2006) proposed that UWES-9 could be measured as a single factor construct.

All versions of the UWES contain only positively worked items. Regardless of factor structure or version of UWES instrument being used, the potential for acquiescence response bias for instruments with only positively phrases items is a limitation (Taras et al., 2009). Literature suggests acquiescence response bias can be mitigated by using both positively and negatively worded items in a single instrument (Schimmack, Oishi, & Diener, 2005; Smith, 2004).
Measurement Invariance of Engagement Related Constructs

Several work constructs related to work engagement are present in the literature. Dalal et al. (2012) provided an empirical relative importance analysis study and included the engagement related constructs of job satisfaction, organizational commitment, job involvement, perceived organizational support, and organizational citizenship behavior. The review below and Table 4 included measurement invariance literature for the aforementioned constructs and also includes measurement invariance literature for the motivation at work and burnout constructs.

In a 2007 study by Watson, Thompson, and Meade, measurement invariance was assessed for patrol officers \( (n = 1,198) \) and administrative officers \( (n = 312) \) sampled from a Southeastern U.S. state police agency. The study utilized the Job Satisfaction Survey (JSS: Spector, 1997) and item response theory (IRT) to assess measurement invariance. In IRT, “when the relationship between the underlying trait(s) of interest and the observed survey scores differs across two groups, the survey instrument is said to display differential item functions (DIF)” (Watson et al., 2007, p. 2 – 3). A total of six of the eight JSS scales demonstrated DIF, and “comparison of mean scores on the JSS between workers employed in different contexts may be specious” (Watson et al., 2007, p. 4).
Table 4

*Measurement Invariance Literature for Engagement Related Constructs*

<table>
<thead>
<tr>
<th>Authors</th>
<th>Study Type</th>
<th>Context</th>
<th>Generations (ns by cohort provided if available)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boles, Dean, Ricks, Short, &amp; Wang (2000)</td>
<td>MI</td>
<td>Educators &amp; Small Business Owners</td>
<td>n/a</td>
<td>Configural, metric, and scalar MI between educators and small business owners for MBI-HS (modified).</td>
</tr>
<tr>
<td>Eisinga, Teelken, &amp; Doorewaard (2012)</td>
<td>MI</td>
<td>Faculty from European Universities (Business Colleges)</td>
<td>n/a</td>
<td>Configural, metric, scalar, and strict invariance across six countries for data from the affective, normative, and continuance commitment instruments.</td>
</tr>
<tr>
<td>Gagne, Forest, Gilber, Aube, Morin, &amp; Malorni (2010)</td>
<td>MI</td>
<td>English and French Speaking Canadians</td>
<td>n/a</td>
<td>Configural, metric, and scalar invariance for data from the Motivation at Work Scale between language groups (i.e. French and English).</td>
</tr>
<tr>
<td>Morin, Madore, Morizot, Boudrias, &amp; Tremblay (2009)</td>
<td>MI</td>
<td>3 Canadian Professional Organizations</td>
<td>n/a</td>
<td>Using WACMQ, found support for configural, metric, and scalar invariance between gender and configural, metric, scalar, and strict invariance for linguistic category (English/French).</td>
</tr>
<tr>
<td>Raineri, Paille, &amp; Morin (2012)</td>
<td>MI</td>
<td>Canadian Government Employees</td>
<td>B(n=444), X(n=238)</td>
<td>Invariance assessments on POS and OCB found support for configural and metric invariance between Boomers and GenXers.</td>
</tr>
<tr>
<td>Stiglbauer &amp; Batinic (2012)</td>
<td>MI</td>
<td>German Online Panel</td>
<td>n/a</td>
<td>Longitudinal study assessing Work Involvement over four waves. Support for configural and metric invariance found.</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Method</td>
<td>Group</td>
<td>Sample Size</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------</td>
<td>--------------------------------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Tayyab (2005)</td>
<td>MI</td>
<td>Public Sector Employees</td>
<td>n/a</td>
<td>Using OCB, configural invariance was found between permanent telecomm employees and involuntary contingent technology employees.</td>
</tr>
<tr>
<td>Vanheule, Rosseel, &amp; Vlerick (2007)</td>
<td>MI</td>
<td>Belgian Healthcare Workers</td>
<td>n/a</td>
<td>Using MBI-HS, found support for only configural invariance between general hospital nurses and welfare healthcare organization nurses and assistants groups.</td>
</tr>
<tr>
<td>Watson, Thompson, &amp; Meade (2007)</td>
<td>MI</td>
<td>US Police Officers and Administrative Officers</td>
<td>n/a</td>
<td>Using IRT, found 6 of 8 job satisfaction scales demonstrated DIF between police officers and administrative officers.</td>
</tr>
</tbody>
</table>

**Notes:** MBI-HS = Maslach Burnout Inventory Health Services, WACMQ = Workplace Affective Commitment Multidimensional Questionnaire, OCB = Organizational Citizenship Behavior, IRT = Item Response Theory, and DIF = Differential Item Functions.
Organizational commitment (OC) was defined through a three-component model by Allen and Meyers (1990) to measure affective, normative, and continuance commitment. Eisinga, Teelken, and Doorewaard (2012) tested the measurement invariance of OC for \( n = 723 \) faculty members of business and economics departments from universities across six European countries (Belgium, Germany, Netherlands, United Kingdom, Finland, and Sweden). After fitting a CFA model across all six countries and finding support for configural invariance, the study proceeded to assess metric, scalar, and strict invariances and found no statistically significant \( \chi^2 \) differences between models, suggesting invariance across countries at the metric, scalar, and strict levels (Eisinga et al., 2012).

A second OC measure, the Workplace Affective Commitment Multidimensional Questionnaire (WACMQ: Madore, 2004), was assessed for measurement invariance using an \( n = 404 \) sample from professionals in various roles across three Canadian organizations (Morin, Madore, Morizot, Boudrias, & Tremblay, 2009). Groups were classified by gender and linguistic categories. Support for strong measurement invariance was found for gender, and support for strict invariance was found for the linguistic category (Morin et al., 2009).

The Motivation at Work Scale (MAWS) was developed within the framework of self-determination theory (SDT: Deci & Ryan, 1985), and measurement invariance of the scale was tested for groups of workers speaking English or French. The English speaking group consisted of pilots (\( n = 881 \)), transportation managers (\( n = 55 \)), and undergraduate commerce students (\( n = 130 \)), and the French speaking group consisted of professional Canadian employees (\( n = 285 \)) and correctional officers (\( n = 249 \)) (Gagne, Forest,
Gilbert, Aube, Morin, & Malorni, 2010). Following the guidelines of Byrne (2006), support for configural, metric, and strong invariance was found between the language groups (Gagne et al., 2010).

Using latent and financial benefits as antecedents to work involvement, a three and a half year longitudinal study evaluated the measurement invariance of work involvement over four time periods (Stiglbauer & Batinic, 2012). The sample consisted of n = 735 participants randomly selected from a German online survey panel (Stiglbauer & Batinic, 2012). The first three waves were assessed in six month intervals, and the fourth wave was assessed 18 months after the third wave (Stiglbauer & Batinic, 2012). Measurement invariance results indicated support for configural and metric invariance across the four waves, but strong invariance was not supported (Stiglbauer & Batinic, 2012).

Organizational citizenship behavior (OCB) was assessed for measurement invariance using the Organizational Citizenship Behavior Scale (OCBS: Podsakoff, MacKenzie, Mooreman, & Fetter, 1990) for a sample of public sector employees (Tayyab, 2005). The sample was collected for permanent employees (n = 508) in the telecommunications industry and for involuntary contingent employees (n = 182) in the technology industry (Tayyab, 2005). Support for configural invariance was found between the groups, but support for metric invariance was lacking (Tayyab, 2005).

In a generational cohort study conducted using a sample from Canadian government employees, perceived organizational support (POS) and organizational citizenship behavior (OCB) were tested for measurement invariance (Raineri, Paille, & Morin, 2012). The sample consisted of Baby Boomers (n = 444) and GenXers (n = 238).
The sample was skewed towards females (75%), nearly half of the Boomers had 20 or more years of public sector experience, and nearly half of GenXers had five or fewer years of public sector experience (Raineri et al., 2012). Reported measurement invariance findings supported configural and metric invariance between the generations, but scalar invariance was not supported between the groups (Raineri et al., 2012).

The human services version of the Maslach Burnout Inventory (MBI-HS; Maslach et al., 1996) survey was used to measure distinct groups of workers at Belgian healthcare organizations (Vanheule, Rosseel, & Vlerick, 2007). The first group consisted of nurses (n = 2,515) from general hospitals, and the second group consisted of nurses and assistants (n = 1,639) from residential welfare institutions (youths with special needs, mentally disabled persons, or geriatrics) (Vanheule, Rosseel, & Vlerick, 2007). The samples were similar demographically by age, gender, and tenure. Upon conducting measurement invariance testing, support for configural invariance was found, but support for metric invariance was not found, suggesting “that the interpretation of the meaning of emotional exhaustion, depersonalization, and reduced personal accomplishment is sample specific” (Vanheule, Rosseel, & Vlerick, 2007, p. 91).

The human survey version of the Maslach Burnout Inventory (MBI-HS; Maslach et al., 1996) was used to assess measurement invariance for samples from educators (n = 182) and small business owners (n = 157) (Boles, Dean, Ricks, Short, & Wang, 2000). The MBI-HS was conceptualized as a three factor model (emotional exhaustion, depersonalization, and personal accomplishment) and modified by the researchers to reflect the varied interactions of the samples with either students or employees (Boles et al., 2000). Support for configural, metric, and scalar measurement invariance was
reported (Boles et al., 2000). However, these reported findings are questionable as the samples received unique instruments, tailored for their specific interactions (students and employees) based upon the nature of the respondent’s employment type.

**Hypotheses Support**

The present study contained two lines of logic for the hypotheses. The first line concerned measurement invariance between Boomer and Millennial generational cohorts, and the second line concerned determining the latent mean differences between the same groups. The four hypotheses were repeated below from Chapter 1 to avoid the need for the reader to search through the document.

**Measurement Invariance Hypotheses**

Much of the employee engagement literature states that engaged employees are important to organizations due to the reported positive outcomes of engaged employees and often continues by describing the workforce under study. Quantitative studies often utilize surveys to gather data about employees, including employee engagement. A derivative of this line of research is to segment the collected sample(s) by demographics or other classifications such as gender, race, occupation, country, age, or more specifically by generational cohort. The studies then proceed to report mean differences between the groups (e.g., generational cohort) of study for the constructs (e.g., work engagement) within the study. Yet, researchers very often fail to perform recommended (Vandenberg & Lance, 2000) assessments of measurement invariance for the groups before proceeding to reporting mean differences by group. Exceptions are for countries (Schaufeli, Martinez, et al., 2002; Schaufeli et al., 2006), race (Storm & Rothmann, 2003), and occupation (Seppala et al., 2009). Several studies have reported mean
differences by generational cohort for engagement within the leisure and hospitality industry without establishing measurement invariance (e.g., Chi et al., 2013; Gursoy et al., 2013; Kralj & Solnet, 2011; Solnet & Hood, 2008). While studies do address all generations within the workforce (Park & Gursoy, 2012), many studies only address the polar ends of the cohorts (i.e., Boomers and Millennials) as prior research has reported statistically significant differences only between the oldest and youngest generations (Kim et al., 2016; Kralj & Solnet, 2011; Solnet & Hood, 2008).

In the studies where configural invariance is assessed for UWES-9 (e.g., Schaufeli et al., 2006; Seppala et al., 2009), configural invariance is supported. CFI is reported between .95 and .98, and RMSEA is reported between .03 and .06 (Schaufeli et al., 2006; Seppala et al., 2009). Thus, H1 was set to hypothesize configural invariance between Baby Boomer and Millennial cohorts for samples within the leisure and hospitality industry.

H1: Evaluation of model fit indices will result in good model fit (i.e., TLI ≥.95, CFI ≥.95, RMSEA ≤.10, and SRMR ≤.08) and support for configural invariance for (a) the three-factor model and (b) the single-factor model.

Studies have found support for metric or weak invariance for groups between countries (Schaufeli, Martínez, et al., 2002) and race (Strom & Rothmann, 2003) using the UWES-17 version. As well, metric invariance has been supported for occupational group comparisons using the UWES-9 version (Seppala et al., 2009). Thus, H2 was set to hypothesize metric invariance between Baby Boomer and Millennial cohorts for samples within the leisure and hospitality industry.
H2: Evaluation of model fit indices, $\chi^2$, and CFI will result in good model fit and support for metric invariance for (a) the three-factor model and (b) the single-factor model.

Building upon the support for H1 and H2, the researcher assessed scalar or strong invariance in H3. Assessment of scalar invariance is supported in the context of this study (i.e., assessing invariance between two statistically determined comparable generational cohort groups from the hospitality industry) by Schaufeli et al. whereas it was recommended that assessment of “invariance of the UWES should include similar occupational groups” (2006, p. 713). This future research direction was proposed by Schaufeli et al. (2006) after citing several studies where invariance for data from UWES had been found across groups (Schaufeli, Martinez, et al., 2002; Strom & Rothmann, 2003) and after discussing the limitations within the local study (i.e., the local study did not have comparable groups). Thus, H3 was set to hypothesize scalar invariance between Baby Boomer and Millenial cohorts for samples within the leisure and hospitality industry.

H3: Evaluation of model fit indices, $\chi^2$, and CFI will result in good model fit and support for scalar invariance for (a) the three-factor model and (b) the single-factor model.

**Latent Mean Difference Hypothesis**

For the same studies that report scores for engagement by various segmentations of the workforce, it is common to report mean differences between the latent constructs (e.g., job satisfaction, organizational commitment, and employee engagement) contained within the study. Reported differences between groups often allow researchers to
formulate recommendations for treatments or interventions within the context of the study (e.g., Chen & Choi, 2008; Chi et al., 2013; Park & Gursoy, 2012). Differences by cohort have been reported between Boomers and Millennials (Chen & Choi, 2008; Chi et al., 2013; Gursoy et al., 2008; Gursoy et al., 2013; Lub et al., 2011; Park & Gursoy, 2012), and specifically mean differences for engagement by generational cohort, with Millennials consistently scoring lower on engagement compared to Boomers (Kralj & Solnet, 2011; Solnet & Hood, 2008; Solnet & Kralj, 2010). Thus, once H1-H3 had been evaluated and if scalar invariance was supported (i.e., H3), H4 was set to hypothesize if there are statistically and practically significant latent mean differences between Boomers and Millennials.

H4: There will be statistically and practically significant latent mean differences for vigor, dedication, and absorption between the Boomer and Millennial groups for (a) the three-factor model and a statistically and practically significant latent mean difference for work engagement for (b) the single-factor model. The Boomers will be more engaged compared to the Millennials.

**Summary of the Chapter**

Chapter 2 reviewed the relevant academic and practitioner literature pertaining to the domains of generational cohorts and employee engagement. The leisure and hospitality supersector and the psychometrics of the UWES scale were included to provide context for the study. The chapter concluded with support for the four hypotheses contained within the study.
Chapter 3 - Methodology

Introduction

Chapter 3 described the methodology and design of the study. Sections of the chapter included the purpose of the study, measurement invariance overview, design of the study, research hypotheses, population and sample, instrumentation for the survey, survey design, data collection procedures, data analysis procedures (i.e., data cleaning, group cohort comparisons and sample representativeness, statistical assumptions, and measurement models), hypotheses testing, descriptive statistics, common method variance, and limitations. A summary of the chapter was also included.

Purpose of the Study

The purpose of the study was to empirically assess measurement invariance by generational cohort using data from the short version of the Utrecht Work Engagement Scale (UWES-9). The population under study was currently employed leisure and hospitality professionals in the United States. Generational cohorts included in the study were Boomers and Millennials. Using confirmatory factor analysis, configural, metric and scalar invariance assessments were tested and reported by generational cohort. Upon assessment of measurement invariance between the Boomer and Millennial groups, latent mean analysis was conducted to determine the latent mean difference of work engagement factors between groups.

Measurement Invariance Overview

The topic of factor structure equivalence was first discussed by Jöreskog (1971), and the concept of measurement invariance was later presented by Byrne, Shavelson, and
Muthen (1989). Social science researchers often utilize self-report questionnaires to collect data regarding individuals’ work and organizational perceptions. Questionnaires are developed to assess underlying phenomena in a consistent fashion and to compare them across time periods or groups. For these comparisons, “a questionnaire should measure identical constructs with the same structure across different groups” (Van de Schoot et al., 2012, p. 486). When this condition is met, the questionnaire is referred to as being measurement invariant (Van de Schoot et al., 2012). When measurement invariance is assessed and confirmed, the researcher can be confident that participants interpret the questions individually the same and also interpret the latent factor the same (Van de Schoot et al., 2012). Thus, “the establishment of measurement invariance across groups is a logical prerequisite to conducting substantive cross-group comparisons” (Vandenberg & Lance, 2000, p. 4).

Testing for measurement invariance includes three procedures, each with an increasingly higher refinement than the previous procedure. The first procedure is to compare reliability estimates for each group of the sample and for the total sample. Reliability estimates at the full sample level may mask measurement issues at the group level. When reliability estimates are calculated and evaluated at the group level, statistically significantly differences between groups may be observed (Nimon & Reio, 2012).

Conducting an exploratory factor analysis (EFA) is the second procedure for measurement invariance testing. This procedure involves performing an EFA and evaluating the pattern and structure coefficients and eigenvalues for differences. Factorial invariance is then assessed by calculating several indices, including the Satisf
Similarity Index (s), the coefficient of congruence (CC), and the correlation between pattern coefficients (rpf) (cf. Brauchle & Azam, 2004; Reynolds & Harding, 1983, Nimon & Reio, 2012).

Confirmatory factor analysis is the third procedure. The procedure starts by fitting a CFA model to each group and evaluating the resulting fit indices. Kline (2016) suggested the following criteria for model evaluation: (a) the comparative fit index (CFI) $\geq .95$; (b) the standardized root mean square residuals (SRMRs) $\leq .08$; and (c) the root mean squared error of approximation (RMSEA) $\leq .10$. The procedure involves four progressively hierarchical levels of assessment: configural invariance, metric (weak) invariance, scalar (strong) invariance, and strict invariance, with each level being required before evaluating the next level in the hierarchy (Meredith, 1993; Van de Schoot et al., 2012). The series of models resulting from the procedure are reported with their corresponding fit indices, and the results of measurement invariance are interrogated by reviewing change in chi-squared tests ($\chi^2$) and change in CFI (CFI) tests for the hierarchical models (Cheung & Rensvold, 2002; Nimon & Reio, 2011; Vandenberg and Lance, 2000). Vandenberg and Lance (2000) through an extensive measurement invariance literature review recommended establishing configural, metric, and scalar invariance before making group comparisons. The Vandenberg and Lance (2000) recommendations were supported in research by Cheung and Rensvold that evaluated the goodness of fit indexes for measurement invariance testing (2002). Strict invariance was determined to be excessively rigorous and related to construct level variance (Byrne, 2001; Cheung & Rensvold, 2002; Teo et al., 2009).
Configural Invariance

Configural invariance is the first of the four procedural steps for CFA measurement invariance assessment and reveals if respondents from groups use the same conceptual framework to respond to scale items (Cheung & Rensvold, 2002). The groups under study (e.g., the Boomer and Millennials groups as in the case with the present study) are fitted to the same factor model. Evidence of configural invariance is confirmed if the groups fit the model the same (Nimon & Reio, 2012). If configural evidence is found, the same construct is being measured across groups (Wu, Li, & Zumbo, 2007; Rusticus, Hubley, & Zumbo, 2008).

Metric (Weak) Invariance

Metric or weak invariance is the second level of the four step procedure for CFA measurement invariance assessment and “tests whether respondents across groups attribute the same meaning to the latent construct” (Van de Schoot et al., 2012, p. 488). The procedure starts with taking the model from the configural invariance step and constraining all factor loadings for like items to be equal but allowing the intercepts to differ between groups. The fit indices of the model are compared with the configural model fit indices (Nimon & Reio, 2012). If metric invariance is confirmed, then correlations across groups can be compared (Wu et al., 2007).

Scalar (Strong) Invariance

Scalar or strong invariance is the third level of the procedure and “implies that the meaning of the construct (the factor loadings) and the level of the underlying items (intercepts) are equal in both groups” (Van de Schoot et al., 2012, p. 488). The strong invariance model has the factor loadings and intercepts constrained to be equal for like
items. The fit indices of the scalar model are compared to the fit indices of the metric invariance model (Nimon & Reio, 2012). If evidence is found for scalar or strong invariance, group mean comparisons are possible (Wu et al., 2007).

**Strict Invariance**

Strict invariance is the fourth level of the CFA procedure. Strict invariance has been determined to be related to construct level invariance instead of measurement invariance (Cheung & Rensvold, 2002) and is considered to be excessively rigorous (Byrne, 2001; Teo et al., 2009). Testing for strict invariance uses the scalar invariance model (where factor loadings and intercepts are fixed for like items) and also constrains the residual variances to be equal for like items (Nimon & Reio, 2012). The results of the strict invariance model are compared to the scalar invariance model. Strict invariance means “the latent construct is measured identically across groups (Van de Schoot et al., 2012, p. 488).

**Design of the Study**

This study was a cross-sectional quantitative research design. Survey methodology was utilized to collect data at a single time period, being operationalized through the Qualtrics® online survey platform. Qualtrics® panel services were used to recruit a prescreened sample of individuals meeting the criteria of the study. At the highest definition, the individuals were Baby Boomers and Millennials working in the United States full-time within the leisure and hospitality industry. Qualtrics® services were used to handle all communications and survey distribution with the participants. Survey participants were asked to complete a survey that insured their anonymity with the ability to opt out at any point. The survey was designed using previously validated
measures. The measures used in the study included the short version of UWES (Schaufeli & Bakker, 2006), Attitudes Toward the Color Blue (ATCB) scale (Miller & Chiodo, 2008), and MBI-GS (Schaufeli et al., 1996). The survey also included prescreening questions, and demographics, and job characteristics.

After the data were collected from the single survey period, the data were cleaned and assessed for statistical assumptions. The demographics and job characteristics were used to evaluate the representativeness of the sample to the population and to determine if the samples (i.e., Boomers and Millennials) were comparable. Data analysis was conducted by means of maximum likelihood structural equation modeling and confirmatory factor analysis.

**Research Hypotheses**

A total of four research hypotheses were tested in this study, and Chapter 2 provided support for each of the four hypotheses. Within the context of previous work engagement research that considers differences between generational cohorts, studies consistently report mean differences between Boomer and Millennial cohorts by work engagement, albeit with variation by the cohorts and context under consideration (e.g., Bano, Vyas, & Rohini, 2015; Hoole & Bonnema, 2015; Kralj & Solnet, 2011; Martins & Ledimo, 2016; Solnet & Hood, 2008).

Stemming from the multidimensional theory and three dimensions of burnout, the original development of UWES-17 (Schaufeli et al., 2002) and subsequent shortened UWES-9 (Schaufeli et al., 2006) instruments defined work engagement as containing the three dimensions of vigor, dedication, and absorption. The literature holds that researchers have used UWES-9 as both a three-factor model (vigor, dedication, and
absorption) and as a total nine-item single indicator to avoid issues with multicollinearity (Schaufeli et al., 2006). The literature review conducted by the researcher confirmed the varied use (i.e., three-factor and single-factor) of UWES-9 scores to measure work engagement. Given the varied use of UWES-9 in the literature, the researcher evaluated the measurement invariance between Boomers and Millennials in both the three-factor and single-factor models of work engagement.

To assess the measurement invariance between Boomer and Millennial generational cohorts using a three-factor and single-factor UWES-9 models, the following three hypotheses were proposed:

H1: Evaluation of model fit indices will result in good model fit (i.e., TLI ≥ .95, CFI ≥ .95, RMSEA ≤ .10, and SRMR ≤ .08) and support for configural invariance for (a) the three-factor model and (b) the single-factor model.

H2: Evaluation of model fit indices, $\chi^2$, and CFI will result in good model fit and support for metric invariance for (a) the three-factor model and (b) the single-factor model.

H3: Evaluation of model fit indices, $\chi^2$, and CFI will result in good model fit and support for scalar invariance for (a) the three-factor model and (b) the single-factor model.

To assess the latent mean differences of work engagement between Boomer and Millennial generational cohorts, the following hypothesis was proposed:

H4: There will be statistically and practically significant latent mean differences for vigor, dedication, and absorption between the Boomer and Millennial groups for (a) the three-factor model and a statistically and practically significant latent
mean difference for work engagement for (b) the single-factor model. The Boomers will be more engaged compared to the Millennials.

**Population and Sample**

A sample frame is a subset of the population and constitutes “those people who have a chance of being included” in the sample (Fowler, 2014, p. 15). Since the study sourced participants using Qualtrics® Panel services, only those individuals who had the opportunity to be included on the Qualtrics® Panel recruitment were within the sample frame of the study. Research panel companies, such as Qualtrics®, maintain “panels of individuals, pre-recruited via a probability-based sampling methodology, from which sub-samples can be drawn according to a researcher’s specification” (Fricker, 2008, p. 204). According to Pollard (2002), research panels have higher response rates than traditional probably sampling methods, and samples can be demographically representative for the intended research criteria. This route of obtaining survey participants is optimized for online surveys, and it is valid for researchers “who require a sample that can be generalized to populations” (Fricker, 2008, p. 204). Therefore, the population of the study consisted of those individuals who had the chance to be recruited by Qualtrics® when the pre-recruited panel was created. The sample frame for the study consisted of those individuals within the pre-recruited panel. The sample was selected from the sample frame using criteria provided by the researcher to Qualtrics.

