AN ASSESSMENT OF GROUP MEASUREMENT INVARIANCE ACROSS FIVE YEARS OF HEALTHCARE EMPLOYEE ENGAGEMENT

A Dissertation

by

CHRISTI A. PHILLIPS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Michael Beyerlein
Committee Members,	Brendan Bartanen
	Khalil Dirani
	Ben Welch
Head of Department,	Mario Torres

December 2020

Major Subject: Educational Human Resource Development

ABSTRACT

Assumptions of measurement invariance are infrequently tested. Establishing the psychometric properties of instruments supports validity and reliability of group research. This study aims to determine factorial invariance of the Employee Engagement Survey (EES) within and between groups through structural equation modeling.

The research design is a cross-sectional analysis of archival data. The EES was administered electronically for five years (2011 - 2015) to employees of a healthcare system made of up five hospitals and a shared services group.

Results confirmed the reliability of the Engagement construct within the instrument, but no other latent factors. The measurement invariance results indicate non-invariance within and between pairings of years and facilities. Although a major M&A occurred in 2013, it is not a primary source of non-invariance in the data set. Organizations relying on psychometric instruments to measure unobserved phenomena may err in results interpretation and decision-making when invariance is not established. The findings contribute to the body of literature on employee engagement and measurement invariance in professional settings.

Keywords – Employee engagement, exploratory factor analysis, multi-group confirmatory factor analysis, healthcare, archival research, mergers & acquisitions

DEDICATION

To Obie L. Phillips Jr. from “Your little Ph.D.”

To Johnnie L. Anderson, Viola “Peaches” Anderson, Sadie L. Phillips, Obie L. Phillips Sr., and Dr. Charles Phillips, Jr. Thanks for taking the first leg. I’ll bring it home. See you at the gate.

ACKNOWLEDGEMENTS

It is no secret that earning a Ph.D. demands a fellowship of people, but no one can precisely say whom you'll need or what the support group will do for you. Now I know. It is unorthodox to thank a university before the long list of people who encouraged and supported my journey. Even so, I used the entire window to complete this crossing—10 years—so, first, I am grateful to Texas A&M University for time limitations.

Thank you, professors. Dr. Homer Tolson, you gave direction. I am not smart enough to know any of this subject matter without the foundations you taught, retaught, and taught me yet again. Drs. Michael Beyerlein and Khalil Dirani, you extended guidance through discernment. Dr. Beyerlein, I am a better professional because you and I reflected on the vicissitudes of organizational development and behavior together without losing conviction that it is possible to get it right. Dr. Dirani, thank you for lending your copy of *Bird by Bird* by Ann Lamott. Not only is it the funniest book I have read as an adult, but its balm was also sufficient inspiration when overwhelmed with how much further I had yet to go. Dr. Ben Welch, thank you for showing what it looks like to live one's purpose so well that your fans will not let you retire from your journey.

I leaned heavily on expert guidance from colleagues I have never met in person like Dr. Alana James from Doctoral.net and Jeff Meyer of the Analysis Factor. Drs. Brad Schuck [The University of Louisville, Department of Educational Leadership, Evaluation and Organizational Development] and Terrence Jorgensen, [The Research

Institute for Child Development and Education, at the University of Amsterdam], our brief exchanges of ideas expanded my understanding of measurement invariance, thus improving this dissertation significantly.

To my friends, Dr. Sandi Murray and Laura Bogard, your reassurance and insight into the mission gave me perspective. BettyAnn Reichter, your hospitality was a much-needed rest stop and place to retreat.

To my family, you remain traveling companions throughout all of my journeys. Mom, Dad, Kem, and Lorenza Jr., thank you for making the expedition a fun-filled celebration. Aunt Hoxi, Uncle Chuck, and Aunt Joycelyn thank you for circling the wagons when I got a little lost. Cheryl, Aunt Blanche, and Uncle John thank you for a peaceful place to focus on the last mile. The Ray family, especially K.C., Karen, and Ralph, your tough love reminded me that completion, not brilliance, was the final destination. We have arrived.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Michael Beyerlein [advisor] of the Department of Educational Administration and Human Resource Development, Assistant Professor Brendan Bartanen of the Department of Educational Administration and Human Resource Development, Associate Professor Khalil Dirani of the Department of Educational Administration and Human Resource Development, and Clinical Professor Emeritus, Ben Welch of the Department of Management in the Mays Business School [Outside Department].

The data analyzed for Chapter IV was provided by Jeff Meyer of the Analysis Factor. All other work conducted for the dissertation was completed by the student independently.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

Page

LIST OF TABLES

Page

Page

CHAPTER I

INTRODUCTION*

Organizations that demonstrate measurement invariance (MI) of instruments can reliably and confidently assert findings arising for their data. This issue is vital because organizational leaders who neglect to establish the validity of the underlying measurement models of scales they employ may err in their grasp of organizational phenomena and make questionable decisions. Currently, survey instrument results are often utilized without rigorous testing of underlying measurement models (Vandenberg & Lance 2000), leading to unfounded or hidden assumptions and potentially questionable interpretation of results. Tests measuring the same constructs permit identical covariance comparisons to be interpreted (Kim & Willson, 2014). Organizational leaders using data from untested instruments to characterize and understand internal dynamics and occurrences may unwittingly and unknowingly arrive at inaccurate, over-simplified conclusions or make dubious decisions with the information.

Measurement of the underlying structure of an instrument can be a difficult idea to follow. Little (2013) provides a metaphor that makes the concept of invariance easier to grasp. Two botanists see two plants located on opposite sides of a hill. The plants' stems and leaves look different, but the outward characteristics and differences do not

---

* Reprinting of The Advisory Board data and resources was allowed with written permission (see Appendix D).

determine whether or not the plants represent the same species. To identify the species, the botanists must dig down to the roots of the plants. If each plant is the same species, it should have a signature root system. If the root systems are the same, then subsequent observation of root patterns, lengths, and thicknesses would further help them identify likenesses (Little, 2013). Mapping the root patterns is a kind of calibration, informing the botanists whether the two species are invariant. MI is a similar process of calibration, identifying if the same constructs are similarly understood and assessed. De Beurs, Fokkema, de Groot, de Keijser, and Kerkhof (2015) described MI as a kind of scale recalibration, also known as response shift, first introduced in the late 1970s.

> Response shift refers to a change in the meaning of one's self-evaluation of a target construct as a result of a) a change in the respondent's internal standards of measurement (scale recalibration); b) a change in the respondent's values (reprioritization); or c) a redefinition of the target construct (reconceptualization). (p. 369, as cited by Schwartz & Sprangers, 1999).

Between 2011 and 2015, leaders from a healthcare organization in Texas utilized mean scores from the Employee Engagement Survey (EES) scale by a healthcare consulting firm, The Advisory Board Company (The Advisory Board), to measure employee engagement. The client organization is referred to as SystemTex throughout the dissertation. SystemTex was comprised of five hospitals and one shared corporate service group. Following each engagement cycle, the leaders leveraged mean scores to implement programs and plans to improve engagement from one year to the next. In 2013, a significant acquisition by a larger, national healthcare organization occurred, and

further complexified the usage of untested constructs when making sense of employee engagement. Mergers and acquisitions (M&A) present landmark junctures in the arch of an organization's life cycle, given the criticality of their role in organizational performance. M&A activity frequently fails to meet expectations (Ismail, Abdou, & Annis, 2011; Weil, 2010), so this research examines if the M&A generated response shifts in the measurement structure of engagement.

This dissertation aims to employ MI to analyze archival data of the employee engagement construct from 2011 to 2015 between and within groups at SystemTex. Through the testing of MI, this study also scrutinizes the stability of the engagement construct by groups to see if the structure holds (Kim & Willson, 2014).

The five years provided data points to analyze a large organization amid complicated transition wherein decision-makers searched for solutions to comprehend, resolve, and simplify complex challenges. Analysis beyond the mean score results offers a more complex or nuanced perspective of SystemTex's engagement. With this study, it may be possible to explain and contextualize engagement results differently from organizational decision-makers at that time.

Terms and Definitions

Throughout the dissertation, the term *engagement* appears frequently. Lower-case engagement refers to the idea of employee engagement in general terms. Capitalized Engagement refers to the construct or factor.

EES stands for Employee Engagement Survey built by The Advisory Board, which measured engagement at the individual level. The purpose of this study is not to

debate the multitude of employee engagement definitions or constructs, but to analyze statistical aspects of the EES instrument built by The Advisory Board. The survey provided focus "…on enhancing workforce culture offered by [The Advisory Board] including employee engagement, physician alignment, and patient safety" (Strumwasser & Virkstis, 2015, p. 179), and was used in this research project with The Advisory Board's written permission (see Appendix D). According to The Advisory Board, engaged employees are "…inspired to do their best work, feel personally motivated to help the larger organization succeed, and to exceed the expected level of effort" (Strumwasser & Virkstis, 2015, p. 179), which is a version of organizational engagement. A way to describe the definition from an employee's perspective could be that one exercises discretionary effort through the personal expense of time, thought, and attitude because of parities between personal and organizational values. An employee is inclined to become less engaged or display less discretionary effort when there is less tangible alignment between personal values and work experiences.

The terms common factor, latent construct, and latent variable refer to a variable representing a theoretical construct that is not directly observed (Bialosiewicz, Murphy, & Berry, 2013). The manifest variables refer to statements or items directly observed (Bialosiewicz et al., 2013).

The terms "participant" and "respondent" are used interchangeably to indicate an eligible employee who completed the EES; there is no difference in their meaning.

MI, the abbreviation for measurement invariance, is used throughout this study; however, MI is sometimes referred to as measurement equivalence (ME) and factorial

invariance (Kim & Willson, 2014; Vandenberg & Lance, 2000). MI is a technique that determines if the same unobserved latent variable is being measured across time or between groups. A change in MI implies that the meaning of the variable has changed for the respondents. When MI failure occurs, it can result in critical theoretical and practical interpretive problems. "If the meaning of the focal construct varies…, mean-level tests will be uninterpretable because they reflect comparisons of different phenomena" (Nye, Brummel, & Drasgow, 2010, p. 1560; Vandenberg, 2002; Vandenberg & Lance, 2000). This subject is vital because organizations neglecting to establish the validity of the underlying measurement models of scales they employ may err in their comprehension of organizational phenomena and make questionable decisions.

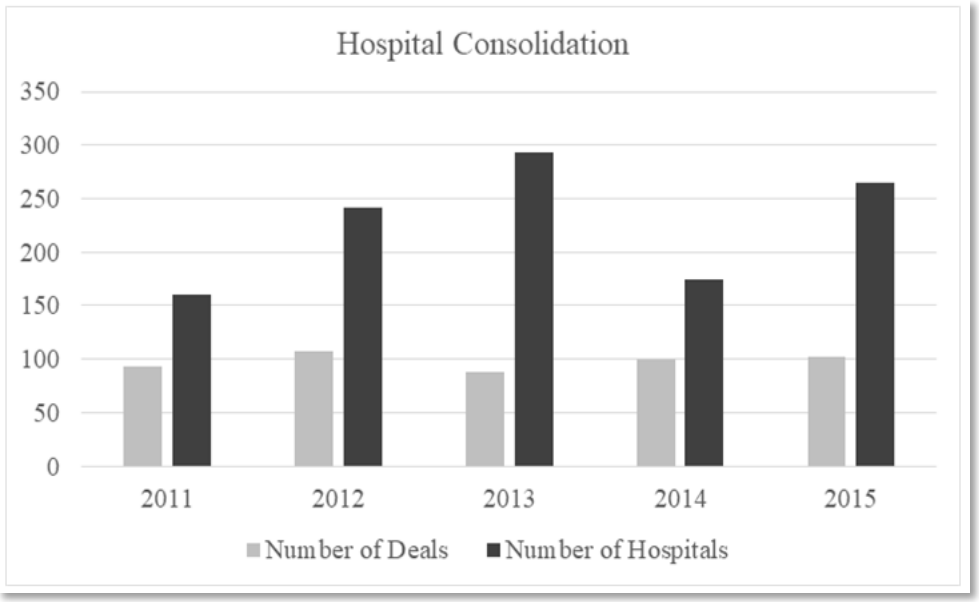
The terms mergers or acquisitions are sometimes employed ubiquitously to intend identical meaning, though it should be understood that SystemTex underwent an acquisition process specifically initiated in 2012 when SystemTex's board of directors solicited bids from multiple healthcare organizations. SystemTex, a non-profit organization, experienced a "friendly" acquisition by another non-profit healthcare company. An acquisition agreement was finally concluded in 2013. This transaction was not a merger because a primary distinction between an acquisition and a merger depends on which organization is making the decisions (Eby, 2005). All the assets and liabilities of SystemTex were bought in total by the acquiring company. Generally, acquisitions are characterized as "friendly," where the acquisition is agreed upon, or "hostile" when there is opposition to the bid for acquisition (Tuch & O'Sullivan, 2007). In this case, the

term “friendly” means that the SystemTex board of directors sought the acquisition. Acquisitions are complex business procedures wherein intended outcomes can take many forms, including favorable positioning within a market or industry, increased market share, and profitability. Acquisition researchers often look at financial inputs and outputs, including ROI, market share, and in more recent times, the procurement of innovation through acquisition (Junni & Sarala, 2013).

A surfeit of hospital consolidation within the U.S. was well underway before the 2012/2013 acquisition, making it no anomaly in the industry (see Figure 1). One hundred sixty hospitals were consolidated in 2011, 242 in 2012, 293 in 2013, 175 in 2014, and 265 hospitals in 2015 (Gaynor, 2016). Nevertheless, there is an acknowledgment of the need for additional theoretical perspectives to explain why M&As so often fail to achieve the desired organizational outcomes (King, Dalton, Daily & Covin, 2004; Stahl & Voigt, 2008). A body of research affirms adverse employee reactions, such as lack of engagement or disengagement in the wake of an acquisition. Other negative reactions include high turnover, employee stress, and internal and external power struggles (Larsson & Finkelstein, 1999; Larsson & Risberg, 1998; Lubatkin, Schweiger & Weber, 1999; Schweiger, Ivancevich, & Power, 1987; Stahl, Kremershof & Larsson, 2004; Stahl & Voight, 2005).

Figure 1

U.S. Hospital Consolidations 2011 - 2015



Note. Adapted from "Consolidation and Competition in US Health Care. Health Affairs Blog," by Gaynor, 2018, http://www.aha.org/research/reports/tw/chartbook/ch2.shtml

Permissions

The Advisory Board designed, disseminated, and collected the data for this study. It is a healthcare think tank and management consulting group, offering clients research-based management services. SystemTex contracted directly with The Advisory Board and permitted using the data with the caveat that hospital and employee information, other than my reported experience, would not be identified.

Background

Throughout, I employ first-person language, recounting of organizational occurrences, and narrative elements from my first-hand full-time employee experience as at SystemTex. I worked in organization development (OD) from 2012 to 2017. My

doctoral studies began in 2010, so there was a near-complete overlap between my doctoral studies and my employment there. My perspective is intended to contribute to the organizational context and insert points of interest in SystemTex. My role as a senior-level OD consultant is foundational to my interest in this topic and mainly why the data was made available to me. As an employee, a great deal of responsibility fell to me for the tactical execution of the EES at SystemTex, including contracting with The Advisory Board, providing the employee data for the survey structure, receiving the results from The Advisory Board, coordinating results distribution, and organizing in post-engagement interventions.

OD is a subset of human resource development (HRD) based on behavioral science and experiential learning. Practitioners engage in long and short-term activities, which have the potential to foster organizational "...knowledge, expertise, productivity, satisfaction, income, interpersonal relationships, and other desired outcomes, whether for personal or group/team gain or the benefit of an organization..." (McLean & Egan, 2008, p. 241).

The EES was used nation-wide during the time of SystemTex's multiple annual contracts. The scale itself, its widespread use in healthcare, and the consulting services attached to the contract made the instrument seem a worry-free option to adopt. The EES scale was easily read and understood; SystemTex also felt assured of the instrument's quality, leaving no room to question the EES design or validity.

During my doctoral studies, I became curious about the validity and reliability of the EES. I began to wonder what tests for reliability and validity had the EES

undergone? How valid was the scale we consistently depended upon to understand employees' challenges and organizational opportunities? The assurances SystemTex received from the vendor had little to do with internal reliability, consistency, or measurement.

Web-based research to locate published Advisory Board articles in peer-reviewed healthcare and social science journals, as well as non-peer-reviewed industry publications concerning the EES, was conducted. An article asserting the validity of The Advisory Board's survey multivariate analysis appears in an article in the Journal of Nursing Administration (Strumwasser & Virkstis, 2015). The authors were employees of The Advisory Board at the time of the publication. I located no other peer-reviewed evidence of the EES, nor was there any peer-reviewed research employing MI analysis of EES. Much of The Advisory Board research was published on its proprietary website or in white papers and made available to clients—which I knew from first-hand experience. However, no studies of instrument validity and reliability were sought by SystemTex. Academic and professional curiosity led to the formation of several theoretical questions for this study

1. How reliable or valid is the EES Engagement construct?
2. What would the factor structure reveal about employee perspectives on engagement?
3. Would the 2013 M&A generate a response shift in engagement?

Further Clarification of the Problem

The measurement process in organizational research attempts to describe the attributes necessary to characterize its people. Vandenberg and Lance (2000) provide a comprehensive examination and literature review on MI in multi-item composite assessments, similar to the EES. They substantiate the necessity of measurement in organizational sciences as evidenced by the number of written articles and journals devoted to measurement topics, the volume and range allotted to methods sections in empirical articles, and the importance placed on appropriate measurement of key variables in publication decisions. As an attempt to describe the organization, "…the measurement process is pivotal because it defines the links between organizational theories and the data utilized to test them" (Vandenberg & Lance, 2000, p. 5).

Measurement is founded in classical test theory (CTT). It examines true and error scores in manifest variables' properties but does not directly address interpretation issues, as does MI (Vandenberg & Lance, 2000). CTT lacks the functionality to understand the mechanics of latent constructs. CTT cannot resolve questions concerning the analysis of latent constructs, employee discernment of a construct, or changes in the grasp of a construct. These kinds of problems cannot be addressed with CTT because (a) measurement models are assumed to be equivalent in CTT; and (b) CTT does not describe the relationship between manifest variables such as items, subscales, and the underlying construct (Bialosiewicz et al., 2013; Vandenberg & Lance, 2000). Little (2013) writes that in CTT scores for indicators or variables comprise different sources of variance. CTT assumes that three sources of score information—the true score, item-

specific variance, and random error are independent of each other, and that variance and error are uncorrelated (Little, 2013). More contemporary test theory suggests that statistical approaches like structural equation modeling (SEM) can explore and explain those elements threatening measurement or construct validity (Little, 2013).

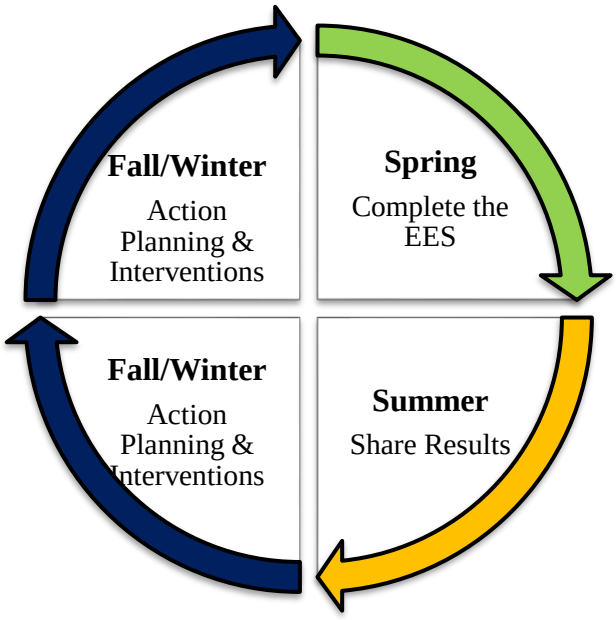
MI is rarely assessed. MI assumptions are infrequently tested (Chungkham, Ingre, Karasek, Westerlund, Theorell & Xia, 2013; Vandenberg & Lance, 2000), though an increasing number of studies focus on MI issues (Kim & Willson, 2014; Vandenberg & Lance, 2000). Observation of constructs, such as "attitudes and beliefs, intentions and motives, and emotional and mental states," is not possible (Bialosiewicz et al., 2013, p. 3). Interpretations of change within an organization or differences are often built upon postulations about the measurement model's underlying parity. However, parity may not be the case. Year over year, differences could vary in terms of the implied meaning of latent variables, making comparison highly questionable at best and impossible at worst (Vandenberg & Lance, 2000). Though this topic is infrequently researched, several factors suggest that MI is relevant, if not foundational, to the body of HRD research, particularly in the areas of item response theory (IRT) and SEM (Kim & Willson, 2014; Vandenberg & Lance, 2000). Validation of a psychometric instrument between groups allows for within and between-group understanding of the same questions provided in a survey (Xu, 2020). The findings of MI studies have implications for industry practice and organization science. Both will be explored more fully in the literature review.

Annual engagement cycle at SystemTex. Between 2011 and 2015, SystemTex offered employees the opportunity to complete the 46-item EES voluntarily. The

engagement survey generated a significant amount of organizational fanfare, and employee participation never fell below 70%. The 46 statements were designed on a 6-point Likert scale (strongly disagree, disagree, slightly disagree, slightly agree, agree, strongly agree). The Engagement construct was posited to be a four-statement index within the EES. The last four statements comprised the Engagement construct: (S43) *I would recommend this organization to my friends as a great place to work*; (S44) *This organization inspires me to perform my best*; (S45*) I am likely to be working for this organization three years from now*; and (S46) *I am willing to put in a great deal of effort in order to help this organization succeed*. The engagement cycle is shown in Figure 2. The full EES scale is printed in Appendix A with The Advisory Board's permission.

Figure 2

SystemTex Annual Employee Engagement Survey Cycle



The survey was open for two weeks permitting participation in the electronically administered survey to all eligible employees across the day, evening, and night shifts. Organizational leaders with the title of Supervisor and above, and HR employees encouraged participation. Once the survey window closed, The Advisory Board required approximately six weeks to analyze the raw data. After their analysis, The Advisory Board produced PowerPoint presentation reports for SystemTex, summarizing the results in the form of means and Percent Engaged scores for the whole organization and each business unit/facility, department, and manager. These reports were the basis of the organization's understanding of employee engagement and where engagement was surging or lagging. The mean score categorized results into four areas and definitions (The Advisory Board Survey Solutions, 2011).

Engaged: Employees who go above and beyond to see the organization's success, tie personal success directly to that of the organization, and are highly loyal and emotionally committed to the organization.

Content: Employees are solid contributors, satisfied with their jobs and the organization; they lack an emotional commitment to the organization.

Ambivalent: Employees would leave if presented with a better offer and see their job as a paycheck more than anything else.

Disengaged: Employees are least satisfied with their jobs and the organization. They tend to be most vocal and actively detract from the quality of the workplace for their peers.

Sharing results. Distribution and broadcasting survey results were part of the yearly EES cycle. The Advisory Board and SystemTex administrators believed the scores required meaningful and clear explanations for the employees. Therefore, the system-level and hospital-level results were interpreted qualitatively first by The Advisory Board and senior-level SystemTex leaders. Both organizations believed it helpful to overlay hospital, market, and industry-based context and circumstances on the EES results. This qualitative translation of scores occurred with the express purpose of sense-making to enable SystemTex to respond to two practical questions: (a) What did the scores mean to SystemTex? (i.e., the system, each hospital, department, unit, and manager); and (b) What should be done about the results to improve them?

As the mean scores shifted from one year to the next, changes were attributed to internal and external business and operational conditions. When added to employee and equipment shortages, these issues were a mixed cocktail of anecdotal evidence offered throughout SystemTex, regardless of whether mean scores improved or declined. A few reasons given for Engagement score shifts included:

- Broad governmental decisions handed down by the Centers for Medicare & Medicaid Services (CMS), the Affordable Care Act (ACA), and the World Health Organization (WHO) were challenging.
- Governmental and regulatory concerns were impacting operational policies and practices, including capital expenditures and reimbursement rates. Information Technology (IT), accounting systems, and medical classification systems required capital expenditures and a lot of training time.

- Dramatic shifts in the local and national healthcare landscape generated movement towards consolidation through M&As to establish economies of scale, cost savings, and risk management.
- Internal organizational challenges such as reduced manager effectiveness, high employee turnover, and revenue cycle management were believed to be antecedents to the manager's inability to affect change, avoid costly quality mistakes, and manage shrinking budgets.

Action planning and interventions. The next step following the sharing of the hospital system-wide results was to identify and focus on poorly performing departments or managers for tailored interventional work. By the time this stage occurred in the process, most employees and their leaders had a priori exposure to the results' interpretive reasoning. OD department team members were responsible for researching Engagement scores, including historical scores for trends. A poorly performing department or leader was identified by having an engagement mean score below the 75th percentile of The Advisory Board's benchmark, so leaders may not have had bad scores per se as will be discussed with the results in Chapter IV. Together, OD consultants and department leaders collaborated, co-architected, and co-implemented organized plans to address issues directly with employees through group processes like brainstorming, unit-level focus groups, appreciative inquiry workshops, and one-on-one interviews (McLean & Egan, 2008). The OD department members led interventional activities with the support of all levels of hospital administration and leadership to (a) identify behaviors within the organization, facility, unit/department, team level over which there was

control or influence; (b) respond to low-scoring survey items with coordinated, agreed-upon tactics; (c) understand what and how leadership and professional developmental could improve the work-lives of employees; (d) deploy training, knowledge sharing, and leadership development to address knowledge and performance gaps; and (e) codify best practices into manager tool kits to improve leadership skills vis a vis engagement and change readiness. These steps took place for six to nine months of the annual engagement cycle.

In the interventions between 2011 and 2012, employees specified issues such as CMS reimbursements being passed down to the hospital, internal quality and safety challenges, and disengagement within the nursing staff—the largest demographic of employees. Other issues included a lack of resources to perform one's job, poor manager ineffectiveness, a need for SystemTex to provide professional development. OD practitioners debriefed the department managers, HR leaders, and other senior administrators on the progress and results of the discussions, decisions, and evolution of the planned interventions.

In my personal/professional experience, the root causes for Engagement results in 2013 began to shift, surfacing new and varied topics than in previous years. One of the concerns that changed was that the acquisition became a continuous discussion topic. Employees mentioned a lack of visibility into the progress, process, and concerns about the acquiring organization's changes, resulting in tension and stress. Employees were concerned about what the acquisition meant for their professional lives, and this unknown permeated daily operations. When the results were issued, interpreted, and

broadcast throughout SystemTex at town halls and other formal and informal meetings, push back began to arise in employee disagreements about differences and changes perceived in the organizational culture. Employee reactions during those meetings, town halls, and focus groups could become accusatory and disparaging of decision-makers' behaviors.

In addition to worries about the culture and Engagement score changes, there were many leadership changes. From 2011 to 2015, SystemTex had three system-level CEO changes, multiple system CFO changes, hired new Chief Nursing (CNO) and Medical Officers (CMO), and made various hospital president changes. Financial strain indicated by two rounds of layoffs and more than $400 million in operating losses deepened employees' worries about the acquisition. Meetings held for any business or clinical purposes could become a kind of pulpit from which leadership could agree with, refute, or even discount employee reasons for their concerns with the acquisition. Front-line employee, manager, and administrator perspectives created tensions at multiple organizational levels.

In 2014 and 2015, leaders and the OD team implemented new approaches across SystemTex to address emergent employee engagement challenges. One program highlighted high-performing department leaders as role-models from whom other leaders could learn. Department leaders were given access to a web-based electronic action-planning tool for developing their action plans. It allowed their up-line managers to review and approve their Engagement-related projects. Poorly performing managers

were mandated to present steps by which they planned to address results, while well-performing leaders would specify how they planned to maintain their results.

You cannot get there from here. As someone steeped in scholastic effort, I became bemused by how much the near-inconspicuous leap from the EES results to the organizational interpretations contributed nothing to theoretical and empirical research in engagement (Alarcon & Edwards, 2011; Barnes & Collier, 2013; Saks, 2006). However, from the perspective of a practitioner/employee, I believe SystemTex deployed appropriate action in applying logic and context to the results because organizations and their employees operate under implicit mandates from internal and external stakeholders to generate results and do so quickly and consistently. As a professional, my primary role included advancing the organization's engagement-related tactics and plans in conjunction with the consultants and vendors like The Advisory Board. Continuous improvement and prompt solution-oriented action meant exploiting data to explain the status quo and suggesting easy-to-implement designs for organizational change. The engagement cycle structure included preplanned steps to delve into the quantitative feedback by employing root cause analysis and interventional measures in a broad and general sense. Thus, there was always the intention of a mixed-methods approach to improve engagement. However, the program's qualitative elements became problematic as we sought artifacts of ideas we already believed to be true.

A second reason for the leap from quantitative scores to interpretive conjecture came down to a need for simplicity. The simplification of the results, sense-making, time, and money allocated to the engagement cycle all needed to ensure that the

multifaceted information in the complex environment was digestible, relevant, and easy. Organization science has affirmed simplicity's usefulness and necessity, particularly when organizational leaders must communicate complex topics with others (Remington & Pollack, 2012). Simplification, in the form of clarity and brevity, is even more essential for leaders than direction, especially when said guidance is not clear (Remington & Pollack, 2012). The problem with simplicity is that it belies the wisdom of rich, diverse, integrated, and complex systems thinking.

Furthermore, as long as the organization was reworking ideas already recognized and assumed, it could not engage in organizational learning. Simplification in the form of management of the narrative tailored towards issues of perceived control may have led to convenient storylines while indulging suboptimal comprehension of organization phenomena. Within complex systems, learning and adapting mechanisms are sometimes buried within employee perspectives only visible with the application of a more elaborate level of analysis.

Significance of the Study

This research represents a unique intersection of concepts. No studies seem to focus on the combination of MI, M&A's in healthcare, and employee engagement. The annual EES cycle of SystemTex created a call to action, launching an urgency for employee engagement and giving leaders a structured program and process by which to talk about the construct. Nevertheless, it is impossible to say whether SystemTex spent its time and financial resources well unless the instrument used to measure the concept did so effectively.

This study may offer fascinating information to the body of HRD, management, and organization science because it explores a large data set from a single organization that can be difficult to acquire in academic investigations.

The first research question is the reliability of the Engagement construct. The underlying concern is whether this survey instrument validates the stability of the engagement factor structure between groups (Chungkham et al., 2013; Vestergren, Rönnlund, Nyberg, & Nilsson, 2012). For SystemTex and other organizations using psychometric instruments, this study may offer valuable perspectives on employee engagement constructs, the healthcare industry, and M&A activity. For vendors and researchers designing similar instruments on behalf of client groups, this analysis of constructs may give evidence as to why instrument validity and reliability is critical to the organizational outcomes they wish to impact.

Implications

Historically speaking, organizations needed employee engagement surveys to test objectively for valuable feedback. Modern concerns with employee engagement are part of the struggle for talent and retention of employees. Changes in dynamics such as national and international mobility, the retirement of an aging workforce, turnover, and talent shortages compel organizations to consider employee engagement to optimize their workforce planning. The study of employee engagement impacts organizational effectiveness, contributing directly to the wealth of organization science, employee behavior, and related management research. These topics become essential elements for

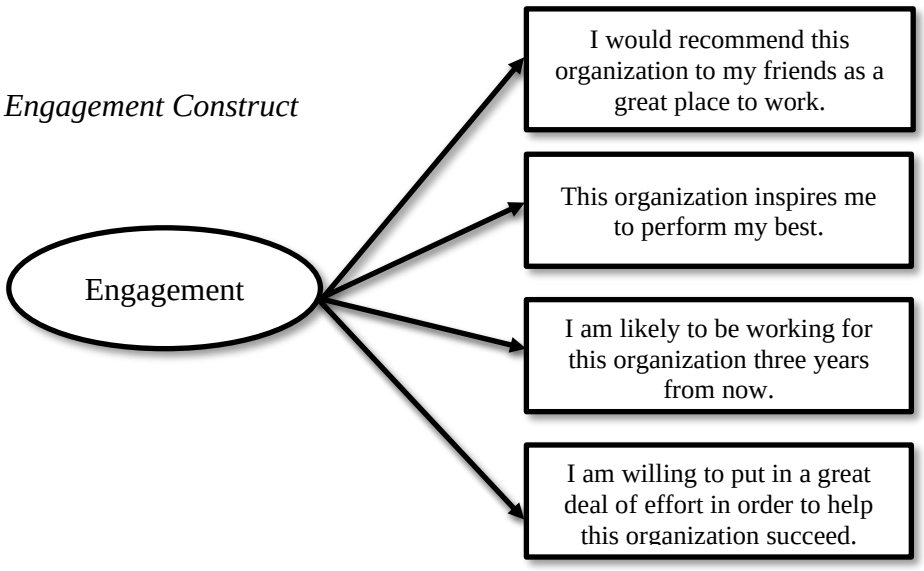
organizations looking to leverage talent and aptitude as a path to business strategies and assert competitive advantage (Gallant & Martins, 2018).

