Development and Initial Validation of the Managerial Success Factors Inventory: Transportation Version (MSFI: TV)

Michael Robert Durr II
University of Denver

Follow this and additional works at: https://digitalcommons.du.edu/etd

Part of the Psychology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu,dig-commons@du.edu.
Development and Initial Validation of the Managerial Success Factors Inventory: Transportation Version (MSFI: TV)

Abstract
The purpose of this study was to develop and initially validate the Managerial Success Factors Inventory: Transportation Version (MSFI: TV), a scale generated to more accurately assess managerial competence in transportation than general managerial competence and leadership scales. The literature on managerial competencies, leadership, and the state of transportation management is summarized. Reigning models of general managerial competency were used with focus group and cognitive interview data to develop the scale. The conventional three-phase scale development research design was followed (exploratory factor analysis, confirmatory factor analysis, and validation phases) by administering the instrument to a sample of 287 managers from a state Department of Transportation (DOT). The hypothesized 4 cluster-11 competency general managerial competency model that the scale was based on could not be found in the exploratory factor analysis. Instead, a five-factor solution was most interpretable; problem solving, emotional competence, initiative, safety leadership, and integrity. This model was then confirmed via confirmatory factor analysis. A higher-order model was also confirmed indicating the measure can be treated as unidimensional. Rasch analysis confirmed these findings. The paper details the use of Rasch and structural equation modeling to analyze the psychometric properties of the scale. Implications for use of the scale in selecting, training, and promoting leaders in transportation and future research needs on the scale are discussed.

Document Type
Dissertation

Degree Name
Ph.D.

Department
Counseling Psychology

First Advisor
Patrick Sherry, Ph.D.

Second Advisor
Kathy Green

Third Advisor
Ruth Chao

Keywords
Competencies, Counseling, Leadership, Management, Performance

Subject Categories
Psychology | Social and Behavioral Sciences

Publication Statement
Copyright is held by the author. User is responsible for all copyright compliance.

This dissertation is available at Digital Commons @ DU: https://digitalcommons.du.edu/etd/175
DEVELOPMENT AND INITIAL VALIDATION OF THE MANAGERIAL SUCCESS FACTORS INVENTORY: TRANSPORTATION VERSION (MSFI: TV)

A Dissertation
Presented to
the Faculty of the Morgridge College of Education
University of Denver

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Michael Robert Durr II
November 2011
Advisor: Patrick Sherry, Ph.D.
Abstract

The purpose of this study was to develop and initially validate the Managerial Success Factors Inventory: Transportation Version (MSFI: TV), a scale generated to more accurately assess managerial competence in transportation than general managerial competence and leadership scales. The literature on managerial competencies, leadership, and the state of transportation management is summarized. Reingning models of general managerial competency were used with focus group and cognitive interview data to develop the scale. The conventional three-phase scale development research design was followed (exploratory factor analysis, confirmatory factor analysis, and validation phases) by administering the instrument to a sample of 287 managers from a state Department of Transportation (DOT). The hypothesized 4 cluster competency general managerial competency model that the scale was based on could not be found in the exploratory factor analysis. Instead, a five-factor solution was most interpretable; problem solving, emotional competence, initiative, safety leadership, and integrity. This model was then confirmed via confirmatory factor analysis. A higher-order model was also confirmed indicating the measure can be treated as unidimensional. Rasch analysis confirmed these findings. The paper details the use of Rasch and structural equation modeling to analyze the psychometric properties of the scale. Implications for use of the scale in selecting, training, and promoting leaders in transportation and future research needs on the scale are discussed.
Acknowledgements

The researcher is thankful for the support and guidance of the faculty chair, Patrick Sherry, Ph.D., as well as the statistical and research methodology coaching by the committee members Kathy Green, Ph.D. and Ruth Chao, Ph.D. The researcher also acknowledges that this study was made possible by Dr. Sherry’s eminent work in vocational psychology. His ideas and network in the area of leadership development in transportation made this study possible. The study was also made possible by the financial support provided by a research grant from the National Center for Intermodal Transportation (NCIT).
Table of Contents

Chapter 1—Introduction ........................................................................................................ 1
  Justification for the Study .......................................................................................... 7
  Leadership versus Management: Clarifying the Target Variable ....................... 9
  Research Questions .................................................................................................. 11
  Summary .................................................................................................................... 12

Chapter 2—Literature Review ...................................................................................... 13
  What is a Competency? .............................................................................................. 13
  Brief Historical Perspective of the Use of Competencies ................................... 18
  Competency Models of Management .................................................................. 24
    Methodology for generating the models ............................................................. 33
  Leadership Theory .................................................................................................... 37
    The Trait approach ................................................................................................ 42
    The behavioral approach ................................................................................... 42
    The contingency approach ............................................................................... 43
  Fiedler’s Contingency Theory ............................................................................... 44
  Charismatic and transformational leadership .................................................... 45
  Managing versus leading ....................................................................................... 47
  Managerial Competency Outcome Research ..................................................... 48
  State of Transportation Leadership ...................................................................... 56
  Benefits of Rasch Analysis .................................................................................. 58
  Absenteeism ........................................................................................................... 61
  Summary .................................................................................................................... 61

Chapter 3—Methodology ............................................................................................ 64
  Sample ......................................................................................................................... 64
  Measures ..................................................................................................................... 66
    Demographic questionnaire ............................................................................... 66
    Absenteeism .......................................................................................................... 66
    Performance questionnaire ............................................................................... 67
    Supervisor’s ratings ............................................................................................. 67
  Managerial Success Factors Inventory: Transportation Version (MSFI:TV) - in development .................................................................................................................. 68
  Procedures .................................................................................................................. 68
    Item development for the MSFI:TV .................................................................... 69
    Study procedures .................................................................................................. 73
  Analysis ....................................................................................................................... 75
    Rasch Modeling .................................................................................................... 76