The sample to be studied consisted of individuals working in selected sub-sectors of the leisure and hospitality supersector. According to the U.S. Bureau of Labor and Statistics, the leisure and hospitality supersector is part of the service-providing industries and includes approximately 15.9 million workers as of June 2017 (BLS, 2017).
Limitations were placed on sub-sectors of the leisure and hospitality supersector, meaning only the hospitality sub-sectors of the supersector were considered, including the convention center, cruise, gaming, lodging, marina, sporting facilities, travel, and tourism sub-sectors (PWC, 2017). The sample excluded other leisure and hospitality sub-sectors that are focused solely on food preparation and food services such as food and beverage (i.e., restaurants, fast food establishments, grills, and drinking establishments) and catering services. As the intent of the study was to evaluate comparable groups, the exclusion of food oriented industries served to provide more comparable groups by occupation and organizational functional area, even considering that some hotels, for example, have food preparation services as part of their operations. The sample consisted of those individuals within the sample frame who are currently employed full-time, to only those individuals in the Millennial or Boomer generations, to those who are living and working in the United States, those who have been working at their place of employment for 5 years or less, and to those who speak English.

Akin to convenience sample concerns, the representativeness of the sample was evaluated once the sample was collected. One method to evaluate the representativeness of the sample is to “compare the sample demographic profile with that of the population” (Kline, 2009, p. 68). To determine the representativeness of the sample to the leisure and hospitality supersector, the researcher compared the sample’s demographic profile to a demographic profile of the leisure and hospitality supersector sourced from the Bureau of Labor Statistics (BLS, 2017). The “Group Cohort Comparison and Sample Representativeness” section in Chapter 3 contained a detailed demographic analysis of
the BLS statistics and an outlined method (Kline, 2016) to perform a comparison of the collected sample data and BLS demographic statistics.

In terms of the general national population, prior research comparing three commercial platforms (i.e., Survey Monkey®, Qualtrics®, and Mechanical Turk®) was conducted to determine the representativeness of samples obtained from each platform. Demographics collected from each national sample frame were compared to the 2010 national census data. Qualtrics® “yielded the lowest average discrepancy rate across categories of acquired demographic characteristics” (Heen, Lieberman, & Miethe, 2014, p. 6).

To assure the sample was aligned to the intended leisure and hospitality sub-sectors, several screening steps were utilized. First, the panel criteria (Boomers, Millennials, convention center, cruise, gaming, lodging [hotel or accommodation], marina, sporting facility, travel, and tourism [entertainment venue or museum] sub-sector employees, full-time workers, living and working in the U.S., tenure of 5 years or less at their place of employment, and English speaker) were provided to Qualtrics®. Qualtrics® used the provided criteria to select potential participants from the pre-recruited panel. A second screening pass was conducted by applying several qualifying questions at the beginning of the survey. Through personal exchanges with Qualtrics® and the researcher, Qualtrics® recommended including the use of screening questions to insure high quality of response data. The qualifying questions were used to screen participants as to whether or not they matched the intended sample criteria. Appendix A contains the survey and associated screening questions. Participants who passed the qualifying questions were allowed to participate in the full survey. Those participants who did not pass the
qualifying questions were not able to participate in the full survey. Layers of demographics were collected for statistical analysis relevant to the study and to evaluate the representativeness of the sample.

**Job Functional Level**

Since the sample was collected from a cross-section of the leisure and hospitality supersector, a process to appropriately categorize employees by their level of work was required. If the sample was collected within a narrower sample frame, such as only from hotel front desk clerks from corporately owned, non-unionized, mid-sized, U.S. based hotels in the Southwestern region of the U.S., the researcher would be reasonably confident that the level of work performed by respondents would be similar regardless of organizational association.

In light of the sample frame that was used for the present study, the researcher triangulated information from empirical literature, review literature, and practitioner based sources to determine a reasonable categorization for representative job levels associated across the sub-sectors. Select empirical literature of the leisure and hospitality industry classifies job levels at either two or three levels. For example, non-supervisor and supervisor/manager (Kralj & Solnet, 2011; Solnet et al., 2012), manager and supervisor (Chen & Choi, 2008), manager and line-level (Chi et al., 2013), coworker, supervisor, and manager (Kim et al., 2009), and non-supervisor, supervisor, and manager (Solnet & Kralj, 2010). Consistent across the selected examples is the concept of employees who manage or supervise others and those who do not.

The literature provided insight into the hospitality industry’s workforce and skillsets that are required to perform the work. According to Baum, “there is little that is
unique about hospitality skills … and … most of the skills that are employed within the sector also have relevance and application to other sectors of the economy” (2002, p. 346). Baum continued by stating it is problematic to rely upon common job titles as they can “certainly mask a very different range of responsibilities, tasks, and skills within jobs in different establishments” (2002, p. 347). Baum recommended using a vertical hierarchy, rather than specific job titles, of job functional level classifications as the hierarchy accounts for organizational diversity such as size of the organization, location, and ownership (2002). Riley (1996) provided the following hierarchy with associated proportions for the hospitality industry: (a) operative (semi-skilled and unskilled) 64%, (b) craft (skilled) 22%, (c) supervisory 8%, and (d) managerial 6%. Riley’s (1996) hierarchy provides a slightly more granular view of the job levels (i.e., four levels) compared to those found in the selected empirical literature containing two or three levels (e.g., Chen & Choi, 2008; Kralj & Solnet, 2011; Solnet et al., 2012). Also noteworthy is the alignment of the proportions to the BLS. For the hospitality industry, the BLS reports supervisory jobs at 16.5%. The combined supervisory (8%) and managerial (6%) levels account for 14% in Riley (1996). The similar proportions found between Riley’s (1996) hierarchy and the BLS is noteworthy.

The American Hotel & Lodging Association (AHLA) is the sole national association representing all segments of the industry, and the American Hotel & Lodging Educational Institute (AHLEI) is a subsidiary of AHLA and is the sole professional certifying body of the hospitality industry (AHLA, 2017; AHLEI, 2017). According to the AHLEI’s A World of Possibilities brochure, the hospitality industry contains three levels of employees (AHLEI, 2017). First are entry level employees that require little or
no education and work experience. Next are skilled level employees who have moved up from entry level positions and may have some relevant education but do not supervise other employees. Last are managerial level employees. The employees in these positions are generally college educated and have responsibilities for other employees. The managerial level employees are executives, managers, division heads, and supervisors. The AHLEI offers professional certification at six different levels: hotel administrator, department head, line employee, managerial, specialty, and supervisor (AHLEI, 2017). The line level certification and specialty level certification represent skilled labor positions rather than entry level positions. The hotel administrator, department head, managerial, and supervisor level certifications represent the managerial level employees. The AHLEI does not offer certifications for unskilled labor positions (AHLEI, 2017).

Based upon the triangulation of information and literature in the prior three paragraphs, the researcher included a job characteristic survey question for the job level categorization at the following four levels: (a) entry level, (b) non-supervisory, (c) supervisory, and (d) managerial. The categorization represents a blended job level to account for the empirical literature, review literature, and practitioner resources. The collected job characteristics also enabled the aggregation of data to supervisory and non-supervisory for comparison to the BLS reported statistics.

**Sample Size**

The targeted sample size was a factor of the design of the survey. The survey consisted of a total of nine items used to capture the full spectrum of work engagement factors and eight items for the attitudes toward the color blue. As this study included CFA to assess measurement invariance between Boomer and Millennials cohorts, the
sample size must be sufficient from the statistical analysis aspect and also sufficient for the general study. From the statistical analysis perspective, Henson and Roberts (2006) recommended a minimum ratio of 10:1 when conducting confirmatory factor analysis (CFA). Therefore, a minimum of 340 respondents was required, 170 for Boomers and 170 for Millennials. From the general survey perspective, a minimum sample size of 210 was determined following the guidelines found in Wolf et al. (2013). Wolf et al. provided guidelines for establishing sample size based upon the “number of indicators and factors” and the “magnitude of factor loadings” (2013, p. 913). Factor loadings for CFAs were classified at the .50, .65, and .80 levels. The UWES has nine indicators loading on three first-order factors, and the ATCB has eight indicators loading on a single first-order factor. Through the cited literature, factor loadings were all near or above the .80 threshold. Based on those parameters, Model B (Figure 3) from Wolf et al. (2013, p. 922) can be used to determine 150 respondents are required for UWES and 60 respondents are required for ATCB, for a total of 210 respondents. Since the researcher considered Boomers and Millennials to be separate samples, a total of 420 responses were required. The researcher assumed a volume of responses would be disqualified for various reasons when conducting the data cleaning and completeness checks. Therefore, a total sample size collection target of 460 was set to insure sufficient generational cohort sample volumes.

Measurement Instruments

As the study was primarily concerned with the assessment of measurement invariance between generational cohorts for work engagement, the UWES-9 scale that measures work engagement was used (Schaufeli et al., 2006). Appendix D contains the
full UWES-9 scale. The UWES-9 scale (Schaufeli et al., 2006) is a well-established scale (Hoole & Bonnema, 2015) and is the most common engagement scale utilized within employee engagement research (Saks & Gruman, 2014). Contextually, prior research supports using UWES-9 for work engagement in the leisure and hospitality supersector (Kim et al., 2009; Park & Gursoy, 2011; Solnet & Kralj, 2011). The UWES-9 is a nine item scale on a seven point Likert scale ranging from 0 (“never”) to 6 (“every day”) (Schaufeli et al., 2006). Example UWES-9 questions are “at my job, I feel strong and vigorous” and “I am enthusiastic about my job” (Schaufeli et al., 2006, p. 714). The UWES-9 scale measures work engagement through the factors or vigor, dedication, and absorption (Schaufeli et al., 2006). Detailed descriptions of the UWES-9 instrument subscales of vigor, dedication, and absorption are provided below.

Vigor was measured using a three item scale on a seven point Likert scale ranging from 0 (“never”) to 6 (“every day”). Reported reliability from empirical studies (Kim et al., 2009; Park & Gursoy, 2012) in the U.S. hospitality industry ranged from .73 to .80. Dedication was measured using a three item scale on a seven point Likert scale ranging from 0 (“never”) to 6 (“every day”). Reported reliability from empirical studies (Kim et al., 2009; Park & Gursoy, 2012) in the U.S. hospitality industry ranged from .83 to .87. Absorption was measured using a three item scale on a seven point Likert scale ranging from 0 (‘never’) to 6 (‘every day’). Reported reliability from empirical studies (Kim et al., 2009; Park & Gursoy, 2012) in the U.S. hospitality industry was .74 for both studies.

A latent marker variable (LMV) was included to assess CMV (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Attitude Toward the Color Blue (ATCB) is an eight item scale based upon a 7-point Likert scale ranging from 1 (‘strongly disagree’) to
7 (‘strongly agree’). Appendix D contains the full ATCB scale. Example ATCB questions are “I prefer blue to other colors” and “I like blue clothes” (Miller & Chiodo, 2008). The ATCB has received support for use in detecting CMV in organizational research (Simmering, Fuller, Richardson, Ocal, & Atinc, 2015). Reported reliability of coefficient alpha ranged from .76 to .85 when tested across samples (Simmering et al., 2015).

The researcher collected additional data from an established measurement instrument (MBI-GS; Maslach et al., 1996) for future research, but the additional data was not analyzed as part of the present study. To avoid issues of response processes and common method variance, it is recommended to include “a temporal, proximal, or psychological separation” (Podsakoff et al., 2003, p. 888). Therefore, the UWES-9 responses were collected first, the ATCB second, MBI-GS third, and demographics last.

**Survey Design**

The survey for the study was designed and developed by the researcher using the Qualtrics® survey interface. The survey is included in Appendix A. The survey included a hyperlink that links to the online survey accessible by commonly available Internet web browsers. The unique hyperlink was only available on the survey and was not publically available or available through any other channel. The single survey contained all of the screening questions, work engagement items, latent marker variable items (i.e., ATCB), instrument for future research (i.e., MBI-GS), demographic variables, and job characteristic variables.

Fan and Yan (2010) suggested only relevant topics be presented to potential participants to increase survey response. Given the participants were a pre-recruited
Qualtrics® panel, the survey topic (i.e., a work perception survey) was relevant to the participants. Fan and Yan (2010) also suggested using academic sponsorship to increase confidence and response rates. To meet this suggestion, banners on the opening page of the survey and on subsequent pages prominently displayed The University of Texas at Tyler colors, logos, and name. To avoid missing data issues (Wolf, Harrington, Clark, & Miller, 2013), all questions were setup as forced responses.

Vandenberg (2002) discussed the susceptibility of measurement invariance testing to such concerns as the local nomological net, sample size, multivariate normality, and common method variance. Preliminary findings by Williams (2001) were discussed in Vandenberg (2002) who suggested common method variance “does affect the interpretability” of measurement invariance analysis (p. 148). Beyond the scope of the present study, Vandenberg (2002) continued by suggesting extensive additional research was needed to address the sensitivity issues. Vandenberg provided direction for reducing sensitivity concerns by focusing measurement invariance research to using “one survey, one construct” (Vandenberg, 2002, p. 147). The present study followed Vandenberg’s suggestion by focusing the measurement invariance assessment to the single construct of UWES-9.

Even considering the focused design of the survey, common method variance was still considered in the study as it is considered as “one of the main sources of measurement errors” (Podsakoff et al., 2003, p. 888). The survey design methodology attempted to mitigate sources of common method variance through six items. First, the next button and status bar features of Qualtrics® were used. Features such as these are suggested to help limit common method variance (Romano & Chen, 2011). Second,
work engagement items were collected before the MBI-GS, demographic, and job characteristic variables to limit “retrieval cues” and avoid potential bias the respondents may have to the generational aspect of the research (Podsakoff et al., 2003, p. 888). Third, participant anonymity was guaranteed in the opening statements on the survey landing page, and participants had the ability to opt out of the survey at any point. Fourth, the survey visuals were kept simple to reduce evaluation apprehension. By implementing a matrix design using Likert scales for each item and simple selection choices for the demographic and job characteristic variables, participants had a simple and non-threatening survey experience (Dillman, 2007; Fowler, 2014). Participants were also informed that there were no correct or incorrect answers, to further produce a non-threatening survey experience (Dillman, 2007). Fifth, to reduce potential survey fatigue, the survey’s estimated completion time was approximately 5 to 7 minutes. A reasonably short survey time has been suggested to limit survey fatigue (Dillman, 2007). Last, a latent marker variable was utilized (Williams, Hartman, & Cavazotte, 2010). The ATCB was originally developed by Miller and Chiodo as a marker variable to capture respondents attitudes as related to the color blue (2008). The ATCB has been cited as being an ideal marker variable (Richardson, Simmering, & Sturman, 2009; Simmering et al., 2015).

Although Qualtrics® provided the pre-recruited panel, the participants were reminded that the survey was completely voluntary and that they could opt out at any time. Participants were asked several initial screening questions before being asked to provide their consent to participate in the survey. The intentions of the screening questions for this survey were multi-fold and with the intention to capture the intended
sample. First, participants were asked if they were currently employed in the United States. Second, participants were asked if they were living in the United States. Although it was assumed the participants are English speakers since they have been pre-recruited by Qualtrics®, all communications were in English, and the survey was only in English, a third qualifying question was asked to confirm the participants were English speakers. Fourth, participants were asked if they are working in specific leisure and hospitality sub-sectors. Convention center, cruise, gaming, lodging (hotel or accommodation), marina, sporting facility, travel, and tourism (entertainment venue or museum) were provided as example industries. Fifth, a question asked if the participants had been working at their current place of employment for 5 years or less. Sixth, participants were asked to identify to which generational cohort they belong. Finally, participants selected from a range of hours (i.e., not currently employed, 1 to 19, 20 to 34, 35 to 45, and 46 or more) that represents the average hours they work per week. A negative response to the screening questions automatically sent the participants to a closing screen that informed them of their ineligibility for the survey and thanked them for their time.

Two common survey design considerations were not applicable to the present study. First, the survey for the present study was reasonably short (5 to 7 minutes) since the survey only included screening questions, the nine work engagement items, the latent marker items, MBI-GS items, the demographics, and job characteristic variables. Second, automated responders or “bot” responders are unfortunately common when using services such as Amazon Mechanical Turk (MTurk®) to source survey participants. The
survey for the present study was deployed by Qualtrics® to a known, pre-recruited panel of participants; thus, the issue of bot responses was not a concern.

Questions asking for respondents’ demographics and job characteristics were included as part of the survey. The demographics collected were gender, generational cohort, race, and education level (Beattie, Kim, Hagen, Egan, Ellinger, & Hamlin, 2014). Job characteristics collected were average hours worked per week, job functional level, their industry within the hospitality sector, tenure at their place of employment, and if they have direct customer contact (Kim & Kuo, 2015). A listing with all of the categories for the variables is available in Appendix C. The variables were used to evaluate the representativeness of the sample to the population and to confirm the sample met the intended sample criteria of the study. The variables included in the survey either directly matched the BLS variables (e.g., gender and race) or were of a lower level of detail (e.g., working hours, job functional level), which allowed for aggregation and summary to compare with BLS level details.

Data Collection

Since the study used Qualtrics® Panel services and allowed for complete voluntary participation, no prior permissions were obtained by the researcher except for Institutional Review Board (IRB) approval through The University of Texas at Tyler. The IRB documentation was submitted upon the committee’s approval of the dissertation proposal. The survey was conducted only after IRB approvals.

All data were collected digitally using an online survey developed by the researcher and provisioned through Qualtrics®. Research has suggested that pre-notification to participants is an effective way to increase response rates when using
online surveys (Baruch & Holtom, 2008). An employee at Qualtrics® sent a pre-notification e-mail one week before the live survey launch and was to distribute the live survey to the Qualtrics® panel by e-mail on a Monday between 12:00pm and 1:30pm local time of the participant. Research suggests response rates for surveys are higher when administered on Mondays during this time period (McCormick & Zheng, 2013). The survey was to remain open until the targeted quantity of responses was obtained, which was estimated to take approximately 2 weeks. The Qualtrics® employee sent a follow up e-mail 1 week after the survey invitation e-mail. All communications to the panel was solely through Qualtrics®, providing a high level of confidentiality for the survey participants. The researcher constructed templates and content for the e-mails Qualtrics® used. Templates for the one-week-prior survey notification, survey invitation, and reminder e-mails are found in Appendix D. Confidentiality was further enhanced in that the data available to the researcher did not contain personally identifiable information of the survey participants. For the study, the researcher paid Qualtrics® directly a flat rate for the number of responses, and Qualtrics® paid respondents directly for successfully completing the survey.

**Data Analysis**

The data analysis section of the study specified the various statistical analyses required for the project. The section covered three areas: (a) data cleaning, (b) group comparison and sample representativeness, (c) statistical assumptions. The section ended with a summary.
**Data Cleaning**

The quality of the response data was critical for the study, inasmuch as invalid or incomplete data could have led to inconclusive or inaccurate findings. The data were collected during a single time period, and it was collected within the online Qualtrics® survey application. Once the survey period was closed, the data were downloaded to the researcher’s local computer. The researcher reviewed the data for patterns, trends, and missing data. The visual check informed the researcher what programmatic data cleansing efforts were required. After the initial visual check was complete, the researcher followed a documented data cleansing plan of action. The researcher utilized the free statistical software of R for all data cleaning activities that required programmatic manipulation or evaluation (R Development Core Team, 2017).

A comprehensive data cleaning process was undertaken, and the data were evaluated for responses that should be removed from the sample. In order to qualify for retention, responses were checked and had to pass all qualifications for complete responses, missing data, range of values, survey duration, and straight lining. Responses where there was an obvious incomplete response due to a participant not completing the survey in full or due to a technology error resulting in a partial response were removed. As each question and demographic variable was a forced response, each individual variable was interrogated for missing data within each field of the sample. If a response was found to have missing data, the entire row was removed from the sample.

Once the two prior steps confirmed all rows of the sample were complete and did not have any missing data, each variable was evaluated for the range of values. The minimum and maximum values for each field were calculated and reviewed. Responses
with data falling outside of the defined minimum and maximum values (e.g., UWES-9 responses will range from “0” to “6” and ATCB items will range from “1” to “7”) for the specific variable were removed. For those variables with categorical data, an inventory calculation reported the complete distribution of counts by variable. If any response was found to have an invalid data point for a variable (e.g., gender will have a distribution of either male or female), it was removed from the sample.

The survey duration was tested once the researcher was finished with the completeness check, missing data check, and range of values checks. According to the Qualtrics® survey tool, the survey required 5 minutes to complete. A pilot test of the survey by the researcher showed the survey required between 2 and 4 minutes to complete. Therefore, responses under 2 minutes were removed from the sample. Responses with durations between 2 and 4 minutes and those with over 4 minutes were retained. These cases were deemed correct when accounting for the individual participants’ abilities and capacities for completing the survey.

The UWES-9 scale contained all positively worded items, and the ATCB scale contained positively and negatively worded items, making the ATCB an ideal straight line test for respondent inattentiveness within this study (Cole, McCormick, & Gonyea, 2012). Straight line testing was not considered for the UWES-9 items, as it was possible a respondent could accurately and honestly respond with valid straight line answers, given the instrument contained all positively worded items. Straight line testing was conducted for ATCB, as the measure contains half positively worded items and half negatively worded items. If a response contained a straight line (i.e., all ATCB items are coded the same), the response were removed.
No items of the work engagement items required reverse coding to accommodate for negatively worded questions. The ATCB did contain four negatively worded items. The scores for these four ATCB items were recoded with a simple algorithm (i.e., \[8 – \text{score}\]) programmed in R. Once the data cleaning operations were complete, the remaining data were ready to be analyzed.

**Group Cohort Comparison and Sample Representativeness**

The central tenet of the study was to assess measurement invariance between Boomers and Millennials for data from UWES-9. The UWES-17 had been found factorially invariant when tested across countries (Schaufeli, Martinez, et al., 2002) and by racial groups (Storm & Rothmann, 2003). Yet, the UWES-9 scale was found to not be factorially invariant when tested across countries (Schaufeli et al., 2006). This difference in findings was addressed by Schaufeli et al. (2006) whereas the respondents were within the same occupational groups when invariance was found (Schaufeli, Martinez et al., 2002; Storm & Rothmann, 2003) and were in different occupational groups when invariance was not found (Schaufeli et al., 2006). Guidance was provided that future invariance research “of the UWES should include similar occupational groups” (Schaufeli et al., 2006, p. 713).

**Group Cohort Comparison**

Following the research recommendation found in Schaufeli et al. (2006), the present study focused to convention center, cruise, gaming, lodging (hotel or accommodation), marina, sporting facility, travel, and tourism (entertainment venue or museum) sub-sectors to obtain respondents working in similarly focused organizations and occupations. As a further step, the samples of Baby Boomers and Millennials were
further interrogated to determine whether or not the samples were statistically
significantly and practically significantly different by the collected demographics and job
characteristics. The goal of the study was to have samples that differ solely by their
generational association. The samples should not be statistically significantly different at
the $p < .05$ level and practically significantly different at the Cramer’s $V < .10$ level due
to imbalances in the demographic or job characteristic variables.

Following the guidance found in Van de Vijver and Leung (1997), the study
considered the two recommended methods to collect and subsequently affirm comparable
groups. First, the study utilized a sampling method that controlled for several of the
potential group differences (i.e., currently work in the United States, currently live in the
United States, primarily speak English at work, 5 years or less of organizational tenure,
full-time employees, and work within the aforementioned leisure and hospitality
industries). This primary method of controlling for group differences, deemed “matching
of subjects” by Van de Vijver and Leung, provided for samples that were “as similar as
possible in their demographic variables” when considering the context of the study (1997,
p. 45). Only subjects who fit the predefined demographic profile of the study were
sampled for the study (Van de Vijver & Leung, 1997).

A series of Pearson’s chi-squared tests were used since the sample definitions
(Boomers and Millennials) could be used as the grouping variable and the related
demographic and job characteristic variables were categorical data. The series of chi-
square tests was conducted on the demographic variables of education level, gender, and
race. Chi-square testing was also performed on the job characteristic variables of
organizational tenure, sub-sectors of convention center, cruise, gaming, lodging (hotel or
accommodation), marina, sporting facility, travel, and tourism (entertainment venue or museum), direct customer contact, average working hours per week, and by job functional level (i.e., entry level, non-supervisory, supervisory, and manager) identified in Chapter 2. Since the results of the chi-square analysis returned negative results and because the samples were statistically and practically significantly different for multiple variables, the researcher utilized propensity score analysis to better balance the samples using the observed covariates (i.e., the aforementioned demographic and job characteristic variables) (Rubin, 1997).

Qualtrics® provided the functionality to manage a soft launch of the survey to provide researchers with an early gauge on the quality of data being collected. The soft launch was managed by Qualtrics® to collect up to 10% of the overall requested quantity of responses. Once a 10% volume was reached, the survey was to be paused and the researcher was to evaluate the collected data by generational cohort as described above. If the evaluated 10% volume yielded a negative outcome (i.e., the Boomers and Millennials were different by demographics and job characteristics), the researcher was to request Qualtrics® to refine the panel selection criteria. Qualtrics® would un-pause and hard launch the survey using the refined panel selection criteria.