Theoretical Framework: Confirmatory Factor Analysis

There are two types of MI, multi-group and longitudinal invariance (Bialosiewicz et al., 2013; Xu, 2020). Multi-group invariance is the domain of interest. Assuming that a model is viable, it is possible to test for differences in the measurement model. *Invariance* is the property verifying if the model fit is as good for the data (Meyers, Gamst, & Guarino, 2016). The model is *non-invariant* if the pattern coefficients do not remain consistent (Meyers et al., 2016). MI design is a widely recognized approach in studies when testing invariance utilizing multivariate items and latent constructs within CFA frameworks (Bialosiewicz et al., 2013; Chungkham et al., 2013; Vandenberg & Lance, 2000). Figure 3 shows the prescribed Engagement construct designed by The Advisory Board and is printed with its written permission (see Appendix D).

Figure 3

EES Prescribed Engagement Construct



Limitations

I had no input on how the survey was theorized, organized, or disseminated. Out of an abundance of caution to protect any methodological intellectual property, The Advisory Board did not provide information on face, or content validity, or reliability. EES data gathering had no qualitative element or open-ended questions associated with the instrument.

Organization of the Study

This dissertation is composed of five chapters and an Appendix. This first chapter introduces the conceptual framework, starting with the background and my personal experience. It presents the significance of the problem, research questions, and definitions and implications of the study.

Chapter II is an integrative literature review on MI, employee engagement, M&As in healthcare, and archival data.

Chapter III focuses on the methodology and procedures required for effectuating MI tests. Descriptions of the employee sample sizes, the EES set up, and data collection and storage are provided.

Chapter IV describes the results of testing the EES engagement construct for model fit, model comparisons of indices, and four levels of stepwise tests for MI.

Chapter V, the final chapter, presents the finding offered in chapter four. Implications of the findings for management, HRD, and healthcare are discussed. Recommendations for additional investigation are also presented.

The Appendix contains the EES survey in total, covariance matrix, configural fit statistics for invariant pairings, and Advisory Board and IRB approvals.

CHAPTER II

LITERATURE REVIEW

Kahn (1990) introduced psychological conditions of work-related personal engagement as being separate from organizational behavior concepts of job involvement and organizational commitment. By analyzing worker behavior and interaction, Kahn (1990) tendered the first grounded theory for personal engagement and disengagement in the workforce. In the 30 years since Kahn's article, employee engagement has come to describe a host of distinctive psychological traits, conditions, behaviors, and outcomes (Gupta & J, 2017; Macey & Schneider, 2008). Researchers seek to concretize employee engagement concepts (Macey & Schneider, 2008), provide frameworks, and validate measurement of the construct based on theory and practice (Shuck, Adelson & Reio, 2017), while practitioners look to operationalize the construct within organizations for its immense implications for business outcomes.

An Integrative Literature Review

This review of the literature brought together contributions of historical and current investigation at the intersections of measurement invariance (MI), engagement, and M&A activity. Each topical area was exponentially vast, so one way to narrow focus and prevent any attempt to "boil the ocean" was to confine the literature review to places where relevant topics overlapped. An integrative literature review permitted the summarization of literature from various emergent fields to create new knowledge **(**Shuck & Wollard, 2010; Torraco, 2005).

Organization and process. This literature review was organized into three major sections (a) key terms used for literature searches, (b) literature intersections (c) a conclusion. The integrative literature review process began with issue selection, then a review of relevant literature, followed by analysis and critique of the literature for insight and synthesis (Shuck & Wollard, 2010). Multiple literature reviews were conducted across the topics to ensure an in-depth and complete evaluation of their contribution to the present study. Several questions guided the analysis of the literature review and were repeated for each search for consistency. What is the (a) topical historical relevance? (b) contemporary topical relevance of the problem to the field (c) relevance to the current theory, and (d) relevance to current practice? The literature review outcome was for the project to be placed in precise historical, contemporary, theoretical, and practical contexts.

Key Terms Used for Literature Searches

Measurement invariance and multi-group measurement invariance: There was a lot of research in MI, but not much in practice because establishing MI can be difficult (Vandenberg & Lance, 2002). When MI issues go unaddressed, ignored, or are constrained by simplistic analysis that compares latent mean scores between groups, then the reason for measurement may not hold, and comparisons become invalid (van de Schoot, Schmidt, De Beuckelaer, Lek, & Zondervan-Zwijnenburg, 2015). The confirmatory factor analysis (CFA) framework is a method for validity testing of the items which comprise the construct “as an indirect measure of the hypothesized latent variable” (Bialosiewicz et al., 2013, p. 4).

Employee engagement: Employee engagement's historical, theoretical relevance started in 1990 with Kahn describing the physical, cognitive, and emotional dimensions people employed in their roles at work. A significant difference between the explanation provided by Kahn (1990) and The Advisory Board's definition was that Kahn did not couple engagement with organizational outcomes. In the decade between Kahn's (1990) and The Advisory Board's (2006) approaches, engagement became widely embraced by practitioners and researchers alike, though the two factions differed on definitions and measurement (Shuck & Wollard, 2010). Table 1 shows how some of the many definitions of engagement have altered. Saks (2006) was credited with dividing employee engagement into organizational and work engagement (Merve Ünal & Turgut, 2015).

Table 1

Employee Engagement Definitions

Category	Definition	Source
Personal engagement	"the simultaneous employment and expression of a person's 'preferred self' in task behaviors that promote connections to work and to others, personal presence, and active full role performances."	Kahn (1990, p. 700)
Absorption	Psychological presence characterized by attention and absorption.	Rothbard (2001)
Work engagement (opposite of burnout)	"A positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption."	Schaufeli et al. (2002, p. 74)
Job satisfaction	Individual involvement, satisfaction, and enthusiasm for work.	Harter, Schmidt, & Hayes, 2002
Organizational engagement (practitioner literature)	The amount of discretionary effort employees put into their work through extra time, brainpower, and energy to go beyond minimum job requirements.	Towers Perrin Talent Report (2003)
	The measure of an employee's emotional and intellectual commitment to their organization and its success.	Hewitt Associates (2004)
Multi-dimensional	Trait, state, and behavioral construct inclusive of work and organizational conditions necessary for state and behavioral engagement.	Macey & Schneider (2008)

In academic literature, engagement consists of cognitive, emotional, and behavioral elements associated with an individual's role performance (Saks, 2006). Distinguishing new ideas and constructs from those that came before is not easy since multiple categories and types of engagement litter the academic and commercial landscapes (Saks, 2006; Shuck & Wollard, 2010). Engagement may be confused or melded with constructs like organizational commitment, organizational citizenship behavior (OCB), job involvement (Saks, 2006). Engagement constructs may include operationalizations of intensity into behavioral, emotional, cognitive energy, and work, job, and social engagement (Shuck et al., 2017). Work engagement is focused on burnout experiences, while job engagement attempts to distinguish investment in employee work performance (Shuck et al., 2017). On the other hand, organizational engagement seeks to clarify and quantify how "present" (p. 4) one feels in their job role. Still, others suggest that engagement is an attitude and emotional attachment of employees towards organizational values and goals (Rasheed, Kahn & Ramzan, 2013). Nevertheless, even with this wide-spread attention, engagement lacks empirical research (Shuck & Wollard, 2010).

Mergers and acquisitions (M&As): This term was defined in Chapter I so that the reader understood the specific nature of the transaction—an acquisition—SystemTex experienced. Here, it is used more broadly for seeking literature overlapping with other topical areas. Febriani and Yancey (2019) summarized the literature on M&As neatly: M&As are complex business procedures wherein intended outcomes can take many forms, including favorable positioning, increased market share, and profitability. In

2015, there were more than 44,000 M&A transactions with a total value beyond U.S.$4.5 trillion (Febriani & Yancey, 2019) M&As represent a significant event in organizational change. Though not all are failures, many studies analyze how and why M&As fail to produce desired outcomes (Febriani & Yancey, 2019; Gleibs, Mummendey, & Noack, 2008). Poor cultural fit and mismanagement of pre- and post-acquisition planning characterizes many failures (Teerikangas & Very, 2006). "If a CEO is enthused about a particularly foolish acquisition, both his internal staff and his outside advisers will come up with whatever projections are needed to justify his stance" (Sirower & Sahni, 2006, p. 83 quoting Warren Buffett from the 1997 Berkshire Hathaway annual report). Teerikangas and Very (2006) point to "inadequate strategic rationale behind the deal" (p. 31), reinforcing Mr. Buffet's notion of weak M&A justification, suggesting that they may not be strategically initiated decisions at all. Adverse downstream outcomes of failed M&As include reduced employee satisfaction, among other financial and cultural consequences of engagement (Febriani & Yancey, 2019). Furthermore, in instances where M&A necessitate organizational restructuring through formal leadership, name, and structural changes, reasons for failure due to employees' feelings have been found. Threats to employee self-esteem, insecurity about how changes might relate to work, and retention of old identities have been identified as source problems. (Bartels, Pruyn, & Jong, 2009).

Commercial viability and ROI of engagement. For practitioners, engagement has become tethered to organizational success and business results, notably cultural harmony, well-being, productivity, and reduction of turnover (Gallant & Martins, 2018;

Harter, Schmidt, & Hayes, 2002), to organizational health, organizational effectiveness, and staff retention (Gallant & Martins, 2018; Seymour & Dupré, 2008; Singh, 2013). The Advisory Board and other consultancies have found a foothold in business, mainly when engagement is applied—often loosely—to organizational theory. The Advisory Board is included in a sizeable group of commercial firms researching and capitalizing on the importance of engagement in business management. Current vendors include but are not limited to Development Dimensions International (DDI), the Australian Institute of Management, the American Society of Talent Development (ATD), the Indian Institute of Management Bangalore, and the Institute of Employment Studies (Zigarmi, Nimon, Houson, Witt, & Diehl, 2009).

In more recent times, HR practices and policies, written to protect and grow an organization's human performance capacities, look for tools said by vendors to secure often-elusive high employee engagement status. *First Break All the Rules* (Buckingham & Coffman, 1999) was a popular source of motivation for organizations, based on industry data and research from the Gallup organization's Workplace Audit (GWA) which further popularized employee engagement. Gallup created and delivered the initial engagement survey in 1999 (Zigarmi et al., 2009) and was engagement's primary purveyor. An example of the scale of Gallup's influence was its 2013 employee engagement poll, which analyzed 49,928 work units or businesses, covering 1.4 million employees, across 49 industries, in 34 countries (Zigarmi et al., 2009).

In other organizations, leaders like General Electric's (GE) former CEO, Jack Welch, cited employee engagement as most critical over other essential management

topics, including customer satisfaction and free cash flow (Welch & Welch, 2006). Welch said, “The most meaningful surveys probe how employees feel about the strategic direction of the company and the quality of their career opportunities” (p. 126). The focus is on the competitive advantage organizations gain when employees bring their whole selves to work (Singh, 2013). In the Society for Human Resource Management's (SHRM) 2006 publication, *Employee Engagement and Commitment*, Vance (2006) wrote about practical aspects of engagement within an organization by describing enhanced safety and cost savings measures due to engagement. Fabick CAT (a subsidiary of Caterpillar in Fenton, MO) hired Gallup for employee engagement and were able to improve engaged employee results from 16% Engaged in 2002 to 45% Engaged in 2006 (Robison, 2006). The company president was recorded as saying that Fabick CAT reduced expenses by $500,000 during the same timeframe, thus proving ROI for the engagement expenses (Robison, 2006).

Molson Coors Brewing Company found relationships between employee engagement, safety incidences, and lost productivity (Vance, 2006). The average safety incident for engaged employees cost the organization $63, compared to an average of $392 for non-engaged employees. The company saved $1,721,760 in safety costs in 2002. Additionally, poorly engaged sales teams dragged behind in sales volumes compared to more engaged teams (Vance, 2006).

Literature Intersections

Integrative literature reviews ideally include both historical and recent publications (Torraco, 2005). The empirical studies historically and currently apropos to

this study are organized by searches for the highest combination of keywords and funneling down to searches of pairs of keywords. No searches for studies of a singular topic were made to preserve the literature review's integrative structure. Regrettably, though not unexpectedly, no published studies included MI to analyze employee engagement in healthcare organizations pre- or post-M&A with archival data. That conglomeration of topics was too narrow by far.

Figures 4 and 5 show the process by which literature was identified through Boolean searches of the keyword combinations through the Texas A&M University Library online collection of 1,146 databases. The database searches were of peer-reviewed journals with the option to apply related terms and search within the full text of an article for keywords. Searchers were also conducted with and without quotation marks ensuring a thorough review. When the search terms were used without quotation marks, a larger number of documents were returned. For example, using *employee engagement* as a keyword search of peer-reviewed abstracts returned 10,568 documents. Adding quotation marks ensured that exact terms were obtained, subsequently reducing reduced the number of search returns. A peer-reviewed search of abstracts of "employee engagement" produced 6,138 results, and for that reason, it was only used with quotation marks in all of the search combinations. In the search for "*mergers and acquisitions*" AND *healthcare* or "*health*…" not confining the search to article abstracts resulted in 13,301,565 documents. For that reason, searches were confined to finding the terms in article abstracts. No results returned for keyword searches combining *MI* and *healthcare*. It unnecessarily complicated the focus since the goal was to avoid research about MI in

scales used to test medical conditions. Some of the searches were limited to academic journals, as some subjects are widely written about in trade publications and non-academic industry journals.

Figure 4

Literature Review Search Process with Three Keywords

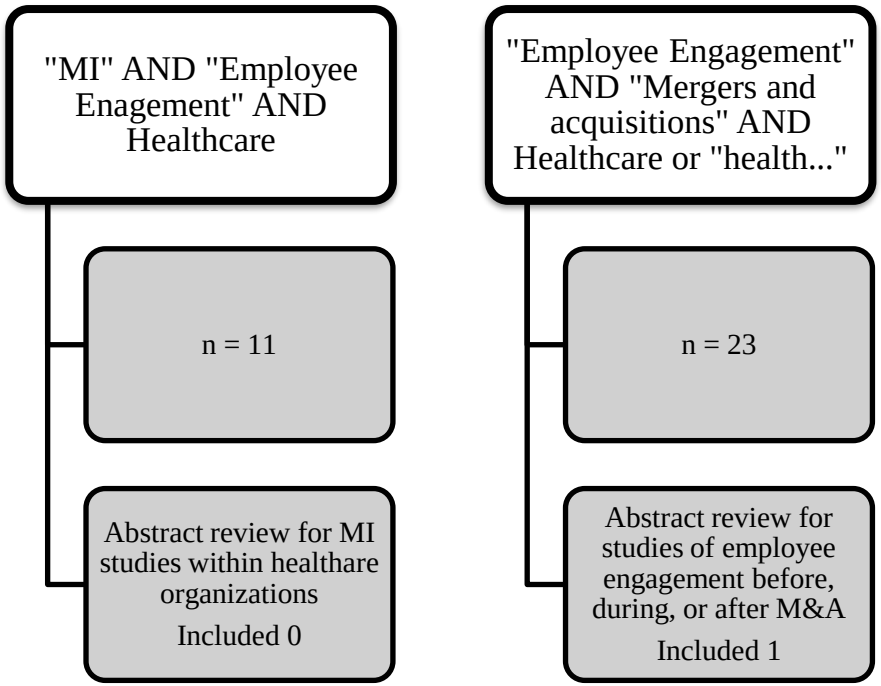
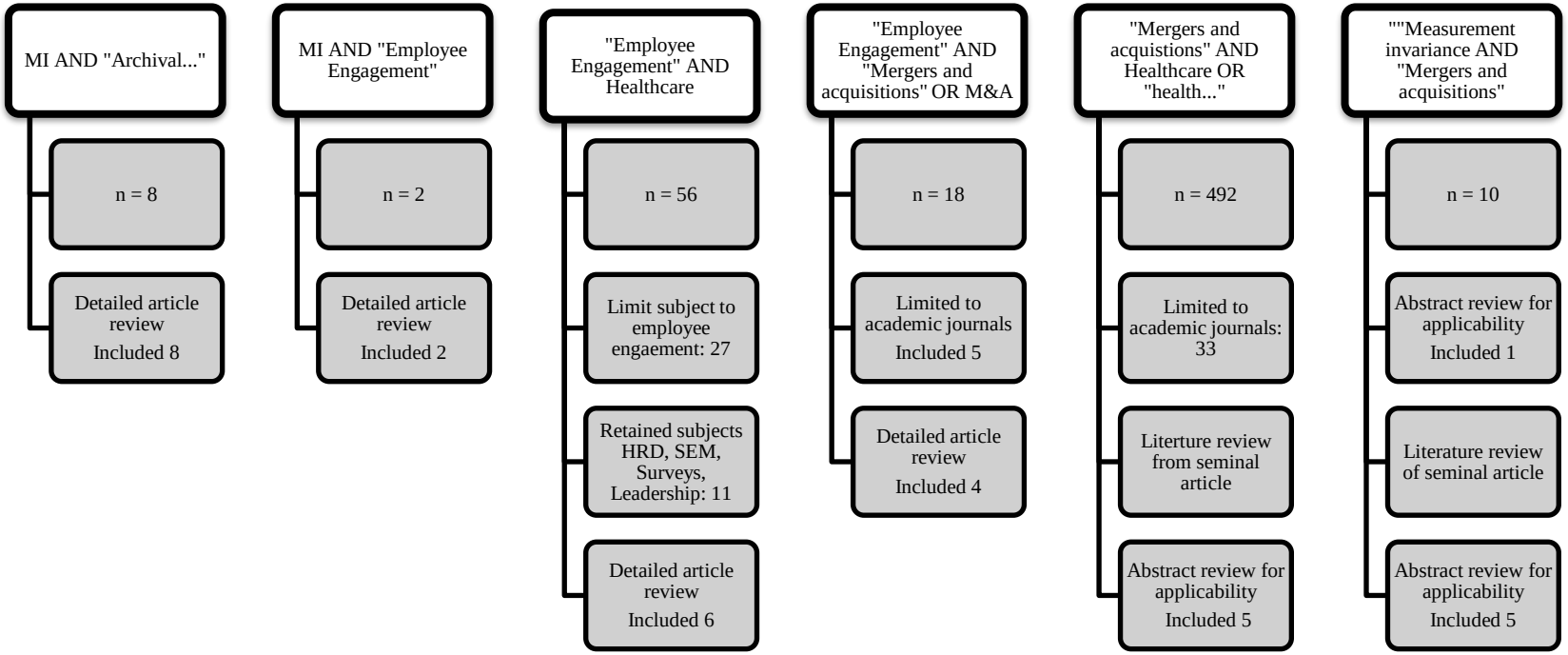


Figure 5

Literature Review Search Process with Two Keywords



MI and Employee Engagement and Healthcare

A search of these three terms resulted in 11 articles, all of which were reviewed and eliminated for applicability because it (a) did not reflect all three search terms, (b) was not empirical, and (c) parsed keywords such that they did not reflect the same concepts of this study. No results were obtained when these same terms were confined to search the abstract of the article.

Employee Engagement and M&A and Healthcare

A particular qualitative study deserves mention for the contribution it makes to current engagement research in organizations going through M&As, considering that it is one of a few attempting to understand the causality of M&A failure. It points out the pallid evidenced-based research of this intersection granted (a) the proclivity of recent M&As in healthcare and (b) a heavy industry focus on employee engagement. Magano and Thomas (2017) employed semi-structured interviews of 60 senior managers of a pharmaceutical company illuminating difficulties that arose from multiple M&As. The senior managers lent their perspective on the breakdown of the psychological contract between the company and its employees during the times of change, uncovering seven themes (1) lack of communication, (2) an absence of planning, (3) ineffectual leadership, (4) lack of employee engagement, (5) less than optimal human resources involvement, (6) lack of preparation of the organizational culture, and (7) imperfect change management process (Magano & Thomas, 2017).

MI and Archival Data

This study advances organization science theory by its methodological approach and research design. Aguinis, Pierce, Bosco, and Muslin (2009) studied the first decade of organizational research methods through content analysis of a popular management journal, Organizational Research Methods (ORM). One objective of the article was to identify areas of opportunity to train doctoral students and topics needing more significant contributions. Between 1998 and 2007, MI articles were published in 20% of the articles from 2002 to 2004. After 2004, MI publications declined steadily to less than 10% of ORM's articles. Aguinis et al. (2009) reported a need for doctoral students and researchers to enhance quantitative methodological practice with validity studies and archival design.

Archival design is an untapped and overlooked resource for organization science, demonstrated by the fact that only 10% of articles published in ORM between 1998 and 2007 used archival data (Aguinis et al., 2009; Das, Jain, & Mishra, 2018). Archival data is that which has been "…gathered and stored prior to the commencement of the researcher, intended for later use" (Das et al., 2018, p. 138) and has distinct advantages. First, these are often valuable data sets existing in abundance because of how organizations survey their populations. Second, procuring archival data can occur without great expense than designing and conducting an original survey (Das et al., 2018; Martin & Friedman, 2000). Third, organizational research is frequently designed to capture trends and measure change (Das et al., 2018) or internal requirements for

record-keeping and document management standards. See Table 2 for the eight studies published from 2000 to 2019 on MI and archival data.

There may be little archival data research because the term “archival” is not designated in the abstract or key terms. In the case of this search, the search's history only began in the year 2000, before which there seemed to be no attempts to use MI to analyze archival data. That said, access to unexplored data may not be easy for researchers to come by. Moreover, practitioners may not care much to have old information analyzed if it cannot be shown to have a bearing on the business's current or future state. However, it is being shown that the analysis of archived studies is validating latent constructs in older studies, thereby giving contemporary scientists reliable findings upon which to build.

Table 2

MI and Archival Data Research Design Studies

No.	Source	Type of archival data used	Method adopted for analyzing data	Key Findings
1	Martin and Friedman (2000)	1,178 males and females averaging 11.2 years old, recruited between 1921-1922, tested for personality traits taken from Terman's Life Cycle Study.	These data were supplemented by contemporary questionnaire responses from samples of 167 children and 203 adults to test for full MI of five scales across all three samples. Rational and correlational techniques measured validity of latent constructs.	They constructed and validated personality scales with archival data, showing consistency of the measurement model between contemporary and archival respondents' understanding of the Big Five personality traits.
2	Biehn, Elhai, Fine, Seligman, and Richardson (2012)	378 Canadian veterans evaluated using the PTSD Checklist (PCL) assessing PTSD symptoms between 2000 and 2008.	Divided sample into PTSD and no-PTSD groups and use CFA to test between-group differences.	There are factor structure differences between individuals with PTSD compared to those without, with some concern that the original 1998 and 2002 study samples included people without PTSD diagnoses.
3	Fox and Mitchum (2013)	6 studies from 1964 to 2012 reporting pass rates of 3,867 non-clinical responses to the Raven Advanced Progressive Matrices Test.	Organize studies into two groups, Cohort 1940 and Cohort 1990 for IRT, specifically differential item functioning (DIF) comparisons where increasing scores on the Ravens Matrix show higher-levels of inference-making capability.	IQ gains can be explained in that those born around 1990 are more capable of understanding higher abstraction levels than those born around 1940 due to newer approaches to analogy.

Table 2 Continued

No.	Source	Type of archival data used	Method adopted for analyzing data	Key Findings
4	Pietschnig, Tran, and Voracek (2013)	5,445 Austrian psychiatric patients from 1978 to 1994 taking the MWT-B intelligence test.	Two parameter model of IRT to analyze aggregated data across four cohorts assessing four or five consecutive years and freeing some non-invariant items improved overall model fit.	The same construct is measured across all four cohorts. Changes in test scores can be attributed to narrowing the ability distribution.
5	Kamody et al. (2014)	266 adolescents completing the 65-item Diabetes Stress Questionnaire (DSQ) since 2003.	These data were supplemented with the Predicting Resiliency in Youth Type 1 Diabetes study. CFA framework used to determine MI across age, sex, and glycemic control (Kamody et al., 2014).	DSQ factor structure is the same across all groups.
6	Kim, Martin, and Nolty (2016)	1,325 Daily Spiritual Experiences Scale (DSES) completions from 2004.	Conduct CFA for differences across men and women.	Full MI was supported. They demonstrated that women have higher levels of daily spiritual experiences than men. The survey wording does not cause gender bias.
7	Klages et al. (2019)	142 youth with Type 1 diabetes completions of the 24-item DSQ Short Form (DSQ-SF) from 2015.	These data were supplemented with an additional 181 youth with Type 1 diabetes. CFA was conducted to assess factor structure of DSQ-SF across groups.	DSQ-SF factor structure is maintained across groups and an independent sample. It also maintains factor structure with the original DSQ across groups.

Table 2 Continued

No.	Source	Type of archival data used	Method adopted for analyzing data	Key Findings
8	McDermott, Hammer, Levant, Borgogna, and McKelvey (2019)	6,774 male and female completions of the Male Role Norms Inventory–Short Form (MRNI-SF) from 2017.	These data were supplemented with a 2nd sample of 484, a 3rd sample of 1,537 completing the MRNI-SF, and a 4th sample of 365 people completing the MRNI-VB. SEM analyzed all 4 samples to see if the MRNI-VB maintained factorial validity with variables as the MRNI-SF.	Traditional male ideology (TMI) is factorially equivalent in both the MRNI-SF completed by the archival dataset and the Male Role Norms Inventory–Very Brief (MRNI-VB).

MI and Employee Engagement

Two studies overlapped the application MI analysis of an employee engagement scale. In 2017, Shuck et al. published a significant contribution to empirical scrutiny of engagement surveys by validating the Employee Engagement Scale, which is not the same EES built by The Advisory Board. The Shuck et al. (2017) EES measured three subfactors—cognitive, emotional, and behavioral engagement. Through four independent studies using CFA, the framework confirmed the scale reliability and validity of the underlying measurement model. The findings revealed the first engagement instrument, "grounded in previously documented theory and definition" (sShuck et al., 2017, p. 972). In a personal email exchange with author Dr. Brad Shuck, I learned that this extensive project took two years to complete due to instrument refinement's iterative nature across four samples. Replicating this kind of effort may be challenging to accomplish due to timeliness and cognitive capacity. Moreover, the research could not connect their findings to turnover, effort, or productivity measures.

In 2018, Gallant and Martins analyzed data from 1,175 employees using exploratory factor analysis (EFA) and CFA for an accurate measure of the engagement construct across race groups. Technically, this study qualified as an archival data design, as the employee data derived from a South African research company's database. I chose not to place it with the archival studies because the authors or publishers had not done so. The assessment—a 5-point Likert scale developed in 2015—measured individual, team, and organizational engagement (Gallant & Martins, 2018). Their endeavor was a cross-sectional design using convenience sampling. The investigators contended that

there is an overrepresentation of white males in their sample, resulting in skewness. A consideration from Gallant and Martins' (2018) investigation was that socioeconomic factors can impact employee engagement and that not all engagement scales fit all populations.

Employee Engagement and Healthcare

Six articles of empirical research were at the intersection of employee engagement and healthcare. Healthcare employee engagement may focus primarily on nurses because they comprise the largest demographic by job title compared to other roles. Excluding burnout as a condition of the Boolean search kept the focus on employee engagement. Many studies documented the prevalence of healthcare workers' burnout experience, and most acknowledged the construct as separate and apart from engagement (Shuck & Reio, 2014). Moreover, burnout has received criticism for overutilization as an inventory of employee engagement (Shuck, Twyford, Reio, & Shuck, 2014). However, other terms, specifically emotional exhaustion, also conveyed burnout and were included in this search for employee engagement in healthcare organizations.

Table 3 displays the six articles recognized in this search. What is striking about the collection is that the first study did not occur until 2011 when Suh, Houston, Barney, and Kwon correlated mission fulfillment with the psychological states of employee engagement, social identity, and emotional exhaustion. Carter and Tourangeau (2012) argued for the strong relationship between turnover intentions and psychological engagement, while Shuck et al. (2014) produced findings leveraging the social exchange

theoretical (SET) framework to suggest that HRD practices could positively impact employee turnover. SET is based on the assumption that (a) work relationships and organizational systems are interdependent, and (b) obligations within work are generated through interactions between interdependent parties (Shuck et al., 2014).

Table 3

Employee Engagement in Healthcare Studies

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
1	Taewon Suh, Houston, Barney, & Kwon (2011)	Analyze how mission fulfillment is correlated to psychological states of employee engagement, social identity, and emotional exhaustion.	3,999 healthcare employees completed a multi-item instrument. CFA validated quality and direct-effects determined variable correlation.	Mission statements are powerful internal marketing levers. Mission fulfillment has a positive impact on engagement and generates commitment.
2	Carter & Tourangeau (2012)	Test eight determinants to observe nursing roles and turnover intentions.	Multi-level SEM analyzed 5-point Likert scale questions of 17,707 nurses from 167 healthcare organizations in England.	There is a strong relationship between turnover intentions and psychological engagement.
3	Granatino, Verkamp, & Stephen Parker (2013)	Understand the relationship between employee engagement and customer service.	Mixed method using phone interview, training, and an engagement survey of 49 staff managers in a healthcare setting	Employee satisfaction depends on the level of customer service employees believe they provide. More time is required to train on customer service specifics.
4	Shuck & Reio (2014)	Explore workplace climate association with personal accomplishment, depersonalization, emotional exhaustion, and physical wellbeing.	Multivariate ANOVA analyzed 216 healthcare workers in the U.S., Canada, and Japan, who completed an online survey.	HR professionals can instruct on managerial skills to improve supervisee engagement and system-level change efforts. Leaders can increase employee engagement and positive individual-level affective outcomes. Positive outcomes can be operationalized, where employee voices do not fall on deaf ears (p. 55).
5	Shuck, Twyford, Reio, & Shuck	Study employee engagement turnover intention using social	207 healthcare workers completing the Job	Organizations should do more to broadcast HRD practices so that

Table 3 Continued

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
	(2014)	exchange theory (SET) and clarify its relationship to employee performance	Engagement Scale (JES) and Perceived Investment in Employee Development (PIED) scales scored separately.	employees perceive equal access. Turnover intentions might be reduced through employee-help perceptions of support for participation in HRD practices.
6	Gupta & J (2017)	Measurement of employee engagement and staff nurse retention.	Cross-section survey of 250 staff nurses completing the NRC picker survey questionnaire over six months.	There is a positive correlation between employee engagement and retention levels. An engagement model based on the findings is proposed using emotional, rational, behavioral dimensions to calculate an engagement scale and predict retention.

Employee Engagement and M&As

To date, empirical studies of employee engagement and M&As have difficulty to home in on employee engagement outcomes in the precise moments before, during, or after an M&A. Only one of the four studies was so specific—Fabriani and Yancey (2019). This search (see Table 4) revealed studies focused on elements or perspectives of employee engagement. This search overlap made clear that research at this intersection is about timing and the instrument. Regarding timing, there must be data collected coinciding with the transaction if a researcher is to react to it.

Baynham (2011) focused on data provided by Mercer, a global business management company. The purpose of the study was to assess and distill organizational learning into repeatable best practices to facilitate the known challenges of maintaining employee engagement before, during, or after an M&A (Baynham, 2011). Like The Advisory Board, Mercer had its proprietary cultural integration survey that it administered throughout its client base to generate "best practice" guidance for organizational leaders of merging cultures. The following four conclusions came from the 2008 survey administration:

- Discover and define: Senior leadership engages in processes to discover and define cultural characteristics required for the M&A transaction.
- Drill down: Get to root-level comprehension of how each organization works and look for alignment and differences, risks, and success "derailers" (Baynham, 2011, p. 13).

- Deploy drivers: Identify common characteristics between organizations and levers to reinforce specific behavioral patterns.
- Determine what is delivered: Track and monitor the success of cultural alignment efforts.

The above best practices came from interviews with an unstated quantity of HR leaders. From my professional perspective, the guidelines typify the guidance sought by practitioners since it is easily understood and shared throughout an organization using a train-the-trainer model. What these best practices did not clearly assert was any theoretical or conceptual rigor a scholar might apply for framing the engagement or M&A problems or their solutions.

In contrast, Waight (2015) signaled that M&As were not identical transactions to integrations and proceeded to use a learning and development (L&D) framework as a lens for learning activity during an M&A. Using a phenomenological design, Waight (2015) interviewed 25 embedded L&D professionals responsible for learning and development within their organizations during the integration phase of M&As between 1996 and 2003. The findings, based on organization learning, fell into the following categories: (a) learning about change; (b) learning how to manage change; (c) learning for work, and (d) learning from experiences. Qualitative comments provided explanatory evidence for the categories and solutions applied by the L&D practitioners in each of the learning areas.