**Sample Representativeness**

To determine the representativeness of the sample to the intended population, the pooled sample demographic statistics were compared to equivalent BLS demographic statistics. The BLS provided statistics specifically for the portions of the hospitality industry that do not include the food oriented business (i.e., accommodations sub-sector and arts, entertainment, and recreation sub-sector), which somewhat corresponded to the
sample data collected for the present study. The BLS provided overall statistics for supervisor and non-supervisor levels of work. For the both hospitality sub-sectors defined by the BLS, supervisor positions represent 16.5% and non-supervisory positions represent 83.5% of the workforce. The pooled sample’s job functional level variable was aggregated to match the supervisor/non-supervisor dyad and the sample’s percentages for supervisor/non-supervisor status were compared to the BLS percentages following Kline (2016).

Other demographics provided by the BLS were race and gender. Converse to the similarities found in the job level, the sub-sectors differed slightly by race and gender. The pooled sample was segmented by the collected hospitality industry of work variable to equate to the two BLS sub-sectors (i.e., accommodation sub-sector and arts, entertainment, and recreation sub-sector). Once the pooled sample was segmented, the percentages by race and gender were compared to the representative demographic percentages provided by the BLS following Kline (2016). For the accommodation sub-sector, women were reported at 55.5% and race was split by 13.8% black or African American, 8.9% Asian, and 28.4% Hispanic or Latino. For the arts, entertainment, and recreation sub-sector, women were reported at 45.2% and race was split by 10.3% black or African American, 4.4% Asian, and 13.2% Hispanic or Latino. For the overall BLS demographic statistics for the entire leisure and hospitality super-sector (including the food and beverage sub-sector), women were reported at 50.6% and race was split by 12.6% black or African American, 6.5% Asian, and 22.9% Hispanic or Latino.
Statistical Assumptions

The IBM® SPSS AMOS version 22.0.0 software package was used for the statistical analysis and structural equation modeling (SEM). “Given that SEM is based on the analysis of covariance structures”, a covariance matrix was used for the present study (Byrne, 2010, p. 101). The statistical analysis included maximum likelihood estimation procedures for SEM, and the statistical assumptions of multivariate normality and multivariate outliers were tested (Byrne, 2010; Kline, 2016). Multivariate normality was assessed using the guidelines (i.e., critical ratio > 5.00 is indicative of nonnormality) proposed by Byrne (2010) and Kline (2016). Based upon the squared Mahalanobis distance ($D^2$), the data were evaluated for multivariate outliers (Kline, 2016). According to Byrne, a multivariate “outlying case will have a $D^2$ value that stands distinctively apart from the other $D^2$ values” (2010, p. 106). If the data do not pass the statistical assumptions testing, remedies found in Byrne (2010) and Kline (2016) would be followed. One such remedy is bootstrapping and the comparison of bootstrapped and non-bootstrapped results (Kline, 2016). If needed, due to a failure of multivariate normality, bootstrapping would be performed using a 2,000 record sampling technique (Kline, 2016). The results would be evaluated and if the bootstrapped and non-bootstrapped resulted were found to be insignificantly different; the non-bootstrapped results would be reported. Cases with missing data would have been removed during the data cleaning stage of the analysis and would not be a concern during the statistical analysis phase.
Hypotheses Testing

Confirmatory factor analysis was used in the present study for assessment of measurement invariance between the Boomer and Millennial generational cohort groups. Configural, metric, and scalar measurement invariance were assessed for hypotheses testing through both the three-factor (H1a - H3a) and single-factor (H1b - H3b) models of work engagement. Following the general outlines in Nimon and Reio (2011), Van de Schoot, Lugtig, and Hox (2012), and Teo et al. (2009), the analyses began by setting measurement models and then proceeded into testing for measurement invariance. If scalar measurement invariance held (H3), a latent mean analysis was to be conducted and used for the assessment of the three-factor (H4a) and single-factor (H4b) models of work engagement.

Measurement Models

The sequence of measurement model testing was determined by examining the literature for empirical studies that tested the factorial structure of UWES-9 and subsequently tested for measurement invariance. Given the UWES-9 instrument is known to have high multicollinearity (Schaufeli et al., 2006), the literature was also investigated for guidance on handling other instruments with high levels of multicollinearity (e.g., Multifactor Leadership Questionnaire [MLQ]).

Antonakis, Avolio, and Sivasubramaniam (2003) provided support for beginning the CFA measurement model process with a pooled sample of data, rather than for independent samples, for an instrument known to have high levels of multicollinearity (i.e., MLQ). The reasoning behind this starting point was that much of the prior MLQ literature was contextually centered and often conflicted regarding the factorial structure.
of the instrument (Antonakis et al., 2003). Multicollinearity was identified as a confounding problem in determining the factorial structure within some studies (Antonakis et al., 2003).

The multicollinearity of UWES-9 was identified within the instrument’s developmental article, and a future research suggestion was that the instrument could be utilized as either a single-factor or three-factor solution (Schaufeli et al., 2006). Later validation studies of UWES-9 continued to echo the same regarding the factorial structure of UWES-9 and stated “research has still to gain a clear understanding of whether VI, DE, and AB have at least partially different nomological networks (Balducci, Fraccaroli, & Schaufeli, 2010, p. 148).

Considering the above literature research and support, the researcher followed the pattern of CFA testing outlined in Lovakov et al. (2017). This study paralleled the current study’s method. Specifically, Lovakov et al. (2017) tested both the three-factor and single-factor solutions by beginning with a pooled sample CFA for UWES-9, undertook a model respecification process using the pooled sample and correlated various item errors to produce a best fitting CFA model, and finally used the respecified three-factor model with error correlations to test for measurement invariance across groups.

The present study followed the same method pattern with the following additions to the method: (a) after a respecification process and finalizing the pooled sample CFA model, the independent samples (i.e., Boomers and Millennials) were fit to the respecified measurement models (i.e., three-factor and single-factor) identified through the pooled sample process in order to report the measurement model fit for each sample
and (b) both the three-factor and single-factor respecified models were used for measurement invariance testing in order to support hypotheses testing.

The measurement model analysis began by fitting the pooled sample (i.e., combined Boomer and Millennials samples) to a CFA model specified based upon the theoretical three-factor correlated (i.e., vigor, dedication, and absorption) operationalization of work engagement (Schaufeli et al., 2006). The observed items were evaluated to determine if they loaded to the correct theoretical latent constructs (Hair et al., 2010). To determine goodness of fit for the measurement model, the following cut-off criteria were used: (a) Tucker-Lewis Index (TLI) ≥ .95, (b) comparative fit index (CFI) ≥ .95 (c) standardized root mean square residuals (SRMRs) ≤ .08, and (d) the root mean squared error of approximation (RMSEA) ≤ .10 (Kline, 2016). Following Kline (2016), standardized residual covariances (SRCs) were evaluated for another model fit indicator, with absolute values greater than |2.58| recorded. The pattern and structure coefficients were reported for the pooled sample and evaluated to determine if the individual items had loaded to their theoretical latent variable (Graham, Guthrie, & Thompson, 2003). Bagozzi and Yi (1988) recommended a minimum factor loading of .5 and a more rigorous level of .7 to demonstrate convergent validity. Bagozzi and Yi (1988) also recommended a composite reliability level ≥ .6 to demonstrate reliability and an average variance extracted (AVE) level of ≥ .5 to demonstrate convergent validity. Both of these statistics were reported and evaluated. Discriminant validity was tested by comparing the square root of the AVE to the correlation of the individual factors (Bagozzi & Yi, 1988).
The analysis next independently fit the Boomer and Millennial samples to the CFA measurement model specified for the pooled sample. Pattern coefficients, structure coefficients, SRCs, correlations between factors, $\chi^2$, degrees of freedom, TLI, CFI, SRMR, and RMSEA statistics were reported for each group. Each sample was evaluated for model fit using the same goodness of fit criteria used for the pooled sample measurement model, including SRCs. Next, pattern and structure coefficients were evaluated to determine if the individual items loaded to their theoretical constructs of vigor, dedication, and absorption. Last, the pattern and structure coefficients were evaluated to determine if differences existed across samples. If the model fit indices and factor loadings were found to be at or above acceptable fit levels for the samples, the samples would be considered ready for measurement invariance testing.

**Measurement Invariance Assessments**

Measurement invariance started with configural invariance and proceeded hierarchically to metric invariance and then to scalar invariance. The same goodness of fit metrics used for the measurement models were used for the measurement invariance assessments. To determine goodness of fit for the models, the following cut-off criteria were used: (a) Tucker-Lewis Index (TLI) $\geq .95$, (b) comparative fit index (CFI) $\geq .95$, (c) standardized root mean square residuals (SRMRs) $\leq .08$, and (d) the root mean squared error of approximation (RMSEA) $\leq .10$ (Nimon & Reio, 2011; Vandenberg & Lance, 2000). Assessment of measurement invariance required evaluation of model fit indices using the guidelines above, as well as changes between models in chi-squared ($\chi^2$) and in CFI (CFI). A statistical significance of $p < .05$ was set as the criteria level to determine a statistically significant change between CFA models (Van de Schoot et al.,
Cheung and Rensvold (2000, 2002) provided guidelines for evaluating CFI at three potential levels. First, a CFI ≤ -.01 would suggest model equivalence. Second, a CFI between -.01 and -.02 would suggest a potential difference in models. Third, a CFI > -.02 would suggest model differences.

For configural invariance, the three-factor correlated CFA specified in the measurement model phase was fitted to the Boomer and Millennial groups. The model (M1) was unconstrained across groups where there were no constraints placed on the factor loadings or intercepts (see Figure 1). Configural invariance was tested by evaluating the how well distinct groups fit the CFA. Assuming each group of data fit the same model “where the number of factor(s) and the pattern of free and fixed loadings are the same across groups, evidence of configural invariance holds” (Nimon & Reio, 2011, p. 205). Factor loadings, parameter estimates, and model fit indices were evaluated to determine if H1a was supported or not. If H1a was supported, the analysis would proceed to metric invariance testing.
Figure 1. Configural Invariance for the Three-Factor Model of UWES-9

Note. VI = Vigor, DE = Dedication, and AB = Absorption

Metric invariance testing was accomplished by constraining the factor loadings of M1 to be the same for like items across the groups and yielding a constrained Model 2 (M2). See Figure 2 for a depiction of M2. Metric invariance “tests whether respondents under study attribute the same meaning to the latent construct” (Van de Schoot et al., 2012, p. 489). The metric invariance model (M2) was evaluated by reviewing the model fit indices for each group and comparing those to the configural model (M1) for differences (i.e., $\chi^2$, CFI, and $p$-value) between models, based upon the guidelines above from Cheung and Rensvold (2000, 2002). Based upon the findings, H2a would be supported or not. If H2a was supported, the analysis would proceed to scalar invariance testing.
Scalar invariance testing was accomplished by further constraining M2 by setting the intercepts for like items across groups to be equal, yielding Model 3 (M3). Figure 3 depicts the constrained model M3. Scalar invariance “implies that the meaning of the construct (the factor loadings), and the levels of the underlying items (intercepts) are equal in both groups” (Van de Schoot et al., 2012, p. 489). The scalar invariance model was evaluated by reviewing the model fit indices for each group and comparing those to the metric model (M2) for differences (i.e., $\chi^2$, CFI, and $p$-value) between models, based upon the guidelines above from Cheung and Rensvold (2000, 2002). Based upon the findings, H3a would be supported or not.
The data analyses outlined above was replicated for the single-factor theoretical conceptualization of UWES-9 (Schaufeli et al., 2006). The analysis began with measurement model testing. The same statistics of pattern coefficients, structure coefficients, SRCs, correlations between factors, $\chi^2$, degrees of freedom, TLI, CFI, SRMR, and RMSEA were reported. Assuming good model fit was found, the analysis would proceed through the measurement invariance testing steps detailed above for the Boomer and Millennial groups using the single-factor conceptualization of UWES-9. Configural invariance for the single-factor model would be evaluated, and H1b would be supported or not. Assuming support was found for H1b, metric invariance would be evaluated, and H2b would be supported or not. Assuming support was found for H2b, scalar invariance would be evaluated, and H3b would be supported or not.
Latent Mean Analysis

The analysis phase of the study continued to a latent mean analysis between Boomer and Millennials for the latent variable means for data from the work engagement scale. As noted by Kline, “in order to formally compare group means on latent variables, strong (scalar) measurement invariance should be established” (2016, p. 462). The requirement of strong (scalar) measurement invariance is required since “group differences in pattern coefficients or intercepts say that the indicators do not measure the factors in the same way” by group (Kline, 2016, p. 462-463). Assuming that hypothesis H3ab held, the Boomer and Millennial groups would be compared using latent mean analysis (LMA).

Following the procedure defined by Byrne (2001), the latent mean analysis was conducted in AMOS by using the scalar invariance model (M3). To conduct the analysis, the Estimate Means and Intercepts option was selected in the Analysis Properties tab. By selecting this option, AMOS assigned a zero followed by a comma to each factor in both groups to signify the groups are equal. However, LMA requires that one group is constrained while the other group is freely estimated (Byrne, 2001). In order to freely estimate one group, the researcher manually removed the mean constraints for the Boomer group by recoding the zero value (0,) assigned by AMOS to a dummy code of mn (mn,). According to Byrne (2001), the group to recode is arbitrary and has no bearing on the outcome of the analysis. Once the model is run, AMOS returns latent means and p-values in the parameter estimates output. Since the Boomers group was freely estimated, positive latent mean scores would indicate the Boomer group was more engaged compared to the Millennial group. Evaluation of the p-values would indicate
whether or not the latent mean difference was statistically significant or not. If the \( p \)-value was statistically significant (i.e., \( p < .05 \)), \( H_{4_ab} \) would be supported; otherwise, \( H_{4_ab} \) would be rejected.

**Descriptive Statistics**

Once the data analysis routines and hypotheses testing were complete, the data were analyzed in R, and descriptive statistics were reported following the outline provided in Teo et al. (2009) for overall sample and sub-samples. The analysis included reporting of the descriptive statistics of means, standard deviations, standard errors, skewness, and kurtosis. Internal reliability and consistency were tested using Cronbach’s alpha (coefficient alpha) and were supported by using empirically tested measurement instrumentation from published studies (i.e., UWES-9 and ATCB).

**Common Method Variance**

The comprehensive confirmatory factor analysis marker technique was developed as one method to identify sources of common method variance and was used in the present study (Williams et al., 2010). The ATCB scale was used as the latent marker variable. The ATCB was an ideal choice as it was developed by Miller and Chiodo (2008) to be used as a marker variable in social science research.

The CFA marker technique aims at “testing for the presence of and equality of method effects associated with the marker latent variable” (Williams et al., 2010, p. 18). Five models were built following the defined technique including the initial CFA model (CFA), baseline model (Baseline), the constrained model (Model-C), the unconstrained model (Model-U), and a restricted parameter model (Model-R) (Williams et al., 2010). The model goodness of fit statistics were reported as \( \chi^2 \), degrees of freedom, and CFI.
Model comparisons following the defined $\chi^2$ comparison to the baseline model technique commenced to determine if any method effects were present (Williams et al., 2010). Using the three-factor correlated model, the CFA marker technique analysis was performed three times (i.e., the pooled sample, the Boomer sample, and the Millennial sample) for a full exploration of CMV at the full sample level and by each cohort. This multi-sample analysis method was similar to the sub-sample validation process for CMV proposed in Craighead, Ketchen, Dunn, and Hult (2011). The sub-sample validation process was not suggested as a remedy for CMV; rather, it was suggested to instill “more confidence in the study’s results” (Craighead et al., 2011, p. 582).

**Limitations**

An effort was made to create a rigorous and generalizable study; however, the study contained the following three limitations. First, all data captured were that of self-reported data and may have introduced the possibility of common method variance (Podsakoff et al., 2003). Six methods to mitigate CMV were outlined in the survey design section of this paper. Second, the researcher relied upon a population sourced from Qualtrics® Panel services. While pre-recruited panels represent a large segment of the general population and are frequently used sources for research (Rao, Kaminska, & McCutcheon, 2010), there was still risk that the sourced sample would not be fully representative of the desired population. This limitation and corresponding risk was mitigated within the survey design by using qualifying questions and the evaluation of gathered demographic data variables. Evaluation and comparison of the demographic variables to published leisure and hospitality population descriptions provided a level of confidence of the sample’s representativeness.
Third, generational cohort was a central component of the study, and participants were asked to indicate their generational cohort association. Age is a sensitive subject for some participants and a concern was that participants would not accurately report their associated generation. The use of a pre-recruited panel of participants who had already indicated their age during recruitment and the use of common generational cohort classifications helped to mitigate this concern.

**Summary of the Chapter**

Chapter 3 provided an outline for the design and methodology of the study. The chapter covered the purpose of the study, a measurement invariance overview, the design of the study, research hypotheses, population and sample, instrumentation for the survey, the survey design, data collection procedures, data analysis procedures, hypotheses testing, descriptive statistics, and common method variance. The chapter concluded with a discussion of the limitations of the study.
Chapter 4 - Results

Introduction

Chapter 4 provided the results for the study. Sections of the chapter included a description of the collected data, results of data cleaning, group cohort comparisons, sample representativeness, statistical assumptions, measurement models, measurement invariance testing, descriptive statistics, common method variance, and a hypotheses discussion. A summary of the chapter was also included.

Data Analysis Results

The purpose of the study was to empirically assess measurement invariance by generational cohort using data from the short version of the Utrecht Work Engagement Scale (UWES-9). In order to obtain the data required for the study, survey methodology was utilized, and Qualtrics® was contracted by the researcher for survey panel services.

Data Collection and Participants

The initial notification e-mail was delivered as scheduled on September 18, 2017, and the survey invitation e-mail was delivered as scheduled on September 25, 2017. The survey link was live at 11:00am central time on September 25, and the survey remained open for data collection until October 9, 2017 at 5:43pm central time. A total of 8,379 survey attempts were initiated by participants.

Counter to the proposed data collection strategy of 230 cases for Boomers and 230 cases for Millennials, only 183 Boomers were collected. The limitation was recognized in that the majority of available Boomers in the Qualtrics® panel groups had job tenures of more than 5 years, which disqualified a large number of Boomers for the present study. Qualtrics® reported this discrepancy and ill-fit of the panel groups when
large numbers of Boomers were being dropped from the survey due to the job tenure screening questions. Once Qualtrics® confirmed that the pool of available Boomers within the panel groups had been exhausted, the Millennial group was oversampled above the proposed 230 number to a total of 386. Since the collected data were not to the specification outlined in the proposal, approval was sought by the researcher from the dissertation committee to use the collected sample, and approval was provided by the full committee.

Using the overall collected sample, the data were investigated for complete cases, missing data, duration checks, and straight-lining. Due to the nature of the survey and screening question constraints, a total of 7,580 cases did not pass the initial screening questions and were removed. Further data evaluation removed an additional 50 cases due to incomplete cases, meaning participants had started the survey but did not complete the full survey. Six cases were removed due to the survey duration check as the six cases fell below the 2 minute survey duration threshold. All remaining cases had a survey duration that was 2 minutes or longer. An additional 196 cases failed secondary screening checks (e.g., tenure and full-time work week) when evaluated for their collected demographic and job characteristic variables. A total of 125 Boomers failed due to tenure of more than 5 years, and one failed for less than a full-time work week, and 58 Millennials failed due to tenure of more than 5 years, and 12 failed for less than a full-time work week. No cases were dropped due to straight-lining of the ATCB instrument. A final pooled sample of 547 was recorded and used for subsequent data analyses, consisting of 178 Boomers and 369 Millennials.
For the remaining 547 cases, each variable was examined for the expected range of values. The table function was utilized in R to print the distribution of values. The reported values for each individual variable were evaluated by the researcher through comparison of the printed range to the range of values predefined within the project data dictionary. No discrepancies were found for the range of values for any variable. As part of this check, it was verified that no variable contained any missing data. This was attributable to the fact that the survey required a forced response for each question. The pooled sample was grouped by generation, and Table 5 contained the distribution by industry and Table 6 contained a chi-square test for the Boomer and Millennial groups.

Table 5

_Distribution of Industries for the Initial Pooled Sample of Boomer and Millennial Groups_

<table>
<thead>
<tr>
<th>Industry</th>
<th>Boomers</th>
<th>Boomers %</th>
<th>Millennials</th>
<th>Millennials %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convention Center</td>
<td>3</td>
<td>1.7%</td>
<td>36</td>
<td>9.7%</td>
</tr>
<tr>
<td>Cruise Line</td>
<td>7</td>
<td>3.9%</td>
<td>21</td>
<td>5.7%</td>
</tr>
<tr>
<td>Gaming or Casino</td>
<td>20</td>
<td>11.2%</td>
<td>62</td>
<td>16.8%</td>
</tr>
<tr>
<td>Lodging</td>
<td>68</td>
<td>38.2%</td>
<td>118</td>
<td>32.0%</td>
</tr>
<tr>
<td>Marina</td>
<td>3</td>
<td>1.7%</td>
<td>9</td>
<td>2.4%</td>
</tr>
<tr>
<td>Sporting Facility</td>
<td>8</td>
<td>4.5%</td>
<td>22</td>
<td>6.0%</td>
</tr>
<tr>
<td>Travel</td>
<td>58</td>
<td>32.6%</td>
<td>63</td>
<td>17.1%</td>
</tr>
<tr>
<td>Tourism</td>
<td>11</td>
<td>6.2%</td>
<td>38</td>
<td>10.3%</td>
</tr>
<tr>
<td>Total by Group</td>
<td>178</td>
<td>100.0%</td>
<td>369</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

*Note.* Percentages were based upon group totals.

Table 6

_Group Comparison Chi-Square Results by Industry for the Initial Pooled Sample of Boomer and Millennial Groups_

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry</td>
<td>31.66</td>
<td>7</td>
<td>&lt; .001</td>
<td>.241</td>
</tr>
</tbody>
</table>

*Note.* $df =$ degrees of freedom.
Group Comparison Results

The group comparison phase was a crucial component in the overall analysis of the study. For measurement invariance testing, the Boomer and Millennial groups were required to be as similar as possible by their associated demographics and job characteristics. By virtue of the screening questions and subsequent secondary screening by demographic and job characteristic variables, the groups were equivalent by (a) all respondents were currently working in the United States, (b) all respondents were currently living in the United States, (c) English was the primary language spoken at their job, (d) all respondents were full-time (i.e., 35 or more hours on average per week) employees, (e) all respondents had a job tenure of 5 years or less, and (f) all respondents were working in the leisure and hospitality industries listed in Table 5.

Acknowledging the group equivalencies drawn from the screening questions, the groups were subsequently evaluated statistically for group differences by (a) gender, (b) race, (c) education, (d) job level, (e) market tier for their place of employment, and (f) whether or not their job required customer contact. A series of chi-square tests were conducted, and the results were reported in Table 7. Table 8 reported the distribution of demographics for the initial pooled sample of Boomers ($n = 178$) and Millennials ($n = 369$).

The $p$-values ranged from less than .001 to .814, resulting in statistically significant ($p \leq .05$) differences by group for each of the comparison variables except for education and customer contact. The Cramer’s $V$ values ranged from .004 to .280, resulting in practically significantly ($V \geq .10$) differences by group for each of the comparison variables, except for customer contact. Due to the results of the chi-square
testing and the study requirement for equivalent groups, propensity score matching was utilized to equate the groups by their covariates (Rubin, 1997).

Table 7

*Group Comparison Chi-Square Results for the Initial Pooled Sample of Boomer and Millennial Groups*

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>20.13</td>
<td>1</td>
<td>&lt; .001</td>
<td>.187</td>
</tr>
<tr>
<td>Race</td>
<td>42.93</td>
<td>4</td>
<td>&lt; .001</td>
<td>.280</td>
</tr>
<tr>
<td>Education</td>
<td>10.88</td>
<td>7</td>
<td>.144</td>
<td>.141</td>
</tr>
<tr>
<td>Job Level</td>
<td>39.70</td>
<td>3</td>
<td>&lt; .001</td>
<td>.270</td>
</tr>
<tr>
<td>Market Tier</td>
<td>6.30</td>
<td>2</td>
<td>.043</td>
<td>.107</td>
</tr>
<tr>
<td>Customer Contact</td>
<td>0.06</td>
<td>1</td>
<td>.814</td>
<td>.004</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom.

**Propensity Score Matching Results**

Following the recommendations for HRD quantitative research found in Keiffer and Lane (2016), the full battery of available (i.e., gender, race, education, job level, market tier, and customer contact) covariates were input into the PSM algorithm, and nearest neighbor matching was used for the matching method with the caliper set to .20. Nearest neighbor matching was used, as it is the “most straightforward matching estimator” (Caliendo & Kopeinig, 2005, p. 9). The caliper setting was used to insure the pairs were well matched and to avoid potential bias of nearest neighbor matching (Lane, To, Shelley, & Henson, 2012). The caliper setting of .20 was specified *a priori* as it is was suggested by Stuart (2010) for reducing bias between groups for nearest neighbor matching. Results from the full battery of covariate matching yielded matched groups ($n_{Boomers} = 167, n_{Millennials} = 167$) that were statistically and practically significantly different by education, job level, and market tier.
Table 8

Distribution of Demographics for the Initial Pooled Sample of Boomer and Millennial Groups

<table>
<thead>
<tr>
<th>Demographic Variable</th>
<th>Boomers</th>
<th>Boomers %</th>
<th>Millennials</th>
<th>Millennials %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>108</td>
<td>60.7%</td>
<td>291</td>
<td>78.9%</td>
</tr>
<tr>
<td>Male</td>
<td>70</td>
<td>39.3%</td>
<td>78</td>
<td>21.1%</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>7</td>
<td>3.9%</td>
<td>19</td>
<td>5.1%</td>
</tr>
<tr>
<td>Black or African American</td>
<td>7</td>
<td>3.9%</td>
<td>57</td>
<td>15.5%</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>3</td>
<td>1.7%</td>
<td>51</td>
<td>13.8%</td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>157</td>
<td>88.2%</td>
<td>239</td>
<td>64.8%</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>2.3%</td>
<td>3</td>
<td>0.8%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some High School</td>
<td>0</td>
<td>0%</td>
<td>6</td>
<td>1.6%</td>
</tr>
<tr>
<td>High School or GED</td>
<td>27</td>
<td>15.2%</td>
<td>76</td>
<td>20.6%</td>
</tr>
<tr>
<td>Some College</td>
<td>47</td>
<td>26.4%</td>
<td>85</td>
<td>23.0%</td>
</tr>
<tr>
<td>Professional or Trade Certificate</td>
<td>8</td>
<td>4.5%</td>
<td>12</td>
<td>3.3%</td>
</tr>
<tr>
<td>Associates Degree</td>
<td>14</td>
<td>7.9%</td>
<td>48</td>
<td>13.0%</td>
</tr>
<tr>
<td>Bachelors Degree</td>
<td>60</td>
<td>33.7%</td>
<td>99</td>
<td>26.9%</td>
</tr>
<tr>
<td>Masters Degree</td>
<td>18</td>
<td>10.1%</td>
<td>33</td>
<td>8.9%</td>
</tr>
<tr>
<td>Doctoral Degree</td>
<td>4</td>
<td>2.2%</td>
<td>10</td>
<td>2.7%</td>
</tr>
<tr>
<td>Job Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td>16</td>
<td>9.0%</td>
<td>68</td>
<td>18.4%</td>
</tr>
<tr>
<td>Non-Supervisory</td>
<td>81</td>
<td>45.5%</td>
<td>100</td>
<td>27.1%</td>
</tr>
<tr>
<td>Supervisory</td>
<td>31</td>
<td>17.4%</td>
<td>135</td>
<td>36.6%</td>
</tr>
<tr>
<td>Managerial</td>
<td>50</td>
<td>28.1%</td>
<td>66</td>
<td>17.9%</td>
</tr>
<tr>
<td>Market Tier</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower</td>
<td>24</td>
<td>13.5%</td>
<td>30</td>
<td>8.1%</td>
</tr>
<tr>
<td>Middle</td>
<td>100</td>
<td>56.2%</td>
<td>244</td>
<td>66.1%</td>
</tr>
<tr>
<td>Upper</td>
<td>54</td>
<td>30.3%</td>
<td>95</td>
<td>25.8%</td>
</tr>
<tr>
<td>Customer Contact</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>158</td>
<td>88.8%</td>
<td>330</td>
<td>89.4%</td>
</tr>
<tr>
<td>No</td>
<td>20</td>
<td>11.2%</td>
<td>39</td>
<td>10.6%</td>
</tr>
</tbody>
</table>

Note. Percentages were based upon each demographic variable.

Table 9 reported the nearest neighbor PSM output after chi-square testing. Table 9 can be compared to Table 7 to determine the results of the nearest neighbor matching.
While the groups were better equated on gender and race after PSM nearest neighbor matching, the groups were still statistically significantly different by education, job level, and market tier. The nearest neighbor matching did not sufficiently equate the groups to where all matched variables reflected no statistically significant differences.

Table 9

*Group Comparison Chi-Square Results by Boomers and Millennials after Nearest Neighbor Match Level Propensity Score Matching*

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.01</td>
<td>1</td>
<td>.909</td>
<td>.000</td>
</tr>
<tr>
<td>Race</td>
<td>1.13</td>
<td>4</td>
<td>.889</td>
<td>.058</td>
</tr>
<tr>
<td>Education</td>
<td>18.92</td>
<td>7</td>
<td>.008</td>
<td>.238</td>
</tr>
<tr>
<td>Job Level</td>
<td>40.12</td>
<td>3</td>
<td>&lt; .001</td>
<td>.347</td>
</tr>
<tr>
<td>Market Tier</td>
<td>8.03</td>
<td>2</td>
<td>.018</td>
<td>.156</td>
</tr>
<tr>
<td>Customer Contact</td>
<td>0.44</td>
<td>1</td>
<td>.509</td>
<td>.027</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom.

As an alternate approach to uncover better matching results, the groups were equated on the variables where $\chi^2$ testing on the pre-matched (i.e., before PSM) sample (Table 7) returned statistically significant differences. The covariates of gender, race, job level, and market tier were input into the PSM algorithm, and genetic matching was used for the matching method with the caliper set to .20. Genetic matching, a computationally intensive algorithm, was used because it has been suggested for use when propensity matching output is required to have highly equivalent groups (Randolph, Falbe, Manuel, & Balloun, 2014). The genetic matching dropped the pooled sample to a total of 264 cases, split evenly by Boomers and Millennials. Table 10 reports a series of chi-square tests that were conducted on the 264 ($n_{Boomers} = 132$, $n_{Millennials} = 132$) cases after the genetic PSM routine was complete.
Table 10

*Group Comparison Chi-Square Results by Boomers and Millennials after Genetic Match Level Propensity Score Matching*

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.00</td>
<td>1</td>
<td>1.000</td>
<td>.000</td>
</tr>
<tr>
<td>Race</td>
<td>0.00</td>
<td>4</td>
<td>1.000</td>
<td>.000</td>
</tr>
<tr>
<td>Education</td>
<td>11.55</td>
<td>7</td>
<td>0.116</td>
<td>.209</td>
</tr>
<tr>
<td>Job Level</td>
<td>0.00</td>
<td>1</td>
<td>1.000</td>
<td>.000</td>
</tr>
<tr>
<td>Market Tier</td>
<td>0.00</td>
<td>1</td>
<td>1.000</td>
<td>.000</td>
</tr>
<tr>
<td>Customer Contact</td>
<td>2.73</td>
<td>1</td>
<td>0.098</td>
<td>.089</td>
</tr>
</tbody>
</table>

Note. $df =$ degrees of freedom.

No statistically significant group differences were present in the genetic match level PSM sample by gender, race, job level, market tier, education, or customer contact. Practically significant group differences ($V = .209$) between the Boomer and Millennial groups continued to exist for education. Table 11 reported the percentage distributions for education by group. Millennials have been suggested to be more educated than Boomers (Eisner, 2005), but there was little evidence of this provided in Table 11. Rather, Boomers had higher percentages for bachelor’s and master’s degrees while Millennials had higher percentages in associate’s and doctoral degrees. Given that education was the only covariate by which the groups were different by a practical significance level (i.e., groups were not practically significantly different by gender, race, job level, market tier, and customer contact, and no groups were statistically significantly different), the groups were considered reasonably equivalent to proceed.
Table 11

Distribution of Demographics for the Genetically Matched Pooled Sample of Boomer and Millennial Groups

<table>
<thead>
<tr>
<th>Demographic Variable</th>
<th>Boomers $n = 132$</th>
<th>Boomers %</th>
<th>Millennials $n = 132$</th>
<th>Millennials %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>92</td>
<td>69.7%</td>
<td>92</td>
<td>69.7%</td>
</tr>
<tr>
<td>Male</td>
<td>40</td>
<td>30.3%</td>
<td>40</td>
<td>30.3%</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>4</td>
<td>3.0%</td>
<td>4</td>
<td>3.0%</td>
</tr>
<tr>
<td>Black or African American</td>
<td>7</td>
<td>5.3%</td>
<td>7</td>
<td>5.3%</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>2</td>
<td>1.5%</td>
<td>2</td>
<td>1.5%</td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>118</td>
<td>89.4%</td>
<td>118</td>
<td>89.4%</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0.8%</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some High School</td>
<td>0</td>
<td>0%</td>
<td>2</td>
<td>1.5%</td>
</tr>
<tr>
<td>High School or GED</td>
<td>18</td>
<td>13.6%</td>
<td>29</td>
<td>21.9%</td>
</tr>
<tr>
<td>Some College</td>
<td>38</td>
<td>28.8%</td>
<td>31</td>
<td>23.5%</td>
</tr>
<tr>
<td>Professional or Trade</td>
<td>7</td>
<td>5.3%</td>
<td>3</td>
<td>2.3%</td>
</tr>
<tr>
<td>Certificate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associate’s Degree</td>
<td>9</td>
<td>6.8%</td>
<td>17</td>
<td>12.9%</td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>43</td>
<td>32.6%</td>
<td>36</td>
<td>27.3%</td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>14</td>
<td>10.6%</td>
<td>9</td>
<td>6.8%</td>
</tr>
<tr>
<td>Doctoral Degree</td>
<td>3</td>
<td>2.3%</td>
<td>5</td>
<td>3.8%</td>
</tr>
<tr>
<td>Job Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td>12</td>
<td>9.1%</td>
<td>12</td>
<td>9.1%</td>
</tr>
<tr>
<td>Non-Supervisory</td>
<td>58</td>
<td>44.0%</td>
<td>58</td>
<td>44.0%</td>
</tr>
<tr>
<td>Supervisory</td>
<td>28</td>
<td>21.2%</td>
<td>28</td>
<td>21.2%</td>
</tr>
<tr>
<td>Managerial</td>
<td>34</td>
<td>25.7%</td>
<td>34</td>
<td>25.7%</td>
</tr>
<tr>
<td>Market Tier</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower</td>
<td>12</td>
<td>9.1%</td>
<td>12</td>
<td>9.1%</td>
</tr>
<tr>
<td>Middle</td>
<td>84</td>
<td>63.6%</td>
<td>84</td>
<td>63.6%</td>
</tr>
<tr>
<td>Upper</td>
<td>36</td>
<td>27.3%</td>
<td>36</td>
<td>27.3%</td>
</tr>
<tr>
<td>Customer Contact</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>123</td>
<td>93.2%</td>
<td>115</td>
<td>87.1%</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td>6.8%</td>
<td>17</td>
<td>12.9%</td>
</tr>
</tbody>
</table>

*Note.* Percentages were based upon each demographic variable.
Sample Representativeness Results

To evaluate the representativeness of the sample, the demographic characteristics were compared to the available demographics for the leisure and hospitality industry sourced from the Bureau of Labor Statistics (BLS, 2017). The survey design specifically collected demographic variables so the resulting data would be comparable to the BLS data. Tables 12 and 13 provided a comparison by percentages of the BLS reported data to the initial sample (n = 547) and to the genetically matched final sample (n = 264) of Boomers and Millennials for the present study.

Table 12

Sample Representativeness Comparison for BLS and the Initial Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>BLS %</th>
<th>Initial Sample %</th>
<th>χ²</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n = 547</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49.4%</td>
<td>27.1%</td>
<td>104.33</td>
<td>1</td>
<td>&lt;.01</td>
<td>.229</td>
</tr>
<tr>
<td>Female</td>
<td>50.6%</td>
<td>72.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>12.6%</td>
<td>11.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>6.5%</td>
<td>4.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>22.9%</td>
<td>9.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian or Other</td>
<td>58%</td>
<td>73.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>72.91</td>
<td>3</td>
<td>&lt;.01</td>
<td>.191</td>
</tr>
<tr>
<td>Supervisor</td>
<td>16.5%</td>
<td>51.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Supervisor</td>
<td>83.5%</td>
<td>48.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Role</td>
<td></td>
<td></td>
<td>272.76</td>
<td>1</td>
<td>&lt;.01</td>
<td>.370</td>
</tr>
</tbody>
</table>

Note. BLS = Bureau of Labor Statistics. Initial Sample % = initial collected sample (n = 547).
Table 13

Sample Representativeness Comparison for BLS and the Final Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>BLS %</th>
<th>Final Sample %</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$n = 264$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49.4%</td>
<td>30.3%</td>
<td></td>
<td></td>
<td>&lt;.01</td>
<td>.195</td>
</tr>
<tr>
<td>Female</td>
<td>50.6%</td>
<td>69.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>12.6%</td>
<td>5.3%</td>
<td>75.3</td>
<td>1</td>
<td>&lt;.01</td>
<td>.195</td>
</tr>
<tr>
<td>Asian</td>
<td>6.5%</td>
<td>3.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>22.9%</td>
<td>1.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian or Other</td>
<td>58%</td>
<td>90.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>16.5%</td>
<td>47%</td>
<td>300.32</td>
<td>3</td>
<td>&lt;.01</td>
<td>.388</td>
</tr>
<tr>
<td>Non-Supervisor</td>
<td>83.5%</td>
<td>53%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Role</td>
<td></td>
<td></td>
<td>52.41</td>
<td>1</td>
<td>&lt;.01</td>
<td>.222</td>
</tr>
</tbody>
</table>

Note. BLS = Bureau of Labor Statistics. Final Sample % = final sample after genetic matching ($n = 264$).

Notable differences existed between the BLS demographic profile and the final sample profile. The final sample profile was skewed toward a more female population with a higher representation of supervisors compared to the BLS. In addition, the final sample was less racially diverse compared to the BLS profile, skewing toward the Caucasian or other category. The final sample was not representative of the general leisure and hospitality industry demographic profile of the BLS; however, the BLS demographic profile included the restaurant and food services subsectors that were specifically excluded from the present study’s sample criteria. The BLS did provide limited demographic profile information on subsectors of the leisure and hospitality supersector, such as the arts and entertainment subsector’s racial profile of 72.1%...
Caucasian or other, but a directly comparable demographic profile from the BLS that matched the final sample’s subsectors was not available.

The final sample skewed more heavily toward females than the BLS profile, but the final sample did retain the perspective that females represented the majority of the gender within the leisure and hospitality industry. The final sample was also more skewed toward supervisors compared to the general BLS profile. Insomuch as the final sample’s market tier variable represented a skew toward the mid and upper tiers, the distribution of more supervisors in the final sample was not surprising. Upper and mid-scale leisure and hospitality establishments offer more services to guests and have a higher staff to customer ratio, requiring more managers and supervisors (Gallup, 2017).

Given the large cross-sectional footprint contained within the BLS profile for the leisure and hospitality industries, it was expectable that the narrowly defined final sample would not be fully representative of the general BLS profile. The final sample did represent a specific segment of the leisure and hospitality industry that was constituted largely of Boomer and Millennial females who were Caucasian and supervisors in predominantly upper and mid-scale market tier establishments with job tenure of 5 years or less.

The combination of survey screening questions, secondary demographic screening, and propensity score matching resulted in a final sample of Boomers and Millennials that was largely similar on multiple characteristics. Overall, the groups were equivalent by a) currently employed in the United States, b) currently living in the United States, c) speaking English as their primary language at work, d) working full-time, e) job tenure of 5 years or less, f) gender, g) race, h) job level, i) market tier, and j) customer
contact. Table 14 reported the distribution of leisure and hospitality industries for the final sample, and Table 15 reported the final group difference for industry. Of note, while the groups maintained a low practical difference \((V = .225)\) by industry, the groups are no longer statistically significantly different \((p = .063)\) by industry.

Table 14

<table>
<thead>
<tr>
<th>Industry</th>
<th>Boomers</th>
<th>Boomers %</th>
<th>Millennials</th>
<th>Millennials %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convention Center</td>
<td>3</td>
<td>2.3%</td>
<td>13</td>
<td>9.9%</td>
</tr>
<tr>
<td>Cruise Line</td>
<td>5</td>
<td>3.8%</td>
<td>5</td>
<td>3.8%</td>
</tr>
<tr>
<td>Gaming or Casino</td>
<td>14</td>
<td>10.6%</td>
<td>21</td>
<td>15.9%</td>
</tr>
<tr>
<td>Lodging</td>
<td>51</td>
<td>38.6%</td>
<td>46</td>
<td>34.8%</td>
</tr>
<tr>
<td>Marina</td>
<td>3</td>
<td>2.3%</td>
<td>4</td>
<td>3.0%</td>
</tr>
<tr>
<td>Sporting Facility</td>
<td>6</td>
<td>4.5%</td>
<td>4</td>
<td>3.0%</td>
</tr>
<tr>
<td>Travel</td>
<td>42</td>
<td>31.8%</td>
<td>26</td>
<td>19.7%</td>
</tr>
<tr>
<td>Tourism</td>
<td>8</td>
<td>6.1%</td>
<td>13</td>
<td>9.9%</td>
</tr>
<tr>
<td>Total by Group</td>
<td>132</td>
<td>100.0%</td>
<td>132</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

*Note*. Percentages were based upon group totals.

Table 15

<table>
<thead>
<tr>
<th>Variable</th>
<th>(\chi^2)</th>
<th>df</th>
<th>p-value</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry</td>
<td>13.41</td>
<td>7</td>
<td>.063</td>
<td>.225</td>
</tr>
</tbody>
</table>

*Note*. \(df\) = degrees of freedom.

**Statistical Assumptions Results**

The sample of 132 Boomers and 132 Millennials was formatted for input into AMOS by use of the upload and format functionality found in SPSS. The formatted SPSS file was reviewed for consistency, both in quantity and overall individual variable consistency, by comparing it to the output file from the data cleaning analysis. No discrepancies were noted.
The data were evaluated for the statistical assumptions of multivariate normality and multivariate outliers before proceeding to the measurement model phase of the analysis. The recommendations of Byrne (2010) and Kline (2016) were followed for the statistical assumption analysis. To test for multivariate normality, the kurtosis and the critical ratio were reviewed. The kurtosis was observed at 43.575, and the critical ratio was observed at 25.158. The critical ratio exceeded 5.0 (Byrne, 2010), suggesting moderate multivariate nonnormality. The Mahalanobis distance ($D^2$) statistic report was reviewed, and no $D^2$ values were observed to be distinctly apart from the other reported $D^2$ values. According to Byrne (2010), if a $D^2$ value “stands distinctively apart from all the other $D^2$ values,” evidence of multivariate outliers exists. Since no $D^2$ values were observed to be distinctly apart, the data were considered to not have multivariate outliers present.

To address the multivariate nonnormality, bootstrapping was performed following Kline (2016). A 2,000 record sampling technique was implemented, and bootstrapped and non-bootstrapped results were analyzed. The bootstrapped and non-bootstrapped results were found to be insignificantly different; therefore, the non-bootstrapped results were reported (Kline, 2016).

**Pooled Measurement Models Results for the Three-Factor Model**

Confirmatory factor analysis using SPSS AMOS v24.0 was used to analyze the data, and a three-factor measurement model was specified for the pooled sample of Boomers and Millennials. Factor loadings, chi-square ($\chi^2$), degrees of freedom ($df$), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), standardized residual covariances (SRC), Tucker-Lewis index (TLI), and
comparative fix index (CFI) statistics were reported. According to Hu and Bentler (1998), SRMR is the recommended fit index to evaluate for maximum likelihood modeling. Further recommendations were to supplement SRMR by reporting TLI, CFI, and RMSEA indices; however, caution was supplied that TLI and RMSEA were “less preferable at small sample sizes” (Hu & Bentler, 1998, p. 424). Due to the reduced sample size of \( n = 264 \) derived from the propensity score matching, the researcher used CFI and SRMR as the fit indices to evaluate model fit. Upon fitting the Original three-factor measurement model (M₀), the fit statistics indicated a moderately poor-fitting model (CFI = .923 and SRMR = .054). Review of the modification indices indicated model misspecification due to error covariances among the items of vigor, dedication, and absorption. This finding was not surprising due to the known high levels of intercorrelations for UWES-9 (Schaufeli et al., 2006). Following the method suggested in Byrne (2010), the measurement model was respecified and reestimated. Theoretical support was found in the literature for each of the item error correlations. The first respecification was the Respecified One Model (M₁) that included correlated errors for items DE4 and AB4. This specific error correlation at the item level between DE4 and AB4 was found in a model respecification exercise within an international validation study (Klassen, Aldhafri, Mansfield, Purwanto, Siu, Wong, & Woods-McConney, 2012). The Klassen et al. (2012) validation study recognized a significant improvement in CFI from .87 to .97 when the DE4 and AB4 errors were correlated along with two other item level correlations. Schaufeli and Bakker commented that absorption plays a “special role” with “both other engagement scales” (2004, p. 305). After Schaufeli and Bakker (2004) inspected modification indices, they correlated errors for dedication and
absorption with a significant improvement in model fit. As was found in Schaufeli and Bakker (2004) and with Klassen et al. (2012), model M1 produced a better fitting model, but the CFI = .939 was below the cutoff criteria and additional investigation was warranted. Next, VI2 and VI3 were correlated. Balducci et al. (2010) correlated the same item errors during a validation study with a resulting improvement in overall model fit for the three-factor solution. The Respecified Two Model (M2) showed a better fit with a CFI = .956. Given that the CFI for model M2 was just over the cut off and that subsequent literature research revealed additional error correlation was merited, AB4 and AB5 item errors were correlated to produce the Final Respecified Model (MF). Review of the fit indices for model MF showed a well-fitting model with CFI = .965 and SRMR = .036. Support for correlation of the errors for these two items was found in three articles as the items are closely worded: AB4 (“I am immersed in my work”) and AB5 (“I get carried away when I am working”) (Balducci et al., 2010; Lovakov et al., 2017; Seppala et al., 2009). Overall, support for the model respecification process and correlation of errors in the three-factor model of UWES-9 was found in the literature as “some additional error correlations are needed to be freed in order to reach appropriate fit” (Hakanen, 2009, p. 24). Results of the model respecification process were reported in Table 16.
The Final Respecified (MF) three-factor measurement model was further interrogated by reviewing the standardized regression weights. As was shown in Table 17, all factors loaded above the minimum recommended cutoff limit of .5, and seven of the nine items loaded above the more rigorous level of .7 (Bagozzi & Yi, 1988). Factor loadings ranged from .532 to .914, no factors loaded above .95, and no factors were required to be deleted from the three-factor measurement model. Review of the structure coefficients in Table 17 demonstrated that all manifest variables loaded correctly to their respective factors (Graham et al., 2003). The results confirmed the theoretical structure of the UWES-9 model as proposed by Schaufeli and colleagues (2006).

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC $\geq 2.58$</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0: Original</td>
<td>132.30</td>
<td>24</td>
<td>.131</td>
<td>.054</td>
<td>0</td>
<td>.884</td>
<td>.923</td>
</tr>
<tr>
<td>M1: Respecified One</td>
<td>108.86</td>
<td>23</td>
<td>.119</td>
<td>.048</td>
<td>0</td>
<td>.904</td>
<td>.939</td>
</tr>
<tr>
<td>M2: Respecified Two</td>
<td>84.17</td>
<td>22</td>
<td>.104</td>
<td>.042</td>
<td>0</td>
<td>.927</td>
<td>.956</td>
</tr>
<tr>
<td>MF: Final Respecified</td>
<td>69.54</td>
<td>21</td>
<td>.094</td>
<td>.036</td>
<td>0</td>
<td>.940</td>
<td>.965</td>
</tr>
</tbody>
</table>

*Note.* df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.
Table 17

*Pattern (P) and Structure (S) Coefficients for the Pooled Sample Measurement Model*

<table>
<thead>
<tr>
<th>Construct Variable</th>
<th>Vigor</th>
<th></th>
<th>Dedication</th>
<th></th>
<th>Absorption</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>S</td>
<td>P</td>
<td>S</td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>VI1</td>
<td>.762</td>
<td>.762</td>
<td>.000</td>
<td>.654</td>
<td>.000</td>
<td>.627</td>
</tr>
<tr>
<td>VI2</td>
<td>.914</td>
<td>.914</td>
<td>.000</td>
<td>.785</td>
<td>.000</td>
<td>.752</td>
</tr>
<tr>
<td>VI3</td>
<td>.815</td>
<td>.815</td>
<td>.000</td>
<td>.700</td>
<td>.000</td>
<td>.670</td>
</tr>
<tr>
<td>DE2</td>
<td>.000</td>
<td>.738</td>
<td>.860</td>
<td>.000</td>
<td>.721</td>
<td></td>
</tr>
<tr>
<td>DE3</td>
<td>.000</td>
<td>.755</td>
<td>.880</td>
<td>.000</td>
<td>.738</td>
<td></td>
</tr>
<tr>
<td>DE4</td>
<td>.000</td>
<td>.546</td>
<td>.636</td>
<td>.000</td>
<td>.533</td>
<td></td>
</tr>
<tr>
<td>AB3</td>
<td>.000</td>
<td>.674</td>
<td>.000</td>
<td>.688</td>
<td>.820</td>
<td>.820</td>
</tr>
<tr>
<td>AB4</td>
<td>.000</td>
<td>.593</td>
<td>.000</td>
<td>.605</td>
<td>.722</td>
<td>.722</td>
</tr>
<tr>
<td>AB5</td>
<td>.000</td>
<td>.437</td>
<td>.000</td>
<td>.446</td>
<td>.532</td>
<td>.532</td>
</tr>
</tbody>
</table>

*Note.* VI = Vigor, DE = Dedication, and AB = Absorption

Table 18 provided statistics for analyzing the internal structure of the pooled measurement model (MF). The composite reliability (CR ≥ .6) and average variance extracted (AVE ≥ .5) were at or above the recommended thresholds (Bagozzi & Yi, 1988). The CR range (.74 - .87) suggested sufficient reliability. The AVE range (.49 - .69) suggested the pooled three-factor measurement model demonstrated convergent validity. Discriminant validity was evaluated by comparing the square root of the AVE statistic to the correlations for each individual factor (Bagozzi & Yi, 1988). Discriminant validity was not confirmed as the individual factor correlations were mostly higher than the square root of AVE. The failure of discriminant validity was due to the known high intercorrelations of the UWES-9 instrument (Schaufeli et al., 2006).
Table 18

*Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Pooled Sample*

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vigor</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Dedication</td>
<td>.86</td>
<td>.80</td>
<td></td>
</tr>
<tr>
<td>3. Absorption</td>
<td>.82</td>
<td>.84</td>
<td>.70</td>
</tr>
<tr>
<td>CR</td>
<td>.87</td>
<td>.84</td>
<td>.74</td>
</tr>
<tr>
<td>AVE</td>
<td>.69</td>
<td>.64</td>
<td>.49</td>
</tr>
</tbody>
</table>

*Note.* Square root of AVE is along the diagonal.