Table 4

Employee Engagement and Mergers and Acquisition Studies

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
1	Baynham (2011)	Describe successful practices for cultural integrations in M&As using Mercer interview data from HR leaders.	Mercer's culture integration survey 2007 – 2008 of an unknown number of HR leaders. Interviews were codified into four best practices.	There are four best practices organizations should do for cultural integrations: Discover and define Drill down Deploy drivers Determine what is delivered.
2	Waight (2015)	Use an L&D framework to examine the work L&D professionals do during an M&A.	Phenomenological study of 25 L&D practitioners from 1996 to 2003 analyzing constructs of organizational learning. Guided interview process included inter-raters using a 4-step analysis to establish themes.	4 themes clarified the work L&D practitioners accomplished during M&As Learning about change is a cornerstone of employee engagement. Learning how to manage change led to training of change management competencies. Learning from work centered on both technical and soft skills. Learning from experience was a critical reflexive process to develop M&A competencies and establish learning circles.
3	Williams (2015)	Describe approaches for banks to maintain customer engagement during M&As.	Phone interviews of 9,9219 members of the Gallup Panel ages 18 and older surveyed (a) June – October 2011, and (b) June 2014.	Leadership should be doing the following to engage customers: Get the affairs of the organization in order. Know the customer engagement of the target bank. Form strategies to emulate customer experiences Communicate early and often about M&A plans.

Table 4 Continued

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
4	Febriani & Yancey (2019)	Analysis of the relationship of five variables (integration approach, organizational culture, employee engagement, organizational commitment, and effectiveness of HR initiatives) in M&As of 2 merged companies.	Survey of 106 employees merged Indonesian companies following their merger. Each integration approach – preservation and transformational - was an independent variable to measure. OCAI measured 4 dimensions of organizational culture 3 Component Model of Commitment measured organizational commitment Utrecht Work Engagement Scale measured employee engagement.	The transformational approach results in employees perceived: • Less change during the merger • Greater organizational commitment and employee engagement after the merger (p. 108). • Employees see the organization becoming a market culture The preservation approach employees: • Observed small cultural changes during the merger • Do not observe the company becoming a market culture • Note that effective HR interventions can mitigate negative cultural changes.

M&As and Healthcare

Generalized problems associated with M&As include employee distress, poor shareholder value, disappointing customer experiences with new systems and processes, and overall failure to meet financial and cultural expectations (Febriani & Yancey, 2019; Li-Ping Tang & Timmer, 2008). To make a distinction from those far-reaching concerns, this intersection of topics highlighted healthcare-specific challenges of M&As. Some scholars suggest that U.S. healthcare M&As began in the 1980s due primarily through the consolidation of freestanding hospitals into larger hospital systems (Angeli & Maarse, 2012), while others claim M&A activity has been part of U.S. healthcare since the 1950s (Creasy & Kinard, 2013). M&As in healthcare generally occur to gain access to market share, expand existing markets, improve capital, prevent lockout from a consolidating market, decrease costs, minimize competition, and control costs (Li-Ping Tang & Timmer, 2008; Vita & Sacher, 2001). However, not-for-profit hospitals and systems are more charity-minded than profit-maximizers (Sloan, Ostermann, & Conover, 2003; Vita & Sacher, 2001). Healthcare M&A failures can pinpoint the sometimes-ignored requirement of medical staff input for hospital admissions. Said differently, no one is admitted to a hospital without a physician's order (Weil, 2010). Other M&A complications include merging redundant medical departments (e.g., pathology, radiology, mother/baby wards) and choosing whom to select as leadership across the merged entity (Weil, 2010). Costs of care resulting from M&As have been fount to rise from 2% to 40%, and often produce the butterfly effect of increasing costs

in hospitals that have not merged by 17% (Weil, 2010). Challenges and conditions that persist in acute-care system M&As include

- Technology advances, permitting shorter hospital stays and ambulatory treatments.
- Constrained hospital reimbursements encourage physicians to release patients as soon as or before they are clinically able to leave.
- Economies of scale whereby larger geographically based healthcare systems can better negotiate with governmental and economic authorities for reimbursement.

Four of these five studies analyzed archived data, but none of them was found in the search for archival design. The included studies ranging from 2001 to 2008 were in-depth analysis of M&A transactions in healthcare (see Table 5). Two of the articles point out a need for policymaker vigilance. Vita & Sacher (2001) evaluated not-for-profit hospitals exclusively. Many transactions are too small to meet the Federal Trade Commission (FTC) filing requirements, deals are executed without federal oversight, and are cause for anti-trust concerns. All the while, competition among not-for-profits is lessened, and costs for merged and competing hospitals grow. Sloan et al. (2003) analyzed ownership changes in 5,089 hospitals from 1986 to 1996, confirming that difficult financial times predicted likely ownership conversions, closures, and mergers. These transactions most likely occurred in markets where a merger would lead to market power for the acquirer. Moreover, it may be policymakers generating the increased competition and, thus, their disruptive effects on local communities.

Dranove and Lindrooth (2003) posited that mergers and system consolidations have distinct and noteworthy differences. They specified that mergers generate savings over time but require a level of commitment wherein the merged organization gives up its operating license and usually its CEO. On the other hand, system consolidations did not lead to savings even after four years; nevertheless, hospitals within a system did not sacrifice their license and leadership (Dranove & Lindrooth, 2003). Touted potential business benefits of healthcare consolidation such as coordination of care, less fragmentation, costs reduction, economies of scale (Gaynor, 2018) deliver problematic results in the form of higher prices, lower quality, less innovation, and anti-competitive practices when placed under scrutiny.

Table 5

Mergers and Acquisitions and Healthcare Studies

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
1	Vita & Sacher (2001)	Analysis of U.S. not-for-profit hospital mergers to determine if cost objectives were successful.	Regressions of inpatient admission prices defined a market. Archived quarterly data from the Office of Statewide Health Planning and Development from 1986 to 1996 was used.	Competition among hospitals was reduced. Post-merger hospital prices increased, but reasons could not be conclusively stated. Mergers among not-for-profit hospitals are a cause for anti-trust concerns. Certain not-for-profit hospitals may derive different logic from for-profits to increase post-merger costs such as charity care. May help policymakers determine if enforcement decisions predict outcomes within acceptable degrees of accuracy.
2	Dranove & Lindrooth (2003)	Investigate if hospital consolidation leads to cost savings.	Modeled probability likelihood to merge using Probit, then matched merged hospitals to 10 closest hospitals that did not merge based on the propensity to merge score. Archival data came from the American Hospital Association Annual Survey of Hospitals 1988 – 2000.	System consolidations do not generate savings even after 4 years. Mergers generate savings likely due to reduced capacity and other synergies. Mergers are critical events requiring that relinquishment of licenses and usually a CEO, while system consolidations require less commitment in the form of license maintenance and CEO retention.
3	Sloan, Ostermann, & Conover (2003)	Investigate the causes of hospital ownership mergers, closures, and conversions.	Ownership change transactions of 5,089 hospitals from 1986 – 1996.	Difficult financial times predicted likely ownership conversions, mergers, and closures. High competition predicted large-scale change. Mergers were most likely in concentrated areas where a merger would increase market power. Policymakers may generate increases in competition whose disruptive effects are the closures or other hospitals' transitions in local communities.

Table 5 Continued

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
4	Kjekshus & Hagen (2007)	Analyze effects on technical and cost efficiency of 7 hospital mergers in Norway.	Efficiency scores generated using archival Data Envelopment Analysis (DEA) from 53 merged and non-merged hospitals from 1992 – 2000. The effect of the merger was estimated through panel data analysis.	Mergers show no significant effect on technical efficiency and show a significant negative effect on cost efficiency. Positive effects on cost and technical efficiency are shown in a merger where 5 hospitals were involved, and administration and acute services were centralized (p. 230).
5	Li-Ping Tang & Timmer (2008)	Examine the effects of organizational changes such as hospital name and CEO on hospital performance measures.	MANOVA of performance differences between rural and urban hospitals. Regression on payroll changes. ANOVA on patient admissions as a function of a hospital name change. Archived data from 1992 to 1995 of 155 hospitals came from the Tennessee Hospital Association.	M&As may lead to a name change, or CEO change, or both. Hospitals with name changes had higher patient to full-time-employee ratios and, therefore, better customer service. Payrolls increased while the number of beds decreased due to multiple causes, including higher overhead costs.

MI and M&As

The literature found overlapping between these two key terms required studies of M&A changes using MI as the methodology of analysis. The significance was the testing and bearing out of MI models—the maintenance of constructs over time, and the implications of change—when considering outcomes of M&As. The current study diverges from the literature in that the EES was not designed to measure the same respondents year over year. Instead, it measures the identical construct year over year, an assessment technique elaborated upon in Chapter III.

Five studies were chosen for the demonstration of MI analysis in M&As (see Table 6). The common ground of this research was the concentration on social identity theory, change, and behaviors of managers of M&As or integrations. It became evident that group identification, dominance, and subordination in group belongingness contributed significantly to longer-term M&A outcomes. Guerrero (2008) explained that threats related to the loss of one's identity are the key variables for understanding an employee's sense of stress (p. 232) in a study where insecurity scores show the presence of long-term stress created by trauma in the years following an M&A. In other words, employees found restructuring painful. Reactions in social identity from employees exposed to a hostile takeover were much lower than those exposed to an acquisition. After an M&A, organizational changes may create feelings of insecurity when changes imply identity discontinuity (Guerrero, 2008). Ingroup favoritism described by Gleibs, Noack, and Mummendey (2010) bespoke the tendency to favor one's group members over other groups related to negative attitudes towards mergers. Edwards, Lipponen,

Edwards, and Hakonen (2017) compared two different longitudinal studies using latent growth modeling for rates of change in reactions to an M&A in the merged organization, finding that employees in acquired organizations showed much lower levels of identification with the post-acquisition firm than those in the acquiring firm, but over time, employees from the acquired firm showed more positive linear increase in identity than to those from the acquiring firm.

Practicable suggestions from this literature highlight the need for managers of M&As and integrations to take notice of perceptions of group membership and to anchor internal communication to improve negative feelings. A suggestion for mixed work teams is made, allowing new employees to learn about each other. Those accountable for M&As must ensure that they are fair and just (Edwards et al., 2017).

Table 6

MI and Mergers & Acquisition Studies

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
1	Gleibs, Mummendey, & Noack (2008)	Post-merger identification changes during a merger. Resolve if poster merger identification can be predicted by pre-merger identification.	Questionnaire was given to 157 students from a newly formed university at 3 waves (4, 6, and 12 months after a merger). The research used multi-level modeling for change with post-merger identification serving as the outcome variable.	Pre-merger organizational identification is a predictor of post-merger identification. Members of dominant organizations perceive the merger as a continuation of the business, while the subordinate partner perceives it as a threat. There are implications for social identity theory and self-categorization theory emphasizing the importance of belonging to different social categories.
2	Guerrero (2008)	Employee reactions in the wake of an acquisition.	Probability sampling of 85 work sites belonging to three organizations over five years: an acquiring firm (Firm A), the firm that it acquired (Firm B, and a firm absorbed in a hostile takeover (Firm C) by Firm A.	Social identity, stress theory, and the importance of preparing for change were reinforced. Social identity scores are antecedents of feelings of insecurity (p. 232). Firm B employee reactions were more positive than those in Firm C.
3	Bartels, Pruyn, & Jong (2009)	Study of determinants of organizational ID.	Self-administered perceived external prestige (PEP) questionnaire to employees of 4 divisions of Dutch university merging into two divisions over two waves. MANOVA determined if T2 respondents differed from T2 non-respondents on variables from T1.	Pre-merger identification influences post-merger identification for employees at the same organizational level. Determinants of employee identification differ by division.

Table 6 Continued

No.	Source	Purpose	Method adopted for analyzing data	Key Findings
4	Gleibs, Noack, & Mummendey (2010)	Study of directional effects of identification and ingroup favoritism towards merger.	Cross-lagged regression analysis of online questionnaires sent to 211 students between the ages of 20 and 34 years old across two waves.	When social identity is under threat, identity is positively linked to ingroup favoritism. The study reinforces ingroup favoritism operating in a feedback loop. Pre-merger identification and T1 increased ingroup favoritism at T2. Ingroup favoritism is related to negative attitudes toward the merger.
5	Edwards, Lipponen, Edwards, & Hakonen (2017)	Comparison of 2 longitudinal studies of reactions to M&As for rates of change in employee organizational identity within the merged entity.	Latent growth modeling analyzed organizational identification change across 3 waves over two years from 3 Finish universities in two different samples. M&A1 = 938 employees M&A2 = 346 employees in a multinational acquisition	Trajectories of identification vary depending on the type of M&A and differ between employees and parties. In both samples, an increase in justice and a decrease in threat perceptions predicted an increase in identification across the post-M&A period. Larger groups show much higher identification with the merged entity than do smaller organizations.

Software Applications

This literature review cannot conclude without acknowledging that software programs are necessary to execute latent model calculations. Consequently, research in this field is bound to the progressive iteration of software built or modified to analyze measurement models. One of the critical functions is accurate estimates of parameters (Waller, 1993) because it is the software that handles "aberrant estimates" (p. 76). Initially, three software packages were most commonly utilized—AMOS (Bialosiewicz et al., 2013; Kline, 1998), EQS (Kline, 1998; Waller, 1993), and LISREL (Kline, 1998), which was employed in SEM by D. Sörbom when he introduced mean vector modeling in multigroup factor analysis in 1974. LISREL was the only program available (Kline, 1998; Meredith, 1993; van de Schoot et al., 2015). Initially, programming code was cumbersome, requiring advanced command of statistical procedures (Kline, 1998). Micro-computing, personal computers, and enlargements in storage and memory on more affordable computers have improved the programs and expanded the radius of researchers who can access them (Kline, 1998). Stata (Cox, 2005), Mplus ("Mplus—General Description," n.d.), R, and its lavaan package (Fox & Leanage, 2016) are widely used software programs that continue to transform SEM data execution. Liu et al. (2017) recommend lavaan, Mplus, and OpenMx statistical software packages for testing MI because of the available estimators.

Summary

Although there no single procedure for how best to formulate an integrative literature review framework, the aim is to generate new knowledge for HRD (Torraco,

2005). This approach's primary rationale is that the keywords inform about historical and current theory and practice better together than they do apart. One outcome of this approach is the funneling of volumes of research by overlapping key terms to analyze for synthesized contributions and meaning. The literature teaches that more empirical research, preferably with large datasets, will add credibility to MI and engagement theories. Secondly, the researchers show that leader behavior and ability have a considerable influence on M&As and employee engagement over short and long terms. Finally, the literature illustrates that the thorough development of validated measurement models is an intensive exercise of time and capacity, consequently leveraging existing volumes of data and scales to assess the reliability of a demonstration of resource management.

A question arising from these literature searches is: If the latent measurement model does not hold does the obligation and charge lie in the instrument design, with the respondents, the organization’s leadership, or the organizational changes? In employee engagement, single instances of measurement can be informative. Nevertheless, time intervals have become essential to understanding organizational phenomena (de Beurs et al., 2015). Academicians call for more rigorous scales and data, even if the broader surveying public does not.

CHAPTER III

METHODS*

This chapter describes the research design, subjects, data collection, and plan for analysis of the uninterpreted raw data from SystemTex. Chapter I provided a descriptive context and background; Chapter II was an integrative literature review of relevant subjects to position the study in historical and contemporary theoretical and practical contexts. Chapter III considers how the data is to be analyzed for measurement invariance (MI). Other investigative opportunities found in this data set may go unaddressed to prevent "scope creep," which is an unmanaged progressive swelling of project boundaries.

Research Design

The decision to undertake MI analysis of this data set was determined many years after The Advisory Board made its initial design decisions. Having the survey results available was a bit of good luck and timing. Because of my exposure to the information, I developed an interest in deeper employee engagement instrument reliability. Carlson and Herdman (2012) stated that a hallmark of good research is demonstrating the robustness of findings. The use of various protocols, samples, measures, and analysis elements assures that findings are not "idiosyncratic components of methodology" (p. 17).

---

*Reprinting of The Advisory Board data and resources was allowed with written permission (see Appendix D).

While there is certainly goodwill between commercial firms and their clients, the academic community demands a considerably different standard than commercial firms for rigorous testing of ideas. Commercial firms create and sell intellectual property that is facile for implementation and provides financially viable solutions and defendable results. However, those ideas are often based on untested methodologies and unconfirmed constructs (Zigarmi et al., 2009). In contrast, academic research requires an exhaustive study of constructs, theoretical frameworks, and methodologies when testing and applying concepts to ensure validity and replication of research. The extensive scholarly endeavors can be cumbersome, time-consuming, expensive, and perceived as out of touch with real-world business operations (Zigarmi et al., 2009).

Research Questions and Hypotheses

By having permission to the raw data acquired by The Advisory Board, careful academic testing can be applied, providing the potential for a meeting of both commercial and academic interests. When there is a mutual benefit between exacting academic standards and commercial viability, academics may enjoy access to difficult-to-acquire, meaningful organizational data and engage in theory making or building. Commercial firms may see where and how their products perform under additional scrupulous inquiry and offer change or improvement options. The purpose of the study is to explore the following questions

1. How reliable or valid is the EES Engagement construct?
2. What would the factor structure reveal about employee perspectives on engagement?

3. Would the 2013 M&A generate a response shift in engagement?

Two hypotheses were developed to answer the research questions.

H_1: The Engagement measurement model meets fit standards for multi-group CFA comparisons in 2011 and 2012.

Resolving this question would force assessment of construct validity and reliability.

H_2: The Engagement measurement model will display response shifts in multi-group CFA comparisons in 2013, 2014, and 2015.

Background of the Sample

The Employee Engagement Survey (EES) is a self-reporting instrument built from 46 variables or statements. The statements are all positively phrased and not stated as questions. There were five hospitals in SystemTex and the Shared Services group in 2011. Shared Services included HR, Finance, Accounting, Payroll, IT, Medical Billing, and Quality. From 2011 to 2015, additional SystemTex hospitals and business units were surveyed, but only the same six facilities were assessed for the present study to maintain continuity of group comparison. Examining collected demographic information such as the participant's job title, work location, department, unit, and the direct supervisor is not a part of this project. The collection of demographic data points, e.g., gender, age, ethnicity, and education level, did not occur.

Respondent participation. The employee population eligible to take the survey included all full-time, part-time, and PRN (pro re nata) SystemTex employees hired by the month preceding the survey launch. PRN staff are employed on an as-needed basis to accommodate unexpected staffing changes. Two examples include a sudden rise in

hospital census and when absences prevent scheduled staff from working their shifts. Contract employees were not eligible to take the survey. Table 7 shows that respondent participation clustered into their primary work location, where n = the number of respondents and % is the percentage of the facility population.

Figure 6 shows participation across the surveyed years. High participation rates are attributed to the importance placed on the engagement survey by senior organizational leadership.

Table 7

Respondent Participation by Facility

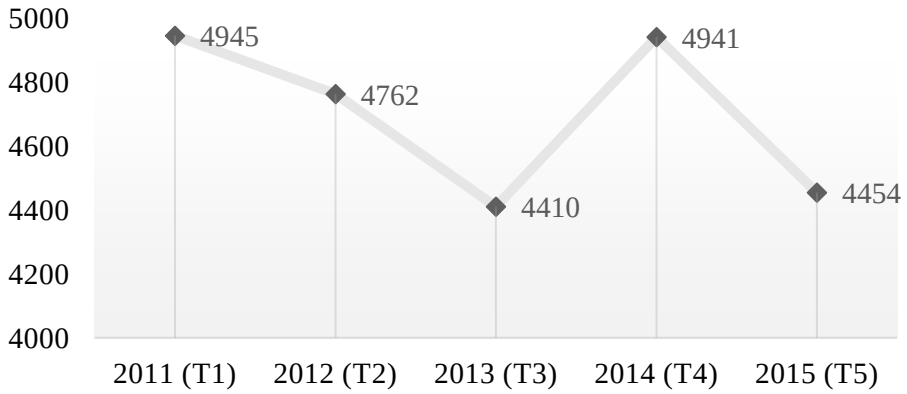
Facility	2011		2012		2013		2014		2015	
	n	%	n	%	n	%	n	%	n	%
Hospital 1	133	100	140	100	127	100	128	89.5	119	90.5
Hospital 2	271	97	329	79	388	93	315	81.4	284	76.9
Hospital 3	681	88	621	75	686	80	725	77.2	714	75.9
Hospital 4	377	82	321	91	351	97	328	90.6	309	80.9
Hospital 5	3,022	77	2,890	72	2,416	62	2,900	74.3	2,497	64.1
Shared Services	461	96	461	87	442	82	545	94.8	531	88.7

Note. n = number of participants and % = percentage of the employee population participating in the survey.

Sample size. Survey sample sizes (see Figure 6) affect the statistical power and parameter estimates (Brown, 2015). Statistical power is one minus the probability of making a Type II error (failing to reject a null hypothesis) and pertains to parameter estimates and the ability to detect model misspecifications. Sample sizes over 1000 observations are considered excellent for mitigating Type I and II errors (Meyers et al., 2016).

Figure 6

SystemTex Respondent Participation by Year



Survey Completion

The OD department managed the internal engagement cycle for SystemTex, including communication plans and technical issues with IT and The Advisory Board. Each engagement cycle, the OD team initiated an organization-wide communication campaign in partnership with The Advisory Board. Employees were given three weeks to complete the survey. Once the survey closed, The Advisory Board conducted a multivariate regression analysis of the 4-statement Engagement index and 42 predictors (Strumwasser &Virkstis, 2015).

Data Verification and Coding

SystemTex team provided The Advisory Board with the employee and organizational information to set up the survey coding. Between 2011 and 2013, the survey was anonymous. The coded information included the facility name, department, workgroup, benchmarking tags, unit names, and a count of employees working in each unit and for each manager. Before launching the survey, the SystemTex OD department

initiated requests for information of the unit/department managers to be gathered and verified by the most direct leader. In instances where employees were mapped incorrectly, it was the unit/department manager's responsibility to make the OD department aware of the need for correction. Updates of employee information were provided back to the OD department, then shared with The Advisory Board ensuring that the survey reporting structure matched the organization.

From 2011 through 2013, the survey was made available to eligible staff via an electronic link posted on the company intranet. The concern brought forward by the OD team and several managers with this delivery method was that there was no control of the number of survey completions by any given employee. As a result, multiple completions could be an unintended outcome of this delivery practice. The Advisory Board responded to the worry with guidance that they observed minimal deception in completion rates as a byproduct of the delivery design. As an employee, I recall that the survey could take 10 to 15 minutes to complete depending on how thoroughly one read and considered each statement. Though the survey was important, there was little personal or professional motivation or benefit to take the survey multiple times in one year.

In 2014 and 2015, the survey design changed from anonymous to confidential. Managers were still asked to verify that employees were correctly mapped to their department. The SystemTex data file provided to The Advisory Board included the employee’s email address, and The Advisory Board generated a custom unique identifier for each respondent. When unique identifiers were required for survey administration,

the practice eliminated the opportunity for multiple survey completions by any one employee and simultaneously protected employee confidentiality since completion required accessing the survey with the unique identifier.

Data storage. Following the survey's close, The Advisory Board compiled and analyzed the data, then it was hosted on The Advisory Board's proprietary servers. Per the contract, the raw data was not shared with SystemTex during the survey years as a precaution and best practice to prevent SystemTex from identifying individual responses. When approached for permission in 2015, The Advisory Board agreed to provide the raw data. The data, stored as Excel files, was emailed directly to me from The Advisory Board and initially stored on a SystemTex network drive. The files were later transferred to an external hard drive and maintained on my Texas A&M Google drive as an additional precaution against data loss.

Communication

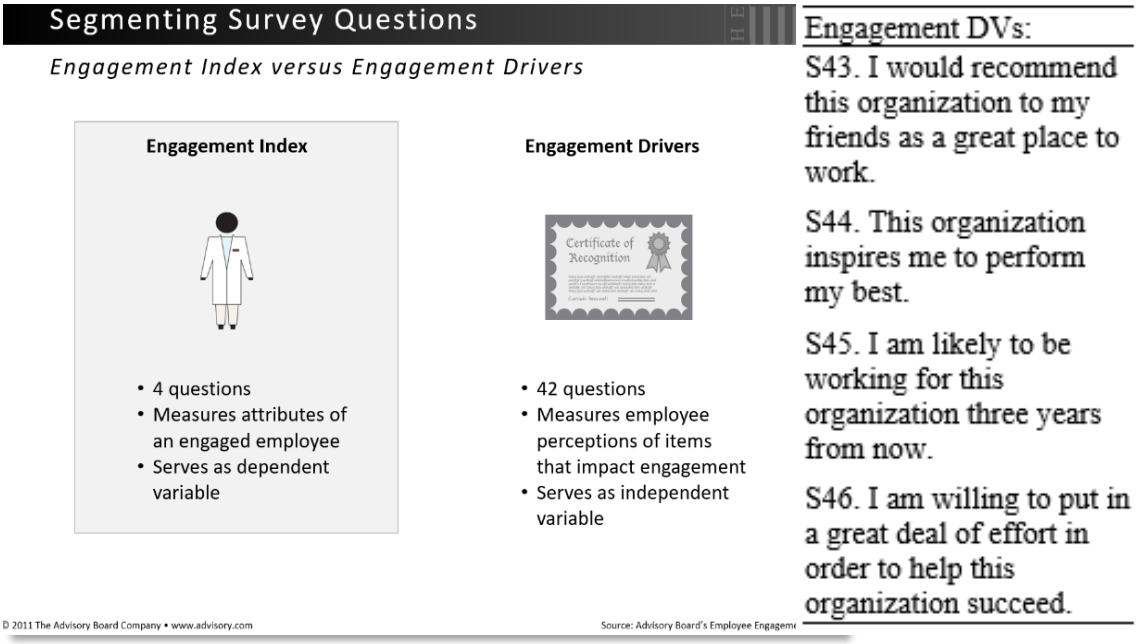
Before and during the survey, system-level and hospital-level leaders communicated messages to employees at each facility via email and meetings. Print and other electronic collateral advertised the arrival of the survey cycle and encouragement for participation. Employees were made aware of the surveys' timing, connection to organizational values and goals, and provided contacts for questions that might arise. Survey promotion was enhanced through marketing on the intranet, wearable badges, writing utensils, and posted on unit/department communication boards.

Constructs and Driver Categories

Figure 7 shows the prescribed Engagement construct. The graphic comes directly from an Advisory Board presentation made for SystemTex. Four items served as the dependent variable, while 42 items measured employee perceptions impacting engagement and served as independent variables (The Advisory Board Survey Solutions, 2011).

Figure 7

Prescribed Advisory Board Engagement Construct



Note. This above image and survey statements were was used with the written permission from The Advisory Board. See Appendix D.

In addition to the Engagement construct, two additional constructs, Change Readiness and Manager Effectiveness, are posited to be housed within the EES. The predictor variables are organized into eight driver categories, but not true subscales. Several variables are shared across multiple categories shown in Figure 8 with *(R)*

designating items repeated across the categories. Frequently called “driver” statements, the predictors function as levers the organization might use to establish a more engaging culture (Rivera, Fitzpatrick & Boyle, 2011). *Drivers* is a term employed by practitioners and scholars, including Popli and Rizvi (2016) and Rivera et al. (2011) when referring to observed or unobserved factors with the potential to influence another variable. Practitioner-built instruments may rely on “drivers” to replace predictor variables. The language of “drivers” helps commercial research because it is easily understood and translated without the often-confounding terminology of independent, dependent, mediating, and moderating variables.

Figure 8

EES Statements Listed by Driver Categories

Communication and Input

I am kept informed of the organization's future plans and direction.
My ideas and suggestions are valued by my organization.
My manager communicates messages that my coworkers need to hear, even when the information is unpleasant (R).
My manager is open and responsive to staff input (R).
My manager stands up for the interests of my unit/department (R).

Employee Support

I have a manageable workload.
My manager helps me balance my job and personal life (R).
My organization does a good job of selecting and implementing new technologies to support my work.
My organization helps me deal with stress and burnout.
My organization supplies me with the equipment I need.
My unit/department has enough staff.

Feedback and Recognition

Executives at my organization respect the contributions of my unit/department.
I have helpful discussion with my manager about my career.
I know what is required to perform well in my job.
I receive regular feedback from my manager on my performance (R).
My organization recognized employees for excellent work.

Professional Growth

I am interested in promotion opportunities in my unit/department.
I have the right amount of independence in my work.
I receive effective on the job training.
My current job is a good match for my skills.
My manager helps me explore other jobs within my organization.
My manager helps me learn new skills (R).
My most recent performance review helped me to improve.
Training and development opportunities within my organization have helped me to improve.

Manager Effectiveness

I have helpful discussions with my manager about my career.
I receive regular feedback from my manager on my performance (R).
My manager communicates messages that my coworkers need to hear, even when the information is unpleasant (R).
My manager helps me balance my job and personal life (R).
My manager helps me learn new skills (R).
My manager is open and responsive to staff input (R).
My manager stands up for the interests of my unit/department (R).

Mission and Values

I believe in my organization's mission.
I understand how my daily work contributes to the organization's mission.
My organization gives back to the community.
My organization provides excellent care to patients.
My organization provides excellent customer service to patients.
Over the past year I have never been asked to do something that compromises my values.
The actions of executives in my organization reflect our mission and values.

Teamwork

Abusive behavior is not tolerated at my organization.
Conflicts are resolved fairly in my unit/department.
I have good personal relationships with coworkers in my unit/department.
I receive the necessary support from employees in my unit/department to help me succeed in my work.
I receive the necessary support from employees in other units/departments to help me succeed in my work.
My coworkers do a good job.

Baseline Satisfiers

I have job security.
My organization pays me fairly for my job.
My organization supports employee safety.
My organization understands and respects differences among employees.
The benefits provided by my organization meet my needs.

Note. The above survey statements were used with written permission from The Advisory Board. See Appendix D. Statements used across multiple subscales are denoted with (R) for Repeated.

Data Analysis Plan

SEM focuses on latent variables built by concepts, allowing measured indicators to represent or estimate scores of underlying constructs. Using the constructs, sense is made of ideas, attitudes, and beliefs about the world (Little 2013). The analysis strategy for this study begins with the instrument's Engagement construct's reliability and validity through EFA and CFA and concludes with the measurement of group invariance of Engagement. Without MI, it is not recommended to make strong statements about a change in the construct due to some exogenous influences (Vandenberg & Lance, 2000).

EFA and CFA factor analysis techniques determine the number of latent constructs accounting for the variance/covariance of measured variables (Espinosa, 2016). As mentioned before, the survey potentially houses two additional constructs—Change Readiness and Manager Effectiveness—comprised of statements from the broader survey. Yet, how The Advisory Board determined the indices is not stated as plainly as the Engagement factor. Where The Advisory Board situates all 42 drivers as predictors of Engagement, the Change Readiness and Manager Effectiveness indices have no stated indications as to their design, to their relationship to Engagement, or each other. Since all of the variables that comprise the Change Readiness and Manager Effectiveness indices have already been designated as predictor variables, it would be easiest to assume that the constructs, too, predict the latent Engagement factor.

As with all the data, the three indices— Engagement, Change Readiness, and Manager Effectiveness—could be accessed through the online portal connected to The

Advisory Board databank site made available to SystemTex through a user interface. Through this interface, all the survey information, including indices, could be filtered, aggregated, and downloaded. The Change Readiness index was comprised of 10 items; the Manager Effectiveness index was comprised of nine. Both constructs shared statements 4, 6, 14, 15, and 28. Table 8 shows the two constructs with items shared across both scales in italics. The two columns to the far right identify the two prescribed factors, and the numbers in those two columns tally the items belonging to the factor. All of the items contained in Table 8 are used with written permission from The Advisory Board (see Appendix D).

Table 8

Speculated EES Change Readiness and Manager Effectiveness Factors

Variables		Change Readiness	Manager Effectiveness
S1	I am kept informed of the organization's future plans and direction.	1	
S4	*My manager communicates messages that my coworkers need to hear, even when the information is unpleasant.*	*2*	*1*
S5	My ideas and suggestions are valued by my organization.	3	
S6	*My manager is open and responsive to staff input.*	*4*	*2*
S12	The actions of executives in my organization reflect our mission and values.	5	
S14	*My manager stands up for the interests of my unit/department.*	*6*	*3*
S15	*Conflicts are resolved fairly in my unit/department.*	*7*	*4*
S18	I have helpful discussions with my manager about my career.		5
S19	My manager helps me learn new skills.		6
S20	My most recent performance review helped me to improve.		7
S24	I receive regular feedback from my manager on my performance.		8
S28	*My organization helps me deal with stress and burnout.*	*8*	*9*
S30	Executives at my organization respect the contributions of my unit/department.	9	
S34	I have a manageable workload.	10	

Content and Construct Validity

Efforts to attain reliability and validity information regarding the EES from The Advisory Board were made via email as recently as June 2020. An Advisory Board representative informed me that the company exited the engagement survey business in the years since SystemTex was a client. Validity designates that an instrument precisely gauges its intended concepts, while reliability ensures consistent measurement precision and minimizes error (Carlson & Herdman, 2012; Heale & Twycross, 2015). It would be ideal to know if the EES (a) accurately measured all aspects of organizational engagement and (b) whether or not the instrument measured organizational engagement compared to some other construct, theme, or idea before proceeding with MI.