After review of the reported fit statistics, measurements of reliability and validity, and pattern and structure coefficients, the pooled three-factor measurement model portrayed in Figure 4.1 was considered satisfactory. The pooled sample was delineated by Boomers and Millennials for independent measurement model testing.

![Figure 4. Three-Factor Measurement Model (MF)](image-url)
Measurement Model Results for the Three-Factor Model Boomer Sample

Confirmatory factor analysis was used again to analyze data for the sample of Boomers \( (n = 132) \). The same final three-factor measurement model (MF) that was specified for the pooled sample was fit to the Boomer sample. Upon fitting the specified three-factor measurement model for the Boomer sample, fit statistics indicated a well-fitting model (CFI = .952 and SRMR = .046). The RMSEA (.118) was over the recommended cutoff criteria \( (\leq 10, \text{Kline, 2016}) \), but the relatively small sample size \( (n_{\text{Boomers}} = 132) \) and cautions provided by Hu and Bentler (1998) regarding the lack of reliability for RMSEA at such a sample size provided a level of confidence to proceed. Similarly, strong theoretical support for the three-factor model (Schaufeli et al., 2006) provided a strong secondary layer of confidence to proceed, even with the RMSEA result. Results of the three-factor measurement model for the Boomer sample were reported in Table 19.

Table 19

<table>
<thead>
<tr>
<th>Model</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC [2.58]</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millenial Sample</td>
<td>59.48</td>
<td>21</td>
<td>.118</td>
<td>.046</td>
<td>0</td>
<td>.917</td>
<td>.952</td>
</tr>
</tbody>
</table>

*Note. df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.*

Using the standardized regression weights, the Boomer sample three-factor measurement model was further interrogated. As was shown in Table 20, all factors loaded above the minimum recommended cutoff limit of .5, and seven of the nine items loaded above the more rigorous level of .7 (Bagozzi & Yi, 1988). Factor loadings ranged
from .567 to .953, no factors loaded above .95 (VI2 was considered at .95 once rounded from .953), and no factors were required to be deleted from the three-factor measurement model for the Boomer sample. Review of the structure coefficients in Table 20 demonstrated that all manifest variables loaded correctly to their respective factors (Graham et al., 2003).

Table 20

*Pattern (P) and Structure (S) Coefficients for the Boomer Sample Measurement Model*

<table>
<thead>
<tr>
<th>Construct Variable</th>
<th>Vigor</th>
<th>Dedication</th>
<th>Absorption</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI1</td>
<td>.840</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>VI2</td>
<td>.953</td>
<td>.760</td>
<td>.000</td>
</tr>
<tr>
<td>VI3</td>
<td>.788</td>
<td>.000</td>
<td>.712</td>
</tr>
<tr>
<td>DE2</td>
<td>.000</td>
<td>.891</td>
<td>.000</td>
</tr>
<tr>
<td>DE3</td>
<td>.000</td>
<td>.862</td>
<td>.000</td>
</tr>
<tr>
<td>DE4</td>
<td>.000</td>
<td>.662</td>
<td>.000</td>
</tr>
<tr>
<td>AB3</td>
<td>.000</td>
<td>.711</td>
<td>.839</td>
</tr>
<tr>
<td>AB4</td>
<td>.000</td>
<td>.611</td>
<td>.721</td>
</tr>
<tr>
<td>AB5</td>
<td>.000</td>
<td>.480</td>
<td>.567</td>
</tr>
</tbody>
</table>

*Note.* VI = Vigor, DE = Dedication, and AB = Absorption

Table 21 provided statistics for analyzing the internal structure of the Boomer measurement model. The composite reliability (CR ≥ .6) and average variance extracted (AVE ≥ .5) were all above the recommended thresholds (Bagozzi & Yi, 1988). The CR range (.76 - .90) suggested sufficient reliability. The AVE range (.52 - .74) suggested the Boomer sample three-factor measurement model demonstrated convergent validity. Discriminant validity was again evaluated by comparing the square root of the AVE statistic to the correlations for each individual factor (Bagozzi & Yi, 1988). Similar to the pooled sample, discriminant validity was not confirmed as the individual factor correlations were mostly higher than the square root of AVE for the Boomer sample.
Table 21

**Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Boomer Sample**

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vigor</td>
<td>.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Dedication</td>
<td>.90</td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td>3. Absorption</td>
<td>.77</td>
<td>.85</td>
<td>.72</td>
</tr>
<tr>
<td>CR</td>
<td>.90</td>
<td>.85</td>
<td>.76</td>
</tr>
<tr>
<td>AVE</td>
<td>.74</td>
<td>.66</td>
<td>.52</td>
</tr>
</tbody>
</table>

*Note.* Square root of AVE is along the diagonal.

**Measurement Model Results for the Three-Factor Model Millennial Sample**

Confirmatory factor analysis was used once again to analyze data for the sample of Millennials (n = 132). The same final three-factor measurement model (MF) that was specified for the pooled and Boomer samples was fit to the Millennial sample. Upon fitting the specified three-factor measurement model for the Millennial sample, fit statistics indicated a well-fitting model (CFI = .980 and SRMR = .039). Results of the three-factor measurement model for the Millennial sample were reported in Table 22.

Table 22

**Fit Indices for the Three-Factor Measurement Model for the Millennial Sample**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boomer Sample</td>
<td>33.73</td>
<td>21</td>
<td>.068</td>
<td>.039</td>
<td>0</td>
<td>.965</td>
<td>.980</td>
</tr>
</tbody>
</table>

*Note.* df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.

Following the path of the preceding analyses, the Millennial sample three-factor measurement model was further interrogated by reviewing the standardized regression weights. As was shown in Table 23, all factors loaded above the minimum recommended
cutoff limit of .5, and seven of the nine items loaded above the more rigorous level of .7 (Bagozzi & Yi, 1988). Factor loadings ranged from .505 to .891, no factors loaded above .95, and no factors were required to be deleted from the three-factor measurement model. Review of the structure coefficients in Table 23 demonstrated that all manifest variables loaded correctly to their respective factors (Graham et al., 2003).

Table 23

**Pattern (P) and Structure (S) Coefficients for the Millennial Sample Measurement Model**

<table>
<thead>
<tr>
<th>Construct Variable</th>
<th>Vigor</th>
<th>Dedication</th>
<th>Absorption</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>S</td>
<td>P</td>
</tr>
<tr>
<td>VI1</td>
<td>.716</td>
<td>.716</td>
<td>.000</td>
</tr>
<tr>
<td>VI2</td>
<td>.888</td>
<td>.888</td>
<td>.000</td>
</tr>
<tr>
<td>VI3</td>
<td>.804</td>
<td>.804</td>
<td>.000</td>
</tr>
<tr>
<td>DE2</td>
<td>.000</td>
<td>.677</td>
<td>.832</td>
</tr>
<tr>
<td>DE3</td>
<td>.000</td>
<td>.725</td>
<td>.891</td>
</tr>
<tr>
<td>DE4</td>
<td>.000</td>
<td>.529</td>
<td>.650</td>
</tr>
<tr>
<td>AB3</td>
<td>.000</td>
<td>.694</td>
<td>.000</td>
</tr>
<tr>
<td>AB4</td>
<td>.000</td>
<td>.610</td>
<td>.000</td>
</tr>
<tr>
<td>AB5</td>
<td>.000</td>
<td>.427</td>
<td>.000</td>
</tr>
</tbody>
</table>

Note. VI = Vigor, DE = Dedication, and AB = Absorption

Table 24 provided statistics for analyzing the internal structure of the Millennial sample measurement model. The composite reliability (CR ≥.6) and average variance extracted (AVE ≥.5) were all above the recommended thresholds (Bagozzi & Yi, 1988) with the exception of the AVE (.48) for Absorption. The CR range (.73 - .85) suggested sufficient reliability. The AVE range (.48 - .65) suggested the Millennial sample three-factor measurement model demonstrated overall convergent validity, with caution being noted for Absorption that was just below the recommended (AVE ≥.5) level at .48. Discriminant validity was once again evaluated by comparing the square root of the AVE statistic to the correlations for each individual factor (Bagozzi & Yi, 1988). Similar to the pooled and Boomer samples, discriminant validity was not confirmed as the
individual factor correlations were all equal to or greater than the square root of AVE for the Millennial sample.

Table 24

*Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Millennial Sample*

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vigor</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Dedication</td>
<td>.81</td>
<td>.80</td>
<td></td>
</tr>
<tr>
<td>3. Absorption</td>
<td>.85</td>
<td>.83</td>
<td>.69</td>
</tr>
<tr>
<td>CR</td>
<td>.85</td>
<td>.84</td>
<td>.73</td>
</tr>
<tr>
<td>AVE</td>
<td>.65</td>
<td>.64</td>
<td>.48</td>
</tr>
</tbody>
</table>

*Note.* Square root of AVE is along the diagonal.

**Measurement Model Comparison between Boomers and Millennials**

Comparisons of the fit indices between the Boomer and Millennial three-factor measurement models revealed that each sample fit the specified measurement model well. The Millennial sample demonstrated an overall better fit (CFI = .980, SRMR = .039) compared to the Boomer sample (CFI = .952, SRMR = .046), and both models met the stringent criteria (CFI ≥ .95, SRMR ≤ .08, Kline, 2006) for the fit indices (i.e., CFI, SRMR) recommended by Hu and Bentler (1998) for maximum likelihood modeling using smaller samples. Table 25 provided the summary of fit indices for both samples.

Table 25

*Fit Indices for Boomer and Millennial Three-Factor Measurement Models*

<table>
<thead>
<tr>
<th>Model</th>
<th>χ²</th>
<th>df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boomer</td>
<td>59.48</td>
<td>21</td>
<td>.118</td>
<td>.046</td>
<td>0</td>
<td>.917</td>
<td>.952</td>
</tr>
<tr>
<td>Millennial</td>
<td>33.73</td>
<td>21</td>
<td>.068</td>
<td>.039</td>
<td>0</td>
<td>.965</td>
<td>.980</td>
</tr>
</tbody>
</table>

*Note.* df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.
Review of the pattern and structure coefficients for the Boomer (Table 20) and Millennial (Table 23) samples uncovered no major differences in the patterns of loadings. Items VI1 and VI2 showed the only notable variation between the samples whereas both had lower factor loadings for the Millennial sample compared to the Boomer sample. Conversely, VI3 had a higher factor loading for the Millennial sample. Generally, the strength of loadings and patterns of loadings were consistent between the samples. Considering the model fit indices were at acceptable levels and the factor loadings had expected patterns, the researcher deemed the data to be sufficient to continue with measurement invariance testing using the combined data set of Boomers and Millennials.

**Measurement Invariance Results for the Three-Factor Model**

The measurement invariance testing proceeded through a set of hierarchically defined models and began with configural invariance testing before moving to metric and scalar models. In the configural invariance model, the three-factor correlated measurement model (MF) was used and fit to the combined data set of Boomers and Millennials while retaining the independent group definitions within the model parameter settings. Defining the groups by the generational variable allowed the AMOS software to utilize the single input data set and retain the group differences for invariance testing.

As elaborated in the measurement model section of this paper, the sample size was reduced through propensity score matching. For the measurement model evaluation and continuing for this measurement invariance section, the researcher utilized CFI and SRMR as the fit indices. This decision was supported by review of the literature and subsequently followed the recommendations of Hu and Bentler (1998). In addition, the measurement invariance was proposed to utilize both the $\Delta$CFI and $\Delta\chi^2$ results to
determine change and invariance between models; however, following the suggested guidance in Cheung and Rensvold (2002), the researcher reported CFI and $\chi^2$ results but based model change differences and hypotheses decisions on CFI. This decision was based upon the sample size outcome from propensity score matching and commentary within Cheung and Rensvold that “models rarely seem to fit by that criterion (i.e., $\chi^2$) due to its well-known dependence on sample size” (2002, p. 239). Cheung and Rensvold continued and stated that CFI was a robust statistic “for testing the between-group invariance of CFA models” since it did not suffer from the weaknesses of the $\chi^2$ statistic (2002, p. 250). Thus, while CFI and $\chi^2$ results were reported, model evaluation and decisions were based upon CFI.

Retaining the measurement model respecification parameters, the configural invariance model (also commonly understood as the unconstrained invariance model) was fitted in the same manner as the measurement model with no constraints to the factor loadings or intercepts. In order to confirm configural invariance, the “pattern of free and fixed loadings” was required to be the same (Nimon & Reio, 2012, p. 205) and the model was required to have good model fit, as determined through evaluation of the model fit indices in Table 26. As expected, the pattern and structure coefficients were the same as the delineated Boomer and Millennial measurement models since the samples were independently fitted to the measurement model, and the same measurement model specification was used for the configural invariance model. As noted in Table 26, the fit statistics represented a well-fitting model (CFI = .964 and SRMR = .046). Provided that the patterns of factor loadings were similar between groups and that the fit statistics revealed a well-fitting model, configural invariance was confirmed and H1a was
supported. Support for H1a allowed for the continuation of measurement invariance testing to metric invariance.

The model specification for metric invariance was determined by using the three-factor configural invariance model and constraining like factor loadings to be equal between groups. Support for metric invariance was determined by evaluating the model fit indices and comparing the metric invariance model to the configural invariance model. As presented in Table 26, the three-factor metric invariance model was well-fitting with a CFI = .965 and SRMR = .047. Comparison of the metric invariance model to the configural invariance model resulted in a ΔCFI of .00. This result provided support for metric invariance for the three-factor solution between Boomer and Millennial groups and support for H2a. Since H2a was supported, scalar invariance testing was conducted as the next phase in the hierarchical modeling.

The hierarchical model testing proceeded to scalar invariance testing. In scalar invariance models, both the factor loadings and intercepts for like items were constrained to be equal across groups. Scalar invariance began by using the metric invariance model and constraining the intercepts. The metric invariance model already had the factor loadings constrained. Model evaluation was the same as for metric invariance testing except the metric invariance and scalar fit indices were compared for invariance. Review of Table 26 showed that the three-factor scalar invariance model was well-fitting with a CFI = .955 and SRMR = .047. In comparing the ΔCFI between the models, there was a -.01 difference. The -.01 difference was directly at the threshold limit where Cheung and Rensvold (2002) recommended invariance between models. Thus, since the ΔCFI did not exceed the -.01 difference, support for scalar invariance between Boomers and
Millennials for the three-factor correlated solution was found, and support for H3a was found. Given that support for H3a was found and scalar invariance was determined between the Boomer and Millennial groups, latent mean analysis could be conducted.

Table 26

*Tests of Measurement Invariance for the Three-Factor Correlated Model*

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>TLI</th>
<th>CFI</th>
<th>$\Delta$ df</th>
<th>$\Delta$ $\chi^2$</th>
<th>$\Delta$ p</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Con-figural</td>
<td>42</td>
<td>93.21</td>
<td>&lt;.01</td>
<td>.068</td>
<td>.046</td>
<td>.938</td>
<td>.964</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metric</td>
<td>48</td>
<td>97.65</td>
<td>&lt;.01</td>
<td>.063</td>
<td>.047</td>
<td>.948</td>
<td>.965</td>
<td>6</td>
<td>4.04</td>
<td>.67</td>
<td>.00</td>
</tr>
<tr>
<td>Scalar</td>
<td>57</td>
<td>121.54</td>
<td>&lt;.01</td>
<td>.066</td>
<td>.047</td>
<td>.943</td>
<td>.955</td>
<td>9</td>
<td>23.89</td>
<td>.00</td>
<td>-.01</td>
</tr>
</tbody>
</table>

*Note.* df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.

**Latent Mean Analysis Results for the Three-Factor Model**

The latent mean analysis for the study followed the recommended procedure outlined by Byrne (2010). The scalar invariance model for the three-factor solution was retained and the mean constraints for the Boomer group were removed for the vigor, dedication, and absorption factors. Removing the mean constraints by recoding the “0,” values to “mn,” allowed the Boomer group to be freely estimated, while the Millennial group constraints were retained at “0,”. Differences in mean estimates were reflective of the latent mean differences between the Boomer and Millennial groups. Positive latent means would reflect the Boomers to be more vigorous, dedicated, and absorbed in their work. Negative latent means would signal the opposite.

Upon processing the model, latent means and $p$-values were returned from AMOS. Table 27 contains the returned statistics from the latent mean analysis. The reported means for each factor were all negative, indicating that Boomers were less
vigorous, less dedicated, and less absorbed in their work compared to Millennials.

Review of the $p$-values showed that while all means were negative, only dedication was statistically significant at $p = .037$. Based upon the latent mean analysis results reported in Table 27 for the three-factor model, H4a was rejected due to the negative means. Boomers were not more vigorous, dedicated, or absorbed in their work.

Table 27

<table>
<thead>
<tr>
<th>Construct</th>
<th>$M$</th>
<th>$p$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigor</td>
<td>-.276</td>
<td>.072</td>
<td>-.222</td>
</tr>
<tr>
<td>Dedication</td>
<td>-.195</td>
<td>.037</td>
<td>-.258</td>
</tr>
<tr>
<td>Absorption</td>
<td>-.062</td>
<td>.563</td>
<td>-.071</td>
</tr>
</tbody>
</table>

*Note.* Means of Boomers were freely estimated.

**Pooled Measurement Model for the Single-Factor Model**

The single-factor analyses closely followed the three-factor model sequence of analyses. Confirmatory factor analysis using SPSS AMOS v24.0 was once again used to analyze the data, and a single-factor measurement model was specified for the pooled sample of Boomers and Millennials. The same reporting mechanisms were used for the single-factor model including factor loadings, chi-square ($\chi^2$), degrees of freedom ($df$), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), standardized residual covariances (SRC), Tucker-Lewis index (TLI), and comparative fix index (CFI) statistics. Since the single-factor analysis utilized the same pooled sample ($n = 264$) as the three-factor analysis, the researcher used CFI and SRMR as the fit indices to evaluate model fit (Hu & Bentler, 1998). Upon fitting the initial original single-factor measurement model (M0S), fit statistics indicated a poorly
fitting model (CFI = .883 and SRMR = .062). Review of the modification indices indicated model misspecification due to error covariances among the items. Again following Byrne (2010), the measurement model was respecified and reestimated. Theoretical support was again used in order to justify the item error correlations. First and similar for the three-factor model, errors for AB4 and AB5 were again correlated to produce the Respecified One Model (M1S). Lovakov et al. (2017) provided support for the error correlations between AB4 and AB5 for the single-factor model, and Seppala et al., (2009) provided support for correlating errors for AB4 and AB5 for both the three-factor and single-factor models. The fit indices improved, but the CFI = .903 was still below acceptable fit level. Next, VI1 and VI2 were correlated to produce the Respecified Two Model (M2S). In a recent psychometric properties study for UWES, VI1 and VI2 were correlated in a single-factor model of UWES (Lovakov et al., 2017). Reasoning for the error correlation of VI1 and VI2 was given that the “covariances could be meaningfully explained (sequence, similar wording)” (Lovakov et al., 2017, p. 151). The Respecified Two Model (M2S) improved the overall fit for CFI = .951, but the fit was just barely met the cut off criteria. Since the Respecified Two Model (M2S) only barely met the cutoff and by using additional literature support, errors for DE4 and AB4 were correlated, similar to the three-factor model. Literature support for correlating these errors was found in Klassen et al (2012) where errors were specifically correlated for items DE4 and AB4. Support was also found in Schaufeli and Bakker due to the “special role that this dimension (absorption) seems to play, as compared to both core characteristics of engagement (i.e., VI and DE)”, and the error correlation recommendation that Schaufeli and Bakker proposed resulted in the Final Respecified
Model (MFS) (2004, p. 309). The resulting Final Respecified (MFS) single-factor measurement model was well-fitting with a CFI = .968 and SRMR = .038. Results of model respecification process for single-factor measurement model were reported in Table 28.

Table 28

Fit Indices for Original and Respecified Single-Factor Measurement Models

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC ($\geq 2.58$)</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0S: Original</td>
<td>190.40</td>
<td>27</td>
<td>.152</td>
<td>.062</td>
<td>1</td>
<td>.844</td>
<td>.883</td>
</tr>
<tr>
<td>M1S: Respecified One</td>
<td>161.21</td>
<td>26</td>
<td>.141</td>
<td>.055</td>
<td>1</td>
<td>.866</td>
<td>.903</td>
</tr>
<tr>
<td>M2S: Respecified Two</td>
<td>93.11</td>
<td>25</td>
<td>.102</td>
<td>.043</td>
<td>0</td>
<td>.930</td>
<td>.951</td>
</tr>
<tr>
<td>MFS: Final Respecified</td>
<td>69.41</td>
<td>24</td>
<td>.085</td>
<td>.038</td>
<td>0</td>
<td>.951</td>
<td>.968</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.

The Final Respecified (MFS) single-factor measurement model was investigated by reviewing the standardized regression weights. As was shown in Table 29, all factors loaded above the minimum recommended cutoff limit of .5, and five of the nine items loaded above the more rigorous level of .7 (Bagozzi & Yi, 1988). Factor loadings ranged from .495 to .867, no factors loaded above .95, and no factors were required to be deleted from the single-factor measurement model. Of note, AB5 had a factor loading of .495 but was considered to have reached the minimum cutoff once rounded to .500, and all manifest variables were retained (Bagozzi & Yi, 1988).
Table 29

Pattern (P) and Structure (S) Coefficients for the Pooled Sample Single-Factor Measurement Model

<table>
<thead>
<tr>
<th>Construct Variable</th>
<th>UWES</th>
<th>P</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI1</td>
<td></td>
<td>.673</td>
<td>.673</td>
</tr>
<tr>
<td>VI2</td>
<td></td>
<td>.758</td>
<td>.758</td>
</tr>
<tr>
<td>VI3</td>
<td></td>
<td>.780</td>
<td>.780</td>
</tr>
<tr>
<td>DE2</td>
<td></td>
<td>.844</td>
<td>.844</td>
</tr>
<tr>
<td>DE3</td>
<td></td>
<td>.867</td>
<td>.867</td>
</tr>
<tr>
<td>DE4</td>
<td></td>
<td>.633</td>
<td>.633</td>
</tr>
<tr>
<td>AB3</td>
<td></td>
<td>.727</td>
<td>.727</td>
</tr>
<tr>
<td>AB4</td>
<td></td>
<td>.656</td>
<td>.656</td>
</tr>
<tr>
<td>AB5</td>
<td></td>
<td>.495</td>
<td>.495</td>
</tr>
</tbody>
</table>

Note. VI = Vigor, DE = Dedication, and AB = Absorption

Table 30 provided statistics for analyzing the internal structure of the pooled measurement model. The composite reliability (CR ≥.6) and average variance extracted (AVE ≥.5) were above the recommended thresholds (Bagozzi & Yi, 1988). The CR = .91 suggested sufficient reliability. The AVE = .52 suggested the pooled single-factor measurement model demonstrated convergent validity. Discriminant validity was not able to be confirmed by comparing the square root of the AVE statistic to the correlations since the solution was a single-factor model (Bagozzi & Yi, 1988).

Table 30

Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Single-Factor Pooled Sample Measurement Model

<table>
<thead>
<tr>
<th>Statistic</th>
<th>UWES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>.91</td>
</tr>
<tr>
<td>AVE</td>
<td>.52</td>
</tr>
<tr>
<td>Sqrt of AVE</td>
<td>.72</td>
</tr>
</tbody>
</table>

Note. UWES = Utrecht Work Engagement Scale.
After review of the reported fit statistics, measurements of reliability and validity, and pattern and structure coefficients, the pooled single-factor measurement model (MFS) portrayed in Figure 4.2 was considered satisfactory. The pooled sample was delineated by Boomers and Millennials for independent measurement model testing under the single-factor solution.

**Figure 5. Single-Factor Measurement Model (MFS)**

**Measurement Model Results for the Single-Factor Model Boomer Sample**

The Boomer sample \( (n = 132) \) was fit to the same single-factor measurement model that was specified for the pooled sample (MFS). Fit statistics for the Boomer sample indicated well-fitting model (CFI = .974 and SRMR = .042). Results of the single-factor measurement model for the Boomer sample were reported in Table 31.
The single-factor measurement model for the Boomer sample was interrogated further by analyzing the standardized regression weights. As was shown in Table 32, all but one factor loaded above the minimum recommended cutoff limit of .5, and six of the nine items loaded above the more rigorous level of .7 (Bagozzi & Yi, 1988). Factor loadings ranged from .471 to .884, no factors loaded above .95. The factor of AB5 loaded just below the minimum cutoff and close to the loading (.495) found in the pooled sample. The factor was retained to maintain the theoretical UWES-9 model structure as proposed by Schaufeli and colleagues (2006).

Table 31

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boomers</td>
<td>44.34</td>
<td>24</td>
<td>.080</td>
<td>.042</td>
<td>0</td>
<td>.962</td>
<td>.974</td>
</tr>
</tbody>
</table>

*Note. df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.*
Table 32

*Pattern (P) and Structure (S) Coefficients for the Boomer Sample Single-Factor Measurement Model*

<table>
<thead>
<tr>
<th>Construct Variable</th>
<th>UWES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>VI1</td>
<td>.743</td>
<td>.743</td>
</tr>
<tr>
<td>VI2</td>
<td>.808</td>
<td>.808</td>
</tr>
<tr>
<td>VI3</td>
<td>.805</td>
<td>.805</td>
</tr>
<tr>
<td>DE2</td>
<td>.884</td>
<td>.884</td>
</tr>
<tr>
<td>DE3</td>
<td>.867</td>
<td>.867</td>
</tr>
<tr>
<td>DE4</td>
<td>.665</td>
<td>.665</td>
</tr>
<tr>
<td>AB3</td>
<td>.734</td>
<td>.734</td>
</tr>
<tr>
<td>AB4</td>
<td>.651</td>
<td>.651</td>
</tr>
<tr>
<td>AB5</td>
<td>.471</td>
<td>.471</td>
</tr>
</tbody>
</table>

*Note.* VI = Vigor, DE = Dedication, and AB = Absorption

Table 33 provided statistics for analyzing the internal structure of the Boomer sample measurement model. The composite reliability (CR ≥ .6) and average variance extracted (AVE ≥ .5) were above the recommended thresholds (Bagozzi & Yi, 1988). The CR = .92 suggested sufficient reliability. The AVE = .56 suggested the Boomer sample single-factor measurement model demonstrated convergent validity.

Discriminant validity was not able to be confirmed by comparing the square root of the AVE statistic to the correlations since the solution was a single-factor model (Bagozzi & Yi, 1988).

Table 33

*Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Single-Factor Boomer Sample Measurement Model*

<table>
<thead>
<tr>
<th>Statistic</th>
<th>UWES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>.92</td>
</tr>
<tr>
<td>AVE</td>
<td>.56</td>
</tr>
<tr>
<td>Sqrt of AVE</td>
<td>.75</td>
</tr>
</tbody>
</table>

*Note.* UWES = Utrecht Work Engagement Scale.
Measurement Model Results for the Single-Factor Model Millennial Sample

The Millennial sample \((n = 132)\) was fit to the same single-factor measurement model (MFS) that was specified for the pooled and Boomer samples. Fit statistics for the Millennial sample indicated well-fitting model (CFI = .964 and SRMR = .044). Results of the single-factor measurement model for the Millennial sample were reported in Table 34.

Table 34

*Fit Indices for the Single-Factor Measurement Model for the Millennial Sample*

<table>
<thead>
<tr>
<th>Model</th>
<th>(\chi^2)</th>
<th>(df)</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>SRC</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millennials</td>
<td>46.59</td>
<td>24</td>
<td>.085</td>
<td>.044</td>
<td>0</td>
<td>.946</td>
<td>.964</td>
</tr>
</tbody>
</table>

*Note. df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.*

For the single-factor measurement model results shown in Table 35, all factors loaded above the minimum recommended cutoff limit of .5, and five of the nine items loaded above the more rigorous level of .7 (Bagozzi & Yi, 1988). Factor loadings ranged from .499 to .859, no factors loaded above .95. The factor of AB5, similar to the pooled sample, loaded just below the minimum cutoff and close to the loading (.495) found in the Boomer sample. The factor was retained to maintain the theoretical UWES-9 model structure as proposed by Schaufeli and colleagues (2006) and since the AB5 factor, once rounded, met the minimum cutoff level (Bagozzi & Yi, 1988).
Table 35

*Pattern (P) and Structure (S) Coefficients for the Millennial Sample Single-Factor Measurement Model*

<table>
<thead>
<tr>
<th>Construct Variable</th>
<th>UWES P</th>
<th>UWES S</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI1</td>
<td>.617</td>
<td>.617</td>
</tr>
<tr>
<td>VI2</td>
<td>.722</td>
<td>.722</td>
</tr>
<tr>
<td>VI3</td>
<td>.747</td>
<td>.747</td>
</tr>
<tr>
<td>DE2</td>
<td>.809</td>
<td>.809</td>
</tr>
<tr>
<td>DE3</td>
<td>.859</td>
<td>.859</td>
</tr>
<tr>
<td>DE4</td>
<td>.646</td>
<td>.646</td>
</tr>
<tr>
<td>AB3</td>
<td>.735</td>
<td>.735</td>
</tr>
<tr>
<td>AB4</td>
<td>.669</td>
<td>.669</td>
</tr>
<tr>
<td>AB5</td>
<td>.499</td>
<td>.499</td>
</tr>
</tbody>
</table>

*Note.* VI = Vigor, DE = Dedication, and AB = Absorption

Table 36 provided statistics for analyzing the internal structure of the Millennial sample measurement model. The composite reliability (CR ≥ .6) and average variance extracted (AVE ≥ .5) were above the recommended thresholds (Bagozzi & Yi, 1988). The CR = .90 suggested sufficient reliability. The AVE = .50 suggested the Millennial sample single-factor measurement model demonstrated convergent validity, albeit the AVE was at the cutoff level for convergent validity. Discriminant validity was not able to be confirmed by comparing the square root of the AVE statistic to the correlations since the solution was a single-factor model (Bagozzi & Yi, 1988).

Table 36

*Implied Correlations, Average Variance Extracted (AVE), and Composite Reliability (CR) for the Single-Factor Millennial Sample Measurement Model*

<table>
<thead>
<tr>
<th>Statistic</th>
<th>UWES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>.90</td>
</tr>
<tr>
<td>AVE</td>
<td>.50</td>
</tr>
<tr>
<td>Sqrt of AVE</td>
<td>.71</td>
</tr>
</tbody>
</table>

*Note.* UWES = Utrecht Work Engagement Scale.
Single-Factor Measurement Model Comparison between Boomers and Millennials

Comparison of the fit indices between the Boomer and Millennial single-factor measurement models revealed the samples both fit the specified measurement model adequately. Conversely to the three-factor solution, the Boomer sample demonstrated an overall better fit (CFI = .974, SRMR = .042) compared to the Millennial sample (CFI = .964, SRMR = .044), and both models met the stringent criteria (CFI ≥ .95, SRMR ≤ .08, Kline, 2006) for the fit indices (i.e., CFI, SRMR) recommended by Hu and Bentler (1998) for maximum likelihood modeling using smaller samples. Table 37 provided the summary of fit indices for both samples.

Table 37

<table>
<thead>
<tr>
<th>Fit Indices for Boomer and Millennial Single-Factor Measurement Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Boomers</td>
</tr>
<tr>
<td>Millennials</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.

Review of the pattern and structure coefficients for the Boomer (Table 32) and Millennial (Table 35) samples uncovered no major differences in the patterns of loadings. Item VI1 had a slight variation between the samples. Similar to the three-factor model, VI1 had a higher factor loading for the Boomer sample compared to the Millennials, and overall, the vigor items all had higher factor loadings on Boomers. This is opposite from the three-factor solution where VI1 and VI2 loaded higher on Boomers, but VI3 loaded higher on Millennials. Also within the single-factor solution, all three absorption items loaded higher for Millennials than Boomers, albeit with only a slight advantage in
loadings. Considering the model fit indices and factor loadings were at acceptable levels and had consistent patterns, the researcher deemed the data to be sufficient to continue with measurement invariance testing using the combined data set of Boomers and Millennials for the single-factor solution.

**Measurement Invariance Results for the Single-Factor Model**

Following the process for the three-factor model, measurement invariance testing proceeded through the same set of hierarchically defined models and began with configural invariance testing before moving to metric and scalar models. The specified single-factor correlated measurement model (MFS) was used and fit to the combined data set of Boomers and Millennials to the configural measurement invariance model. The \( \Delta CFI \) and \( \Delta \chi^2 \) statistics were reported, and model evaluation and decisions were based upon CFI (Cheung & Rensvold, 2002).

The configural invariance model for the single-factor solution was fitted to the single-factor measurement model with no constraints on any factor loadings or intercepts. As expected, the pattern and structure coefficients were the same as the independently fitted Boomer and Millennial measurement models. As noted in Table 38, the fit statistics represented a well-fitting model (CFI = .970 and SRMR = .042). Provided the pattern of factor loadings were similar between groups and that the fit statistics revealed a well-fitting single-factor model, configural invariance was confirmed and H1b was supported. Since support for H1b was found, the testing continued to metric invariance.

The model specification for metric invariance was determined by using the single-factor configural invariance model and constraining like factor loadings to be equal between groups. Support for metric invariance was determined by evaluating the model
fit indices and comparing the metric invariance model to the configural invariance model for the single-factor solution. As presented in Table 38, the single-factor metric invariance model was well-fitting with a CFI = .971 and SRMR = .043. Comparison of the metric invariance model to the configural invariance model resulted in a ΔCFI of .00. This result provided support for metric invariance for the single-factor solution between Boomer and Millennial groups and support for H2b. Since H2b was supported, scalar invariance testing was conducted as the next phase in the hierarchical modeling.

Testing moved to scalar invariance testing within the final phase of the single-factor process. Using the metric invariance single-factor model, the factor loadings and intercepts for like items were constrained to be equal across groups. Review of Table 38 showed that the single-factor scalar invariance model was well-fitting with a CFI = .961 and SRMR = .043. Comparison of the ΔCFI between the models showed there was a -.01 difference, within the limit defined by Cheung and Rensvold (2002) to confirm invariance between models. Since the ΔCFI did not exceed the -.01 difference, support for scalar invariance between Boomers and Millennials for the single-factor correlated solution was found, and H3b was supported. Latent mean analysis for the single-factor model between Boomers and Millennials was now allowable due to the support of H3b.

Table 38

Tests of Measurement Invariance for the Single-Factor Correlated Model

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>p</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>TLI</th>
<th>CFI</th>
<th>Δdf</th>
<th>Δ$\chi^2$</th>
<th>p</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Con-figural</td>
<td>48</td>
<td>90.93</td>
<td>&lt;.01</td>
<td>.058</td>
<td>.042</td>
<td>.955</td>
<td>.970</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metric</td>
<td>56</td>
<td>96.72</td>
<td>&lt;.01</td>
<td>.053</td>
<td>.043</td>
<td>.963</td>
<td>.971</td>
<td>8</td>
<td>5.79</td>
<td>.67</td>
<td>0.00</td>
</tr>
<tr>
<td>Scalar</td>
<td>65</td>
<td>120.17</td>
<td>&lt;.01</td>
<td>.057</td>
<td>.043</td>
<td>.957</td>
<td>.961</td>
<td>9</td>
<td>23.45</td>
<td>.01</td>
<td>-.01</td>
</tr>
</tbody>
</table>

*Note. df = degrees of freedom, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, SRC = standardized residual covariance, TLI = Tucker-Lewis Index, and CFI = comparative fit index.*

156
Latent Mean Analysis Results for the Single-Factor Model

The latent mean analysis for the single-factor model replicated the procedure followed for the three-factor model which was outlined by Byrne (2010). The scalar invariance model for the single-factor solution was used and the mean constraint for the Boomer group was removed for the UWES factor. Similar to the three-factor latent mean analysis, a positive latent mean would reflect the Boomers to have a higher level of work engagement, and a negative latent mean would indicate the Millennials had a higher level of work engagement.

The latent mean and $p$-value were returned from AMOS, and Table 39 contained the returned statistics from the latent mean analysis. The reported mean was negative, indicating that Boomers were less engaged in their work compared to Millennials. The $p$-value was not statistically significant at $p = .056$; although, the value was just above the required cutoff of $p \leq .05$ for statistical significance. Based upon the latent mean analysis results reported in Table 39 for the single-factor model, H4b was rejected due to the negative mean value for UWES. Boomers were not more engaged in their work compared to Millennials.

Table 39

*Latent Mean Difference between Boomers and Millennials for the Single-Factor Scalar Invariance Model*

<table>
<thead>
<tr>
<th>Construct</th>
<th>$M$</th>
<th>$p$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UWES</td>
<td>-.177</td>
<td>.056</td>
<td>-.236</td>
</tr>
</tbody>
</table>

*Note.* Means of Boomers were freely estimated.
Descriptive Statistics

The data were analyzed in R to provide tables of descriptive statistics. Following Teo et al. (2009), the pooled sample was analyzed first, followed by analysis of the sub-samples. Table 40 reported the descriptive statistics for the pooled sample of $n = 264$, Table 41 reported the Boomer sub-sample, and Table 42 reported the Millennial sub-sample.

Table 40

Descriptive Statistics for the Pooled Sample

<table>
<thead>
<tr>
<th>Construct</th>
<th>$n$</th>
<th>$M$</th>
<th>SD</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>UWES</th>
<th>VI</th>
<th>DE</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>UWES</td>
<td>264</td>
<td>5.17</td>
<td>0.98</td>
<td>-0.27</td>
<td>-0.34</td>
<td>.87</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VI</td>
<td>264</td>
<td>4.91</td>
<td>1.15</td>
<td>-0.27</td>
<td>-0.15</td>
<td>.90</td>
<td>.90</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DE</td>
<td>264</td>
<td>5.44</td>
<td>1.08</td>
<td>-0.31</td>
<td>-0.75</td>
<td>.91</td>
<td>.76</td>
<td>.91</td>
<td>-</td>
</tr>
<tr>
<td>AB</td>
<td>264</td>
<td>5.15</td>
<td>1.07</td>
<td>-0.32</td>
<td>0.14</td>
<td>.86</td>
<td>.64</td>
<td>.68</td>
<td>.86</td>
</tr>
</tbody>
</table>

*Note. $n$ = Sample size, $M$ = Mean, $SD$ = Standard Deviation, $VI$ = Vigor, $DE$ = Dedication, and $AB$ = Absorption. Cronbach’s alpha is on the diagonal.*

Table 41

Descriptive Statistics for the Boomer Sample

<table>
<thead>
<tr>
<th>Construct</th>
<th>$n$</th>
<th>$M$</th>
<th>SD</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>UWES</th>
<th>VI</th>
<th>DE</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>UWES</td>
<td>132</td>
<td>5.06</td>
<td>0.94</td>
<td>-0.05</td>
<td>-0.52</td>
<td>.88</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VI</td>
<td>132</td>
<td>4.78</td>
<td>1.11</td>
<td>-0.20</td>
<td>-0.05</td>
<td>.92</td>
<td>.92</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DE</td>
<td>132</td>
<td>5.34</td>
<td>1.00</td>
<td>-0.12</td>
<td>-0.78</td>
<td>.93</td>
<td>.83</td>
<td>.93</td>
<td>-</td>
</tr>
<tr>
<td>AB</td>
<td>132</td>
<td>5.08</td>
<td>1.01</td>
<td>-0.06</td>
<td>-0.10</td>
<td>.85</td>
<td>.62</td>
<td>.69</td>
<td>.85</td>
</tr>
</tbody>
</table>

*Note. $n$ = Sample size, $M$ = Mean, $SD$ = Standard Deviation, $VI$ = Vigor, $DE$ = Dedication, and $AB$ = Absorption. Cronbach’s alpha is on the diagonal.*
Table 42

Descriptive Statistics for the Millennial Sample

<table>
<thead>
<tr>
<th>Construct</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>UWES</th>
<th>VI</th>
<th>DE</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>UWES</td>
<td>132</td>
<td>5.27</td>
<td>1.02</td>
<td>0.49</td>
<td>0.11</td>
<td>0.86</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VI</td>
<td>132</td>
<td>5.05</td>
<td>1.18</td>
<td>0.37</td>
<td>-0.20</td>
<td>0.89</td>
<td>0.88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DE</td>
<td>132</td>
<td>5.54</td>
<td>1.15</td>
<td>-0.50</td>
<td>-0.70</td>
<td>0.90</td>
<td>0.70</td>
<td>0.90</td>
<td>-</td>
</tr>
<tr>
<td>AB</td>
<td>132</td>
<td>5.21</td>
<td>1.12</td>
<td>0.54</td>
<td>0.34</td>
<td>0.87</td>
<td>0.64</td>
<td>0.69</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Note. n = Sample size, M = Mean, SD = Standard Deviation, VI = Vigor, DE = Dedication, and AB = Absorption. Cronbach’s alpha is on the diagonal.

Boomers consistently had lower observed mean scores ($M_{UWES} = 5.06$, $M_{VI} = 4.77$, $M_{DE} = 5.33$, and $M_{AB} = 5.08$) compared to Millennials ($M_{UWES} = 5.27$, $M_{VI} = 5.05$, $M_{DE} = 5.53$, and $M_{AB} = 5.21$). The consistently lower observed mean scores for Boomers correspondingly resulted in negative observed mean differences between Boomers and Millennials ($MD_{UWES} = -.21$, $MD_{VI} = -.28$, $MD_{DE} = -.20$, and $MD_{AB} = -.13$). The calculation of effect sizes resulted in small levels of practical significance ($d_{UWES} = -.208$, $d_{VI} = -.241$, $d_{DE} = -.185$, and $d_{AB} = -.126$).

Common Method Variance Results

The comprehensive CFA marker technique was implemented for the study. Due to the sample size constraints derived from the propensity score matching, four of the eight ATCB items were used. The approach to use the four positively worded ATCB items followed the method prescribed in Simmering et al (2015) where ATCB and other latent marker variables were evaluated. The ATCB was an ideal latent marker variable as it was developed to be used as such within social science research (Miller & Chiodo, 2008).

Following the method developed by Williams et al (2010), the comprehensive CFA marker technique tested a series of models and compared the resulting fit statistics.
and changes in fit statistics between models to detect evidence of CMV. The first model that was built was the CFA model. This model included the three-factor correlated model of UWES-9 and added the four positively worded ATCB items loading to the ATCB latent variable. The latent variables were correlated without constraint to develop a picture of the overall model structure. Using the unstandardized results, the latent marker paths and latent marker variables were constrained to the regression weights from the CFA model to create a baseline model. The third model in the sequence was the Model-C or constrained model. In this model, a direct path was added from the ATCB latent variable to each of the substantive variables and constrained to the same value. Model-U, or the unconstrained model, retained the paths added in Model-C but removed the constraints from the paths and allowed the paths to be freely estimated. The final model was Model-R or the restricted model. For this last step in the process, Model-C was replicated except that the covariances between the substantive factors were forced to the covariance values determined from the baseline model.

The comprehensive CFA marker technique was processed for the pooled sample of Boomers and Millennials, the Boomer only sample, and the Millennial only sample. This multi-sample method followed the suggestions of Craighead et al. (2011) and helped to determine whether any CMV was present in the individual sub-samples that may have been masked when only evaluating the pooled sample. The results for the pooled sample were reported in Table 43, for the Boomer sample in Table 44, and for the Millennial sample in Table 45. The results of the CFA marker testing were highly consistent for the three samples, suggesting that the sub-samples did not deviate from the pooled sample findings and that the results from all three samples could be interpreted the same.
Following interpretation guidance in Williams et al (2010), the baseline model was compared to Model-C, and a statistically significant result between the models was noted, which suggested there was shared CMV between the latent marker variable and substantive variables. Comparison of Model-C to the unconstrained Model-U showed a negligibly better fit for Model-U (ΔCFI = .002) and a non-statistically significant p-value (p = .12). This indicated that the CMV was constant for the indicators. A final comparison between Model-C and Model-R, the restricted model, showed no statistically significant difference, indicating that the presence of CMV was not biasing the relationships among the substantive variables (Williams et al., 2010).

Table 43

*Common Method Variance Testing with the Pooled Sample*

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>p</th>
<th>vs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFA</td>
<td>56</td>
<td>139.53</td>
<td>.952</td>
<td>.075</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Base</td>
<td>67</td>
<td>174.76</td>
<td>.939</td>
<td>.078</td>
<td>11</td>
<td>35.23</td>
<td>.013</td>
<td>.003</td>
<td>&lt;.01</td>
<td>CFA</td>
</tr>
<tr>
<td>M-C</td>
<td>66</td>
<td>122.43</td>
<td>.968</td>
<td>.057</td>
<td>1</td>
<td>52.33</td>
<td>.029</td>
<td>-.021</td>
<td>&lt;.01</td>
<td>Base</td>
</tr>
<tr>
<td>M-U</td>
<td>58</td>
<td>109.72</td>
<td>.970</td>
<td>.058</td>
<td>8</td>
<td>12.71</td>
<td>.002</td>
<td>.001</td>
<td>.12</td>
<td>M-C</td>
</tr>
<tr>
<td>M-R</td>
<td>69</td>
<td>123.33</td>
<td>.969</td>
<td>.055</td>
<td>3</td>
<td>0.90</td>
<td>.001</td>
<td>-.002</td>
<td>.83</td>
<td>M-C</td>
</tr>
</tbody>
</table>

*Note.* df = degrees of freedom, RMSEA = root mean square error of approximation, CFI = comparative fit index, CFA = confirmatory factor analysis, M-C = method constrained, M-U = method unconstrained, and M-R = method restricted.
Table 44

*Common Method Variance Testing with the Boomer Sample*

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>$p$</th>
<th>vs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFA</td>
<td>56</td>
<td>102.24</td>
<td>.951</td>
<td>.079</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Base</td>
<td>67</td>
<td>115.99</td>
<td>.948</td>
<td>.075</td>
<td>11</td>
<td>13.75</td>
<td>.003</td>
<td>-.004</td>
<td>.25</td>
<td>CFA</td>
</tr>
<tr>
<td>M-C</td>
<td>66</td>
<td>102.87</td>
<td>.961</td>
<td>.065</td>
<td>1</td>
<td>13.12</td>
<td>.013</td>
<td>$.010</td>
<td>&lt;.01</td>
<td>Base</td>
</tr>
<tr>
<td>M-U</td>
<td>58</td>
<td>85.64</td>
<td>.970</td>
<td>.060</td>
<td>8</td>
<td>17.23</td>
<td>.009</td>
<td>-.005</td>
<td>.03</td>
<td>M-C</td>
</tr>
<tr>
<td>M-R</td>
<td>69</td>
<td>102.97</td>
<td>.964</td>
<td>.061</td>
<td>3</td>
<td>10.10</td>
<td>.003</td>
<td>-.004</td>
<td>.99</td>
<td>M-C</td>
</tr>
</tbody>
</table>

*Note.* $df =$ degrees of freedom, RMSEA = root mean square error of approximation, CFI = comparative fit index, CFA = confirmatory factor analysis, M-C = method constrained, M-U = method unconstrained, and M-R = method restricted.

Table 45

*Common Method Variance Testing with the Millennial Sample*

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>$p$</th>
<th>vs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFA</td>
<td>56</td>
<td>115.56</td>
<td>.932</td>
<td>.090</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Base</td>
<td>67</td>
<td>142.26</td>
<td>.915</td>
<td>.093</td>
<td>11</td>
<td>26.70</td>
<td>.017</td>
<td>.003</td>
<td>.01</td>
<td>CFA</td>
</tr>
<tr>
<td>M-C</td>
<td>66</td>
<td>101.39</td>
<td>.960</td>
<td>.064</td>
<td>1</td>
<td>40.87</td>
<td>.045</td>
<td>.029</td>
<td>&lt;.01</td>
<td>Base</td>
</tr>
<tr>
<td>M-U</td>
<td>58</td>
<td>86.18</td>
<td>.968</td>
<td>.061</td>
<td>8</td>
<td>15.21</td>
<td>.008</td>
<td>.003</td>
<td>.06</td>
<td>M-C</td>
</tr>
<tr>
<td>M-R</td>
<td>69</td>
<td>104.03</td>
<td>.960</td>
<td>.062</td>
<td>3</td>
<td>2.64</td>
<td>.000</td>
<td>.002</td>
<td>.45</td>
<td>M-C</td>
</tr>
</tbody>
</table>

*Note.* $df =$ degrees of freedom, RMSEA = root mean square error of approximation, CFI = comparative fit index, CFA = confirmatory factor analysis, M-C = method constrained, M-U = method unconstrained, and M-R = method restricted.

**Hypotheses Summary**

A total of four hypotheses were predicted for the present study. Overall, three of the four hypotheses were supported, and measurement invariance was confirmed for both the three-factor and single-factor solutions at the configural, metric, and scalar levels.

When evaluating the latent mean differences between Boomers and Millennials, hypotheses for both the three-factor and single-factor solutions were rejected. The hypotheses and testing results are summarized in Table 46.
Hypotheses H1a – H4a predicted the three-factor model. Specifically, H1a predicted that configural invariance would be found between the Boomer and Millennial groups for the three-factor model. Support was evident for H1a as the fit indices of the configural model were well-fitting (CFI = .964, SRMR = .046) and the factor loadings demonstrated adequate patterns between the groupings. Metric invariance was also found for the three-factor model which supported H2a. Once the factor loadings were constrained across the Boomer and Millennial groups, the model fit for the metric invariance model continued to be well-fitting (CFI = .965, SRMR = .047), and the ΔCFI was .00 when the configural and metric invariance models were compared. Scalar invariance and H3a were supported due to the good model fit indices (CFI = .955, SRMR = .047) and a ΔCFI = -.01 between the metric and scalar invariance models, which was within the limits proposed by Cheung and Rensvold (2002). Only since support for H3a was found, the project proceeded to test hypothesis H4a. Counter to the supporting literature (e.g., Hoole & Bonnema, 2015; Park & Gursoy, 2012), Boomers were found to be less vigorous, dedicated, and absorbed in their work. While only dedication was found to have statistical significance (p = .037), all three resulting latent means were negative, indicating that Millennials were consistently more vigorous, dedicated, and absorbed at work. Although statistical significance was found for dedication, practical significance effect sizes were small for vigor (d = -.222), dedication (d = -.258), and absorption (d = -.071). The opposite mean score than predicted resulted in rejecting H4a.