Construct validity may be ascertained through the convergence of the EES to other engagement surveys; however, this process requires a deep level of statistical analysis across numerous instruments and is best left for another time or scholar. No single test can produce construct validity, so it is difficult to obtain (Carlson & Herdman, 2012). Instruments obtain construct validity through homogeneity, theory-based evidence, or convergence (Heale & Twycross, 2015). Homogeneity would indicate that an instrument effectively measures only one construct (Heale & Twycross, 2015), but the EES is purported to comprise at least three constructs. Theoretical-based evidence of an instrument is present when participant "behavior is similar to the theoretical propositions of the construct" (Heale & Twycross, 2015, p. 66). Showing that a model is meaningful based on prior investigation and theory usually occurs before model evaluation (Brown, 2015). Since that information was not available, that could have been

a significant enough reason to halt the procedure, but I chose to press on. The EES was developed in 2006, employing qualitative and quantitative methods, academic literature, survey assessments, and industry experts (Strumwasser & Virkstis, 2015). Furthermore, the authors reported that survey validation occurred through an initial group of participants (p. 179), which could indicate face validity, though additional evidence was still required for rigorous credibility.

Model Estimation

Classical estimation methods in SEM assume that observed variables are continuous (Rhemtulla, Brosseau-Liard, & Savalei, 2012). However, when using Likert scales (strongly agree, disagree, neutral, agree, strongly agree), the responses are coded numerically in ascending order, making them ordered-categorical variables, not continuous. Ignoring the variables' categorical nature contributes to biased parameter estimates, incorrect standard errors, and model test statistics (Rhemtulla et al., 2012). These errors arise due to continuous CFA models applied to ordinal variables, which is corrected when ordinal variables are treated as such (Rhemtulla et al., 2012). Ordered-categorical indicators would normally not use maximum likelihood (ML) as an estimator in CFA models (Brown, 2015, p. 353). However, ML can be used and the data treated as continuous if the model can recreate the variance/covariance matrix to be very similar to the data as determined by the goodness of fit statistics (Meyer, 2020). ML rests on an assumption of adequate sample size and multivariate normality (Brown, 2015). Moreover, if the data are suitable for ML, a variance-covariance matrix may be used as input data for the CFA (Brown, 2015, p. 92).

Goodness of fit. There are three model fit classes used across EFA, CFA, and MI—absolute, parsimony, and comparative—that each provides different information about fit solutions (Brown, 2015). The fit indicates how well a model represents the data. It is shown in how closely the observed data from the covariance matrices match the relationships identified in the hypothesized model (Meyer, 2020). When a model is well fit, there is no statistical difference between the model and sample variance/covariance matrix (Meyer, 2020). The indices evaluate model acceptability. Good fitting models show support that a model is properly specified, while poor-fitting models may require diagnosing model misspecification (Brown, 2015). Sometimes, conflicting information about model fit shows that caution is required when determining a solution (Brown, 2015). Model fit will first be assessed through all three classes starting with chi-square ($\chi 2$) and then through the AIC, BIC, RMSEA, CFI, TLI, and SRMR indices. Detailed descriptions of the indices follow in Chapter IV.

ML EFA. EFAs are often conducted before a CFA to refine the measurement model (Brown, 2015). EFA and CFA are based on the common factor model and often rely on the same estimation methods, but EFA is an exploratory technique due to freely estimated parameters (Brown, 2015). This test validated the presence of any constructs, not just those prescribed by The Advisory Board. EFA is urged when there is no hypothesized model to substantiate or when there is some doubt of its validity (Meyers et al., 2016). This technique divides the variance of each indicator into common and unique variance (Espinosa, 2016). Shared variance refers to the common variance among factors, while unique variance is specific to an indicator; therefore, not accounted for by

common variance (Espinosa, 2016). Unique variance can be due to the indicator's uniqueness and random error (Espinosa, 2016). The EFA will produce a rotated factor solution to maximize the magnitude of primary loadings (Brown, 2015) to evaluate the model.

ML CFA. Conducing a CFA to test if items grouping under latent factors are good indicators of the construct (Gallant & Martins, 2018) will further validate the factors found within the measurement structure. CFAs often analyze a non-standardized variance/covariance matrix, and indicators are specified to load only on one factor, and unlike EFAs, measurement error can be correlated (Brown, 2015).

Tests for MI

Assuming that an Engagement construct converges in the EFA and is confirmed through the CFA, the final step will determine if the Engagement construct is invariant. Vandenberg and Lance (2000) published a highly cited summation of MI suggestions, practices, and recommendations. In it, they remind us that the invariance assumptions can render comparisons dubious. Attempting to liken measures which have a particular meaning for one group to another for whom the measure means something different "may be tantamount to comparing apples and spark plugs" (Vandenberg and Lance, 2000, p. 9). Thus, MI is important for establishing group likeness and differences. *Groups* refer to independent groups, not longitudinal ones, though the survey contains sequential survey administration years. The smallest unit of measurement within the SystemTex dataset is the individual level, but respondents were not tracked from one administration to the next. Individual responses are aggregated to the group level by the

facility at which the respondent works. As cited by Vandenberg and Lance (2000), the equation below is a simplified equation of testable aspects within MI. Measurement of the relationship between *k* items in the *g*th group is characterized as

$$X^g_k = \tau^g_k + \Lambda^g_k \xi^g + \delta^g_k$$

X^g_k is the vector of items of the latent response, Λ^g_k is the matrix of (factor loadings) regression slops relating the latent response (X^g_k) to the common factor score (ξ^g), τ^g_k is the vector of regression intercepts, and δ^g_k is the vector of unique factors (Vandenberg & Lance, 2000, p. 10).

Assessing factorial invariance. In testing unobservable constructs like engagement, observable indicators serve as measures of the construct; therefore, a comparison of constructs across groups necessitates that each indicator is related to the construct across all groups (Milfont & Fischer, 2010). MI is a series of nested evaluations of model fit conducted through multi-group confirmatory factor analyses (MGCFA) (Horn & McArdle, 1992; Milfont & Fischer, 2010). The MI tests sequence has become a recommended practice in a CFA framework (Vandenberg & Lance, 2000). Each level of group analysis is more restrictive than the previous one. The invariance test sequence assesses the variances, covariances, and means of latent variables between groups (Meyer, 2020; Milfont & Fischer, 2010). Each model is a pair-wise, cross-group comparison tested for invariance. Because the step-wise tests establish the presence or lack of equivalence between means, slopes, intercepts, and unique factors, MI is interchangeably referred to as measurement equivalence. However, it should be noted

that separate and distinct processes are effectuated for equivalence testing (Yuan & Chan, 2016).

Configural invariance. ($\xi^{g} = \xi^{g'}$) First comes the test for equivalence of the underlying construct (the latent ξ) for each group. This test substantiates whether the sample covariance matrix in a group can be fitted by the same factor model (Yuan & Chan, 2016). In configural invariance, regression slopes (Λ), intercepts (τ) and unique factors (δ) are freely estimated. This model must be tested, regardless of model fit, as it is considered a baseline (Milfont & Fischer, 2010). Acceptance of this hypothesis is a demonstration that the same number of latent variables (ξ) with the same patterns of factor loadings, intercepts, and unique variances underlie the indicator set (Horn & McArdle, 1992; Liu et al., 2017; Vandenberg & Lance, 2000).

Metric invariance. ($\Lambda^{g}_{k} = \Lambda^{g'}_{k'}$) If configural invariance is established, metric invariance is tested through the analysis of cross-group factors (Yuan & Chan, 2016). Metric invariance tests equivalence of factor loadings or slopes (Λ), with the observed response (X) and ξ for each group. The level is sometimes referred to as weak factorial or pattern invariance (Millsap, 2012). It builds from configural invariance by requiring that factor loadings are equivalently constrained to be the same across groups (Horn & McArdle, 1992). Slopes and means of the latent variable are constrained to 0; intercepts and residuals are free to change. Factor loadings reflect the degree to which differences among participants' item responses arise from differences among their levels of the underlying construct being assessed by that item (Bialosiewicz et al., 2013). Each item's contribution to the factor remains constant across groups showing identical relationships

between the construct and responses to observed variables measuring the construct (Assunção et al., 2020).

***Scalar invariance*.** ($\tau^{g}_{k} = \tau^{g'}_{k'}$) Sometimes referred to as intercept, or strong factorial invariance, scalar invariance builds on metric invariance through comparison of group means (Assunção et al., 2020; Milfont & Fischer, 2010). Scalar tests the equivalence of intercepts (τ) and loadings (Λ), for each group (Vandenberg & Lance, 2000) and is assumed when invariance between intercepts and factor loadings is present (Bialosiewicz et al., 2013). Intercepts are constrained to be the same across groups (Chakraborty, 2017; Milfont & Fischer, 2010), while unique factors and the latent factor's mean are free to change. This test surfaces important, though unseen information cannot be reacted to since it is unobserved (Bialosiewicz et al., 2013). More about this will be discussed in Chapter V.

***Strict invariance*.** ($\delta^{g}_{k} = \delta^{g'}_{k'}$) This level tests the equivalence of unique factors as well as slopes (Λ), and intercepts (τ) for each group. Sometimes called invariant uniqueness (Vandenberg & Lance, 2000) or error variance (Milfont & Fischer, 2010), this test is considered to be overly restrictive by many scholars as it tests whether residual error is equivalent between groups or administrations (Assunção et al., 2020; Milfont & Fischer, 2010). It is not often achieved in practice and will not be pursued in the present study.

Full and partial invariance. Vandenberg and Lance (2000) made distinctions between full and partial invariance. Establishing full invariance may not hold in practice; therefore, partial invariance, wherein a "subset of parameters in a model is constrained to

be invariant while another subset of parameters is allowed to vary" (Milfont & Fischer, 2010, p. 117), may be considered. Partial invariance may be exhibited when measures are invariant in some groups or when some measures are invariant in all groups (Chakraborty, 2017). There are no requirements for the implementation of partial invariance.

Chi-square Difference Test

The chi-square difference ($\Delta\chi^2$) test assesses factorial invariance between groups, identifying instances of invariance between the models. Each level compares the previous test of invariance with the next by using the $\Delta\chi^2$ test to tell how good the model is. The next more restrictive test is made if the current model is not significant ($p > .05$) (Yuan & Chan, 2016). Each model begins with a base model, which is compared to the nested model. When the base model is correctly specified, a significant $\Delta\chi^2$ statistic indicates that the nested model is non-invariant; a non-significant result is due to invariance at the present level (Chakraborty, 2017; Yuan & Chan, 2016). Type 1 and Type II errors have been produced in mispecified base models (Yuan & Chan, 2016).

Constraining parameters. Parameters (loadings, intercepts, and error) are estimated using a *df*; when parameters are constrained, no *df* is used (Meyer, 2020). The decision to constrain a parameter is based on the change in the χ^2 statistic between the model with constrained parameters and the model without constrained parameters. Determining which indicators to constrain can be made by looking at the modification indices. When the $\Delta\chi^2$ is significant, then it is statistically ideal to use more *df* to

constrain the model; however, if the $\Delta\chi^2$ is not significant, there is no reason to constrain the model (Meyer, 2020).

CHAPTER IV

RESULTS*

The following analysis details the processes and results from the exploratory factor analysis (EFA), confirmatory factor analysis (CFA), measurement invariance (MI) tests. Finally, the chapter summary debriefs the results but leaves the discussion, repercussions, and implications of the final chapter. The reader should recall that the organizational Engagement construct is the factor of interest. That said, this chapter probes the instrument for the validity of the EES instrument thoroughly rather than assuming the prescribed structural model's veracity.

EFA

EFA's usefulness includes defining and determining the number of latent constructs used to explain the correlations among a set of observed variables (Gallant & Martins, 2018).

I first attempted to test the construct validity of the three posited factors of Engagement, Change Readiness, and Manager Effectiveness factors by conducting an EFA. Two rotation methods—promax, varimax—were attempted to observe any significant differences between an orthogonal or oblique rotation. Orthogonal rotations (varimax) assume uncorrelated factors, and oblique rotations (promax) assume correlated factors, both of which help attain simple structure (Giannoulis, 2008). Both rotation strategies

---

* Reprinting of The Advisory Board data and resources was allowed with written permission (see Appendix D).

were used since no information revealed if or how these two factors should be expected to correlate. The promax rotation result showed cross-loadings of an extremely high correlation of 0.90. A split-half analysis was the second approach taken with no discriminate validity demonstrated between Change Readiness and Manager Effectiveness.

Spit-Half Reliability & Model Fit Evaluation

Split-half analysis can increase an instrument's relative reliability when working with large sample sizes because it ensures that the model will not be overfit. Instrument reliability can be exhibited through internal consistency (Heale & Twycross, 2015), which the split-half analysis accomplished through strong correlations showing high reliability. The desired result was a model that fitted the Engagement construct data as verified through the chi-square statistic and goodness of fit indices. In the split-half, the 42 non-engagement items functioned as predictors of the Engagement construct, which is the model designed by The Advisory Board.

A split-half using varimax rotation was run for Group 1 (year 2011, n = 4,945) using ML estimation (Meyers et al., 2016). The analysis was conducted on 2,472 observations extracting seven factors as a precaution against assumptions of the dataset's posited factors. Items S43 through S46 supposedly comprised the Engagement index, but no apriori hypothesis was assumed for the split-half analysis. Thus, all variables and latent factors could correlate freely. ML factor extraction calculated the sample data to attempt to "directly estimate the population covariance matrix," which helps to yield

more replicable results" (Meyers et al., 2016, p. 424). The output of the EFA is the goodness of fit statistics, which is discussed below.

Missing data. 200 observations were missing from the split-half, just 8% of the sample size—considered a minimal amount of missing data (Little, 2013). It is believed that the data are missing at random, though it is not possible to certain. Missing values were not imputed. In the ML estimation process, missing data do not impact parameter estimates or standard errors because the software estimates a likelihood function for each individual based on the variables that are present so that all the available data are used (Newsom, 2020; Straatmann, Almquist, Oliveira, Rostila, & Lopes, 2018, p. 6).

Goodness-of-fit indices. Methodological approaches for fit evaluation are established on differences between the actual and hypothesized covariance and mean structures (Meyers et al., 2016). Chi-square (χ^2), a global fit test, aids in drawing conclusions about the population samples, offering evidence of the generalizability of findings (Millsap, 2012; Shuck, 2010). χ^2 analysis tests the null hypothesis for significant differences between the expected (hypothesized covariance matrix) and an observed result of a specified variable distribution (Shuck, 2010, p. 82). The degrees of freedom (df) indicate the number of data points used to calculate a statistic (Shuck, 2010). The *p-value* is the probability that the deviation of the observed matrix of data from the expected or calculated matrix is due to chance (Shuck, 2010). Deviation is due solely to chance 5% of the time or less ($\alpha = .05$). Preferably, the χ^2 is insignificant —the p-value is greater than .05 (Shuck, 2010)—in which case, other tests may not be conducted (Espinosa, 2016). Failing to reject the null hypothesis indicates a good fit.

(Bialosiewicz et al., 2013). Table 9 lists the χ^2 results. Better fitting models resulted in smaller residuals using the χ^2 statistic.

Table 9

Chi-square Results for Split-Half

Model	Number of Parameters	χ^2	df	p-value
1-factor	138	15860.374	989	0.0000
2-factor	183	9505.299	944	0.0000
3-factor	227	7749.240	900	0.0000
4-factor	270	5883.255	857	0.0000
5-factor	312	4875.272	815	0.0000
6-factor	353	4051.124	774	0.0000
7-factor	393	3250.853	734	0.0000

Table 10 shows comparisons of fit results against one another. Evaluating model fit is challenging because multiple options may be plausible where no one model is ideal or should be assumed for extraction into the population. In addition to the χ^2 statistic, model fit decreases as power increases; as loading variances increase, power increases; and as the number of strong indicators increases, power increases (Meyer, 2020).

Table 10

Split-Half Factor Comparison for Model Estimation

Models Compared	χ^2	df	p-value
1-factor against 2-factor	6355.075	45	0.0000
2-factor against 3-factor	1756.059	44	0.0000
3-factor against 4-factor	1865.984	43	0.0000
4-factor against 5-factor	1007.984	42	0.0000
5-factor against 6-factor	824.148	41	0.0000
6-factor against 7-factor	800.271	40	0.0000

Identifying a non-significant model with the most df, the most variance explained, and the fewest parameters is not always easy.

One of the weaknesses of χ^2 index associated with ML estimations is an assumption of normality (Millsap, 2012), making it prone to error, particularly with large sample sizes. Thus it is necessary for other indices to scrutinize model fit (Bialosiewicz et al., 2013). Identifying the best model is a blend of art and science because of the complexity of rightly assessing multiple elements. See Table 11 for additional goodness of fit index results.

AIC and BIC. AIC (Akaike's information criterion) and BIC (Bayesian information criterion) are relative measures estimating the distance between the proposed model and the data. They help determine which model minimizes relative distance; therefore, no fixed points for good fit exist (Millsap, 2012, p. 100).

RMSEA. RMSEA (root mean square error of approximation) is a parsimony correction indicating how well the hypothesized model fits the sample data. Parsimony, in this instance, refers to model misspecification (Little, 2013). RMSEA focuses on error in the fitted population (Σ_{0Xk}) as an approximation of the true covariance matrix (Σ_{Xk}) (Millsap, 2012). A lower coefficient is a better result (Bialosiewicz et al., 2013). It considers the *df* within a model and addresses misspecification by incorporating correction through a penalty for poor model parsimony (Rausch, 2009). It is also not sensitive to sample size and ranges from 0 to 1: an unacceptable fit is >.10, and good fit is ≤ .06 (Bialosiewicz et al., 2013; Little, 2013; Meyer, 2020).

CFI and TLI. These indices are incremental fix indices. CFI (Comparative fit index) compares the hypothesized model to a more restrictive baseline model (Millsap, 2012). It ranges between 0 and 1. A fit closer to 1 is better. The acceptable range is > .90 good fit ≥ .95 (Meyer, 2020). The TLI (Tucker Lewis index) penalizes models with freely estimated parameters that do not meaningfully improve the model fit (Brown, 2015). It is interpreted similar to CFI where values closer to one indicate better model fit, though because it is non-normed, values can fall outside of the range of zero to one (Brown, 2015).

SRMSR. SRMR (standardized root mean square residual**)** is a measure of absolute fit, ranging from 0 to 1. Smaller coefficients mean smaller differences between residuals. An acceptable range would be < .05, and closer to 0 is better (Bialosiewicz et al., 2013; Millsap, 2012).

Table 11

Goodness of Fit Indices for Model Estimation

Number of Factors	AIC BIC	RMSEA	CFI TLI	SRMR
1	268009.171 268811.335	0.078	0.835 0.828	0.047
2	261744.096 262807.835	0.061	0.905 0.896	0.029
3	260076.037 261395.539	0.055	0.924 0.913	0.025
4	258296.052 259865.504	0.049	0.944 0.933	0.021
5	257372.069 259185.657	0.045	0.955 0.943	0.018
6	256629.921 258681.833	0.041	0.964 0.951	0.016
7	255909.650 258194.073	0.037	0.972 0.961	0.015

Eigenvalues and scree plot. Eigenvalues (see Table 12) demonstrate how much of the variation of the original group of variables is accounted for by a specific factor. They are standardized to one and equal the total number of variables (Giannoulis, 2008). The varimax rotation attempted maximization of the variance of each factor, distributing the variance across all extracted factors, and giving each of the 46 factors a variance of one (UCLA Statistical Consulting Group, n.d.). The orthogonal rotation approach permitted rotated solutions to be less prone to sampling error though it had the drawback of forcing factors to be orthogonal to each other (Matsunaga, 2010). The digit sitting atop top of the line is the number of factors. The number below shows the number of units of variance accounted for by that factor; the first factor accounted for 24.863 units

of variance. Because the total variance was 46, dividing 24.863 by 46 gave the percentage of variance extracted by the first component, which was 54%; the second component explained 4.6%; the third explained 2.5%, and the fourth component explained 2.4% of the variance. Each successive factor explained smaller and smaller amounts of the total variance (UCLA Statistical Consulting Group, n.d.).

Table 12

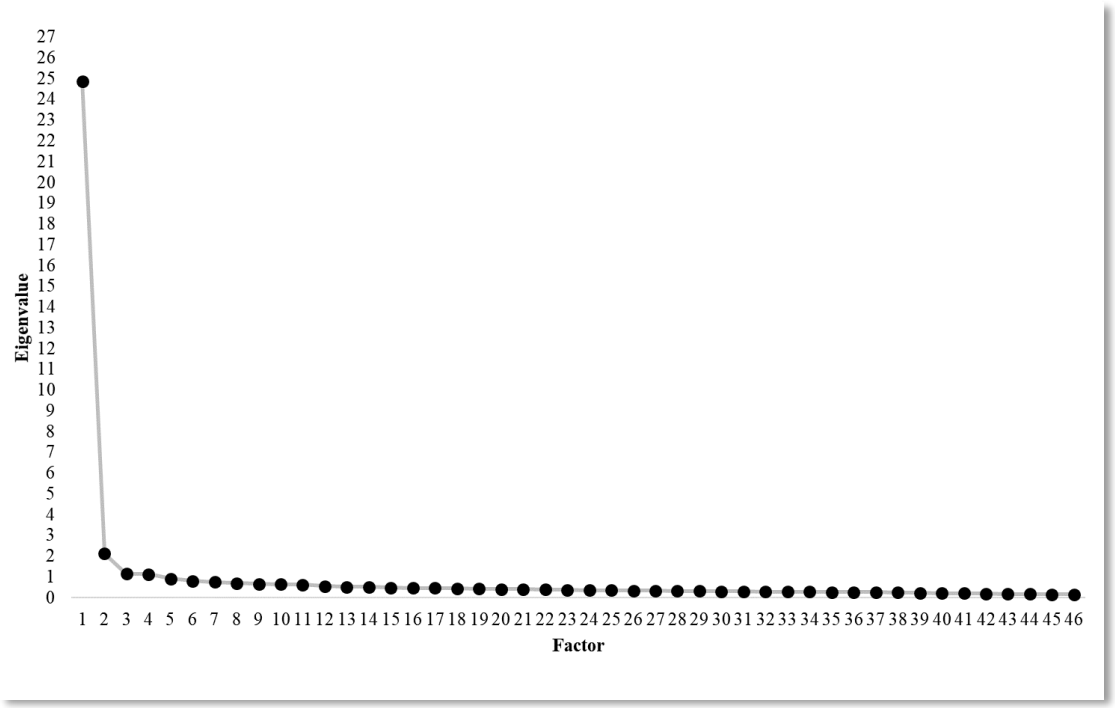
Eigenvalues for Split-Half ML EFA

1	2	3	4	5	
24.863	2.126	1.155	1.128	0.916	
6	7	8	9	10	
0.818	0.769	0.698	0.664	0.648	
11	12	13	14	15	
0.628	0.550	0.529	0.520	0.490	
16	17	18	19	20	
0.471	0.462	0.444	0.428	0.417	
21	22	23	24	25	
0.412	0.401	0.384	0.362	0.351	
26	27	28	29	30	
0.345	0.335	0.325	0.314	0.305	
31	32	33	34	35	
0.296	0.290	0.281	0.279	0.273	
36	37	38	39	40	
0.271	0.261	0.248	0.241	0.221	
41	42	43	44	45	46
0.208	0.198	0.185	0.175	0.163	0.152

The visual of eigenvalues, called a scree plot (see Figure 9), was first introduced by Raymond Cattell (Meyers et al., 2016). After the fourth factor, the line begins to flatten.

Figure 9

Scree Plot of Variables 1 Through 46



Eigenvalues show the "number of components, but not common factors than can but should not *necessarily* be extracted in the population" (Giannoulis, 2008, p.1). Because the fourth factor in the eigenvalues had a coefficient of 1.128, it suggested that a four-factor solution was plausible when the 42 variables are treated as predictors of the Engagement construct. Even so, it was not sufficient information to deduce that a four-factor solution was the best-fitting model. Relying heavily on eigenvalues greater than one is to be cautioned against as the statistic is a theoretical lower bound, which has been

found to be inaccurate (Giannoulis, 2008). There was enough covariance to account for four factors, but one was the clearly the dominant factor. Given that the first factor explained so much of the variance, it was concluded that the EES extracts one factor, and no more than one.

Rotated factor matrices. Varimax is an orthogonal rotation introduced by Kaiser in 1958 focused on the factors, capable of keeping factors independent of each other during rotation (Meyers et al., 2016, p. 429). This rotation functions by having some factors correlated very strongly, while others remain weakly correlated. The magnitude of structure coefficients (see Table 13) was helpful for further model-fit interpretation. Meyers et al. (2016) recommended using .40 as the lower bound when working larger sample sizes. The rotated factor structure matrix discloses correlations between the variables and factors (UCLA Statistical Consulting Group, n.d.). Meyers et al. (2016) suggested that coefficients below .5 may reduce scale reliability, especially when built from "a combination of close to a dozen variables" (p. 436). Since the instrument had 46 variables, .5 was chosen as the lower threshold for evaluating the factor matrix's pattern coefficients to preserve scale reliability best.

Table 13

Magnitude of Coefficients

Coefficient Range	
≥ .7	Excellent
.63	Very Good
.55	Good
.45	Fair

Meyers et al. (2016)

In the EFA, seven factors were extracted, and the variables of interest remained the Engagement index (S43 - S46) plus any other variables which might improve or further explain the Engagement construct. Other indicators belonging to the Engagement construct would be shown in how the factor loads in the rotated factor matrix.

Consideration of the χ^2, goodness of fit indices, eigenvalues, and factor loadings helped identify the best fitting model for the data. A varimax rotation with a 0.50 threshold was used to preserve scale reliability. Examining the rotated factor structure loadings using a 0.50 coefficient threshold indicated that a three-factor solution was equally plausible because it isolated the variables of interest with the highest factor loadings (see Table 14) and showed that Engagement index variables are statistically different from other items. Only the variables loading at or above .5 remain in the matrix. Variables loading at or below .499 were removed. In this table, the variables occupy the rows, and the factors occupy the columns. The choice not to round up did not make a difference in which variable loadings were removed.

Table 14

2, 3, and 4-Factor Rotated Matrix

2-Factor Solution		3-Factor Solution			4-Factor Solution			
1	2	1	2	3	1	2	3	4
0.615		0.597			0.546			
	0.673		0.609			0.614		
	0.598		0.56			0.581		
0.722		0.708			0.715			
0.73		0.708			0.674			
0.835		0.821			0.844			
0.533	0.54	0.51						
	0.529		0.561			0.556		
	0.52		0.573			0.562		
0.524	0.514	0.504	0.527					
	0.524							
0.571	0.573	0.547			0.505			
	0.503							
0.815		0.797			0.814			
0.74		0.723			0.711			
0.603		0.584			0.539			
0.852		0.839			0.82			
0.825		0.811			0.787			
0.706		0.688			0.658			
0.621		0.603			0.551			
0.772		0.759			0.747			
	0.52							
0.812		0.799			0.789			
0.518								
0.696		0.677			0.617			
0.682		0.661			0.617			
0.597	0.525	0.573			0.518			
	0.725		0.603			0.62		
	0.71		0.606			0.629		
	0.565		0.505					
0.508								
0.519		0.506						
	0.691		0.593			0.602		
0.513	0.535							
0.777		0.763			0.746			
0.552	0.557	0.531						
	0.652		0.562			0.545		
0.528	0.561	0.507	0.532			0.505		
	0.637			0.726				0.703
0.509	0.619			0.734				0.711
				0.569				0.559
	0.644			0.571				0.563

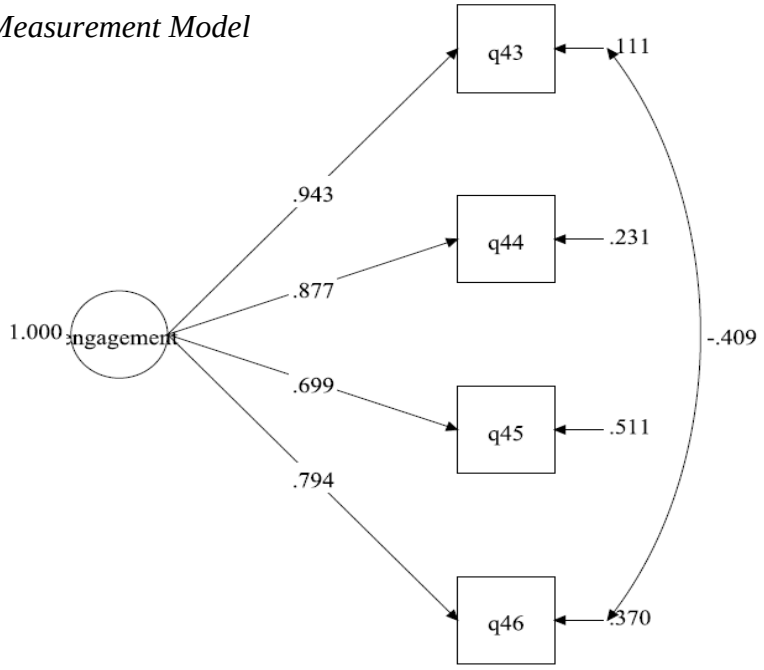
EFA Model Summary

The χ^2, factor rotation and the goodness of fit statistics made the two, three, or four-factor solutions plausible regarding the Engagement construct. All the items were highly correlated, as shown in the covariance matrix (see Appendix B), though none of the rotations clearly established a tie in any way with the prescribed Change Readiness and Manager Effectiveness constructs proposed by the EES. Ultimately, the presence of more than one factor was rejected due to the variance explained in one factor shown in the eigenvalues. Rejection of the four-factor solution, despite a fourth factor with an eigenvalues above one, occurred because it decreased all other items' factor loadings. Furthermore, though the Engagement index factor loadings for S43 through S46 were 0.703, 0.711, 0.559, 0.563, respectively, in the four-factor solution, none of the 42 other items loaded above .499. It would have been easy to accept the three-factor solution with the following model fit results: $\chi^2 = 7749.240$, $p < .05$; CFI = .924; TLI = .913; AIC = 260076.037; RMSEA = .055; and SRMR = .025.

The clearest information from the statistical analysis was that the rotated factor structure set out the Engagement construct distinctly, and along with the eigenvalues, it simultaneously indicated that it is the only construct determined with the data. Figure 10 shows the confirmed Engagement measurement model. Statements 43 (*I would recommend this organization to my friends as a great place to work*) and 46 (*I am willing to put in a great deal of effort in order to help this organization succeed*) have correlated error terms. It seems legitimate that these two indicators might further explain each other.

Figure 10

Verified Engagement Measurement Model



CFA

Cronbach's alpha for the four-item Engagement construct ($\alpha = 0.868$) established scale reliability (Brown, 2015). As a second measure of scale validity, randomization of the CFA was accomplished by drawing three random samples of 1000 respondents. Of the 4,945 observations, 2,522 observations were not in any of the random samples. 1,892 observations were in only one random sample; 485 were in two random samples, and 46 were found in all three random samples.

Stata software drew random numbers with its internal seeding mechanism. The randomization technique, called Monte Carlo randomization, or a Monte Carlo draw, is not the same as the Monte Carlo algorithm (Meyer, 2020). Each random number was generated using the day of the year and time of day down to the millisecond, resulting in

a sequence that used a normal distribution curve. The data set was then sorted based on the random number's value, and an identifier was generated for each observation. The first 1000 observations were used in the SEM, and the process was repeated. When drawing the second and third random samples from the population, previously included observations were not omitted to mean that each sample was not mutually exclusive from another sample, but each was randomized. P-values were calculated by comparing the observed statistic to the reference distribution (Meyer, 2020; “Monte Carlo Randomization,” n.d.).

Standard information across all three samples included:

- the number of sample groups: 3
- observations: 1000
- observed dependent variables: S43, S44, S45, S46
- independent variables: 0
- latent variables: 1
- estimator: ML

Goodness of fit statistics were suitable for all three samples (see Table 15), confirming the Engagement measurement model fit with the SystemTex data. Only one pair of errors could be correlated due only four indicators in construct (ten df available and nine df used).