The statistical analyses and hypotheses testing were repeated using a single-factor solution with the sole factor being UWES. In staying with the trend for the three-factor hypotheses, H1b was supported as configural invariance was confirmed for the single-
factor model. A good model fit (CFI = .970, SRMR = .042) and consistent pattern of factor loadings provided evidence in support of configural invariance. The testing of metric invariance again continued the trend, and support was found for H2b. The metric invariance model had good fit (CFI = .971, SRMR = .043), and the CFI between the configural and metric invariance models mirrored the three-factor result at .00. Scalar invariance was also found, which provided support for H3b. The scalar invariance model for the single-factor solution also had good fit (CFI = .961, SRMR = .043). The CFI between the metric and scalar invariance models was at the limit but within tolerance for measurement invariance at -0.01 (Cheung & Rensvold, 2002). Latent mean analysis for the single-factor model returned a negative mean of $M = -0.177$ with a $p$-value of .056 and a small effect size of $d = -0.236$. Due to the negative mean, H4b was rejected. The rejection and negative mean indicated that Boomers were less engaged in their work compared to Millennials. The result again was counter from much of the leisure and hospitality literature (e.g., Park & Gursoy, 2012; Solnet & Kralj, 2010) that purports Boomers to be more engaged than Millennials.

Table 46

Results of Predicted Hypotheses

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
<th>Model</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a</td>
<td>Configural Invariance</td>
<td>Three-Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>H2a</td>
<td>Metric Invariance</td>
<td>Three-Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>H3a</td>
<td>Scalar Invariance</td>
<td>Three-Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>H4a</td>
<td>Positive Latent Mean</td>
<td>Three-Factor</td>
<td>No</td>
</tr>
<tr>
<td>H1b</td>
<td>Configural Invariance</td>
<td>Single-Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>H2b</td>
<td>Metric Invariance</td>
<td>Single-Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>H3b</td>
<td>Scalar Invariance</td>
<td>Single-Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>H4b</td>
<td>Positive Latent Mean</td>
<td>Single-Factor</td>
<td>No</td>
</tr>
</tbody>
</table>

Note. Three-factor = three factor solution of vigor, dedication, and absorption. Single-factor = single factor solution of UWES.
Summary of the Chapter

Chapter 4 described the results for the present study. The outcomes from the multiple analyses used to analyze the data were reported and used to evaluate the hypotheses. The upcoming Chapter 5 of the document provides a discussion of findings, a set of recommendations for future research, and implications for practice.
Chapter 5 - Discussion

Introduction

Chapter 5 contains five sections. The first section constitutes a discussion of the results from Chapter 4 and the relationships to relevant literature. Implications to theory, research, and implications for HRD practitioners are found in the second section. The limitations of the study are included in the third section, and the fourth section provides suggestions for future research. The fifth section is a summary of the chapter.

Discussion of the Results

This section provides a discussion of the results for each of the four research hypotheses. Comparisons to the literature are included to discuss the similarities and differences found herein. Any notable impacts to the literature are also discussed. In order to present a logical sequence of discussion, this section is grouped into four parts according to the research hypotheses (a) configural measurement invariance (H1a, H1b), (b) metric measurement invariance (H2a, H2b), (c) scalar measurement invariance (H3a, H3b), and (d) latent mean analysis (H4a, H4b).

Configural Measurement Invariance (H1a, H1b)

Hypotheses H1a and H1b predicted configural measurement invariance would be found for the three-factor model and single-factor model of the UWES-9 instrument between Boomers and Millennials. Prior literature has shown support for configural invariance for UWES-9 across countries (Schaufeli et al., 2006) and for occupational groups (Seppala et al., 2009). In line with and confirming the findings of prior literature, configural invariance was found in the current study for both the three-factor (H1a) and
single-factor (H1b) models of UWES-9. The current study extends the body of literature by confirming configural measurement invariance between Boomers and Millennials within the selected sub-sectors of the leisure and hospitality industry. The confirmation of configural invariance also provided support to advance the analysis to metric invariance testing.

Configural invariance was established by evaluating the pattern of factor loadings for the nine items of UWES and by evaluating the fit indices for the CFA. The factor loadings for the three-factor solution were largely consistent between the samples, with slight variances in the loadings for vigor. The items VI1 and VI2 loaded higher for Boomers, and VI3 loaded higher for Millennials. The single-factor model also had a noted variance for the vigor items, but all three of the items loaded higher for Millennials than Boomers. The variations in loadings were minor, and the factor loadings were overall well above the cutoff criteria of ≥5 (Bagozzi & Yi, 1988), supporting configural invariance from this aspect of evaluation.

The fit indices were also evaluated to confirm configural invariance. Using the fit indices recommended by Hu and Bentler (1998) for smaller sample sizes, CFI and SRMR were evaluated for the models. The three-factor model returned a well-fitting model with a CFI = .964 and SRMR = .046. Similarly, the single-factor model returned a well-fitting model with a CFI = .970 and SRMR = .042. The combination of factor loading and fit indices evaluation provided ample evidence of configural measurement invariance for both the three-factor and single-factor model solutions, supporting both H1a and H1b.
Metric Measurement Invariance (H2a, H2b)

Whereas configural invariance has been frequently supported in the empirical literature, metric invariance has been more difficult to obtain for researchers. For example, during the original development of UWES-9, configural invariance was established across countries, but metric invariance was not supported across the countries (Schaufeli et al., 2006). It was suggested that metric invariance was not found due to the differences in occupational groups for each country in the sample (Schaufeli et al., 2006). Notwithstanding the previous example, empirical research has confirmed metric invariance for UWES for countries (UWES-17: Schaufeli, Martinez et al., 2002), race (UWES-17: Strom & Rothmann, 2003), and by occupational group (UWES-9: Seppala et al., 2009).

Hypotheses H2a and H2b predicted metric measurement invariance would be found for the three-factor model and single-factor model of the UWES-9 instrument between Boomers and Millennials. Metric invariance is determined by evaluating the fit indices of the metric model, which constrains the factor loadings of like items across groups, and by comparing the change in CFI between the configural model and metric model fit indices. Overall, the metric invariance model was well-fitting for the three-factor model with a CFI = .965 and SRMR = .047. The single-factor model was also well-fitting with a CFI = .971 and SRMR = .043. When comparing the metric model to the configural model, there was a ΔCFI of .00 for both the three-factor and single-factor solutions. This CFI and the well-fitting fit statistics provided evidence to support metric invariance between Boomers and Millennials for data from UWES-9 (Cheung & Rensvold, 2002) and support for H2a and H2b. The evidence of metric invariance
supported the previous literature that found metric invariance between countries (e.g., Seppala et al., 2009). It is also counter to the literature that did not find metric invariance support between occupational groups for UWES-9 (e.g., Schaufeli et al., 2006).

**Scalar Measurement Invariance (H3a, H3b)**

Scalar invariance is the strictest form of measurement invariance this study assessed. At the scalar level, the factor loadings and intercepts of the like items are constrained to be equal. Additionally, group mean comparisons are only feasible once scalar invariance has been determined (Wu et al., 2007). Comparison of group means without establishing scalar invariance threatens the interpretability and validity of empirical results (Nimon & Reio, 2012; Vandenberg & Lance, 2000).

Hypotheses H3a and H3b predicted scalar measurement invariance would be found for the three-factor model and single-factor model of the UWES-9 instrument between Boomers and Millennials. Similar to the steps used for metric invariance, scalar invariance was determined by evaluating the model fit indices and comparing the change in CFI from the metric model to the scalar model. The scalar CFA for the three-factor solution returned a well-fitting model with a CFI = .955 and SRMR = .047. Of note, the ΔCFI was -.01 for the three-factor solution. This -.01 CFI was directly at the limit Cheung and Rensvold (2002) established for invariance between models. While there was a distinguishable CFI, the CFI was within tolerance, scalar invariance was established for the three-factor UWES-9 model, and support for H3a was found. The single-factor model followed the trend of the three-factor model and had a CFI = .961 and SRMR = .043. The CFI between the single-factor metric model and single-factor scalar model was -.01, again directly at the limit for invariance between models (Cheung and
Rensvold, 2002). Based upon these results for the single-factor model, scalar invariance support was found and H3b was supported.

The support of scalar invariance is an important empirical contribution of this study. Other studies have found scalar invariance for engagement related constructs. For example, scalar invariance was found between educators and small business owners for the MBI-HS (Boles et al., 2000), and faculty in business colleges across countries for organizational commitment (Eisinga et al., 2012). The current study’s evidential support of H3a and H3b affords an extension to the literature and establishes scalar invariance between Boomers and Millennials for data from UWES-9 for a sample within the leisure and hospitality industry.

Particularly important is that the acceptance of H3a and H3b supports the directional research proposed by Schaufeli and colleagues (2006) whereas measurement invariance assessment “of the UWES should include similar occupational groups” (2006, p. 713). This research direction was recommended within the seminal empirical article reporting on the development of UWES-9, wherein only configural invariance was established across countries. Counter to the developmental article for UWES-9, configural and metric invariance had been established for UWES-17 across countries (Schaufeli, Martinez et al., 2002) and racial groups (Strom & Rothmann, 2003) previously to 2006. The lack of metric invariance for UWES-9 was attributed to the sample data available for testing rather than attributing fault to the instrument. The proposed future research direction by Schaufeli et al. (2006) toward using more similar groups of comparison (i.e., occupational groups) signaled the need for appropriate research design and sample definition for measurement invariance assessment. The
empirical findings of this study answered the research direction and confirmed the accuracy of the recommendations made by Schaufeli and colleagues (2006).

**Latent Mean Analysis (H4a, H4b)**

With scalar invariance confirmed, the study could feasibly continue to latent mean analysis for the differences between cohorts for the three-factor model (H4a) and single-factor model (H4b) of UWES-9. Hypothesis H4a predicted there would be statistically and practically significant mean differences for vigor, dedication, and absorption between Boomers and Millennials for the three-factor UWES model. Hypothesis H4b predicted there would be a statistically and practically significant mean difference for work engagement between Boomers and Millennials for the single-factor UWES model. The direction for both hypotheses was set that Boomers would be more engaged compared to Millennials. This direction of Boomers being more engaged was positioned using literature support for both within the leisure and hospitality industry (e.g., Kralj & Solnet, 2011; Park & Gursoy, 2012; Solnet et al., 2012; Solnet & Kralj, 2010) and outside the leisure and hospitality industry (Bano et al., 2015; Hoole & Bonnema, 2015; Lovakov et al., 2010).

The results of the latent mean analysis for the three-factor model rejected H4a. Boomers were not more vigorous, dedicated, or absorbed in their work compared to Millennials. Latent mean analysis returned a negative latent mean for vigor \((M = -.276)\), non-statistically significant \(p\)-value \((p = .072)\), and a small effect size \((d = -.222)\). The analysis returned a slightly different result for dedication with a negative latent mean \((M = -.195)\), a statistically significant \(p\)-value \((p = .037)\), and a small effect size \((d = -.258)\). The statistics for absorption showed a negative latent mean \((M = -.062)\), a non-
statistically significant $p$-value ($p = .563$), and a small effect size ($d = -.071$). The single-factor model also returned a negative mean ($M = -.177$), insignificant $p$-value ($p = .056$), and small effect size ($d = -.236$) for work engagement, effectively causing the researcher to reject $H_{4b}$. Although the study’s hypothesis $H_{4ab}$ was based upon latent mean differences, the observed mean scores reported in Chapter 4 are relevant for literature comparison since comparative literature examples have generally reported observed mean scores rather than latent mean scores. Observed mean differences between Boomers and Millennials were $MD_{UWES} = -.21$, $MD_{VI} = -.28$, $MD_{DE} = -.20$, and $MD_{AB} = -.13$. The calculation of effect sizes resulted in small levels of practical significance ($d_{UWES} = -.208$, $d_{VI} = -.241$, $d_{DE} = -.185$, and $d_{AB} = -.126$) for each of the observed mean differences.

This study contradicts the position that Boomers are more engaged in their work than Millennials. As these findings differ from much of the empirical literature, they necessitate deeper empirical exploration. For example, Hoole and Bonnema (2015) reported statistically significant mean differences with Boomers being more engaged than Millennials ($p < .001$) and GenXers ($p = .017$). The authors made recommendations regarding selection, reward, and development policies based upon the generational differences (Hoole & Bonnema, 2015). As a component of their limitations section, the authors stated that measurement invariance may have biased their results, and they recommended future research assess measurement invariance to overcome this limitation of group comparison research (Hoole & Bonnema, 2015).

A second example was found in the empirical study grounded in the leisure and hospitality industry by Park and Gursoy (2012). This study utilized the UWES-9
instrument and reported Boomers to be statistically and practically significantly more vigorous ($p < .05, d = .27$), dedicated ($p < .01, d = .62$), and absorbed ($p < .01, d = .35$) in their work. The authors reported that “employees in the older generations are likely to be more dedicated to, engrossed in, and even vigorous at work” (Park & Gursoy, 2012, p. 1198). Similar to the prior example, this study lacked the prerequisite statistical analysis of measurement invariance assessment and confirmation of scalar invariance for the generational cohorts before making comparisons between those groups. Even with the statistical limitation, Park and Gursoy (2012) has been influential to subsequent literature in the leisure and hospitality industry. Specifically, it influenced the following studies for generational differences (Dimitriou & Blum, 2015; Gursoy et al., 2013; Kim et al., 2016), psychological capital (Paek et al., 2016), and turnover (Brown et al., 2015; Chi et al., 2013; Lu & Gursoy, 2016).

A series of hospitality studies comparing a combined cohort of Boomers and GenXers to Millennials by levels of engagement provides a set of comparative examples on observed means. In all three studies within the series, the combined Boomer/GenXer group was reported to have a higher level of engagement than the Millennial group. Solnet and Kralj (2010) reported a practical significance of $d = .36$, Kralj and Solnet (2011) reported a practical significance of $d = .45$, and Solnet et al. (2012) reported a practical significance of $d = .45$. Comparing the current study’s reported practical significance for the single-factor model ($d_{UWES} = -.208$) shows that the level of significance is much lower for the present study and confirms the contradictory finding of Boomers being less engaged than Millennials within a similar industry scope.
The next two comparative examples both utilized UWES-9 and found the older generations to be more engaged than the younger. Bano et al. (2015) reported a strong practical significance of $d = 1.36$ when comparing GenXers to Millennials. This effect size is much larger than what the present study found within any aspect of the results. Lovakov et al. (2017) did not specifically categorize a sample by generational cohort, but the study did define older and younger generations that were close to the Boomer and Millennial generational cohort age ranges. Within the study, the Over 50 years of age group was found to be more engaged than the younger Under 30 age group with a mean difference of $M = .49$ and an effect size of $d = .47$ (Lovakov et al., 2017). These results are in contrast with the current study’s mean difference of $M = -.21$ and the effect size of $d = -.208$ for UWES. Even considering the differences between sample group age ranges, the Lovakov et al. (2017) study did provide a relevant comparative since the UWES-9 instrument was used and the sample was constrained to a single energy company in Russia.

A study within the South African context found Millennials to be more engaged than Boomers, consistent with the current study’s findings (Martins & Ledimo, 2016). Even though this comparative study found consistent results (i.e., Millennials are more engaged than Boomers), the effect size was small ($d = .31$), and the engagement instrument used was the Employee Engagement Instrument (EEI) (Martins & Ledimo, 2016). Contextually, the sample consisted of various governmental sectors instead of the leisure and hospitality industry. Considering all three of these factors, the results of Martins and Ledimo (2016) may not be greatly comparable to the present study, but the
study does provide an example of generational cohort research finding Millennials to be more engaged than Boomers.

While the current study’s results generally differ from the comparative literature, the differences are not fully comparable (not apples to apples) since the prior studies lacked measurement invariance assessments by generational cohort. The different outcomes in the comparable literature are perhaps explained by the potential fact that the generational cohort differences are a function of measurement variance instead of true generational cohort differences. As eloquently stated by Vandenberg and Lance, “the establishment of measurement invariance across groups is a logical prerequisite to conducting substantive cross-group comparisons” (2000, p. 4).

Vandenberg and Lance’s argument is perhaps a central source feeding the calls for more methodological rigor in generational cohort studies (Costanza et al., 2012; Lyons & Kuron, 2013) and HRD empirical research in general (Nimon & Reio, 2012; Reio, 2010). Measurement variance across comparison groups and the ensuing loss of statistical validity could be a principal reason for the inconsistent empirical findings in generational cohort research. Researchers should endeavor to apply research design and set sample criteria that support measurement invariance assessment and report the findings of measurement invariance analyses before reporting group mean differences. Without first establishing scalar invariance for the research groupings, readers cannot be confident that reported mean differences are due to true group differences rather than confounding measurement variance effects.
Implications

The implications are categorized into three sections: implications to theory, implications to research, and implications to business practice.

Implications to Theory

Generational cohort theory (Strauss & Howe, 1991) was a centerpiece of the current study. While the general notion of generational differences is commonly held, meta-analytic literature that synthesized generational cohort empirical research reported that an inconsistent and inconclusive landscape existed for empirical evidence of such differences (Costanza et al., 2012; Lyons & Kuron, 2013). The leisure and hospitality industry consistently reported Boomers to be more engaged than Millennials (e.g., Park & Gursoy, 2012; Solnet & Kralj, 2011). Yet, appropriate and required statistical methodology was lacking in the literature examples. Lyons and Kuron (2013) specifically called for more rigorous research design and methodology for both longitudinal and cross-sectional research in generational cohort empirical research.

Perhaps the inconsistencies of empirical findings for generational differences are partly attributable to the lack of appropriately applied research design and the lack of rigorous methodological approaches.

The current study applied a rigorous research design and allocated the appropriate statistical methodologies to derive the empirical evidence of the group mean generational differences. In fact, the study contradicted findings of empirical studies within the leisure and hospitality industry since Boomers were not more engaged than Millennials. This important finding was substantiated through the research design, sample equivalency, and measurement invariance assessment work that was completed before statistically
evaluating the group mean differences for the generations. Propensity score matching was used to equate the groups of Boomers and Millennials by their collected demographics. However, the groups could not be fully equated by their education level and maintained a practically significant difference of $V = .209$ after genetic PSM matching equated the groups by other demographic and job characteristics (i.e., gender, race, job level, and market tier). Generational cohort theory was leveraged to position the concept that employees could be categorized into comparable groups from different generations and that those groups could have different work perceptions (e.g., work engagement) simply due to their generational association. Using the present study as an example, the generational cohorts could be statistically equated by several demographics, but education levels remained different between the cohorts. Where generational cohort theory may inform researchers on how employees can be classified into comparable groups and equated by some characteristics (e.g., gender and race), the implication from the current study is that generational cohort theory may not fully inform where the generational cohorts will necessarily vary by other characteristics (i.e., education level). Additional research is needed to explore this implication for generational cohort theory and whether or not supporting theories may need to be considered for measurement invariance studies.

The multidimensional theory of burnout was the second major theory for the study (Maslach, 1999). The study specifically utilized the work engagement portion of burnout and the pervasive UWES-9 instrument as the foci of study. Much in the same light as the generational cohort theory, the contradictory finding that Boomers are not more engaged than Millennials within the leisure and hospitality industry was an
important implication for work engagement. An additional theoretical implication for work engagement is the confirmation of the factorial structures, for both the three-factor and single-factor models. Establishment of the factorial structures for the highly specified sample and also confirming the structure through independent samples, provides important evidential support for UWES-9 as a measure for work engagement in the leisure and hospitality industry.

The establishment of configural, metric, and scalar invariance for data from the UWES-9 instrument for a sample within the leisure and hospitality industry is perhaps the most significant theoretical implication for burnout and work engagement. Prior studies that attempted to confirm measurement invariance for data from UWES-9 often did not find support for metric or scalar invariance, citing group differences as a likely cause for the group measurement variance. This includes the original UWES-9 development article (Schaufeli et al., 2006). The finding of scalar invariance confirmed that Boomers and Millennials within the leisure and hospitality industry did interpret, understand, and attribute the same meanings for the work engagement instrument of UWES-9.

**Implications to Research**

This study has three implications for HRD research. First, the calls for more rigorous research methodology (Reio, 2010; Nimon & Reio, 2012) were partially answered. The study demonstrated a statistical analysis method that met the prerequisite statistical requirements for making group mean comparisons. The comprehensive nature of the study included considerations and statistical analyses for research design, sample criteria, sample equivalency, measurement invariance assessments, and latent mean
analysis. Notably, propensity score matching was utilized to equate the generational cohort samples before assessing measurement invariance. Using PSM as a statistical analysis to equate generational cohort groups is a methodological extension to prior comparable research within both the generational cohort and HRD literature areas. The future research section of this chapter discusses the possibility of creating a comparable measurement invariance study assessing the differences between Boomers and Millennials for data from UWES-9 with non-equated samples. Common method variance was also tested using the comprehensive CFA latent marker technique (Williams et al., 2010).

Second and related to the first implication, the study provided empirical evidence that Boomers and Millennials do interpret and find the same meaning for UWES-9. The findings should be constrained to interpretations within the leisure and hospitality industry and to within the narrower selection of sub-sectors the study included (i.e., convention center, cruise line, gaming or casino, lodging, marina, sporting facility, travel, and tourism). Whereas the current study partially confirmed findings of previous studies that found support for configural invariance (Schaufeli et al., 2006) and metric invariance (Seppala et al., 2009; Strom & Rothmann, 2003) for data from UWES, the current study found support for scalar invariance between generational cohorts for data from UWES-9 for the leisure and hospitality industry. This extension into the body of literature will support future research that concerns generational comparisons and the prerequisite statistical requirements for establishing scalar invariance before undertaking group mean comparisons.
Third and most importantly, the study found contradicting results for generational differences. Within the context of the study parameters and design, Boomers are not more engaged than Millennials. Except for Millennials having a statistically significantly higher level of dedication within the three-factor model of UWES, no other statistically significant mean differences were found. Also, for the three-factor and single-factor models of UWES, practically significant results were all at low significance levels. These results generally contradict other studies within the leisure and hospitality literature (e.g., Park & Gursoy, 2012; Solnet & Kralj, 2011) and outside the leisure and hospitality industry (Bano et al., 2015; Hoole & Bonnema, 2015). This finding was important for HRD researchers who are leveraging prior literature for hypothesis development within their studies. The present study indicated that we may know less than we thought we knew about work engagement differences between generational cohorts and that additional contextual research is needed.

The current study cannot be held as evidence that the previous literature results are incorrect because the studies are not directly comparable. The current study’s sample was in a specific and narrow slice of the leisure and hospitality industry, and the current study employed a more rigorous methodological approach compared to the literature examples. Notwithstanding, the current study does present findings, based upon a firm methodological foundation, that contradict established norms within the literature that Boomers are the more engaged cohort. A plausible explanation of this contradiction might be due to the nature of the collected sample. The tenure (i.e., tenure of 5 years or less) screening question eliminated a large number of Boomer respondents, meaning that many Boomers within the survey sample pool had tenure of more than 5 years. The
current study specifically limited tenure and did not allow employees from either

generational cohort to be different by the tenure category. Other comparable studies may

not constrain the sample collection for groups of study by tenure or may not report on

such constraints in the methods section. The implications for research include the need

for researchers to adhere to prerequisite methodological procedures before making group

mean comparisons and the need to consider the potential confounding impact job tenure

may have when interpreting results.

**Implications to Business Practice**

Four implications are noted for managers and business practice. First, the

landscape of employees is evolving as more Boomers are working later into their careers.

This fact has increased the breadth of working age for the workforce. With more

Boomers staying in the workforce for longer, the potential for conflicts due to

generational differences increases. The Silent generation has largely retired from the

workforce, but the Homeland generation is on the cusp of entering the workforce as

many are nearing the end of their secondary education careers. In a few short years,

managers will be dealing with four distinct generations within the workforce: Boomers,

GenXers, Millennials, and Homelanders. Gaining an understanding and clarity on the

current generational differences (Boomers, GenXers, and Millennials) will be beneficial

for managers before the Homelanders hit the workforce and cause further disruption.

Second, generational differences do exist and are a commonly addressed topic of

concern for organizations. Internal and external consultants provide services to help

organizations and managers determine and understand the generational differences within

their given workforce. Regardless of the desired organizational outcomes they are hired
to influence, consultants require data upon which to make recommendations, but the well
of current empirical literature is inconsistent and inconclusive (Costanza et al., 2013;
Lyons & Kuron, 2013). The current study found support for scalar measurement
invariance for data from the UWES-9 instrument between Boomers and Millennials
within the leisure and hospitality industry. This finding supports the subsequent analysis
of making group mean comparisons between generational groups. Consultants can now
take some stock into using UWES-9 as a measure for work engagement, especially within
the leisure and hospitality industry comparing when Boomers and Millennials and when
proper methodological protocol is followed.

Third, managers should take caution that Boomers may not be more engaged than
Millennials. The current study found that Millennials are overall more engaged than
Boomers considering the context and parameters of the study. It is important for
managers to understand that while the Boomers consistently had lower mean scores for
both the three-factor model (vigor, dedication, and absorption) and the single-factor
model (work engagement) of UWES-9, the mean differences were not universally
statistically significant. Only for dedication did the Millennials show a statistically
significant mean difference, and only low levels of practical effect sizes were found.
Based upon this result, Millennials could be considered more dedicated than Boomers.
Until such time that further empirical research finds contradictory statistically and
practically significant support, managers and leaders of leisure and hospitality
organizations should consider Boomers and Millennials to be similarly engaged in their
work, with the exception of Millennials being slightly more dedicated.
Fourth, managers who implement interventions to engage their employees should strongly consider the notion that Boomers may not be the most engaged generation at work. It cannot be universally stated or accepted that Boomers are more engaged than Millennials and subsequently that the different generations require individualized interventions. The opposite may be true for the manager’s organization or division within the overall organization. Interventions should be designed and based upon empirically discernible group characteristics and not based upon generalized generational cohort group differences. Employee engagement interventions can be a significant investment for organizations in both the level of effort and in monetary terms.