Table 15

Goodness of Fit Statistics for Randomized CFA

	Sample 1	Sample 2	Sample 3
Missing Data Patterns	1	1	1
Free Parameters	13	13	13
χ^2 Model Fit	4.089	3.497	2.769
df	1	1	1
p-value	0.0432	0.0615	0.0961
AIC	9727.404	9889.828	9954.066
BIC	9791.205	9953.629	10017.866
RMSEA Estimate	0.056	0.050	0.042
CFI	0.999	0.999	0.999
TLI	0.993	0.993	0.995
SRMR	0.007	0.007	0.007

Measurement Invariance

Within-group measurement invariance by year are the first MI results. The tests' output show within-group invariance perspectives by testing the parameter estimates of two or more independent groups (Marsh & Byrne, 1993). The first perceptive is the analysis of invariance by year.

Within-Group Invariance

Table 16 shows the $\Delta\chi^2$ results for group invariance within the SystemTex organization by year. There were four models for each facility—2011-2012, 2012-2013, 2013-2014, and 2014-2015 ($\alpha = .05$). A non-significant result suggested no difference in the base and nested model. Invariant models are highlighted in grey for easier viewing.

- Hospital 1 was invariant across all waves, though it was only partially invariant from 2013-2014 due to not constraining S46.

- Hospital 2 was non-invariant 2013-2014; however, it was partially invariant 2012-2013 and 2014-2015 due to not constraining S45 and partially invariant 2014 – 2015 due to not constraining S46.
- Hospital 3 was invariant for the first three comparison models and non-invariant for the final comparison in 2014-2015.
- Hospital 4 was invariant in all comparison years but showed non-invariance in the metric-scalar comparisons in 2011-2012.
- Hospital 5 was non-invariant in the first and fourth comparison groups while being invariant in the third pairing.
- The Shared Services group was invariant in 2013 – 2014 and invariant in all other pairings.

Table 16

Within-Group Invariance by Year

Hospital 1													
Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar			
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig	
2011 – 2012	1.38	4	9.133	7	16.473	11	7.753	3	0.05	7.34	4	0.1190	
2012 – 2013	3.849	4	4.935	7	8.24	11	1.086	3	0.78	3.305	4	0.5081	
2013 – 2014	1.974	4	7.506	7	14.344	11	5.532	2	0.14	6.838	4	0.1447	Partial invariance due to not constraining S46
2014 – 2015	7.179	4	10.5	7	11.788	11	3.321	3	0.34	1.288	4	0.8634	
Hospital 2													
Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar			
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig	
2011 – 2012	3.91	2	4.795	5	44.941	9	0.885	3	0.83	40.146	4	0.0000	
2012 – 2013	5.089	2	5.446	5	12.158	9	0.357	3	0.95	6.712	4	0.1519	Partial invariance due to not constraining S45
2013 – 2014	8.978	2	19.281	5		9	10.303	3	0.02				
2014 – 2015	2.606	2	3.075	5	5.253	9	0.469	3	0.93	2.178	4	0.7031	Partial invariance due to not constraining S46
Hospital 3													
Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar			
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig	
2011 – 2012	5.892	2	6.215	5	13.36	9	0.323	3	0.96	7.145	4	0.1284	
2012 – 2013	7.779	2	9.068	5	15.973	9	1.289	3	0.73	6.905	4	0.1410	
2013 – 2014	13.376	2	15.862	5	20.68	9	2.486	3	0.48	4.818	4	0.3065	
2014 – 2015	6.874	2	20.242	5	80.344	9	13.368	3	0.00	60.102	4	0.0000	

Note. Models where $\Delta\chi^2$ indicates measurement invariance are highlighted. Selection of the unconstrained variable was determined by modification indices.

Table 16 Continued

Hospital 4												
Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
2011 – 2012	3.459	2	10.273	5	27.768	9	6.814	3	0.08	17.495	4	0.0015
2012 – 2013	0.711	2	4.795	5	6.72	9	4.084	3	0.25	1.925	4	0.7496
2013 – 2014	1.501	2	6.895	5	14.559	9	5.394	3	0.15	7.664	4	0.1047
2014 – 2015	4.615	2	9.939	5	13.761	9	5.324	3	0.15	3.822	4	0.4306

Hospital 5												
Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
2011 – 2012	12.288	2	23.465	5	29.715	9	11.177	3	0.01	6.25	4	0.1812
2012 – 2013	14.677	2	19.58	5	35.883	9	4.903	3	0.18	16.303	4	0.0026
2013 – 2014	39.114	2	41.383	5	45.523	9	2.269	3	0.52	4.14	4	0.3874
2014 – 2015	65.563	2	75.01	5	122.454	9	9.447	3	0.02	47.444	4	0.0000

Shared Services												
Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
2011 – 2012	25.33	2	35.433	5	70.954	9	10.103	3	0.02	35.521	4	0.0000
2012 – 2013	15.982	2	27.777	5	46.105	9	11.795	3	0.01	18.328	4	0.0011
2013 – 2014	3.916	2	5.487	5	10.858	9	1.571	3	0.67	5.371	4	0.2513
2014 – 2015	12.201	2	20.113	5	26.784	9	7.912	3	0.048			

Note. Models where $\Delta\chi^2$ indicates measurement invariance are highlighted. Selection of the unconstrained variable was determined by modification indices.

Between-Group Invariance

The second analysis of invariance was by facility. First, configural invariance testing occurred individually for the facility. The findings for configural invariance varied widely from very good to adequate fit, impacting the model comparison results. Eleven models were found to be "completely" group invariant in the facility-based group analysis across EES's five years. "Complete" invariance means that the cross-pairing was invariant throughout all MI tests. Table 17 shows the goodness of fit statistics for facility pairings that were completely invariant throughout. As with the individual facility assessments for configural invariance, the goodness of fit statistics for the pair-wise tests varied greatly from very good fits to poor but acceptable fit models. Appendix C shows the configural invariance that must first be established with each pairing before metric and scalar invariance models can be tested.

Tables 18 through 22 show between-facility comparisons as models for group invariance. This view pinpoints the areas of model differences by location. As with the time-based invariance models, p-values greater than .05 indicated a non-significant result and suggested no difference in the base and nested model. Invariant models are highlighted in grey for easy viewing.

Table 17

Goodness of Fit Statistics for Complete Group Invariance

Year	Model	χ^2	Df	Sig	CFI/ TLI	RMSEA/ 90% CI	SRMR
2012	Hosp. 5 – Hosp. 3	1.369	1	0.0000	1.000 1.000	0.010 0.000	0.002
	Shared Services – Hosp. 4	10.664	2	0.0048	0.995 0.985	0.074 0.035	0.015
	Hosp. 4 – Hosp. 3	7.209	2	0.0272	0.998 0.993	0.053 0.015	0.007
2013	Hosp. 2 – Hosp. 4	0.894	1	0.3445	1.000 1.000	0.000 0.000	0.004
2014	Hosp. 5 – Hosp. 2	23.624	1	0.000	0.997 0.981	0.084 0.057	0.011
	Hosp. 5 – Hosp. 4	24.751	1	0.000	0.997 0.980	0.086 0.059	0.011
	Hosp. 2 – Hosp. 4	0.604	1	0.4371	1.000 1.000	0.000 0.000	0.003
2015	Shared Services – Hosp. 2	12.720	1	0.0004	0.993 0.960	0.120 0.067	0.015
	Shared Services – Hosp. 4	8.022	1	0.0046	0.996 0.978	0.091 0.041	0.011
	Hosp. 2 – Hosp. 4	1.877	1	0.1707	0.999 0.996	0.038 0.000	0.007
	Hosp. 2 – Hosp 3	0.562	1	0.4536	1.000 1.000	0.000 0.000	0.003

Table 18

2011 Between-Group $\Delta\chi^2$ Test

Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
Shared Services - Hosp. 5	14.983	2	26.887	5	46.538	9	11.904	3	0.01	19.651	4	0.0006
Shared Services - Hosp. 2	3.799	2	24.899	5			21.1	3	0.00			
Shared Services - Hosp. 1	3.375	2	10.005	5			6.63	3	0.08			
Shared Services - Hosp. 4	7.319	2	22.262	5			14.943	3	0.00			
Shared Services - Hosp. 3	7.319	2	31.264	5			23.945	3	0.00			
Hosp. 5 - Hosp. 2	12.105	2	79.928	5			67.823	3	0.00			
Hosp. 5 - Hosp. 1	11.68	2	33.898	5			22.218	3	0.00			
Hosp. 5 - Hosp. 4	15.624	2	50.296	5			34.672	3	0.00			
Hosp. 5 - Hosp. 3	32.224	2	48.278	5			16.054	3	0.00			
Hosp. 2 - Hosp. 1	0.497	2	17.531	5			17.034	3	0.00			
Hosp. 2 - Hosp. 4	4.441	2	6.822	5	43	9	2.381	3	0.50	36.178	4	0.0000
Hosp. 2 - Hosp. 3	21.041	2	44.505	5			23.464	3	0.00			
Hosp. 1 - Hosp. 4	4.017	2	19.572	5			15.555	3	0.00			
Hosp. 1 - Hosp. 3	20.617	2	30.439	5			9.822	3	0.02			
Hosp. 4 - Hosp. 3	24.56	2	35.976	5			11.416	3	0.01			

Note. Group comparisons where $\Delta\chi^2$ indicates measurement invariance are highlighted. Scalar models were not tested if the metric model did not hold.

Table 19

2012 Between-Group $\Delta\chi^2$ Test

Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
Hosp. 5 – Shared Services	14.333	2	21.317	5	49.62	9	6.984	3	0.07	28.303	4	0.0000
Hosp. 5 – Hosp. 2	0.74	2	8.198	5	54.689	9	7.458	3	0.06	46.491	4	0.0000
Hosp. 5 – Hosp. 1	0.912	2	13.965	5		9	13.053	3	0.00			
Hosp. 5 – Hosp. 4	2.457	2	4.839	5	21.804	9	2.382	3	0.50	16.965	4	0.0020
Hosp. 5 – Hosp. 3	1.887	2	4.617	5	9.014	9	2.73	3	0.44	4.397	4	0.3549
Shared Services – Hosp. 2	13.785	2	17.495	5	40.391	9	3.71	3	0.29	22.896	4	0.0001
Shared Services – Hosp. 1	13.956	2	32.584	5		9	18.628	3	0.00			
Shared Services – Hosp. 4	15.502	2	19.444	5	25.017	9	3.942	3	0.27	5.573	4	0.2334
Shared Services – Hosp. 3	14.932	2	18.383	5	34.105	9	3.451	3	0.33	15.722	4	0.0034
Hosp. 2 – Hosp. 1	0.364	2	19.877	5		9	19.513	3	0.00			
Hosp. 2 – Hosp. 4	1.909	2	3.371	5	25.752	9	1.462	3	0.69	22.381	4	0.0002
Hosp. 2 – Hosp. 3	1.339	2	3.578	5	41.19	9	2.239	3	0.52	37.612	4	0.0000
Hosp. 1 – Hosp. 4	2.08	2	14.865	5		9	12.785	3	0.01			
Hosp. 1 – Hosp. 3	1.511	2	15.998	5		9	14.487	3	0.00			
Hosp. 4 – Hosp. 3	3.056	2	3.334	5	12.562	9	0.278	3	0.96	9.228	4	0.0556

Note. Group comparisons where $\Delta\chi^2$ indicates measurement invariance are highlighted. Scalar models were not tested if the metric model did not hold.

Table 20

2013 Between-Group $\Delta\chi^2$ Test

Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
Hosp. 5 – Shared Services	16.326	2	51.197	5		9	34.871	3	0.00			
Hosp. 5 – Hosp. 2	22.795	2	59.481	5		9	36.686	3	0.00			
Hosp. 5 – Hosp.1	17.032	2	50.44	5		9	33.408	3	0.00			
Hosp. 5 – Hosp. 4	16.621	2	27.319	5		9	10.698	3	0.01			
Hosp. 5 – Hosp. 3	20.568	2	30.223	5		9	9.655	3	0.02			
Shared Services – Hosp. 2	11.056	2	26.152	5		9	15.096	3	0.00			
Shared Services – Hosp. 1	5.293	2	60.538	5		9	55.245	3	0.00			
Shared Services – Hosp. 4	4.881	2	11.923	5	39.84	9	7.042	3	0.07	27.917	4	0.0000
Shared Services – Hosp. 3	8.829	2	34.499	5		9	25.67	3	0.00			
Hosp. 2 – Hosp. 1	11.762	2	64.729	5		9	52.967	3	0.00			
Hosp. 2 – Hosp. 4	11.35	2	18.567	5	23.959	9	7.217	3	0.07	5.392	4	0.2494
Hosp. 2 – Hosp. 3	15.298	2	40.615	5		9	25.317	3	0.00			
Hosp. 1 – Hosp. 4	5.588	2	44.265	5		9	38.677	3	0.00			
Hosp. 1 – Hosp. 3	9.536	2	33.546	5		9	24.01	3	0.00			
Hosp. 4 – Hosp. 3	9.124	2	16.477	5	31.488	9	7.353	3	0.06	15.011	4	0.0047

Note. Group comparisons where $\Delta\chi^2$ indicates measurement invariance are highlighted. Scalar models were not tested if the metric model did not hold.

Table 21

2014 Between-Group $\Delta\chi^2$ Test

Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
Hosp. 5 – Shared Services	26.704	2	58.124	5		9	31.42	3	0.00			
Hosp. 5 – Hosp. 2	25.297	2	27.714	5	36.94	9	2.417	3	0.49	9.226	4	0.0557
Hosp. 5 – Hosp.1	25.23	2	44.352	5		9	19.122	3	0.00			
Hosp. 5 – Hosp. 4	25.477	2	31.257	5	36.235	9	5.78	3	0.12	4.978	4	0.2896
Hosp. 5 – Hosp. 3	31.922	2	39.703	5	65.017	9	7.781	3	0.05	25.314	4	0.0000
Shared Services – Hosp. 2	1.839	2	12.616	5		9	10.777	3	0.01			
Shared Services – Hosp. 1	1.772	2	4.419	5	74.849	9	2.647	3	0.45	70.43	4	0.0000
Shared Services – Hosp. 4	2.018	2	16.14	5		9	14.122	3	0.00			
Shared Services – Hosp. 3	8.463	2	28.1	5		9	19.637	3	0.00			
Hosp. 2 – Hosp. 1	0.365	2	13.37	5		9	13.005	3	0.00			
Hosp. 2 – Hosp. 4	0.611	2	1.438	5	3.991	9	0.827	3	0.84	2.553	4	0.6352
Hosp. 2 – Hosp. 3	7.056	2	8.209	5	36.97	9	1.153	3	0.76	28.761	4	0.0000
Hosp. 1 – Hosp. 4	0.545	2	14.255	5		9	13.71	3	0.00			
Hosp. 1 – Hosp. 3	6.99	2	21.783	5		9	14.793	3	0.00			
Hosp. 4 – Hosp. 3	7.236	2	7.757	5	27.909	9	0.521	3	0.91	20.152	4	0.0005

Note. Group comparisons where $\Delta\chi^2$ indicates measurement invariance are highlighted. Scalar models were not tested if the metric model did not hold.

Table 22

2015 Between-Group $\Delta\chi^2$ Test

Model	Configural		Metric		Scalar		Config-Metric			Metric-Scalar		
	χ^2	*df*	χ^2	*df*	χ^2	*df*	$\Delta\chi^2$	Δdf	sig	$\Delta\chi^2$	Δdf	sig
Hosp. 5 – Shared Services	51.061	2	64.757	5		9	13.696	3	0.00			
Hosp. 5 – Hosp. 2	42.872	2	51.806	5		9	8.934	3	0.03			
Hosp. 5 – Hosp.1	44.91	2	55.403	5		9	10.493	3	0.01			
Hosp. 5 – Hosp. 4	40.568	2	43.715	5	56.938	9	3.147	3	0.37	13.223	4	0.0102
Hosp. 5 – Hosp. 3	40.515	2	50.274	5		9	9.759	3	0.02			
Shared Services – Hosp. 2	12.968	2	20.411	5	23.974	9	7.443	3	0.06	3.563	4	0.4684
Shared Services – Hosp. 1	15.006	2	22.777	5	105.739	9	7.771	3	0.051	82.962	4	0.0000
Shared Services – Hosp. 4	10.664	2	11.855	5	20.651	9	1.191	3	0.76	8.796	4	0.0664
Shared Services – Hosp. 3	10.612	2	17.562	5	30.315	9	6.95	3	0.07	12.753	4	0.0125
Hosp. 2 – Hosp. 1	6.818	2	16.081	5		9	9.263	3	0.03			
Hosp. 2 – Hosp. 4	2.476	2	6.06	5	15.25	9	3.584	3	0.31	9.19	4	0.0565
Hosp. 2 – Hosp. 3	2.423	2	5.747	5	13.453	9	3.324	3	0.34	7.706	4	0.1030
Hosp. 1 – Hosp. 4	4.513	2	12.083	5	53.91	9	7.57	3	0.06	41.827	4	0.0000
Hosp. 1 – Hosp. 3	4.461	2	10.551	5	93.816	9	6.09	3	0.11			
Hosp. 4 – Hosp. 3	0.119	2	2.074	5	17.978	9	1.955	3	0.58	15.904	4	0.0032

Note. Group comparisons where $\Delta\chi^2$ indicates measurement invariance are highlighted. Scalar models were not tested if the metric model did not hold.

Engagement Means and Percentages

The reader may recall that the Engagement scale resulted in four categories, *Engaged, Content, Ambivalent,* and *Disengaged,* each with its definition. Each year, The Advisory Board presented SystemTex with an Engagement mean and percentage of employees who fell into the *Engaged* category by each facility. To be considered Engaged, respondents had to answer "Strongly Agree" to at least two of the four Engagement independent variable (indicators) and no less than "Agree" to any of them. As stated earlier, the organizational priority was improvement of engagement at SystemTex by an increasing mean score and Percent Engaged from one year to the next and one facility to the next. The mean and percentages of Engagement were the sole markers that SystemTex understood the Engagement phenomena by and triggered the activity to change them.

For the present study, percentages of Engagement were recalculated using the same logic prescribed by The Advisory Board. Figure 11 shows that the Engagement percentage calculations were very similar. Though the Percent Engaged recalculations resulted in slightly lower percentages than those of The Advisory Board, the differences were consistently within three percentage points or less of each other, regardless of respondent sample sizes. The percentages offered Engagement for each facility based on its specific respondent population. However, the percentage alone did not clarify how the Engagement mean score results were related to the percentage and how the mean score was interpreted.

The Advisory Board, having clients across the U.S., aggregated its mean scores across all respondents meaning that benchmarking of the four engagement categories—Engaged, Content, Ambivalent, and Disengaged—could fluctuate annually due to changes within the whole database.

Generally speaking, the higher the mean score, the greater the percentage of engaged employees and vice versa. However, the Percent Engaged reflected how much of the employee population of any given facility fell into the Engaged category, while the mean score reflected an average of all EES respondents across The Advisory Board falling into the Engaged category. To clarify how The Advisory Board benchmark database profile changed while SystemTex was a client, consider the following data points provided in SystemTex from Advisory Board presentations:

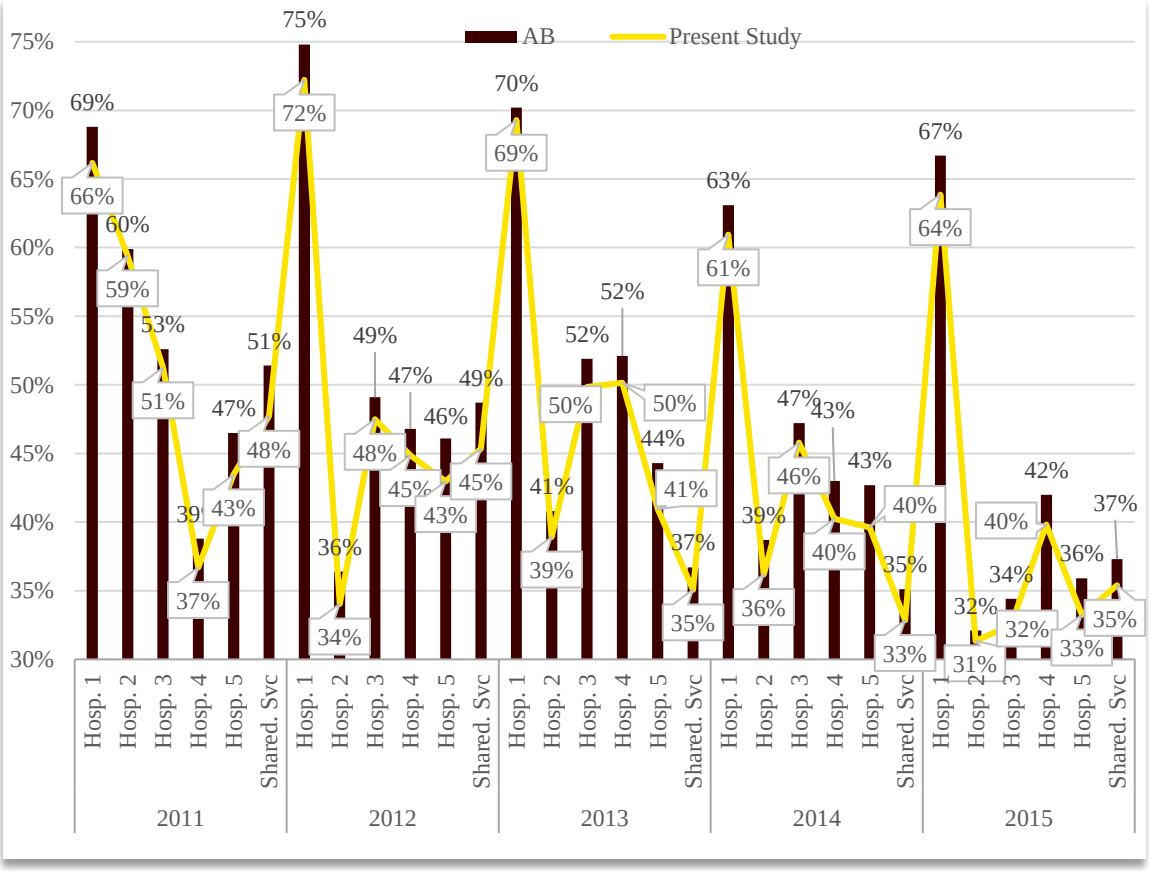
- In 2011, with 104,000 respondents participating in the EES, 35% of The Advisory Board's benchmark was 35% Engaged, 45% Content, 15% Ambivalent, and 5% Disengaged (The Advisory Board Survey Solutions, 2011).
- In 2013, 37.1% of the database was Engaged, 40% were Content, 15% were Ambivalent, and 5.9% of the benchmark was Disengaged. Between 2013 and 2014, the number of respondents participating in the EES grew to 342,000 (The Advisory Board Survey Solutions, 2014).
- In 2014, 39.3% of Advisory Board clients were Engaged, 40.6% were Content, 13.6% were Ambivalent, and 4.8% were Disengaged (The Advisory Board Survey Solutions, 2014). By June of 2014, The Advisory Board reported more

than 650,000 respondents participating in the EES (The Advisory Board Survey Solutions Presentation to WellSpan Health, 2015).

- In 2015, The Advisory Board reported 750,000 respondents participating in the EES benchmark (The Advisory Board Survey Solutions, 2015). The percent engaged is unknown.

Figure 11

Facility Percent Engaged Calculations



Note. The Advisory Board mean score data was reprinted with written permission from The Advisory Board. See Appendix D.

Unlike Percent Engaged, the mean scores had a range, given that responses were based on a 6-point Likert scale and subsequently divided into the four engagement categories (see Table 23).

Table 23

Defining Engagement Categories

Category	Range	Definition
Engaged	5.5 - 6	• Go above and beyond to see the organization succeed, tying personal success directly to that of organization • Highly loyal and emotionally committed to the organization
Content	4.5 - 5.5	• Solid contributors, satisfied with their jobs and the organization • Lacking emotional commitment to organization
Ambivalent	3.5 - 4.5	• Would leave if presented with a better offer • See job as paycheck more than anything else
Disengaged	< 3.5	• Least satisfied with their job and organization • Tend to be most vocal, actively detracting from quality of workplace for peers

Note. The above material is reprinted with written permission from The Advisory Board. See Appendix D (The Advisory Board Survey Solutions, 2015).

See Figures 12 through 17 for charting of SystemTex facility mean scores against the Percent Engaged. The Advisory Board calculations were used for the charts since they were within five percentage, and a more liberal calculation of the data. What is evident within SystemTex in Figures 12 through 17 was that the overall downward movement in the percentage of engaged employees tracked with shifts in the engagement mean. Yet, due to known non-invariance within the dataset, I refrain from referring to the movement as a trend. Data shown in Figures 12 through was reprinted with written from The Advisory Board (see Appendix D).

Figure 12

Hosp. 1 Mean Score and % Engaged

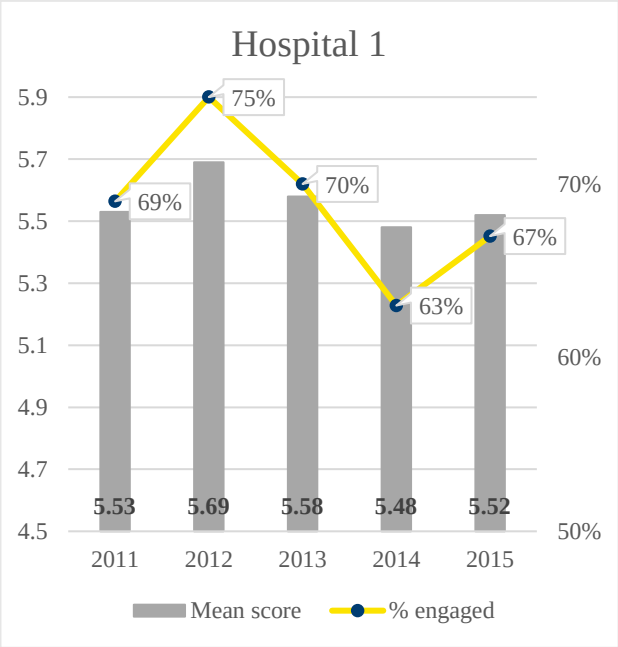


Figure 13

Hosp. 2 Mean Score and % Engaged

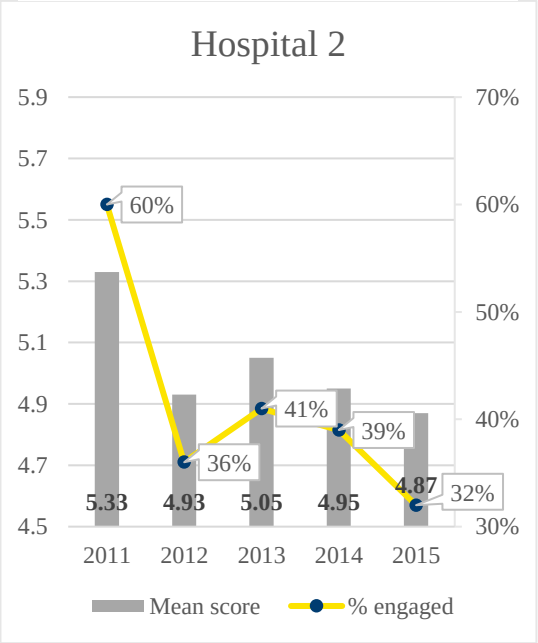


Figure 14

Hosp. 3 Mean Score and % Engaged

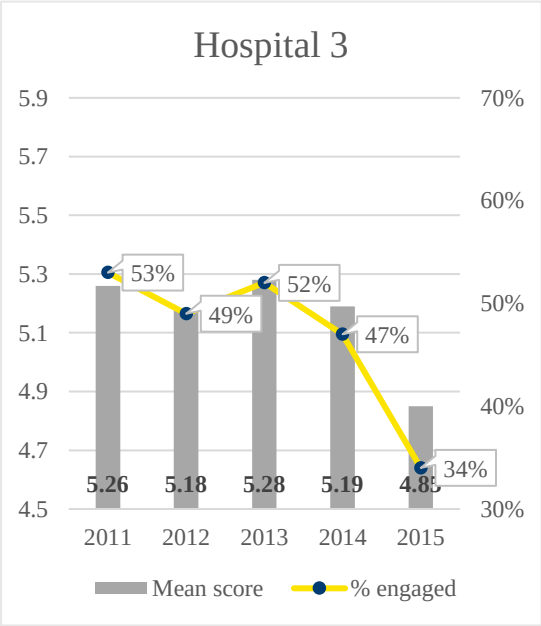


Figure 15

Hosp. 4 Mean Score and % Engaged

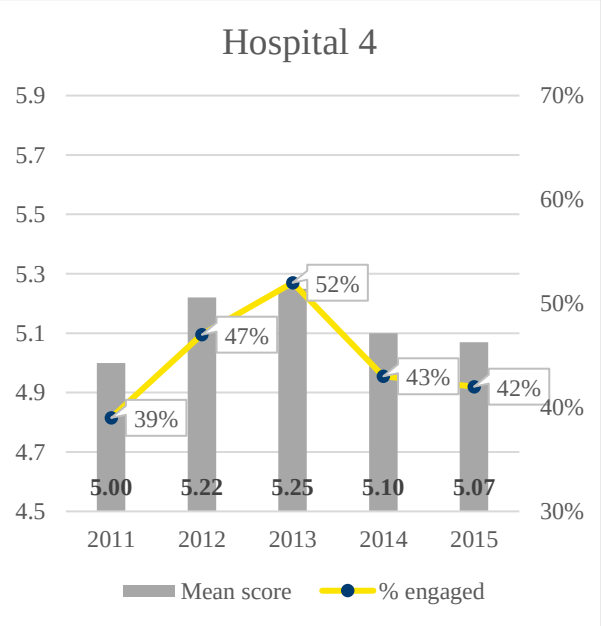


Figure 16

Hosp. 5 Mean Score and % Engaged

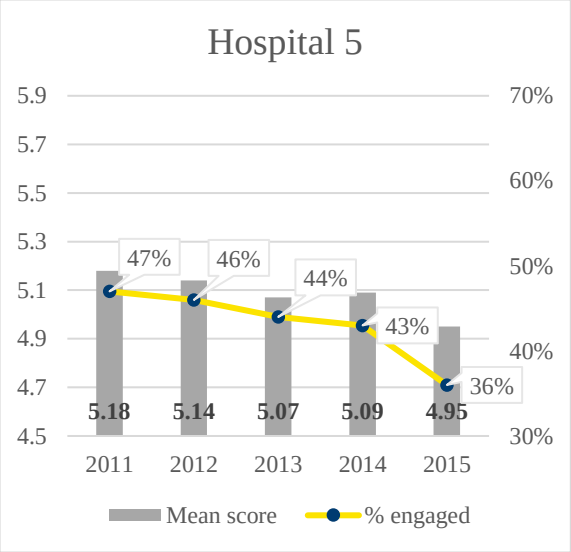
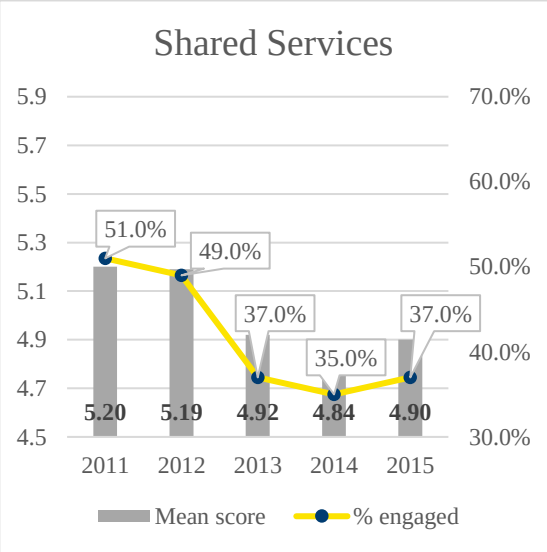


Figure 17

Shared Svcs. Mean Score and % Engaged



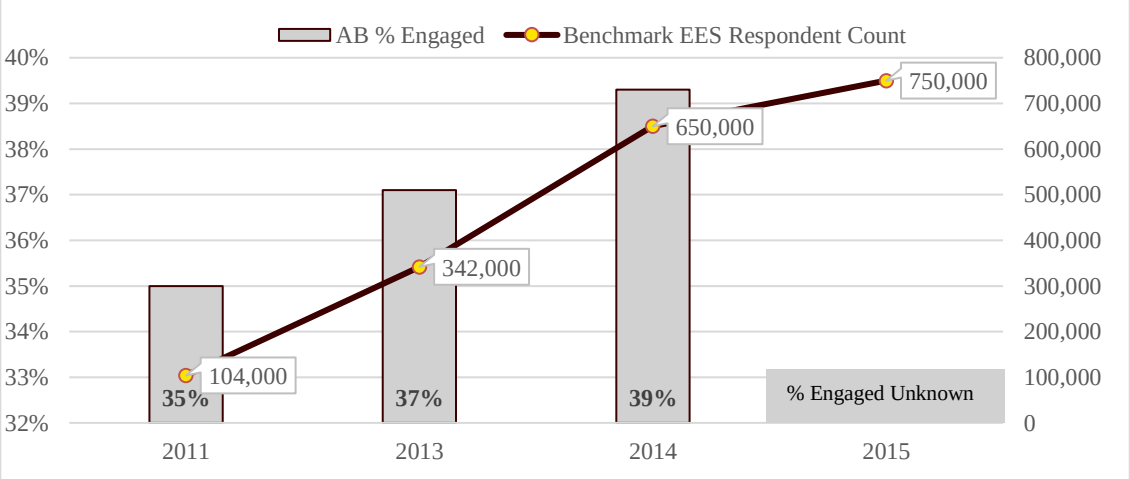
Summary

The results did not support either of the two hypotheses. The first hypothesis was that the merger and acquisition with a larger national organization in 2013 contributed noticeable measurement non-invariance within SystemTex's engagement response, which was not the case. The second hypothesis was that non-invariance response shifts would not occur until 2013, which was not the case, as there was non-invariance present in each facility and within each survey completion wave. Configural invariance existed within each facility, which was appropriate, as there would be no reason to explore further differences if the Engagement construct did not exist. Lack of metric invariance showed different Engagement conceptualizations within a facility from one wave to the next and between some facilities. Lack of scalar equivalence showed that some groups differed in how they scored the indicators. In other words, a score of "3" on the response

scale at one facility did not consistently mean the same thing as a "3" at another facility (Meyer, 2020), meaning that the scores were not related to each other (Chakraborty, 2017; Milfont & Fischer, 2010). Figure 18 plots Advisory Board data for Percent

Figure 18

Advisory Board Benchmark % Engaged & Respondent Count



Note. The above data is printed with the written permission of The Advisory Board. See Appendix D.