Organizations generally expect a positive outcome or a positive return on investment for such programs and interventions; otherwise, organizations would not allocate funding. Establishing an employee engagement program on a faulty foundation (i.e., non-empirically based and non-contextually based) may lead to wrongly designed or misapplied interventions that may have the opposite of the intended effect.

**Limitations**

Ten limitations should be maintained when considering the contributions found in this study. First, the study limited the sample participants to those who were working in and living in the United States. This precludes extrapolating the results outside of the U.S. context. Second, the sample was also narrowed to a specific set of leisure and hospitality sub-sectors (i.e., convention center, cruise line, gaming or casino, lodging, marina, sporting facility, travel, and tourism). This limitation has a double consideration impact as the study results should not be generalized to other sub-sectors of the leisure and hospitality industry that are outside of the study’s sample constraints (e.g.,
restaurants) or generalized outside of the leisure hospitality industry. Third, the sample was sourced solely using Qualtrics® panel services and bias may exist due to the nature of the participants being associated with the panel and Qualtrics®. An alternative data source, perhaps directly from a leisure and hospitality organization, would serve to challenge this limitation.

Fourth, common method bias (Podsakoff et al., 2003) is a common concern for social science empirical studies as it can introduce variance. The study did utilize the comprehensive CFA marker analysis technique (Williams et al., 2010) to test for CMV and found limited impact, but a more substantive research design (i.e., longitudinal or random sampling) may aid to limit CMB concerns. Fifth, CFA was used as the measurement invariance analysis method. While the CFA technique is robust, it is not the only methodological option to conduct measurement invariance assessments. Other techniques, such as IRT, could perhaps be employed to complement the findings of the CFA analysis. Sixth, item errors were correlated for both the three-factor and single-factor models of UWES-9. While correlating the items errors was supported by prior literature, not all comparative studies may have correlated item errors or correlated the same item errors as this study. Comparisons to other studies should consider the context of which item errors were correlated for each study. Seventh, the current study did not consider the possibility of a two-factor model of UWES-9. This limitation precludes comparing results to other studies utilizing the two-factor model of UWES-9.

Eighth, the study used the short version (UWES-9) of the work engagement instrument and did not consider the original, long version of the instrument (UWES-17). As such, the findings of the study cannot be generalized to the long version of the
instrument. Ninth, the study utilized PSM to equate the Boomer and Millennial samples. From a methodological viewpoint, the PSM process returned largely equivalent groups based upon the reviewed statistics, but the process could have confounded the results since the samples were manipulated from their original construction. Until more comparable empirical studies utilize PSM to equate groups within measurement invariance studies, the study’s results cannot be directly compared with prior empirical research without accounting for this difference in methods. Last and most importantly, the study only included two of the four generations found in the current professional workforce. Boomers and Millennials were included in the study, but the Silent and GenX generations were excluded. Thus, the findings are strictly limited to the generations included in the study and cannot be generalized beyond that generational cohort context.

**Suggestions for Future Research**

The study creates at least nine directions for future research. First, the suggestions found in the literature (c.f., Schaufeli et al., 2006) to narrow measurement invariance assessments to more comparable groups (e.g., professionals within a single industry) are supported since the study utilized a focused industry approach to define the sample and subsequently found evidence for scalar invariance for data from UWES-9. This direction is multifaceted in that one path can lead to an even narrower focus to a single sub-sector of the leisure and hospitality industry (i.e., lodging) or an alternate path of evaluating various economic aspects of the industry, such as the luxury versus budget tiers (Kamery, 2004).

Second, the parameters of the study could be replicated to evaluate the measurement invariance between generational cohorts for industries other than the leisure
and hospitality industry. While the leisure and hospitality industry has a wide view of
generations represented within its workforce (Solnet & Hood, 2008), Boomers are
working longer and are extending their careers in multiple industries (Fry, 2015).
Exploring the measurement invariance by generational cohort for data from UWES-9
through methodologically rigorous empirical research in other industries will extend our
understanding of the complexities and nature of generational cohorts and work
engagement in the workforce.

Third, the effort to capture a statistically sufficient number of Boomers who
matched the sample criteria proved to be difficult. One of the main limiting criteria for
the Boomer group was the requirement for the participant’s tenure to be 5 years or less at
their place of employment. Many Boomers failed the screening questions as their tenure
was more than 5 years. Thus, understanding the role of tenure in work engagement
perceptions is a viable route for future research. The current study limited the view to
tenure of 5 years or less. Future research could recruit a sample with participants of
greater than 5 years or delineate a sample into smaller tenure bands (e.g. 0-2 years and 3-
5 years) and empirically test the differences in perceptions of work engagement.

Fourth, similar to the previous future research direction, the study controlled for
job level (i.e., supervisor versus non-supervisor) by equating the generational cohort
groups on the job level variable, producing equivalent groups for this variable. Job level
has been cited as a moderator of HRM practices and work engagement (Alfes, Shantz, &
Truss, 2013); thus, it merits additional empirical research within the generational cohort
context, as it may result in different research outcomes. At least two possibilities are
available for this future research direction. One direction could compare supervisors and
non-supervisors within a generational cohort (e.g., only Boomers) to determine the differences in work engagement due to the respondents’ job level classification. Outcomes from this direction could inform as to which job level (supervisors or non-supervisors) is more engaged at work for the generational cohort under study. A second possible direction could retain the generational cohort perspective and utilize only supervisors or only non-supervisors between two or more generational cohorts. For example, a constructed study of this nature that only compared Boomer supervisors to Millennial supervisors could perhaps clarify the levels of work engagement between generational cohorts for employees within a distinct job level.

Fifth, generational cohorts are generally defined in roughly 20 year range bands (Strauss & Howe, 1991). The individual cohorts can also be further segmented into early and late cohorts (Sessa, Kabacoff, Deal, & Brown, 2007). For example, Boomers can be split from the overall 1946-1964 range into Early Boomers (1946-1955) and Late Boomers (1956-1964) (Markert, 2004). Research has suggested the early and late generations can have unique perceptions and “distinct lifestyles” (Markert, 2004, p. 11). An examination of early and late cohort generational differences is pertinent as the Early Boomers are either directly in or are beginning their retirement phase compared to the Late Boomers many of whom are still active in their careers.

Sixth, the use of PSM to equate groups for the study provides an opportunity to create a comparison study. The comparative study could replicate the statistical analyses without equating the Boomer and Millennial groups by PSM. The sample for the comparative study would be the original, unmatched samples of Boomers (n = 178) and Millennials (n = 369) that were derived after the data cleaning routines. Using these non-
equivalent data samples and conducting the same analyses would result in a comparative statistical output that could be evaluated against the PSM equated data results.

Comparison of results for non-equated samples (i.e., with no PSM) and equated samples (i.e., with PSM) could inform researchers about the efficacy of utilizing PSM for future studies. Seventh, the factor structures (i.e., three-factor and single-factor) for UWES-9 established within the study for the leisure and hospitality sample could be replicated across alternate samples for the leisure and hospitality industry (i.e., different sub-sectors). Eighth, and following the seventh future research suggestion, assessments of measurement invariance could be conducted on the alternate leisure and hospitality samples derived from surveying the different sub-sectors.

Last, this study could be enhanced by adding the GenX generation into the sample. The GenXers could replace either the Boomer and Millennial component to maintain the binary generational cohort approach, or the GenXers could be added as the third generation. Other studies (e.g., Hoole & Bonnema, 2015; Park & Gursoy, 2012) included all three generations and empirically evaluated the generational differences across all three groups. It is important to retain the measurement invariance assessment component within the enhanced research. Even though this study found measurement invariance at the scalar level for Boomers and Millennials, which provided statistical grounding for latent mean analysis, it does not provide assurances that measurement invariance will be obtained between other generations.

**Summary of the Chapter**

Chapter 5 contained five sections. The Chapter 4 results and relationships to relevant literature were discussed and followed by a summary of implications for theory,
research, and practice. The chapter proceeded to list the limitations of the study and recommendations for future research. The chapter concluded with a summary.
References


Appendix A. Survey

Are you currently employed within the United States?
○ Yes
○ No

Are you currently living within the United States?
○ Yes
○ No

Do you speak English?
○ Yes
○ No
Do you currently work in the leisure and hospitality industries listed below?

Please select the industry that best describes your place of employment. If your place of employment does not match the industries listed below, please select "Not Listed". If you do not work in the leisure and hospitality industry, please select "I do not work in the leisure and hospitality industry".

- Convention Center
- Cruise Line
- Gaming or Casino
- Lodging (hotels or other accommodation facilities)
- Marina
- Sporting Facility
- Travel
- Tourism (entertainment venues or museums)
- Not Listed
- I do not work in the leisure and hospitality Industry

Have you been at your current place of employment for five years or less?

- Yes
- No
In what year range were you born?
- 1910 - 1945
- 1946 - 1964
- 1965 - 1980
- 1981 - 2000

On average, about how many hours per week do you work?
- Not currently working
- 1 to 19 hours per week
- 20 to 34 hours per week
- 35 to 45 hours per week
- 46 or more hours per week
Welcome to the Survey of Work Perceptions. The purpose of this study is to explore employee perceptions about their work in the leisure and hospitality industry. Your participation is completely voluntary, and you may choose not to participate. If you decide to participate, you may withdraw at any time without consequence by closing your browser.

The following provides information about the study and what to expect as a participant. Please review the information below to make sure you are fully informed about the survey and your expectations as a participant.

About the study:
- The study has been approved by The University of Texas at Tyler Institutional Review Board.
- The principal researcher is a doctoral candidate at The University of Texas at Tyler, and this survey is related to his dissertation.
- Only aggregated summary data from this study will be collected and included in published results for scholarly purposes.
- Personal information such as names, phone numbers, and e-mail addresses will not be collected or published.

What to expect as a participant:
- All information you provide will remain confidential.
- All responses are anonymous.
- The survey is expected to require between 5 and 7 minutes of your time.
- The survey will require an answer to each question.
- There are no right or wrong answers to the questions, and your first response to a question is often your most accurate response.

There are no known risks of participation in this study, other than becoming a little tired of answering the questions or perhaps a little distressed when answering some of the questions. If this happens, you are free to take a break and return to the survey to finish it, or you can discontinue participation without any problems. Potential benefits to you as a participant include the chance to reflect on your perceptions of work and to contribute to a study that analyzes relationships between work-related variables in the leisure and hospitality industry.

If you need to ask questions about this study, please contact the principal researcher, Gregory Kieffer, at Gkieffer@patriots.uttler.edu. If you have any questions about your rights as a research participant, please contact Dr. Gloria Duke, Chair of the UT Tyler Institutional Review Board, at Gduke@uttler.edu or 903-662-7023. After reading and understanding what has been explained above and if you choose to participate in this study, click the "Yes" option below to proceed to the survey. If you decline to participate in this study, please click the "No" option below to exit the survey.

Electronic Consent:

Clicking on the "Yes" button below indicates that:
- You have read the above information.
- You voluntarily agree to participate.
- You are at least 18 years of age.

☐ Yes, I choose to participate in this study.
☐ No, I decline to participate in this study.

Survey Completion: 0%
The following 9 statements are about how you feel at work. Please read each statement carefully and decide if you ever feel this way about your job. If you have never had this feeling, select "Never" for the statement. If you have had this feeling, indicate how often you feel it by selecting the answer that best describes how frequently you feel that way.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Never</th>
<th>Almost Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Very Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>At my work, I feel bursting with energy.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At my job, I feel strong and vigorous.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am enthusiastic about my job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My job inspires me.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When I get up in the morning, I feel like going to work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel happy when I am working intensely.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am proud of the work I do.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am immersed in my work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I get carried away when I'm working.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(C) 2006 Yilmaz B. Schaufeli, Arnold B. Bakker, & Manse Salarova. All rights reserved. Do not duplicate.
Use the response scale below from Strongly Disagree to Strongly Agree to indicate your level of agreement or disagreement with each statement. Please click the button corresponding with your response.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Neither Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I prefer blue to other colors.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I think blue cars are ugly.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I like the color blue.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I don’t think blue is a pretty color.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I like blue clothes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I don’t like blue clothes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I hope my next car is blue.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I really don’t like the color blue.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(C) 2006 Brian K. Miller & Beverly Chodo. All rights reserved. Do not duplicate.
In the section below are 16 statements of job-related feelings. Please read each statement carefully and decide if you ever feel this way about your job. If you have never had this feeling, select “Never” for the statement. If you have had this feeling indicate how often you feel it by selecting the answer that best describes how frequently you feel that way.

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>A few times a year</th>
<th>Once a month or less</th>
<th>A few times a month</th>
<th>Once a week</th>
<th>A few times a week</th>
<th>Every day</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel emotionally drained from my work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel used up at the end of the day.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel tired when I get up in the morning and have to face another day on the job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working all day is really a strain for me.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can effectively solve the problems that arise in my work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel burned out from my work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel I am making an effective contribution to what this organization does.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have become less interested in my work since I started this job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have become less enthusiastic about my work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In my opinion, I am good at my job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel exhilarated when I accomplish something at work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have accomplished many worthwhile things in this job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I just want to do my job and not be bothered.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have become more cynical about whether my work contributes anything.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I have accomplished many worthwhile things in this job.

I just want to do my job and not be bothered.

I have become more cynical about whether my work contributes anything.

I doubt the significance of my work.

At my work, I feel confident that I am effective at getting things done.

Copyright (C) 1996 Wilmur B. Schaufeli, Michael P. Leiter, Christine Maslach, Susan E. Jackson, & Richard L. Schwab. All rights reserved in all media. Published by Mind Garden, Inc., www.mindgarden.com

What is your gender?
- Male
- Female

What is your race?
- Black or African American
- Asian
- Hispanic or Latino
- White or Caucasian
- Other

In what year range were you born?
- 1946 - 1955
- 1956 - 1964
- 1961 - 1980
- 1991 - 2000
In what year range were you born?
- 1946 - 1955
- 1956 - 1964
- 1981 - 1990
- 1991 - 2000

What is your highest level of education completed?
- Some High School
- High School or GED
- Some College
- Professional or Trade Certificate
- Associates Degree
- Bachelors Degree
- Masters Degree
- Doctoral Degree

On average, about how many hours per week do you work?
- Not currently working
- 1 to 19 hours per week
- 20 to 34 hours per week
- 35 to 45 hours per week
- 46 or more hours per week

Select the best answer that describes your work role:
- Entry Level
- Non-supervisory
- Supervisory
- Managerial
Select the best answer that describes your work role:
- Entry Level
- Non-supervisory
- Supervisory
- Managerial

What is the market or price segment of your place of employment?
- Upper Tier
- Middle Tier
- Lower Tier

What is your tenure at your current place of employment?
- Less than a year
- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- More than 5 years

Does your job require direct contact with customers?
- Yes
- No

We thank you for your time spent taking this survey. Your response has been recorded.
Appendix B. Permissions Documentation

Dr. Miller and Chiche: 

Good afternoon! My name is Gregg Koffar, and I am a doctoral candidate at The University of Texas at Tyler in the Human Resource Development program.

I am currently writing my dissertation and evaluating various options to detect CMV. The blue attitude marker you developed and presented in a 2008 Southern Management Association paper has been cited frequently and is considered an excellent option. I am unable to source a copy of the paper through Google Scholar searches or through the campus library resources. Thus, I am reaching out to you both directly as the authors to request a feasible access to the blue attitude scale and/or the presentation paper. I would be very grateful if you are able to share the materials.

Thank you for your time and consideration!

Best,
Gregg

Greggory Koffar
Doctoral Candidate in HRD/ODC
The University of Texas at Tyler
College of Business and Technology
gkoffar@uttyler.edu
grcge.koffar@gmail.com
832-700-0599

Hi Gregg,

Dr. Chiche retired a few years ago after 48 years as a professor on our faculty. She’s probably harder to obtain something really cool on her ranch. We all miss her here at Texas State. She’s a legend.

I’ve moved forward with this project that seems to be useful and of interest to a large number of researchers. My new co-author Dr. Simmering [joined here and whom I affectionately call the “Queen of CMV”], and I are working on a scale development paper for this instrument. We’d be interested in having you take a look at it in a few months. In the meantime, the conference paper and the scale are attached.

Best Regards,
Becky

Becky K. Miller, Ph.D., M.Ed.
Professor of Management
Texas State University
445 McCoy Hall
San Marcos, TX 78665
Tel: 512-245-7179
Fax: 512-245-2850
Associate Editor of the Journal of Managerial Psychology
Tests

Notice for potential users of the UWES and the DUWAS

- You are welcomed to use both tests provided that you agree to the following two conditions:

1. The use is for non-commercial educational or research purposes only. This means that no one is charging anyone a fee.

2. You agree to share some of your data, detailed below, with the authors. We will add these data to our international database and use them only for the purpose of further validating the UWES (e.g., updating norms, assessing cross-national equivalence).

- Data to be shared:
  For each sample, the raw test-scores, age, gender, and (if available) occupation. Please adhere to the original answering format and sequential order of the items.
  For each sample a brief narrative description of its size, occupation(s) covered, language, and country.

- Please send data to: w.schaufeli@uu.nl. Preferably the raw data file should be in SPSS or EXCEL format.

- No explicit, personal permission is required — and will be given — as long as both previously mentioned conditions are fulfilled.

- By continuing to the Test Forms you agree with the above statement.
Maslach Burnout Inventory™
Instruments and Scoring Keys
Includes MBI Forms:
- Human Services - MBI-HSS
- Medical Personnel - MBI-HSS (MP)
- Educators - MBI-ES
- General - MBI-GS
- Students - MBI-GS (S)

Christina Maslach
Susan E. Jackson
Michael P. Leiter
Wilmar B. Schaufeli
Richard L. Schwab

Published by Mind Garden, Inc.
info@mindgarden.com
www.mindgarden.com

Important Note to Licensee
If you have purchased a license to reproduce or administer a fixed number of copies of an existing Mind Garden instrument, manual, or workbook, you agree that it is your legal responsibility to compensate the copyright holder of this work — via payment to Mind Garden — for reproduction or administration in any medium. Reproduction includes all forms of physical or electronic administration including online survey, handheld survey devices, etc.

The copyright holder has agreed to grant a license to reproduce the specified number of copies of this document or instrument within one year from the date of purchase.

You agree that you or a person in your organization will be assigned to track the number of reproductions or administrations and will be responsible for compensating Mind Garden for any reproductions or administrations in excess of the number purchased.

This instrument is covered by U.S. and international copyright laws. Any use of this instrument, in whole or in part, is subject to such laws and is expressly prohibited by the copyright holder. If you would like to request permission to use or reproduce the instrument, in whole or in part, contact Mind Garden, Inc.
Appendix C: Measurement Scales, Demographics, and Job Characteristics

Utrecht Work Engagement Scale: 9-item scale, 7-point Likert scale “0 = Never; 6 = Always” (Schaufeli et al., 2006)

Vigor

1. At my work, I feel bursting with energy.
2. At my job, I feel strong and vigorous.
3. When I get up in the morning, I feel like going to work.

Dedication

1. I am enthusiastic about my job.
2. My job inspires me.
3. I am proud of the work that I do.

Absorption

1. I feel happy when I am working intensely.
2. I am immersed in my work.
3. I get carried away when I am working.

Attitudes toward the color blue: 8-item scale, 7-point Likert scale “1 = Strongly Disagree; 7 = Strongly Agree” (Miller & Chiodo, 2008)

1. I prefer blue to other colors
2. I think blue cars are ugly (R)
3. I like the color blue
4. I don’t think blue is a pretty color (R)
5. I like blue clothes
6. I don’t like blue clothes (R)
7. I hope my next car is blue
8. I really don’t like the color blue (R)

Demographics

1. What is your gender?
   a. Female
   b. Male
2. What generation are you a member of?
   d. Baby Boomers (1946 – 1964)
   e. Traditionalists (1928 – 1945)
3. What is your race?
   a. Black or African American
   b. Asian
   c. Hispanic or Latino
   d. White or Caucasian
   e. Other
4. What is your highest level of education?
   a. Some High School
   b. High School or GED
   c. Professional or Trade Certificate
   d. Some College
   e. Associates Degree
   f. Bachelor’s Degree
   g. Master’s Degree
   h. Doctoral Degree

Job Characteristics

1. What is your current employment status?
   a. Not currently working
   b. 1 to 19 hours per week
   c. 20 to 34 hours per week
   d. 35 to 45 hours per week
   e. 46 to 60 hours per week
2. What is your current job level?
   a. Entry Level
   b. Non-supervisory
   c. Supervisory
   d. Managerial
3. What sector of the hospitality industry do you work in?
   a. Convention Center
   b. Cruise
   c. Gaming
   d. Lodging (e.g., Hotel or Accommodation)
   e. Marina
   f. Sporting Facility
   g. Travel
   h. Tourism (e.g., Entertainment venue or Museum)
   i. Other Hospitality Industry not listed above
4. What is your tenure at your current place of employment?
   a. Less than 2 years
   b. 2 to 5 years
   c. 6 to 10 years
   d. 11 to 15 years
   e. 16 to 20 years
   f. 21 or more years
5. Does your job require direct contact with customers?
   a. Yes
   b. No
Appendix D. E-mail Communications

Survey Recruitment E-mail:

Dear [FirstName of Participant],

Greetings! My name is Greggory Keiffer, and I am a doctoral candidate at The University of Texas at Tyler in the Human Resource Development program. I am contacting you to ask for your assistance and voluntary participation in an online survey which is an integral part of my dissertation. Your perceptions of work are very important and will inform the study, which investigates the meaning of work engagement for employees in the leisure and hospitality industry.

The survey invitation will include a personalized Qualtrics link and will be sent to you on Monday, [Insert Date], at Noon. The e-mail will originate from Qualtrics.

For questions about the study or survey, please contact the principal researcher, Greggory Keiffer at gkeiffer@patriots.uttyler.edu. For questions regarding your rights as a research participant, please contact Dr. Gloria Duke, Chair of the UT Tyler Institutional Review Board at gduke@uttyler.edu or 903-566-7023.

Thank you!

Greggory Keiffer
Doctoral Candidate
UT Tyler
Department of Human Resource Development

Survey Invitation E-mail:

Dear [FirstName of Participant]

Greetings! Thank you again for your consideration to participate in a work perceptions survey related to my dissertation research as a doctoral candidate at UT Tyler. The survey titled Survey of Work Perceptions will help to develop an understanding of how employees are engaged and how to potentially improve the work environments within the leisure and hospitality industry.

Your participation is completely voluntary, confidential, anonymous, and will require the completion of an online survey which will require about 5 to 7 minutes for you to complete in full. The survey has been approved by the UT Tyler Institutional Review Board and will be active from [Insert Time:am] CST today, Monday, [Insert Date] through 11:59pm on Monday, [Insert Date].

In order to participate, click on the personalized access link below.
Alternately, you may copy and paste the URL below into your web browser to access the survey:

[Full text link to Qualtrics survey]

Thank you again!

Greggory Keiffer
Doctoral Candidate
UT Tyler
Department of Human Resource Development

Survey Reminder E-mail:

Dear [FirstName of Participant]

Greetings! On Monday [Insert prior date], a survey invitation was sent directly to you for the Survey of Work Perceptions. The e-mail contained a hyperlink to a survey that considers your perceptions about your work and is related to my doctoral dissertation research. Your participation is important to help develop and understanding of engaged employees in the leisure and hospitality industry and how to improve the work environments within the industry.

Today, I am following up and providing a hyperlink for easy and convenient access to the survey. As a reminder, your participation in the survey is completely voluntary and will require approximately 5 to 7 minutes to complete. The survey will close at 11:59pm on Monday, [Insert Date].

To begin the survey, please click on the link below:

[Qualtrics Hyperlink]

Alternately, you may copy and paste the URL below into your web browser to access the survey:

[Full text link to Qualtrics survey]

Thank you again!

Greggory Keiffer
Doctoral Candidate
UT Tyler
Department of Human Resource Development
Appendix E. IRB Approval

THE UNIVERSITY OF TEXAS AT TYLER
INSTITUTIONAL REVIEW BOARD

EXPEDITED and EXEMPT RESEARCH APPLICATION

IRB: F2017-09
Approved by: G Duke
Date: September 15, 2017

Attach (electronically) to gduke@utttyler.edu with this application, the following:

- Written consent form using the UT Tyler Consent Template unless a waiver of written informed consent is requested
- Signature page of Thesis or Dissertation Committee members showing proposal approval for graduate students
- Brief research proposal that outlines background and significance, research design, research questions/hypotheses, data collection instruments and related information, data collection procedures, data analysis procedures. Most of this can be copied and pasted to relevant parts of the application but please keep Background & Significance brief for the application.
- CITI certifications for PI, co-investigators, and research assistants participating in recruitment, data collection, data analysis, or, if they have any exposure to identifiable data (if training has not been completed at UT Tyler within a 3 year period of time)
- Tool/instrument/survey; if copyright or other issues prohibit electronic form, submit one hard copy

COMPLETE ALL ITEMS TO AVOID DELAY IN IRB APPROVAL

<table>
<thead>
<tr>
<th>Principal Investigator</th>
<th>Keiffer</th>
<th>Gregory</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI Title and Credentials</td>
<td>☐ Assistant Professor</td>
<td>☐ Associate Professor</td>
<td>☐ Professor</td>
</tr>
<tr>
<td>☐ Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty Sponsor Name and Email if PI is Student</td>
<td>Kim Nimon, Ph.D. - <a href="mailto:knimon@utttyler.edu">knimon@utttyler.edu</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI Phone</td>
<td>832-760-8989</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI Email</td>
<td><a href="mailto:gkeiffer@patriots.utttyler.edu">gkeiffer@patriots.utttyler.edu</a></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DATE: 09/11/2017