Engaged against the respondents in the benchmark database. The data for 2012 is unknown. Each year, SystemTex's means were compared against higher and higher performing benchmark, while its employees became less engaged.

An increase in overall EES participation moved the Engagement mean score upward incrementally as more Advisory Board clients fell into the Engaged range, possibly higher than SystemTex. Nevertheless, SystemTex did not have bad mean or Percent Engaged scores. In fact, in 2011, all SystemTex facilities were more engaged than The Advisory Board benchmark. In 2013, all but the Shared Services facility were

above The Advisory Board's benchmark. In 2014, two facilities fell below The Advisory Board's benchmark, even using the more conservative recalculations; however, if the same Percent Engaged were employed for 2015, five of the six facilities would have fallen below the benchmark. The Engagement means and percentages were not the focal points of the MI analysis, but they complexified the findings' practical and academic implications to be discussed in Chapter V.

CHAPTER V

DISCUSSION*

In Chapter I, a metaphor was introduced of two botanists exploring plant biology to determine if the plants were the same species as a way to understand the practice and purpose of measurement invariance (MI). Through the results, it should be understood that because the presence of invariance and non-invariance, many, but not all of the "plants" have a different number of roots of different lengths (factor loadings) and thickness (intercepts). The botanists would conclude that many of the plants (groups) are fundamentally different.

The first hypothesis was that MI would hold in 2011 and 2012. This hypothesis was not only not true but also has little merit based on the presence of measurement non-invariance in the first year of the EES. The second hypothesis predicted response shifts in 2013 – 2015, which, while not wholly inaccurate, was imprecise given that non-invariance existed throughout the dataset. If my hypotheses had borne out, MI would become non-invariant in 2013 to align with the employee experience of perpetual change and upheaval in business processes, systems, and practices I believed others to be experiencing along with me.

In retrospect, my employee experience was largely positive; furthermore, I believed that the high engagement I observed at my point of entry into the organization

---

* Reprinting of The Advisory Board data and resources was allowed with written permission (see Appendix D).

(2012) was ubiquitous throughout SystemTex. My perspective was simplistically biased towards my individual employee experience and disassociated with pre-existing group differences. This belief was enhanced by encouragingly high Engagement mean scores presented by The Advisory Board. But I had not considered the M&As that occurred prior to my tenure. M&As persist as a hallmark of insufficient financial returns, unrealized efficiencies, and other detrimental organizational and cultural changes (Appelbaum, Karelis, Henaff, & McLaughlin, 2017; Baynham, 2011; Dranove & Lindrooth, 2003; Kjekshus & Hagen, 2007). However, this study is not so clear-cut as to lay response shifts at the feet of one M&A transaction.

This chapter considers the findings and implications those findings have for measurement invariance, employee engagement, OD, and HRD. Engagement is a construct with many varying theoretical and lay definitions. All definitions attempt to operationalize the latent variable through measurable indicators (Little 2013). However, instruments striving to show a latent idea must be built on models that can effectively specify the nature of measurement how that idea is to be understood (Little 2013). The EES offered a model whose purpose was to allow SystemTex to see or infer the reality of organizational engagement; perhaps there remains some efficiency and utility of that model. So, how does one assess the serviceability of a project that shows inconsistent findings?

Discussion of Results

Little (2013) says that because models are simplifications of reality, they are "necessarily wrong" (p. 3). Even so, the whole enterprise of their analysis need not be

abandoned. SEM models, relying on strong statistical theory, can clarify untidy processes (Little, 2013). Despite the haphazard results, does the EES model and the research project allow reality and knowledge to LEAP forward?

LEAP stands for

> **L**ogical and internally consistent,
> **E**mpirically testable (falsifiable),
> **A**ccounts for extant findings,
> **P**arsimonious (sufficient verisimilitude).

LEAP suggests that models should be (a) internally consistent, (b) they should be free of "circularities" (c) they should fit within existing theoretical fields of study, and (d) good models should explain the observed data across most contexts and populations while "maintaining the ability to depict reality" (Little, 2013, p. 5).

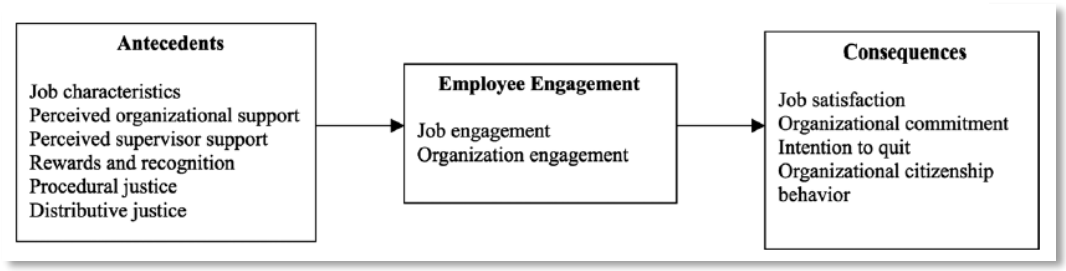
Logical and Internally Consistent

The first research question of this survey asked, *How reliable is the Engagement construct?* Internal consistency and reliability of the Engagement construct was assured through the split-half analysis in the EFA, and Cronbach's α and three sample draws in the CFA. The EES seems to have face validity with other commercial instruments such as Towers Perrin's (2003), which in Advisory Board presentation materials. The workforce study of 35,000 employees did not lay out its measurement model. Rather, it offered a commercially viable definition of Engagement—discretionary effort—for companies attempting to clutch the "prize" of Engagement and "move the needle" (p. 4) within their employee populations. Not coincidently, it is the same definition offered by the Advisory Board.

The EES could be internally consistent with other scales, but it remains unexplored. Measurement of employee engagement and its various categories (e.g., burnout, work, job, intellectual/social, or organizational engagement) is not homogenous, nor does it have to be in order to be viable or rigorous. An assumption that the EES bears no internal consistency with more scholastically thorough scales is faulty without further data analysis. Saks (2006) presented a model of antecedents and consequences for organizational engagement, which, at least at a glance, appears to be a similar model to The Advisory Board's EES's design (see Figure 11). The Saks model's antecedents predicted the mediating construct of employee engagement distinguished into job and organizational engagement subcategories. Finally, Saks' model showed the consequences of employee engagement, which are not completely dissimilar to the four indicators of the EES statements 43-46. Statement 43 (*I would recommend this organization to my friends as a great place to work*) could be consistent with job satisfaction and organizational citizenship behavior (OCB). Statement 44 (*This organization inspires me to perform my best*) may be consistent with job satisfaction, organizational commitment, and OCB. Statement 45 (*I am likely to be working for this organization three years from now*) may be consistent with the intention to quit. Statement 46 (*I am willing to put in a great deal of effort in order to help this organization succeed*) may be consistent with organizational commitment and OCB.

Figure 19

Saks Model of Antecedents and Consequences of Employee Engagement



Logical implications. For most deploying MI analysis, the goal is to assert the equivalence of the underlying measurement model (Vandenberg & Lance, 2000). In MGCFA (multi-group CFA), invariance is proof that the scale can do one very important function—indicate that groups are invariant and can therefore be compared to one another. MI tests pinpoint when construct models are a good fit, maintain consistent factor loadings, and have the same measurement relationships across groups (Meyer, 2020). First, the configural model must hold across groups as an entrance condition for MI. When configural invariance does not hold, there is no basis for further testing of MI because it means that the construct is not present in the group—you cannot measure what is not there. When it does hold, the meaning is that respondents held similar conceptual ideas of the construct(s).

The next level of MI tests is metric invariance compared against the configural model. Factor loadings (λ) are constrained to be equal across groups, and when metric invariance does not hold, then a construct has different meaning across groups (Vandenberg & Lance, 2000). Though independent of one another, group invariance

would have meant that the organization held Engagement as the same underlying conceptual idea.

Scalar non-invariance is that of lack of equivalence in intercepts. It is pernicious because it suggests larger forces influencing how participants respond across time and group administrations (Bialosiewicz et al., 2013). It results in scales meaning one thing for a group and something different for another group, though the specification of those meanings requires further research. Within-group metric and scalar non-invariance summary tables are shown in Tables 24 and 25, isolating the non-invariant group pairings of the present study.

Invariance tests are critical to any group attempting to use latent constructs to test unobserved phenomena, which helps resolve the second research question, "*What would the factor structure reveal about employee perspectives on engagement*?" It is evident that SystemTex employees had some similar but many dissimilar conceptualizations of engagement and differed in some groups as to how they assessed and evaluated the construct.

Given that MI is inconsistent nature throughout, it is equally relevant to consider why failure to demonstrate invariance matters to organizations, researchers, and instrument designers. One not-to-be-ignored consequence of the non-invariance is that SystemTex Engagement trends from one year to the next cannot be assumed, nor should one group be compared to another group with which it is non-invariant (Vandenberg & Lance, 2000). Hospital 1 was the only invariant hospital when compared to itself across all five years, but it exhibited non-invariance with Hospitals 2, 3 4, 5, and Shared

Services at some point. This information did not offer the organization a realistic opportunity to deliberate about its engagement performance.

Table 24

Within Group Metric Non-invariance

Models	**Metric**		
	$\Delta\chi^2$	Δdf	*sig*
Hospital 2			
2011 – 2012	0.885	3	0.83
2013 – 2014	10.303	3	0.02
Hospital 3			
2014 – 2 015	13.368	3	0.00
Hospital 5			
2011 – 2012	11.177	3	0.01
2014 – 2015	9.447	3	0.02
Shared Services			
2011 – 2012	10.103	3	0.02
2012 – 2 013	11.795	3	0.01
2014 – 2015	7.912	3	0.048

Table 25

Within-Group Scalar Non-invariance

Models	**Scalar**		
	$\Delta\chi^2$	Δdf	*sig*
Hospital 2			
2011 – 2012	40.146	4	0.0000
Hospital 3			
2014 – 2015	60.102	4	0.0000
Hospital 4			
2011 – 2012	17.495	4	0.0015
Hospital 5			
2012 – 2013	16.303	4	0.0026
2014 – 2015	47.444	4	0.0000
Shared Services			
2011 – 2012	35.521	4	0.0000
2012 – 2013	18.328	4	0.0011

The MI results answered the third research question, "*Would the 2013 M&A generate a response shift in engagement*?" with a resounding "No." However, non-invariance should not be interpreted as the instrument or this project not serving a practical purpose.

Action research (AR) approaches for improving results from within an organization (e.g., enhancing manager effectiveness through training, ensuring that employees have appropriate decision-making authority, providing ample resources for job completion, and opportunities to learn and grow in job roles) do not appear practicable when results cannot be trended. Respondents lacked equivalent definitions,

conceptualizations, or measurements of constructs to which they are responding. In broad terms, leaders, decision-makers, and those given the authority to act as agents of the organization did not vigorously understand, much less address precisely what changed, what is changing or, or what should be changed.

Chapter I pointed out an average of 90 healthcare M&As in the US each year between 2011 to 2015 (Gaynor, 2016). These deals affected no less than 150 hospitals each year during the same years. In 2012, 2013, and 2015, the number of hospitals impacted by an M&A annually was more than 200. While the current research could not demonstrate that a single acquisition was detrimental to Engagement, it does assert that non-invariance may grow unobserved and unattended in organizations that will be pushed and pulled together into yet other newly formed organizations. The issues attributed to difficulties of M&A may be dormant non-invariance within unexplored latent constructs.

Empirically Testable

Little's (2013) second stipulation for LEAP is that models must not lack specificity. Engagement scale reliability was established through the split-half EFA approach and confirmed through the randomized CFA. Goodness of fit indices were logically considered and provided good model fit making a three or four-factor solution plausible. However, both solutions introduce unresolved circularities that Little (2013) warns against. Poor discriminate validity (0.90) between prescribed Change Readiness and Manager Effectiveness constructs registered that the structures were too much

related to each other to specify distinct construct domains. Fundamentally, the EES neglected to link each item to each construct and each construct to each other.

The categorical drivers prescribed by The Advisory Board remain unaddressed. At a glance, the categories mimic antecedents following other scholarly observations highlighting the dynamics of communication, employee support, feedback, manager behavior, professional development, and teamwork (Macey & Schneider, 2008; Merve Ünal & Turgut, 2015; Saks, 2006). Nevertheless, valid questions persist as to "driver" viability, relationships, and function. Though the Engagement factor showed scale reliability, too many other problems confound the whole instrument's veracity.

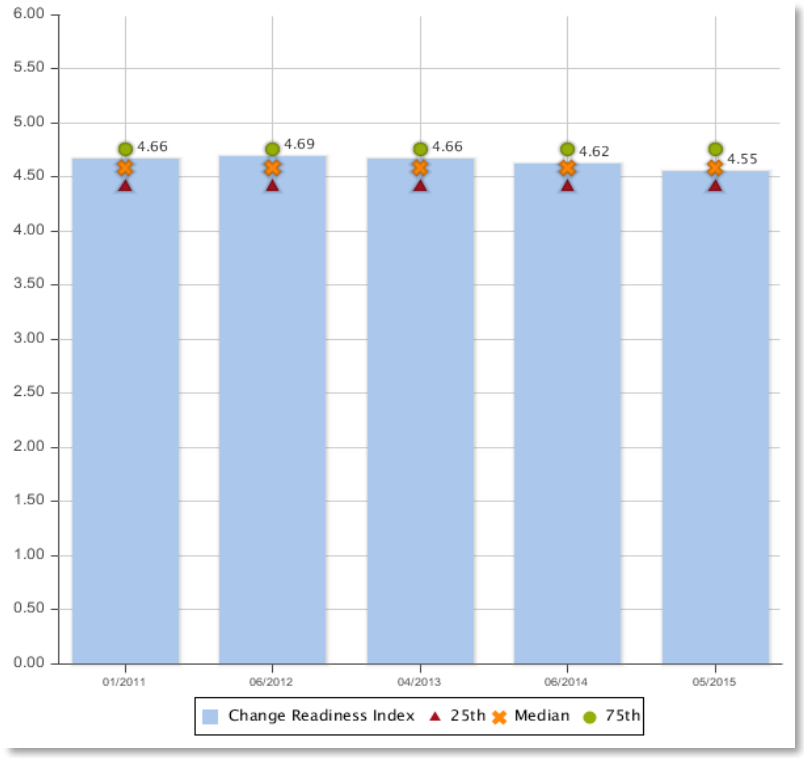
Empirically testable implications. Who bears responsibility for factorial non-invariance? From my professional experience, a client—once assured of instrument quality through marketing, clients, and sales executives—does not have time, capacity, or skill for in-depth statistical analysis when purchasing analytic tools. Nor, in most cases, would a vendor grant permission for such scrutiny. Due diligence in these instances usually involves little more than inquiring of other clients how content they were with the vendor's service, consultant's performance, and company responsiveness. There was ample demonstration by The Advisory Board that their instrument was used by many well-respected healthcare organizations. Therefore, assurances of factorial invariance must fall to those who design and architect psychometric tools.

Figure 20 is a screenshot of the Change Readiness Index made available to SystemTex leaders from the online databank. What is shown is an aggregate of individual, department, and hospital SystemTex change readiness. The information

filtered for the screenshot does not show from what hospital(s), department(s), or leader(s) the data is filtered. The concern is what meaning can be inferred from the data?

Figure 20

Change Readiness Index Screen Shot



Note. The above chart is printed with written permission from The Advisory Board. See Appendix D.

If any single SystemTex leader sought to grasp the Change Readiness Index results from a leader or department in a single year, say 2011, the mean scores might be true. However, imagine that someone led multiple hospital departments (e.g., Therapy, Diagnostic Imaging, Surgical Services), how can the results be interpreted given the presence of non-invariance? Alternatively, if an agent of the organization used the EES data as a performance metric, how confusing might that be? The lowest-performing item

from this output was S28 ("*My organization helps me deal with stress and burnout*."), which SystemTex leaders set out to address every year of the survey because it was believed to be within our control to improve.

Improving the EES model. Improvement of models is a worthwhile endeavor. The prescribed model comprised 42 predictor variables, which seemed like a lot of variables pointing to a four-indicator construct! Leveraging the current dataset to discover which variables have the most meaning to Engagement makes valuable contributions to research and the practical application in workplaces.

Mplus produced a model using the split-half results. By removing 32 non-significant items, goodness-of-fit statistics improved, leaving ten significant predictors (see Table 26). In the improved model (see Figures 21 and 22), Engagement mediates the four indicators. Statements 44 (*This organization inspires me to perform my best*), 45 (*I am likely to be working for this organization three years from now*), and 46 (*I am willing to put in a great deal of effort in order to help this organization succeed*) have correlated error terms, resulting in different relationships between the variables than with the original measurement model. While these correlations could be due to randomness, other reasons could include common factor variance or acquiescent responses from agreeable attitudes (Meyer, 2020). Without more information on the instrument's theoretical background, it cannot be determined if the improved model occurs at the expense of theory (Cole, Ciesla, & Steiger, 2007).

Table 26

Predictors that Improve the EES Engagement Construct

Predictor Variables
S5. My ideas and suggestions are valued by my organization.
S11. I receive the necessary support from employees in other units/departments to help me succeed in my work.
S14. My manager stands up for the interests of my unit/department.
S16. I am interested in promotion opportunities in my unit/department.
S23. My current job is a good match for my skills.
S25. My organization pays me fairly for my job.
S29. My organization recognizes employees for excellent work.
S30. Executives at my organization respect the contributions of my unit/department.
S39. My organization does a good job of selecting and implementing new technologies to support my work.
S42. Over the past year I have never been asked to do something that compromises my values.

Note. The above EES statements are printed with written permission from The Advisory Board. See Appendix D.

Figure 21

1st Half of Improved Engagement Model Results

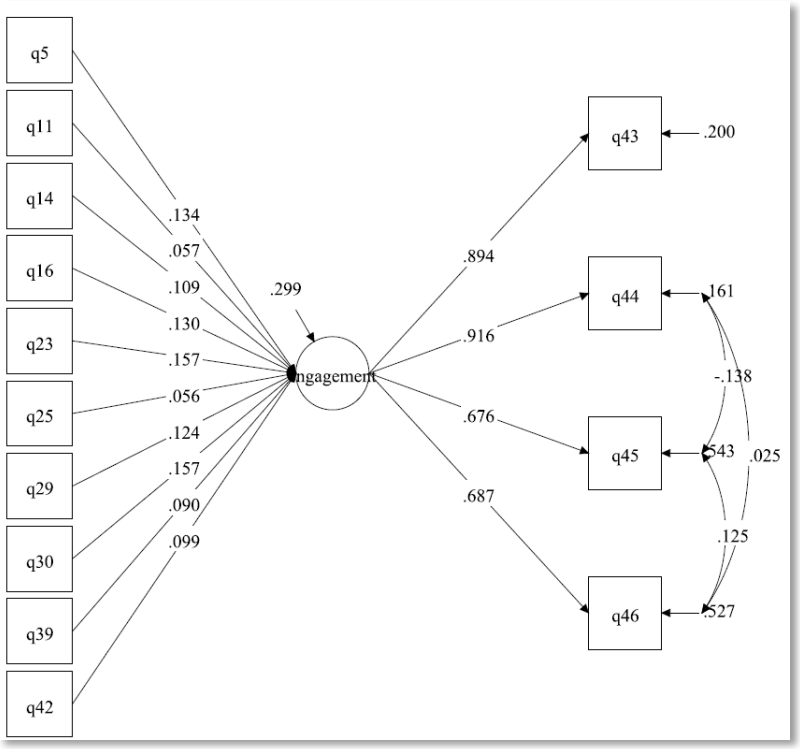
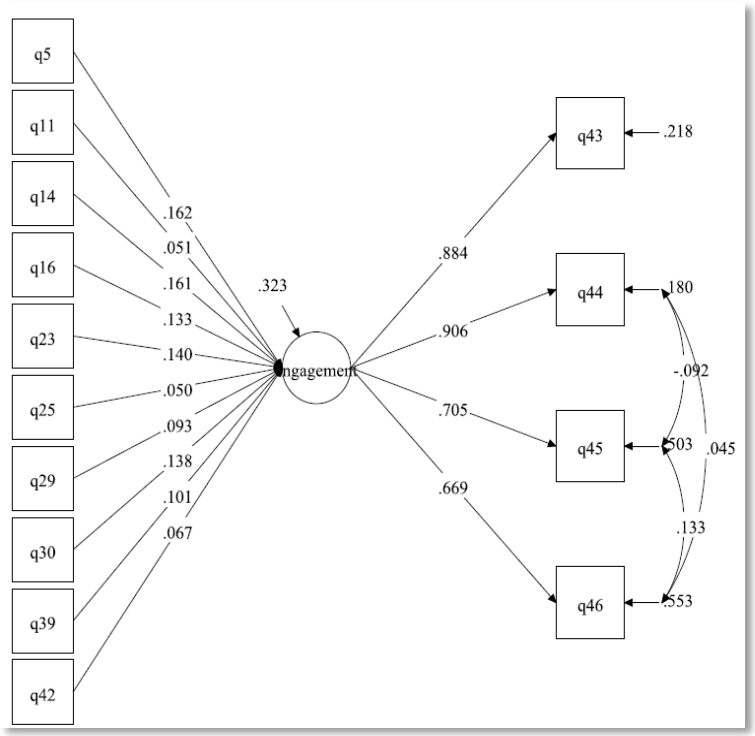


Figure 22

2nd Half of Improved Engagement Model Results



Note. Split-half analysis was conducted in Mplus. Thirty-two non-significant predictors have been removed from the diagrams, which improved goodness-of-fit statistics

Accounts for Present Findings

The third LEAP condition is that a model accounts for constructs that fit "within the scope of the theoretical realm under study" (Little, 2013, p. 5), which is Cronbach's and Meehl's (1955) view of construct validity. It requires theoretical and empirical frameworks for measurement, and specific links between the two (Cronbach & Meehl, 1955; Trochim, 2020), neither of which I had from The Advisory Board. Instrument designers attempting to build well-behaving models will link conceptual ideas with observable ideas so that laws can explain when differences occur (Trochim, 2020). The good news for The Advisory Board and any organization that bought the survey is that analysis through EFA and CFA proved the Engagement construct. The improved model likely improves Engagement's nomological network, making it a better instrument all around.

Implications of accountability for present findings. Good theory describes how things work, and its absence means that processes for problem-solution formulation have no backdrop against which to inform knowledge frameworks (Lynham, 2002). Because The Advisory Board's model is theoretically problematic, both industry and academia should tread carefully to infer meaning from it. The reality is that such caution is not often shown when the goal is to leverage data for organizational change.

MI theory does what it was designed to do by explaining the meaning, nature, and challenges of invariance in real-world contexts so that knowledge and understanding allow us to "act in more informed and effective ways" (Lynham, 2002, p. 222). Non-invariance disregards good theory. If group-based psychometric instruments cannot

allow meaningful cross-group comparisons, they are scholastically superfluous, though maybe not completely void of practical purpose. However, the reliance on mean scores to explain Engagement resulted in overlooking meaningful, substantive information, and also perhaps in misapplication of effort to results that were not particularly terrible.

Expenditure on resources. Organizations may choose to utilize unverified instruments out of an abundance of concerns for cost, capacity, time to implementation, and lack of knowledge, resulting in unaccounted siphons of time, attention, and financial resources. Organizational projects undertaken to respond to survey findings are misplaced in their focus, bringing up a theoretical musing too vague to be one of the formal research questions but still very cogent *"If SystemTex agents and decision-makers had known of the non-invariance, would that knowledge impact applied behavior?*" Possibly, but I do not know for certain. In my opinion, an OD approach for impacting local conceptualizations of Engagement would have been less resource-intensive than the change management approaches already underway. Awareness of dissimilarities in how Engagement is defined, understood, and measured might have invited a different organization to look beneath the surface at constructs, but that is not the perspective from which The Advisory Board approached the declining engagement concerns; therefore, it was not a strategy the organization was going to assume.

Good AR, one of the core skills of OD work (Coghlan & Shani, 2014), is not just to try and explain things, but also to change them. AR is based on two assumptions, (a) involving learners in their learning produces better learning, and (b) it is only possible to understand a system when one is trying to change it since changing human systems

involves “variables which cannot be controlled by traditional research methods.” (Coghlan & Shani, 2014, p. 524). SystemTex used the EES cycles in 2012 – 2015 to understand the organizational system and evidence its ability to affect AR work within that system. In the half-year between survey cycles, some 400+ executives, managers, department leaders, and HR contributors sought to impact organizational engagement measured by a psychometric assessment that was not likely not calibrated to assess groups. Dedicated leaders and front-line employees led webinars, joined appreciative inquiry focus groups, attended townhall and department meetings, as well as one on one coaching interventions to drive change across the system. There is no visibility for the interventions' effectiveness, not only because it was not measured, but more importantly, because the initiatives were disconnected from the source problem.

Finally, there is the financial cost of the survey. The Advisory Board and other vendors proffering engagement and culture surveys charge their clients. In some cases, the survey may be a capital expenditure because of the scale and scope. The number of vendors and instruments is likely to continue to grow. The wide selection of options is helpful for competition since there will be greater diversity in price point. An ROI on an engagement survey may be difficult to prove, especially if an organization is not tracking antecedents such as perceived organizational support, rewards and recognition, and manager effectiveness, or measuring consequences like turnover, AR, and training.

Parsimonious

The final element of LEAP is that a model is deemed “sufficiently good and optimally useful” (Little, 2013, p. 5). This level of accountability begs the questions,

useful for whom, and to what purpose? Precisely how much use should the model be? Little (2013) asserts that scientific endeavors must find the level of parsimony that best explains the observed data in most contexts, circumstances, populations while representing reality. Without insight into the criticality of measurement models underlying an instrument, the SystemTex decision-makers believed the EES model fit for assessing Engagement. Furthermore, the model was simple enough to disseminate throughout the organization and generate accountability to the supervisor level. Results were easily explained, and corresponding actions for correction were implemented throughout.

When deprived of metrics like an Engagement score, it can be difficult to generate shared, focused, structured organizational activity. From the SystemTex perspective, a global assessment of Engagement was the purpose. The Annual EES —though leveraged simplistically—motivated and propelled the organization into action; the model and its design were good enough.

Parsimonious implications. There is an implication for vendors pledging the performance of instruments that can create true change within organizations that implement their tools, but they make that pledge without rigorous scholastic proof of concept. Vendors offer clients support through the design, management, and change processes tied to a contract. There is little impetus for a vendor or external consultant to provide a rigorous, theoretically valid scale if organizations do not request them. Moreover, there is little purpose for a company to deploy an instrument that reports more

sophisticated or complex information than to which it can respond by a timeline and in an environment that meets organizational constraints.

Nevertheless, a pernicious implication for theory and practice comes in the form of a simplistic perspective. 14th-century thinker Franciscan friar William of Ockham (Ockham's Razor) wrote the Summa Logicae in 1323 wherein he said, "It is futile to do with more what can be done with fewer" (Ball, 2016, para. 3). Researchers and practitioners alike use his postulation to imply that the simplest explanations are best, which is not what Ockham meant. A version of Ockham's Razor was repeated by Newton in 1687, who said theories and hypotheses should be as simple as possible while accounting for all observable facts. Newton further implies that the simplest theories are probably also the correct ones (Ball, 2016).

Over-simplicity. A misapplication of Ockham's nod to simplicity introduces an opportunity to get any survey purpose and process very wrong at SystemTex and any other organization with a partiality for simplicity. When leaders want complex organizational phenomena explained very simplistically, in ways that are easy to understand, the risk of over-simplicity is presented. The want of uncomplicatedness, if not satisfaction with it, may blind us to other truths; simpler explanations may be easier to grapple with but can offer erroneous or incomplete understanding. Interpretations and meanings are challenged, and reality is overlooked. Over-simplification can let us off the hook of working harder, looking deeper for dilemmas, and resolving their challenges.

If I may be so bold as to expand Little's (2013) perspective of model parsimony to a broader idea about parsimony of the process, consider the alternate viewpoint that

the EES is effective for an organization. With its searchable online features and ease of explanation, a real-time dashboard was an intoxicating enticement a vendor can offer to large organizations needing to share substantial amounts of information. Companies the size of SystemTex depend on sophisticated feedback systems to (a) benchmark outcomes against competitors, (b) enable leader self-service, and (c) communicate quickly with impressive visuals. Certainly, it matters that the data misrepresented reality, but maybe not enough to demand more rigor of the process. It may seem trite, but the use of data that has the veneer of accuracy shown in a colorful display might be "parsimonious" enough to meet the need, especially if that need is galvanizing people. One of the arguments practitioners level against academia is that scholars do not often realize the value of swift, actionable information, and time as a resource. The previous statement should generate an energetic dialog of organizational priorities and values clarification, but it is best saved for another discussion.

This investigation offered a thought-provoking rhetorical question: How dangerous was the assumption of over-simplification of invariance when it was not there? MI is neither simple nor easy. Invariant surveys are intensive to design and test; predictors could change across the years or groups. When invariance is assumed, then the relationship of predictors to a construct is taken at face value. If it does not hold, then there are changes in definitions and underlying measurements. Any observed or estimated cross-group difference may be due to differences in populations' attributes, not differences in the same attribute between populations intended to be captured (Yuan & Chan, 2016). A lot is going on because the organization's people had been changing, are

changing, and will change. SystemTex is not alone using a simplistic framework; Vandenberg and Lance (2000) informed us that the lack of MI testing means that most who implement psychometric surveys are not typically looking at measurement models.

Strengths and Limitations

This research endorses and informs group MI theorists who insist MI is a critical prerequisite for cross-group comparisons (Assunção et al., 2020; Gallant & Martins, 2018; Shuck et al., 2017; Vandenberg & Lance, 2000;). Organization science relies on the assignment of numbers on variables that represent characteristics like constructs to describe individuals, groups, and organizations, and violations undermine reliability and validity.

Strengths

This study supports MI theory and bolsters it with credible empirical analysis. The reliable testing of MI is the primary strength. It shows how difficult MI can be to establish. This aids practitioners with information and critical questions to ask vendors promising survey instruments that will purportedly and dynamically impact their organization. The findings are also relevant for instrument builders to know how intensive such an undertaking is.

The second strength is that employee engagement has been difficult to examine from substantive datasets. The large data set included multiple years of observations from a single healthcare organization. This study was important from an engagement perspective because it added perspective on how engagement was leveraged commercially. Because engagement is latent, changes in practitioner and scholarly

definitions may shift, disagreements continue about how it should be measured, but the validity and reliability of any psychometric instruments built to test Engagement is critical to its commercial credibility. Testing for the three prescribed constructs adds rigor and eliminated assumptions about the model. It is unclear if this study's reinforcement of commercial instrumentation's precariousness is a strength for everyone concerned. Nevertheless, empirically-based clarification of why instrument rigor matters in a practical sense is essential.

Permission from The Advisory Board and SystemTex to utilize the dataset is a strength of this investigation. Were it not for my prior working relationships with the Advisory Board and SystemTex it would be difficult to ascertain as much about the EES. Academics can experience difficulty obtaining sufficient samples of this size. The survey's reach and five-year duration could not have been undertaken without corporate involvement due to the limited access to employees, The Advisory Board consultants, and finances at the university level. Leveraging archival data meant that minimal expense was incurred. Future investigations based on archival data will enjoy reliable evidence on engagement, healthcare, and MI.

Limitations

The short-falls of this study are considerable. The survey was based on self-reported data, making common method variance (CMV) a concern. CMV is a bias towards high correlations, skewness, and kurtosis due to the measurement method, rather than constructs (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). CMV can generate instrument problems when statements are written in the same format, or responses are

measured in a single instrument. One explanation is that observed values may share variance beyond the actual covariation between them (Malhotra, Schaller, & Patil, 2017). A second reason could be that respondents do not read thoroughly for comprehension due to the repeated common scale anchors (e.g., strongly disagree, agree, strongly agree) (Podsakoff et al., 2003). Respondents get into a rut and reply to statements in an automated manner, making the response itself yet another variable (Malhotra et al., 2017). It is not possible to know if the high correlation between the Change Readiness and Manager Effectiveness constructs is due to the statements' format, or the response anchors, or poorly designed constructs.

A second limitation was that the groups are independent of each other; therefore, their comparisons may not have been equal. It was never the intention of SystemTex to enable the survey for academic scrutiny. Although there is plenty of data to analyze, it was not designed for a longitudinal project. Moreover, I would have asked more questions of The Advisory Board survey consultants about the design and theory when I had them available had I been more forward-thinking. This limitation clarifies the significance of rigorous research design (Nye et al., 2010).

Future Research

The SystemTex dataset would permit the examination of departments within each facility. Since these results indicated that various between-facilities comparisons cannot be made from one year to the next, future research should see if and where MI between departments within a facility holds. It would be thought-provoking to distinguish each department’s differences and similarities regarding the definition and

measurement of engagement. The manager's information would also be interpretable, thus offering more data on leader influences about construct interpretation.

The second area is to measure the EES's improved engagement construct against engagement scales for convergent validity. Potential instruments for comparison are the job engagement scale (JES; Rich, Lepine, & Crawford, 2010), the Utrecht Work Engagement Scale (UWES; Schaufeli, Salanova, González-romá, & Bakker, 2002), the Organizational Engagement Scale (Merve Ünal & Turgut, 2015), or the Shuck et al. (2017) EES. While the UWES and JES differ in their theoretical foundations and capture engagement differently, they are correlated at scale, factor, and item levels to one another (Drake, 2012, p. 29). The UWES is constructed from burnout literature, while the JES has its origins in Kahn's (1990) conception of physical, cognitive, and affective engagement (Drake, 2012). Convergent validity with these scales could improve the academic and commercial credibility of the improved EES engagement model.

Magano and Thomas (2017) studied the breaking of the psychological contract, an unspoken, dynamic exchange agreement between employee and employer. The psychological contract may shift throughout employee experiences, but usually involves a sense of "predictability, job security, and control" (Magano & Thomas, 2017, p. 3). Perhaps it was multiple fissures from M&As, not a single breach that generated the pervasive non-invariance. It would be deeply fascinating to isolate the moments when a group's measurement moves from invariance to non-invariance and to ask the questions of how and why a response shift occurred.

Implications for HRD Research and Practice

Throughout the dissertation, assertions about OD's role in the project have been very explicit. OD functions as a subset of HRD, though it may be unclear what HRD practitioners and academicians might further extrapolate from this research project; therefore, this section offers an unambiguous take on how this study supports HRD research and practice. HRD practitioners hold the core belief that organizations are made by humans and rely on their expertise to achieve desired outcomes. Professionals and researchers believe that expertise can be developed through short-term and long-term interventions. Both groups regularly advocate for individuals, groups, work processes, and organizational integrity. Because of the role of HRD within organizations, there is often access to rich organizational information requiring the responsibility to utilize it judiciously on behalf of many stakeholders (Swanson, Holton E. & Holton E.F., 2001). Practitioners and scholars alike must realize that all parties would prefer instruments and the groups they measure to show invariance. When surveys do not, value is still earned in the process and even the results, but caution is advised when using the findings to effectuate strategic and cultural change.

Implications for Practice

Engagement is a topic of great theoretical and commercial interest. Commercial firms capitalize on that interest by offering clients instruments to analyze staff engagement, and thus the organization. Professionals supporting an organization as an internal consultant must realize their dual, sometimes dueling roles to deliver results and present reality for decision-makers to consider and react to. Moreover, it is rarely an

easy choice to do both simultaneously within organizational resource and cultural constraints. Practitioners should know that the scales and instruments developed internally or externally are often flawed, making them poor choices to drive meaningful change. The counterbalance is to use caution when interpreting and enacting interventions since the EES and many other instruments can be employed appropriately to give a broad perspective and gain valuable feedback from employees. A localized approach for leveraging data from generic instruments would be to surface what a construct like engagement means locally, rather than reacting to item-level responses. An unintended consequence is that over-reliance on an unvalidated scale to enact deep organizational shifts is not likely to conclude the desired effect, resulting in poor ROI and waste. The incumbency for the deployment of instruments and intervention design lies with the organization and its gate-keepers.

External HRD consultants and vendors' services and products are necessary as internal consultants may not have time, capacity, or knowledge to design instruments. Moreover, external consultants and vendors are well-intentioned when offering their wares, even if they do not develop deeply academic work products. So, perhaps they should temper their sales pitches about the validity, veracity, and effectiveness of their products to the degree to which their instruments have undergone rigorous scholarly testing, which is not likely to be popular with the sales team. External consultants can use this study to see how their instruments, when built appropriately and interpreted correctly, clarify organizational phenomena. Surveys, when built poorly, over-simplify organizational phenomena and can confound meaning and even purpose. The increasing

decline witnessed in the mean and Percent Engaged scores were an indication that SystemTex was not likely to get back its engagement track without some organization-level calibration of what it means to be engaged and what the score means to each facility and the whole system. SystemTex might have redirected their intervention efforts with specific and appropriate guidance from the vendor. While that outcome is not guaranteed, SystemTex will never have the chance to find out. An unintended consequence of the deployment of instruments that do not support organizational improvement in the phenomena they measure is further disillusionment with the instrument, the survey process, and its outcomes. The incumbency for well-architected, rigorous instrumentation belongs to those who construct them.

Implications for Theory

HRD researchers should comprehend academic/business partnerships' value, not perpetuate the state of affairs as two impossible dichotomies. No one wants bad data, and the measurement process is the fulcrum between organizational theory and data (Vandenberg & Lance, 2000). Measurement invariance of multi-item composite scales like the EES ensures the presence, equivalent conceptualization, and comparable scaling of latent constructs to assure accurate group comparisons. Scientists must find ways to work with practitioners (internal and external consultants) who desire measurement, though less capable of ascertaining it. What is at stake are the broader scientific inferences that can be made (Vandenberg & Lance, 2000), the generalizability of our discoveries, and the credibility of HRD.

Conclusion

HRD is an applied discipline where theory must be practical (Lynham, 2002). Golembiewski, Billingsley, and Yeager (1976) offer three types of transformations resulting from self-reporting behavior. At SystemTex, the Engagement mean scores were central to an organizational narrative intended to drive practical, measurable change; the EES was not a thought experiment. The first, alpha change, shows the change in the observed means' actual level, measured perhaps before and after an intervention. It is also the focus of most evaluations and interventions (Nye et al., 2010).

Beta and gamma fluctuations represent non-invariant measurement changes and are analyzed using configural, metric, and scalar invariance, though the sequences do not pair up exactly (Nye et al., 2010). Beta change happens when responses to survey items are "recalibrated, and the intervals between the response options are changed" (p. 1560). In MGCFA, beta changes represent scalar differences. Gamma variations occur when an idea is reconceptualized or when respondents shift their reference frames after experiencing an intervention (p. 1560). Golembiewski et al. (1976) long ago pointed out that we cannot consistently or reasonably distinguish between the three types of transformations consistently (p. 143). Contemporary invariance and equivalence scholars might add that researchers cannot distinguish those changes without rigorously practiced measurement science.

If the organizational activity between survey cycles at SystemTex is framed as interventional, we have yet another way of explaining how the difference observed in configural, metric, and scalar MI may have occurred. Each EES administration was a

pre-assessment of the upcoming interventions, and the subsequent survey was a post-assessment of the interventional work that preceded it. 2011 was a starting point for the EES but not the starting point for Engagement, which was already complexified by differences in construct and scale meanings.

Engagement does not appear to be a fad for academicians or practitioners (Saks, 2006). Volumes of engagement research indicate the topic is many years past its point of origin; industry and academics continue to discuss, search, and develop procedures to operationalize it. It may be time to ensure that the employee engagement construct can fulfill its promise of competitive advantage for organizations and leaders (Gallant & Martins, 2018; Shuck & Rose, 2013). To do so, over-simplified, mean-level understanding will not suffice in many instances.

Complexities like MI, Engagement, and organizational change are difficult. Complexity renders interpretation of organizational phenomena obscure, communication fragmented, and leaders overwrought. It can be difficult to unify people around a strategy when it is difficult to determine what direction the organization requires. Nevertheless, complexity is also more nuanced, dynamic, and comprehensive. Organization science will benefit when organizational leaders, HRD professionals, and instrument architects recognize what MI research realizes—that organizational effectiveness lies not in making measurement simpler, but in measuring what best illuminates reality. That will be a true leap forward.

REFERENCES

Aguinis, H., Pierce, C. A., Bosco, F. A., & Muslin, I. S. (2009). First decade of Organizational Research Methods: Trends in design, measurement, and data-analysis topics. *Organizational Research Methods*, *12*(1), 69–112. https://doi.org/10.1177/1094428108322641

Alarcon, G. M., & Edwards, J. M. (2011). The relationship of engagement, job satisfaction and turnover intentions. *Stress and Health*, *27*(3), e294–e298. https://doi.org/10.1002/smi.1365

Angeli, F., & Maarse, H. (2012). Mergers and acquisitions in Western European health care: Exploring the role of financial services organizations. *Health Policy*, *105*(2–3), 265–272. ScienceDirect. Retrieved from ScienceDirect.

Appelbaum, S. H., Karelis, C., Henaff, A. L., & McLaughlin, B. (2017). Resistance to change in the case of mergers and acquisitions: Part 2. *Industrial and Commercial Training*, *49*(3), 139–145. edsemr. https://doi.org/10.1108/ICT-05-2016-0033

Assunção, H., Lin, S.-W., Sit, P.-S., Cheung, K.-C., Harju-Luukkainen, H., Smith, T., … Marôco, J. (2020). University Student Engagement Inventory (USEI): Transcultural Validity Evidence Across Four Continents. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02796

Ball, P. (2016). The Tyranny of Simple Explanations. Retrieved from Pocket website: https://getpocket.com/explore/item/the-tyranny-of-simple-explanations

Barnes, D., & Collier, E. (2013). Investigating work engagement in the service environment. *Journal of Services Marketing*, *27*(6), 485–499. https://doi.org/10.1108/JSM-01-2012-0021

Bartels, J., Pruyn, A., & Jong, M. de. (2009). Employee identification before and after an internal merger: A longitudinal analysis. *Journal of Occupational and Organizational Psychology*, *82*(1), 113–128. https://doi.org/10.1348/096317908X283770

Baynham, P. (2011). Engagement in M&A deals: A best practice guide. *Strategic HR Review*, *10*(3), 12–18. Human Resources Abstracts. Retrieved from Human Resources Abstracts.

Bialosiewicz, S., Murphy, K., & Berry, T. (2013). *An introduction to measurement invariance testing: Resource packet for participants. Do our measures measure up? The critical role of measurement invariance.* (p. 37) [American Evaluation Assn. Demonstration Session]. Washington DC: Claremont Graduate University.

Brown, T. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford publications.

Buckingham, M., & Coffman, C. (1999). *First, break all the rules: What the world's greatest managers do differently*. Simon & Schuster.

Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, *15*(1), 17–32. https://doi.org/10.1177/1094428110392383

Carter, M. R., & Tourangeau, A. E. (2012). Staying in nursing: What factors determine whether nurses intend to remain employed? *Journal of Advanced Nursing (John Wiley & Sons, Inc.)*, *68*(7), 1589–1600. Retrieved from asn.

Chakraborty, R. (2017). Configural, metric and scalar invariance measurement of academic delay of gratification scale. *IJARS International Journal of Humanities and Social Studies*, *3*(3). https://doi.org/10.20908/ijarsijhss.v3i3.7903

Chungkham, H. S., Ingre, M., Karasek, R., Westerlund, H., & Theorell, T. (2013). Factor structure and longitudinal measurement invariance of the demand control support model: An evidence from the Swedish Longitudinal Occupational Survey of Health (SLOSH). *PLoS ONE*, *8*(8), e70541. https://doi.org/10.1371/journal.pone.0070541

Coghlan, D., & Shani, A. B. (2014). Creating action research quality in organization development: Rigorous, reflective and relevant. *Systemic Practice and Action Research*, *27*(6), 523–536. https://doi.org/10.1007/s11213-013-9311-y

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, *12*(4), 381–398. https://doi.org/10.1037/1082-989X.12.4.381

Cox, N. J. (2005). A brief history of Stata on its 20th anniversary. *The Stata Journal*, *5*(1), 2–18.

Creasy, T., & Kinard, J. (2013). Health care mergers and acquisitions: Implications of Robbers Cave Realistic Conflict Theory and Prisoner's Dilemma Game Theory.

The Health Care Manager, *32*(1), 58–68. APA PsycInfo (2013-10099-005). https://doi.org/10.1097/HCM.0b013e31827edadd

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281.

Das, R., Jain, K. K., & Mishra, S. K. (2018). Archival research: A neglected method in organization studies. *Benchmarking: An International Journal*, *25*(1), 138–155. https://doi.org/10.1108/BIJ-08-2016-0123

de Beurs, D. P., Fokkema, M., de Groot, M. H., de Keijser, J., & Kerkhof, A. J. (2015). Longitudinal measurement invariance of the Beck Scale for Suicide Ideation. *Psychiatry Research*, *225*(3), 368–373. Retrieved from edselp.

Drake, T. J. (2012). *Assessing employee engagement: A comparison of the job engagement scale and the Utrecht Work Engagement Scale*. Colorado State University, Fort Collins, Colorado.

Dranove, D., & Lindrooth, R. (2003). Hospital consolidation and costs: Another look at the evidence. *Journal of Health Economics*, *22*(6), 983–997. https://doi.org/10.1016/j.jhealeco.2003.05.001

Eby, R. (2009, August). Aronson LLC [Industry]. Retrieved January 26, 2020, from Guidance on non-profit mergers and acquisitions website: https://aronsonllc.com/guidance-on-non-profit-mergers-and-acquisitions/

Edwards, M. R., Lipponen, J., Edwards, T., & Hakonen, M. (2017). Trajectories and antecedents of integration in mergers and acquisitions: A comparison of two

longitudinal studies. *Human Relations*, *70*(10), 1258–1290. Business Source Ultimate. Retrieved from Business Source Ultimate.

Espinosa, J. E. (2016). *Assessing the Factorial Validity, Measurement Invariance, and Latent Mean Differences of a Second-Order, Multidimensional Model of Academic and Social College Course Engagement: A Comparison Across Course Format, Ethnic Groups, and Economic Status* (University of California). University of California, Santa Barbara. Retrieved from http://proxy.library.tamu.edu/login?url=https://search.proquest.com/docview/1867755400?accountid=7082

Febriani, D. M., & Yancey, G. B. (2019). The effect of integration approaches and human resources initiatives on changes in organizational culture and employee attitudes during a merger. *The Psychologist-Manager Journal*, *22*(2), 108–131. APA PsycInfo (2019-32851-004). https://doi.org/10.1037/mgr0000084

Fox, J., & Leanage, A. (2016). R and the Journal of Statistical Software. *Journal of Statistical Software*, *73*(2). https://doi.org/10.18637/jss.v073.i02

Gallant, W. H., & Martins, N. (2018). Measurement invariance of employee engagement across race groups. *International Journal of Productivity & Performance Management*, *67*(9), 1463. Complementary Index. Retrieved from Complementary Index.

Gaynor, M. (2018). *Consolidation and competition in US health care*. Presented at the Leadership Consortium for a Value & Science-Driven Health System National Academy of Medicine, Washington DC.

Giannoulis, C. (2008, 2020). Four Common Misconceptions in Exploratory Factor Analysis. Retrieved from https://www.theanalysisfactor.com/four-common-misconceptions-efa/

Gleibs, I. H., Mummendey, A., & Noack, P. (2008). Predictors of change in postmerger identification during a merger process: A longitudinal study. *Journal of Personality and Social Psychology*, *95*(5), 1095–1112. (2008-14857-007). https://doi.org/10.1037/0022-3514.95.5.1095

Gleibs, I. H., Noack, P., & Mummendey, A. (2010). We are still better than them: A longitudinal field study of ingroup favouritism during a merger. *European Journal of Social Psychology*, *40*(5), 819–836. https://doi.org/10.1002/ejsp.652

Golembiewski, R. T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *The Journal of Applied Behavioral Science, 12*(2), 133–157.

Guerrero, S. (2008). Changes in employees' attitudes at work following an acquisition: A comparative study by acquisition type. *Human Resource Management Journal*, *18*(3), 216–236. https://doi.org/10.1111/j.1748-8583.2008.00068.x

Gupta, M., & J, A. (2017). Development of employee engagement model in a tertiary care hospital. *International Journal of Nursing Education*, *9*(4), 77. Complementary Index. Retrieved from Complementary Index.

Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology*, *87*(2), 268.

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, *18*(3), 66–67. https://doi.org/10.1136/eb-2015-102129

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3), 117–144. https://doi.org/10.1080/03610739208253916

Ismail, T. H., Abdou, A. A., & Annis, R. M. (2011). Review of literature linking corporate performance to mergers and acquisition. *The Review of Financial and Accounting Studies*, *1*(1), 89–104.

Junni, P., & Sarala, R. M. (2013). The role of absorptive capacity in acquisition knowledge transfer. *Thunderbird International Business Review*, *55*(4), 419–438. Retrieved from bth.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, *33*(4), 692–724. Retrieved from bth.

Kim, E. S., & Willson, V. L. (2014). Measurement invariance across groups in latent growth modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 408–424. https://doi.org/10.1080/10705511.2014.915374

King, D. R., Dalton, D. R., Daily, C., & Covin, J. G. (2004). Meta-analyses of post-acquisition performance: Indications of unidentified moderators. *Strategic Management Journal*, *25*(2).

Kjekshus, L., & Hagen, T. (2007). Do hospital mergers increase hospital efficiency? Evidence from a National Health Service country. *Journal of Health Services*

Research & Policy, *12*(4), 230–235. MEDLINE Complete (17925075). Retrieved from MEDLINE Complete. (17925075)

Kline, R. B. (1998). Software review: Software programs for structural equation modeling: Amos, EQS, and LISREL. *Journal of Psychoeducational Assessment*, *16*(4), 343–364.

Larsson, R., & Risberg, A. (1998). Cultural awareness and national versus corporate barriers to acculturation. *Cultural Dimensions of International Mergers and Acquisitions*, 39–56.

Larsson, R., & Finkelstein, S. (1999). Integrating strategic, organizational, and human resource perspectives on mergers and acquisitions: A case survey of synergy realization. *Organization Science*, *10*(1), 1–26. https://doi.org/10.1287/orsc.10.1.1

Little, T. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.

Liu, Y., Millsap, R. E., West, S. G., Tein, J.-Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods*, *22*(3), 486–506. https://doi.org/10.1037/met0000075

Lubatkin, M., Schweiger, D., & Weber, Y. (1999). Top management turnover in related M&A's: An additional test of the theory of relative standing. *Journal of Management*, *25*(1), 55–73. https://doi.org/10.1016/S0149-2063(99)80003-8

Lynham, S. A. (2002). The general method of theory-building research in applied disciplines. *Advances in Developing Human Resources*, 21.

Macey, W. H., & Schneider, B. (2008). The meaning of employee engagement. *Industrial and Organizational Psychology*, *1*(1), 3–30. https://doi.org/10.1111/j.1754-9434.2007.0002.x

Magano, K. D., & Thomas, A. (2017). Organizational change and the psychological contract at a pharmaceutical company. *South African Journal of Human Resource Management*, *15*(1), 1–10. Retrieved from bsu.

Malhotra, N. K., Schaller, T. K., & Patil, A. (2017). Common method variance in advertising research: When to be concerned and how to control for it. *Journal of Advertising*, *46*(1), 193–212. https://doi.org/10.1080/00913367.2016.1252287

Marsh, H. W., & Byrne, B. M. (1993). Confirmatory factor analysis of multitrait-multimethod self-concept data: Between-group and within-group invariance constraints. *Multivariate Behavioral Research*, *28*(3), 313–449. https://doi.org/10.1207/s15327906mbr2803_2

Martin, L. R., & Friedman, H. S. (2000). Comparing personality scales across time: An illustrative study of validity and consistency in life-span archival data. *Journal of Personality*, *68*(1), 85–110. Retrieved from asn.

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, *3*(1), 97–110. https://doi.org/10.21500/20112084.854

McLean, G. N., & Egan, T. M. (2008). Applying organization development tools in scenario planning. *Advances in Developing Human Resources*, *10*(2), 240–257. https://doi.org/10.1177/1523422307313328

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Merve Ünal, Z., & Turgut, T. (2015). The buzzword: Employee engagement. Does person-organization fit contribute to employee engagement? *Iranian Journal of Management Studies*, *8*(2), 157–179. https://doi.org/10.22059/ijms.2015.52320

Meyer, J. (2020, April). *April 2020 Structure Equation Modeling Workshop*. Retrieved from https://www.theanalysisfactor.com/

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*. Sage publications.

Milfont, T., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111–130. https://doi.org/10.21500/20112084.857

Monte Carlo randomization. (n.d.). Retrieved from BioMedware Geospatial Research and Software website: https://www.biomedware.com/files/documentation/clusterseer/MCR/Monte_Carlo.htm

Mplus—General Description. (2020, April 3). Retrieved April 3, 2020, from https://www.statmodel.com/features.shtml

Newsom. (2020). *Missing Data and Missing Data Estimation in SEM*. Retrieved from http://web.pdx.edu/~newsomj/semclass/ho_missing.pdf

Nye, C. D., Brummel, B. J., & Drasgow, F. (2010). Too good to be true? Understanding change in organizational outcomes. *Journal of Management*, *36*(6), 1555–1577. https://doi.org/10.1177/0149206310376326

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Popli, S., & Rizvi, I. A. (2016). Drivers of employee engagement: The role of leadership style. *Global Business Review*, *17*(4), 965–979. https://doi.org/10.1177/0972150916645701

Rasheed, A., Kahn, S., & Ramzan, M. (2013). Antecedents and consequences of employee engagement: The case of Pakistan. *Journal of Business Studies Quarterly*, *4*(4), 183.

Rausch, J. R. (2009). Investigating change in intraindividual factor structure over time. *Applied Psychological Measurement*, *33*(4), 266–284.

Remington, K., & Pollack, J. (2012). Complexity, decision-making and requisite variety. Retrieved March 2, 2020, from IDOCPUB website: https://idoc.pub/documents/complexity-decision-making-and-requisite-variety-remington-and-pollack-qvndx2xy7j4x

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. https://doi.org/10.1037/a0029315

Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal*, *53*(3), 617–635. https://doi.org/10.5465/amj.2010.51468988

Rivera, R., Fitzpatrick, J., & Boyle, S. (2011). Closing the RN engagement gap: Which drivers of engagement matter? *The Journal of Nursing Administration*, *41*(6), 265–272. Retrieved from edswsc.

Robison, J. (2006). A caterpillar dealer unearths employee engagement. *Gallup Management Journal*, *3*, 1–6.

Rothbard, N. P. (2001). Enriching or depleting? The dynamics of engagement in work and family roles. *Administrative Science Quarterly*, *46*(4), 655–684.

Saks, A. M. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology*, *21*(7), 600–619. https://doi.org/10.1108/02683940610690169

Schaufeli, W. B., Salanova, M., González-romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, *3*, 71–92.

Schweiger, D. M., Ivancevich, J. M., & Power, F. R. (1987). Executive actions for managing human resources before and after acquisition. *Academy of*

Management Perspectives, *1*(2), 127–138. https://doi.org/10.5465/ame.1987.4275830

Seymour, A., & Dupré, K. (2008). Advancing employee engagement through a healthy workplace strategy. *Journal of Health Services Research & Policy*, *13*(1), 35–40.

Shuck, B. (2010). *Employee engagement: An examination of antecedent and outcome variables*. Florida International University.

Shuck, B., Adelson, J. L., & Reio, T. G. (2017). The employee engagement scale: Initial evidence for construct validity and implications for theory and practice. *Human Resource Management*, *56*(6), 953–977. Business Source Ultimate. Retrieved from Business Source Ultimate.

Shuck, B., & Reio, T. G. (2014). Employee engagement and well-being: A moderation model and implications for practice. *Journal of Leadership & Organizational Studies*, *21*(1), 43–58. https://doi.org/10.1177/1548051813494240

Shuck, B., & Rose, K. (2013). Reframing employee engagement within the context of meaning and purpose: Implications for HRD. *Advances in Developing Human Resources*, *15*(4), 341–355. https://doi.org/10.1177/1523422313503235

Shuck, B., Twyford, D., Reio, T. G., & Shuck, A. (2014). Human resource development practices and employee engagement: Examining the connection with employee turnover intentions. *Human Resource Development Quarterly*, *25*(2), 239–270. Retrieved from bsu.

Shuck, B., & Wollard, K. (2010). Employee engagement and HRD: A seminal review of the foundations. *Human Resource Development Review*, *9*(1), 89–110. https://doi.org/10.1177/1534484309353560

Singh, J. (2013). Importance of employee engagement: A literature review. *International Journal of Computer Science and Communication Engineering*, 1–4.

Sirower, M. L., & Sahni, S. (2006). Avoiding the "synergy trap": Practical guidance on M&A decisions for CEOs and Boards. *Journal of Applied Corporate Finance*, *18*(3), 83–95.

Sloan, F. A., Ostermann, J., & Conover, C. J. (2003). Antecedents of hospital ownership conversions, mergers, and closures. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *40*(1), 39–56. https://doi.org/10.5034/inquiryjrnl_40.1.39

Stahl, G. K., Kremershof, I., & Larsson, R. (2004). *Trust dynamics in mergers and acquisitions: A case survey*. Insead.

Stahl, G. K., & Voigt, A. (2008). Do cultural differences matter in mergers and acquisitions? A tentative model and examination. *Organization Science*. Retrieved from pubsonline.informs.org

Straatmann, V. S., Almquist, Y. B., Oliveira, A. J., Rostila, M., & Lopes, C. S. (2018). Cross-lagged structural equation models for the relationship between health-related state and behaviours and body bullying in adolescence: Findings from longitudinal study ELANA. *PLOS ONE*, *13*(1), e0191253. https://doi.org/10.1371/journal.pone.0191253

Strumwasser, S., & Virkstis, K. (2015). Meaningfully incorporating staff input to enhance frontline engagement. *JONA: The Journal of Nursing Administration*, *45*(4), 179–182. https://doi.org/10.1097/NNA.0000000000000179

Suh, T., Houston, M. B., Barney, S. M., & Kwon, I.-W. G. (2011). The impact of mission fulfillment on the internal audience: Psychological job outcomes in a services setting. *Journal of Service Research*, *14*(1), 76–92. Retrieved from bsu.

Swanson, R. A., Holton, E., & Holton, E. F. (2001). *Foundations of human resource development.* Berrett-Koehler Publishers.

Teerikangas, S., & Very, P. (2006). The culture–performance relationship in M&A: From yes/no to how. *British Journal of Management*, *17*(S1), S31–S48. https://doi.org/10.1111/j.1467-8551.2006.00477.x

Tang, T. L. P., & Timmer, L. S. (2008). All in the name of change: Effects of organizational change on performance measures of customer service in the health care industry. *International Journal of Organization Theory & Behavior (PrAcademics Press)*, *11*(2), 220–239. Retrieved from SocINDEX with Full Text.

The Advisory Board Survey Solutions. (2011, June). *Advisory Board Employee Engagement Survey Findings; June 22, 2011*.

The Advisory Board Survey Solutions. (2014, July). *Overview of 2014 Employee Perspective Survey Findings*. Presented at the Sr. Leadership Meeting.

The Advisory Board Survey Solutions. (2015, June). *2015 Employee Engagement Survey System-Wide Results*.

The Advisory Board Survey Solutions Presentation to WellSpan Health. (2015, June). *WellSpan Health 2015 Employee Engagement Results*. Retrieved from https://content.wellspan.org/employeeengage/Presentation%20-%20Overview%20of%20WS%202015%20Employee%20Engagement%20Results.pdf

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, *4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Towers Perrin Talent Report. (2003). *The 2003 Towers Perrin talent report: Working today: Understanding what drives employee engagement* (p. 35). Stamford, CT.

Trochim, W. M. K. (2020, March). What is the Nomological Net? Retrieved from Research Methods Knowledge Base website: https://conjointly.com/kb/nomological-network/

Tuch, C., & O'Sullivan, C. (2007). The impact of acquisitions on firm performance: A review of the evidence. *International Journal of Management Reviews*, *9*(2), 141–170.

UCLA Statistical Consulting Group. (n.d.). Factor Analysis: SPSS Annotated Output. Retrieved from UCLA Statistical Consulting Group website: https://stats.idre.ucla.edu/spss/output/factor-analysis/#:~:text=The%20factor%20structure%20matrix%20represents,done%20is%20an%20oblique%20rotation.

van de Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Measurement invariance. *Frontiers in Psychology*, *6*, 1064. https://doi.org/10.3389/fpsyg.2015.01064

Vance, R. J. (2006). Employee engagement and commitment. *SHRM Foundation*, 1–53.

Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, *5*(2), 139–158. https://doi.org/10.1177/1094428102005002001

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Vestergren, P., Rönnlund, M., Nyberg, L., & Nilsson, L.-G. (2012). Multigroup confirmatory factor analysis of the cognitive dysfunction questionnaire: Instrument refinement and measurement invariance across age and sex. *Scandinavian Journal of Psychology*, *53*(5), 390–400. https://doi.org/10.1111/j.1467-9450.2012.00970.x

Vita, M. G., & Sacher, S. (2001). The competitive effects of not-for-profit hospital mergers: A case study. *Journal of Industrial Economics*, *49*(1), 63–84. https://doi.org/10.1111/1467-6451.00138

Waight, C. L. (2015). Learning During the Integration Phase of Mergers and Acquisitions: Perspectives From Learning and Development Professionals. *Performance Improvement Quarterly*, *28*(1), 7. Retrieved from edb.

Waller, N. G. (1993). Software review: Seven confirmatory factor analysis programs: EQS, EzPATH, LINCS, LISCOMP, LISREL 7, SI MPLIS, and CALIS. *Applied Psychological Measurement*, *17*(1), 73–100. https://doi.org/10.1177/014662169301700112

Weil, T. (2010). Hospital mergers: A panacea? *Journal of Health Services Research & Policy*, *15*(4), 251–253. https://doi.org/10.1258/jhsrp.2010.010035

Welch, S., & Welch, J. (2006). How Healthy Is Your Company? To take the pulse of any business, there are three vital signs that really matter. *Business Week*, (3983). Retrieved from http://proxy.library.tamu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsbig&AN=edsbig.A145343086&site=eds-live

Xu, K. (2020, February). *Multiple group measurement: Invariance analysis in Lavaan*. Presented at the Cambridge Mass. Cambridge Mass. Retrieved from https://slideplayer.com/slide/2332783/

Yuan, K.-H., & Chan, W. (2016). Measurement invariance via multigroup SEM: Issues and solutions with chi-square-difference tests. *Psychological Methods*, *21*(3), 405–426. https://doi.org/10.1037/met0000080

Zigarmi, D., Nimon, K., Houson, D., Witt, D., & Diehl, J. (2009). Beyond engagement: Toward a framework and operational definition for employee work passion. *Human Resource Development Review*, *8*(3), 300–326. https://doi.org/10.1177/1534484309338171

APPENDIX A

ADVISORY BOARD EMPLOYEE ENGAGEMENT SURVEY

Statement ID	Statement Text
S1	I am kept informed of the organization's future plans and direction.
S2	I understand how my daily work contributes to the organization's mission.
S3	I know what is required to perform well in my job.
S4	My manager communicates messages that my coworkers need to hear, even when the information is unpleasant.
S5	My ideas and suggestions are valued by my organization.
S6	My manager is open and responsive to staff input.
S7	I have the right amount of independence in my work.
S8	I have good personal relationships with coworkers in my unit/department.
S9	My coworkers do a good job.
S10	I receive the necessary support from employees in my unit/department to help me succeed in my work.
S11	I receive the necessary support from employees in other units/departments to help me succeed in my work.
S12	The actions of executives in my organization reflect our mission and values.
S13	Abusive behavior is not tolerated at my organization.
S14	My manager stands up for the interests of my unit/department.
S15	Conflicts are resolved fairly in my unit/department.
S16	I am interested in promotion opportunities in my unit/department.
S17	I receive effective on the job training.
S18	I have helpful discussions with my manager about my career.
S19	My manager helps me learn new skills.
S20	My most recent performance review helped me to improve.
S21	Training and development opportunities offered by my organization have helped me to improve.
S22	If I wanted to explore other jobs within the organization, my manager would help me do that.
S23	My current job is a good match for my skills.
S24	I receive regular feedback from my manager on my performance.
S25	My organization pays me fairly for my job.
S26	The benefits provided by my organization (such as health care, retirement savings, etc.) meet my needs.
S27	I have job security.
S28	My organization helps me deal with stress and burnout.
S29	My organization recognizes employees for excellent work.
S30	Executives at my organization respect the contributions of my unit/department.
S31	My organization provides excellent care to patients.
S32	My organization provides excellent customer service to patients.
S33	My organization gives back to the community.

Statement ID	Statement Text
S34	I have a manageable workload.
S35	My unit/department has enough staff.
S36	I believe in my organization's mission.
S37	My organization supplies me with the equipment I need.
S38	My manager helps me balance my job and personal life.
S39	My organization does a good job of selecting and implementing new technologies to support my work.
S40	My organization supports employee safety.
S41	My organization understands and respects differences among employees (gender, race, age, religion, etc.).
S42	Over the past year I have never been asked to do something that compromises my values.
S43	I would recommend this organization to my friends as a great place to work.
S44	This organization inspires me to perform my best.
S45	I am likely to be working for this organization three years from now.
S46	I am willing to put in a great deal of effort in order to help this organization succeed.

Note. The survey items are reprinted with permission from The Advisory Board. See Appendix D.

APPENDIX B

EFA COVARIANCE MATRIX

	S1	S2	S3	S4	S5
S1	0.966				
S2	0.956	0.968			
S3	0.955	0.956	0.966		
S4	0.956	0.957	0.956	0.968	
S5	0.952	0.954	0.953	0.954	0.965
S6	0.956	0.956	0.956	0.957	0.955
S7	0.957	0.959	0.958	0.959	0.956
S8	0.952	0.954	0.954	0.955	0.953
S9	0.954	0.956	0.956	0.958	0.953
S10	0.955	0.956	0.956	0.956	0.952
S11	0.956	0.956	0.957	0.959	0.953
S12	0.950	0.951	0.953	0.952	0.948
S13	0.952	0.953	0.953	0.955	0.951
S14	0.954	0.956	0.955	0.956	0.953
S15	0.954	0.956	0.956	0.956	0.955
S16	0.954	0.956	0.955	0.956	0.953
S17	0.953	0.953	0.955	0.954	0.953
S18	0.953	0.953	0.956	0.956	0.951
S19	0.953	0.955	0.952	0.954	0.951
S20	0.949	0.949	0.949	0.951	0.948
S21	0.953	0.954	0.953	0.953	0.951
S22	0.951	0.951	0.950	0.953	0.948
S23	0.957	0.960	0.958	0.959	0.956
S24	0.955	0.955	0.956	0.956	0.953
S25	0.956	0.958	0.960	0.959	0.955
S26	0.955	0.956	0.957	0.956	0.953
S27	0.950	0.951	0.952	0.952	0.949
S28	0.951	0.952	0.951	0.954	0.949
S29	0.953	0.953	0.953	0.955	0.950
S30	0.955	0.956	0.957	0.956	0.953
S31	0.953	0.954	0.954	0.956	0.952
S32	0.951	0.953	0.955	0.956	0.953
S33	0.948	0.951	0.949	0.951	0.948
S34	0.955	0.956	0.956	0.956	0.953
S35	0.954	0.955	0.957	0.955	0.953
S36	0.952	0.953	0.952	0.952	0.950
S37	0.955	0.957	0.956	0.957	0.953
S38	0.955	0.957	0.955	0.957	0.952
S39	0.953	0.955	0.954	0.955	0.951
S40	0.956	0.959	0.958	0.958	0.955
S41	0.955	0.956	0.956	0.956	0.952
S42	0.954	0.955	0.957	0.956	0.953
S43	0.966	0.968	0.966	0.968	0.965
S44	0.966	0.968	0.966	0.968	0.965
S45	0.966	0.968	0.966	0.968	0.965
S46	0.966	0.968	0.966	0.968	0.965

	S6	S7	S8	S9	S10
S6	0.969				
S7	0.958	0.968			
S8	0.953	0.956	0.963		
S9	0.956	0.959	0.954	0.966	
S10	0.956	0.958	0.953	0.956	0.967
S11	0.958	0.958	0.955	0.956	0.956
S12	0.952	0.953	0.949	0.952	0.952
S13	0.953	0.956	0.954	0.952	0.954
S14	0.956	0.958	0.954	0.955	0.954
S15	0.955	0.959	0.954	0.956	0.956
S16	0.956	0.958	0.954	0.956	0.954
S17	0.956	0.956	0.952	0.955	0.952
S18	0.953	0.955	0.953	0.953	0.953
S19	0.956	0.956	0.952	0.955	0.954
S20	0.950	0.953	0.947	0.950	0.950
S21	0.953	0.956	0.953	0.953	0.955
S22	0.954	0.953	0.948	0.951	0.949
S23	0.959	0.960	0.957	0.957	0.958
S24	0.956	0.958	0.953	0.955	0.955
S25	0.958	0.960	0.956	0.958	0.957
S26	0.958	0.959	0.956	0.958	0.957
S27	0.951	0.953	0.949	0.951	0.951
S28	0.951	0.954	0.951	0.951	0.950
S29	0.955	0.956	0.953	0.954	0.953
S30	0.956	0.958	0.953	0.956	0.956
S31	0.956	0.957	0.953	0.956	0.955
S32	0.953	0.956	0.953	0.954	0.952
S33	0.949	0.951	0.948	0.949	0.948
S34	0.957	0.958	0.955	0.956	0.958
S35	0.956	0.956	0.952	0.955	0.955
S36	0.952	0.956	0.949	0.953	0.953
S37	0.958	0.959	0.956	0.957	0.957
S38	0.956	0.956	0.953	0.953	0.955
S39	0.954	0.956	0.952	0.955	0.953
S40	0.959	0.960	0.955	0.958	0.957
S41	0.957	0.956	0.953	0.953	0.954
S42	0.957	0.959	0.953	0.957	0.956
S43	0.969	0.968	0.963	0.966	0.967
S44	0.969	0.968	0.963	0.966	0.967
S45	0.969	0.968	0.963	0.966	0.967
S46	0.969	0.968	0.963	0.966	0.967

	S11	S12	S13	S14	S15
S11	0.967				
S12	0.953	0.961			
S13	0.954	0.949	0.964		
S14	0.956	0.951	0.953	0.966	
S15	0.955	0.953	0.955	0.956	0.965
S16	0.955	0.949	0.953	0.954	0.954
S17	0.954	0.949	0.950	0.953	0.954
S18	0.956	0.949	0.952	0.953	0.953
S19	0.954	0.949	0.951	0.954	0.954
S20	0.950	0.947	0.946	0.949	0.949
S21	0.952	0.951	0.951	0.953	0.953
S22	0.952	0.947	0.948	0.952	0.951
S23	0.959	0.953	0.956	0.958	0.958
S24	0.956	0.949	0.953	0.955	0.954
S25	0.960	0.953	0.957	0.958	0.958
S26	0.957	0.953	0.955	0.956	0.957
S27	0.953	0.947	0.948	0.951	0.951
S28	0.951	0.948	0.949	0.950	0.952
S29	0.955	0.950	0.951	0.954	0.953
S30	0.956	0.953	0.954	0.957	0.958
S31	0.956	0.952	0.952	0.955	0.956
S32	0.956	0.951	0.952	0.953	0.955
S33	0.949	0.945	0.947	0.948	0.949
S34	0.956	0.951	0.954	0.956	0.956
S35	0.955	0.952	0.951	0.953	0.953
S36	0.953	0.947	0.949	0.953	0.953
S37	0.956	0.951	0.954	0.956	0.956
S38	0.955	0.951	0.951	0.953	0.953
S39	0.956	0.952	0.953	0.955	0.956
S40	0.957	0.951	0.955	0.957	0.957
S41	0.956	0.953	0.953	0.955	0.956
S42	0.956	0.951	0.953	0.955	0.956
S43	0.967	0.961	0.964	0.966	0.965
S44	0.967	0.961	0.964	0.966	0.965
S45	0.967	0.961	0.964	0.966	0.965
S46	0.967	0.961	0.964	0.966	0.965

	S16	S17	S18	S19	S20
S16	0.966				
S17	0.953	0.965			
S18	0.952	0.952	0.962		
S19	0.954	0.952	0.952	0.965	
S20	0.947	0.948	0.947	0.947	0.960
S21	0.953	0.952	0.950	0.953	0.948
S22	0.950	0.948	0.950	0.950	0.945
S23	0.958	0.956	0.957	0.957	0.951
S24	0.956	0.954	0.954	0.955	0.948
S25	0.956	0.957	0.957	0.955	0.951
S26	0.956	0.954	0.953	0.956	0.950
S27	0.949	0.950	0.949	0.949	0.946
S28	0.951	0.951	0.950	0.951	0.946
S29	0.954	0.951	0.952	0.951	0.948
S30	0.956	0.954	0.955	0.956	0.949
S31	0.954	0.953	0.951	0.954	0.948
S32	0.954	0.953	0.952	0.952	0.948
S33	0.949	0.948	0.947	0.949	0.943
S34	0.956	0.954	0.953	0.955	0.949
S35	0.953	0.953	0.951	0.953	0.949
S36	0.951	0.949	0.949	0.952	0.947
S37	0.958	0.954	0.953	0.956	0.949
S38	0.953	0.952	0.954	0.952	0.950
S39	0.953	0.953	0.952	0.952	0.949
S40	0.957	0.957	0.955	0.956	0.950
S41	0.955	0.951	0.953	0.953	0.948
S42	0.953	0.953	0.953	0.953	0.949
S43	0.966	0.965	0.962	0.965	0.960
S44	0.966	0.965	0.962	0.965	0.960
S45	0.966	0.965	0.962	0.965	0.960
S46	0.966	0.965	0.962	0.965	0.960

	S21	S22	S23	S24	S25
S21	0.963				
S22	0.947	0.960			
S23	0.955	0.953	0.970		
S24	0.952	0.951	0.958	0.965	
S25	0.956	0.952	0.960	0.958	0.968
S26	0.954	0.951	0.959	0.957	0.959
S27	0.949	0.947	0.953	0.951	0.954
S28	0.950	0.947	0.954	0.951	0.953
S29	0.951	0.950	0.956	0.953	0.954
S30	0.954	0.951	0.960	0.956	0.959
S31	0.952	0.949	0.957	0.953	0.957
S32	0.951	0.947	0.956	0.954	0.958
S33	0.946	0.944	0.952	0.950	0.951
S34	0.955	0.950	0.959	0.957	0.959
S35	0.953	0.948	0.955	0.953	0.956
S36	0.950	0.947	0.955	0.953	0.954
S37	0.955	0.950	0.958	0.956	0.959
S38	0.953	0.951	0.957	0.953	0.956
S39	0.951	0.949	0.956	0.952	0.957
S40	0.955	0.952	0.961	0.958	0.960
S41	0.952	0.951	0.959	0.954	0.957
S42	0.952	0.950	0.958	0.954	0.958
S43	0.963	0.960	0.970	0.965	0.968
S44	0.963	0.960	0.970	0.965	0.968
S45	0.963	0.960	0.970	0.965	0.968
S46	0.963	0.960	0.970	0.965	0.968

	S26	S27	S28	S29	S30
S26	0.967				
S27	0.953	0.960			
S28	0.951	0.947	0.961		
S29	0.955	0.949	0.949	0.964	
S30	0.957	0.952	0.951	0.953	0.966
S31	0.955	0.951	0.952	0.953	0.956
S32	0.956	0.949	0.951	0.952	0.956
S33	0.949	0.946	0.948	0.946	0.951
S34	0.957	0.951	0.953	0.954	0.957
S35	0.956	0.950	0.951	0.952	0.954
S36	0.952	0.947	0.948	0.950	0.953
S37	0.958	0.952	0.952	0.954	0.956
S38	0.955	0.951	0.953	0.954	0.953
S39	0.955	0.951	0.951	0.953	0.956
S40	0.958	0.953	0.953	0.954	0.958
S41	0.955	0.951	0.951	0.953	0.957
S42	0.957	0.951	0.950	0.953	0.955
S43	0.967	0.960	0.961	0.964	0.966
S44	0.967	0.960	0.961	0.964	0.966
S45	0.967	0.960	0.961	0.964	0.966
S46	0.967	0.960	0.961	0.964	0.966

	S31	S32	S33	S34	S35
S31	0.966				
S32	0.955	0.964			
S33	0.947	0.950	0.958		
S34	0.956	0.955	0.950	0.966	
S35	0.953	0.954	0.948	0.954	0.966
S36	0.952	0.951	0.946	0.954	0.951
S37	0.956	0.955	0.949	0.957	0.956
S38	0.953	0.952	0.948	0.954	0.953
S39	0.953	0.953	0.947	0.953	0.955
S40	0.957	0.955	0.952	0.958	0.956
S41	0.954	0.953	0.948	0.956	0.953
S42	0.953	0.953	0.949	0.956	0.955
S43	0.966	0.964	0.958	0.966	0.966
S44	0.966	0.964	0.958	0.966	0.966
S45	0.966	0.964	0.958	0.966	0.966
S46	0.966	0.964	0.958	0.966	0.966

	S36	S37	S38	S39	S40	
S36	0.962					
S37	0.953	0.968				
S38	0.950	0.953	0.965			
S39	0.950	0.955	0.953	0.966		
S40	0.954	0.958	0.956	0.956	0.968	
S41	0.952	0.956	0.953	0.954	0.958	
S42	0.951	0.954	0.953	0.953	0.957	
S43	0.962	0.968	0.965	0.966	0.968	
S44	0.962	0.968	0.965	0.966	0.968	
S45	0.962	0.968	0.965	0.966	0.968	
S46	0.962	0.968	0.965	0.966	0.968	

	S41	S42	S43	S44	S45	S46
S41	0.966					
S42	0.954	0.964				
S43	0.966	0.964	1.000			
S44	0.966	0.964	1.000	1.000		
S45	0.966	0.964	1.000	1.000	1.000	
S46	0.966	0.964	1.000	1.000	1.000	1.000

APPENDIX C

CONFIGURAL GOODNESS OF FIT FOR INVARIANT PAIRINGS

Figure 23

2012 Hosp. 3 and 5 Model Fit

MODEL FIT INFORMATION

Number of Free Parameters	13
Loglikelihood	
H0 Value	-17328.305
H1 Value	-17327.621
Information Criteria	
Akaike (AIC)	34682.610
Bayesian (BIC)	34762.690
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	34721.382
Chi-Square Test of Model Fit	
Value	1.369
Degrees of Freedom	1
P-Value	0.2419
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.010
90 Percent C.I.	0.000 0.048
Probability RMSEA <= .05	0.963
CFI/TLI	
CFI	1.000
TLI	1.000
Chi-Square Test of Model Fit for the Baseline Model	
Value	8669.235
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value	0.002

Figure 24

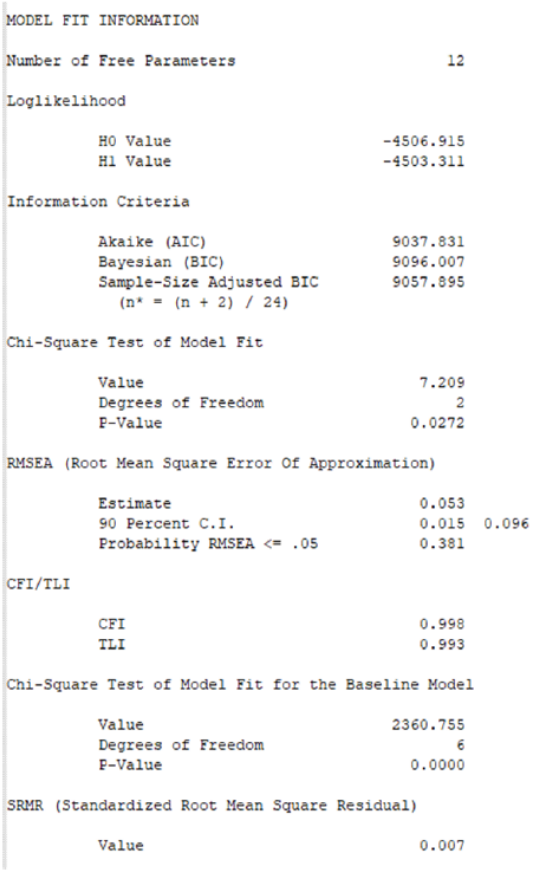
2012 Shared Services and Hosp. 4

MODEL FIT INFORMATION

Number of Free Parameters	12
Loglikelihood	
H0 Value	-3816.079
H1 Value	-3810.747
Information Criteria	
Akaike (AIC)	7656.158
Bayesian (BIC)	7712.100
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	7673.994
Chi-Square Test of Model Fit	
Value	10.664
Degrees of Freedom	2
P-Value	0.0048
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.074
90 Percent C.I.	0.035 0.121
Probability RMSEA <= .05	0.138
CFI/TLI	
CFI	0.995
TLI	0.985
Chi-Square Test of Model Fit for the Baseline Model	
Value	1768.259
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value	0.015

Figure 25

2012 Hosp. 3 and 4 Model Fit



MODEL FIT INFORMATION

Number of Free Parameters 12

Loglikelihood

H0 Value	-4506.915
H1 Value	-4503.311

Information Criteria

Akaike (AIC)	9037.831
Bayesian (BIC)	9096.007
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	9057.895

Chi-Square Test of Model Fit

Value	7.209
Degrees of Freedom	2
P-Value	0.0272

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.053	
90 Percent C.I.	0.015	0.096
Probability RMSEA <= .05	0.381	

CFI/TLI

CFI	0.998
TLI	0.993

Chi-Square Test of Model Fit for the Baseline Model

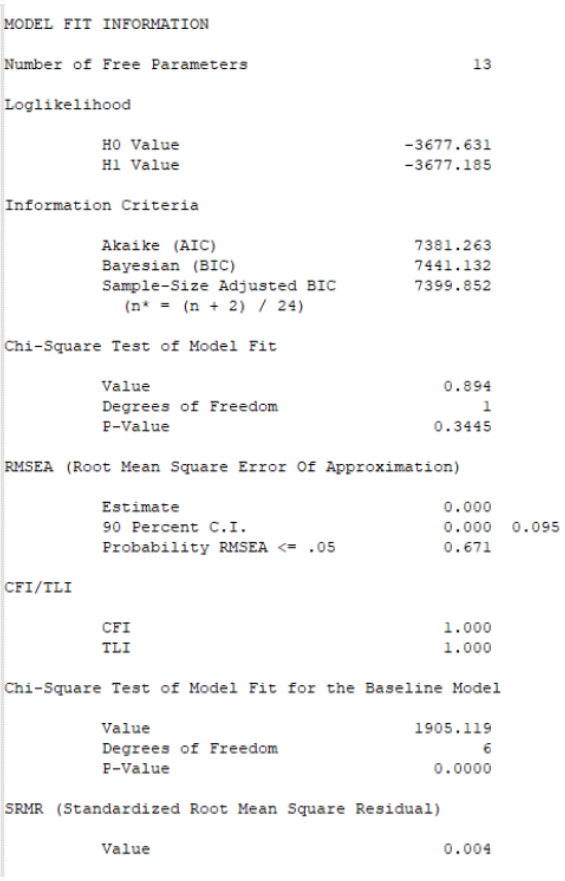
Value	2360.755
Degrees of Freedom	6
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.007

Figure 26

2013 Hosp 2 and Hosp. 4 Model Fit



MODEL FIT INFORMATION

Number of Free Parameters 13

Loglikelihood

H0 Value	-3677.631
H1 Value	-3677.185

Information Criteria

Akaike (AIC)	7381.263
Bayesian (BIC)	7441.132
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	7399.852

Chi-Square Test of Model Fit

Value	0.894
Degrees of Freedom	1
P-Value	0.3445

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.000	
90 Percent C.I.	0.000	0.095
Probability RMSEA <= .05	0.671	

CFI/TLI

CFI	1.000
TLI	1.000

Chi-Square Test of Model Fit for the Baseline Model

Value	1905.119
Degrees of Freedom	6
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.004

Figure 27

2014 Hosp. 2 and 5 Model Fit

MODEL FIT INFORMATION

Number of Free Parameters	13
Loglikelihood	
H0 Value	-16416.873
H1 Value	-16405.061
Information Criteria	
Akaike (AIC)	32859.746
Bayesian (BIC)	32938.729
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	32897.422
Chi-Square Test of Model Fit	
Value	23.624
Degrees of Freedom	1
P-Value	0.0000
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.084
90 Percent C.I.	0.057 0.1
Probability RMSEA <= .05	0.021
CFI/TLI	
CFI	0.997
TLI	0.981
Chi-Square Test of Model Fit for the Baseline Model	
Value	7017.394
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value	0.011

Figure 28

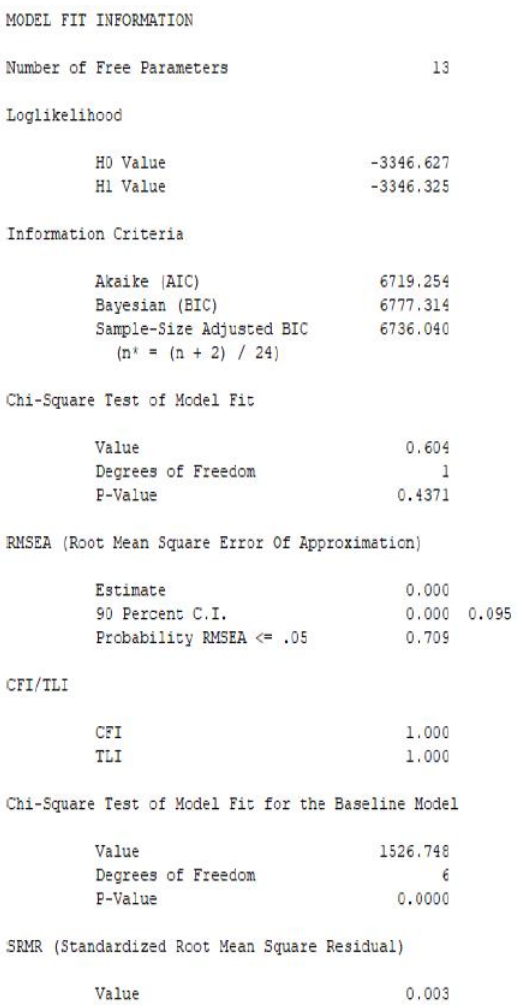
2014 Hosp. 4 and 5 Model Fit

MODEL FIT INFORMATION

Number of Free Parameters	13
Loglikelihood	
H0 Value	-16423.878
H1 Value	-16411.503
Information Criteria	
Akaike (AIC)	32873.757
Bayesian (BIC)	32952.792
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	32911.485
Chi-Square Test of Model Fit	
Value	24.751
Degrees of Freedom	1
P-Value	0.0000
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.086
90 Percent C.I.	0.059 0.117
Probability RMSEA <= .05	0.016
CFI/TLI	
CFI	0.997
TLI	0.980
Chi-Square Test of Model Fit for the Baseline Model	
Value	7000.730
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value	0.011

Figure 29

2014 Hosp. 2 and 4 Model Fit



MODEL FIT INFORMATION

Number of Free Parameters 13

Loglikelihood

H0 Value	-3346.627
H1 Value	-3346.325

Information Criteria

Akaike (AIC)	6719.254
Bayesian (BIC)	6777.314
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	6736.040

Chi-Square Test of Model Fit

Value	0.604
Degrees of Freedom	1
P-Value	0.4371

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.000	
90 Percent C.I.	0.000	0.095
Probability RMSEA <= .05	0.709	

CFI/TLI

CFI	1.000
TLI	1.000

Chi-Square Test of Model Fit for the Baseline Model

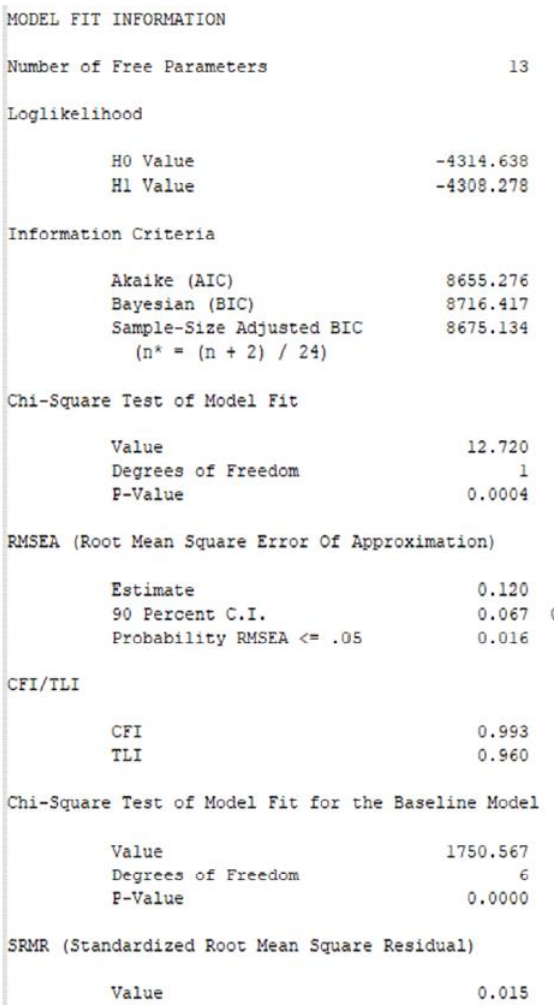
Value	1526.748
Degrees of Freedom	6
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.003

Figure 30

2015 Shared Services and Hosp. 2 Model Fit



MODEL FIT INFORMATION

Number of Free Parameters 13

Loglikelihood

H0 Value	-4314.638
H1 Value	-4308.278

Information Criteria

Akaike (AIC)	8655.276
Bayesian (BIC)	8716.417
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	8675.134

Chi-Square Test of Model Fit

Value	12.720
Degrees of Freedom	1
P-Value	0.0004

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.120
90 Percent C.I.	0.067
Probability RMSEA <= .05	0.016

CFI/TLI

CFI	0.993
TLI	0.960

Chi-Square Test of Model Fit for the Baseline Model

Value	1750.567
Degrees of Freedom	6
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.015

Figure 31

2015 Shared Services and Hosp. 4 Model Fit

MODEL FIT INFORMATION

Number of Free Parameters	13
Loglikelihood	
H0 Value	-4361.010
H1 Value	-4356.999
Information Criteria	
Akaike (AIC)	8748.019
Bayesian (BIC)	8809.553
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	8768.270
Chi-Square Test of Model Fit	
Value	8.022
Degrees of Freedom	1
P-Value	0.0046
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.091
90 Percent C.I.	0.041 0.154
Probability RMSEA <= .05	0.083
CFI/TLI	
CFI	0.996
TLI	0.978
Chi-Square Test of Model Fit for the Baseline Model	
Value	1918.487
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value	0.011

Figure 32

2015 Hosp. 2 and 4 Model Fit

MODEL FIT INFORMATION

Number of Free Parameters	13
Loglikelihood	
H0 Value	-3089.655
H1 Value	-3088.716
Information Criteria	
Akaike (AIC)	6205.309
Bayesian (BIC)	6262.317
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	6221.046
Chi-Square Test of Model Fit	
Value	1.877
Degrees of Freedom	1
P-Value	0.1707
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.038
90 Percent C.I.	0.000 0.124
Probability RMSEA <= .05	0.444
CFI/TLI	
CFI	0.999
TLI	0.996
Chi-Square Test of Model Fit for the Baseline Model	
Value	1252.772
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value	0.007

Figure 33

2015 Hosp. 2 and 3 Model Fit

```
MODEL FIT INFORMATION

Number of Free Parameters                       13

Loglikelihood

          H0 Value                       -5353.618
          H1 Value                       -5353.337

Information Criteria

          Akaike (AIC)                   10733.236
          Bayesian (BIC)                 10797.011
          Sample-Size Adjusted BIC       10755.722
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              0.562
          Degrees of Freedom                     1
          P-Value                           0.4536

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.000
          90 Percent C.I.                    0.000  0.076
          Probability RMSEA <= .05           0.807

CFI/TLI

          CFI                                1.000
          TLI                                1.000

Chi-Square Test of Model Fit for the Baseline Model

          Value                           2235.248
          Degrees of Freedom                     6
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.003
```

APPENDIX D

ADVISORY BOARD PERMISSION

From: Walker, Mary [mailto:WalkerMar@advisory.com]
Sent: Monday, August 31, 2015 7:43 AM
To: Phillips, Christi A.
Subject: Dissertation

Hi Christi,

I wanted to follow up with you about using the Advisory Board material for your dissertation. You are allowed to use any content now as long as you cite The Advisory Board in your resources and wherever our data is used.

We no longer need a form, just citations! I hope this helps and am happy to help as much as I can through your process.

Thanks,

Mary

Mary Walker
Dedicated Analyst, Research and Insights
The Advisory Board Company
202-266-5682 direct | 202-266-5700 fax
walkermar@advisory.com | www.advisory.com

TEXAS A&M UNIVERSITY

Christi Phillips <achanda1@tamu.edu>

FW: Dissertation Permission from Advisory Board

1 message

Phillips, Christi A. <[redacted] Tue, Mar 14, 2017 at 2:03 PM
To: "achanda1@tamu.edu" <achanda1@tamu.edu>, "[redacted]

From: Michael Beyerlein [mailto:beyerlein@tamu.edu]
Sent: Monday, August 31, 2015 9:41 AM
To: Walker, Mary
Cc: Phillips, Christi A.
Subject: Re: Dissertation

Thank you, Mary.

Data collection is one of the biggest challenges in doing a dissertation. Your support will enable Christi to focus more on the other components of the process.

Enjoy your day.

Mike

On Mon, Aug 31, 2015 at 9:38 AM, Walker, Mary <WalkerMar@advisory.com> wrote:

Hi Dr. Beyerlein,

I hope you are doing well. We would be happy to provide a letter informing the IRB board of the permission from Advisory Board.

Thanks,
Mary

This email is redacted to protect organizational anonymity requested by SystemTex and The Advisory Board.

APPENDIX E

IRB APPROVAL FORM

DIVISION OF RESEARCH

TEXAS A&M UNIVERSITY

NOT HUMAN RESEARCH DETERMINATION

May 28, 2020

Any study that requires in person or face-to-face study visits may not begin or resume until your site has an approved plan that adheres to the re-opening guidelines posted on the Division of Research's VPR website: https://vpr.tamu.edu/covid-19. This plan is to be sent to your Department Chair and Dean, then forwarded to the Clinical Research, Education and Service Advisory Committee for approval.

Type of Review:	Submission Response for Initial Review Submission Form
Title:	Multi-group Measurement Invariance Pre- and Post-Acquisition
Investigator:	Michael Beyerlein
IRB ID:	IRB2020-0600
Reference Number:	111276
Funding:	Internal Funds
Documents Received:	Advisory Board Data Permission Email Exchange 1.1

Dear Michael Beyerlein:

The Institution determined that the proposed activity is not research involving human subjects as defined by DHHS and FDA regulations.

Further IRB review and approval by this organization is not required because this is not human research. This determination applies only to the activities described in this IRB submission and does not apply should any changes be made. If changes are made you must immediately contact the IRB about whether these activities are research involving humans in which the organization is engaged. You will also be required to submit a new request to the IRB for a determination.

Please be aware that receiving a 'Not Human Research Determination' is not the same as IRB review and approval of the activity. IRB consent forms or templates for the activities described in the determination are not to be used and references to TAMU IRB approval must be removed from study documents.

If you have any questions, please contact the IRB Administrative Office at 1-979-458-4067, toll free at 1-855-795-8636.

Sincerely,
IRB Administration

750 Agronomy Road, Suite 2701

1186 TAMU
College Station, TX 77843-1186

Tel. 979.458.1467 Fax. 979.862.3176
http://rcb.tamu.edu