Chapter 4—Results ....................................................................................................... 78
  Focus Group Content Analysis ............................................................................... 78
  Phase 1 ....................................................................................................................... 80
    Exploratory Factor Analysis ............................................................................... 80
    Internal consistency and correlations among the subscales ......................... 84
Validity ......................................................... 85
Phase 2 .......................................................... 88
Confirmatory Factor Analysis ................................ 88
Rasch modeling results ....................................... 94
Summary ................................................................ 110

Chapter 5—Discussion ........................................ 112
Brief Literature Review ........................................ 113
Key Findings ...................................................... 114
  Research Question 1 .......................................... 114
  Research Question 2 .......................................... 115
  Research Question 3 .......................................... 117
  Research Question 4 .......................................... 119
  Research Questions 5, 6, and 7 ......................... 121
  Research Question 8 .......................................... 122
General Conclusions from Rasch Modeling ............ 123
Theoretical Implications ..................................... 125
Usefulness of the Scale ....................................... 129
  Factor analysis ................................................. 132
  Implications for future research from Rasch Modeling 132
Limitations and Future Research ......................... 133
Summary .......................................................... 136

References .......................................................... 139

Appendix A ........................................................ 144
Appendix B ........................................................ 145
Appendix C ........................................................ 146
List of Tables

Table 1  Prominent Leadership and General Managerial Competency Models .......... 26

Table 2  Primary Managerial Competence Themes Identified in Focus Groups .......... 80

Table 3  Items, Factor Loadings, and Standard Deviations ........................................ 83

Table 4  Leadership Scales and Associated Cronbach’s Alpha, Eigenvalue, and Percentage of Variance Accounted for by the Factors ......................................................... 85

Table 5  Correlations Among the Leadership Subscales, Performance Measure, and Demographic Variables ........................................................................................................ 87

Table 6  Goodness-of-Fit Statistics for the Models Compared in CFA ......................... 94

Table 7  Summary of Category Structure ....................................................................... 98

Table 8  Item Statistics for the MSFI: TV Total Scale .................................................. 101
Chapter 1—Introduction

The method by which we select and train managers is a pressing issue not only in business and the social sciences, but in the general public and popular press (Hollenbeck, 2009). It seems companies are spending more time and money today than ever on leadership development systems to place themselves ahead of the competition and establish a competitive advantage through highly effective leadership (Boyatzis & Saatcioglu, 2007). The use of managerial and/or leadership competency systems has become the gold standard design to improve leadership performance (Boyatzis & Saatcioglu, 2007). Most organizations rely on competency-based assessment tools to select, train, promote, and reward managers (Boyatzis, 2007). Prior to the work of psychologist David McClelland (1973), who first proposed competencies as critical differentiating factors of leadership performance, most looked to traditional human resource factors such as education and experience to select, promote and train leaders (Levenson, Van der Stere, & Cohen, 2006).

Since competencies were first proposed as key differentiating leadership performance determinants during the height of the behavioral movement, there has been an outpouring of leadership competency research in business (Boyatzis, 2007; Boyatzis, 1982; Calhoun, Dollett, Wainio, Butler, Griffith, & Warden, 2008; Garman & Johnson, 2006; Levenson, et al., 2006). Today both practitioners and scholars seem to agree that effective organizations have behavioral-competency systems in place (Calhoun et al.,
Over the past thirty years most industries have grown to rely on competency-based systems and accompanying assessment tools to improve individual, unit, and organizational performance (Calhoun et al., 2008). Indeed, it is difficult to find an organization larger than 300 employees today without a behavioral-competency model (Boyatzis, 2007). Human resource (HR) processes such as selecting, training, promoting, and compensating managers that are based on competency assessment instruments have been found to be more effective than traditional measures (Levenson et al., 2006).

Managerial and leadership competencies have been empirically linked to performance and organizational success (Kowske & Anthony, 2007; Levenson et al., 2006). Despite widespread use of various forms of competency-based performance scales (e.g., 360s) to improve HR processes, as well as mounting empirical evidence to support such scales, it is still difficult to locate psychometrically sound leadership competency scales in the empirical literature for certain key industries. Transportation is one such industry. The scales that have surfaced mostly did so in corporate America as opposed to academia. As a result, the methodology, psychometrics, and science supporting the practice have lagged behind practice (Boyatzis, 2007). Thus, the purpose of this study was to develop a psychometrically sound managerial competency scale for use in the transportation industry.

Leadership and managerial competency definitions are as widespread as the scholars and practitioners who have investigated the topic (Cheng, Dainty, & Moore, 2002). Nevertheless, most generally agree that managerial competency refers to the behavioral and technical characteristics (competencies) that discriminate outstanding
managerial or leadership performance from typical performance (Calhoun et al., 2008). The competency research and practice movement is largely charged by the well documented notion that leadership can make a difference (Holton & Lynham, 2000), and that competency systems can aid organizational efforts to recruit, train and develop leaders (Levenson et al., 2006). Competency-based scales for managers lend themselves nicely to leadership development and performance improvement programs because by definition competencies in business are considered to be the “right” behaviors (Ricciardi, 2005). The “right behaviors,” or competencies, allow organizations to reach their desired outcomes (Ricciardi, 2005). Thus, competencies are at their core considered to be specific behaviors applied to successfully complete a task directly linked to a desired outcome (Ricciardi, 2005; Tett, Guterman, Bleier, & Murphy, 2000). Identification of core competencies for specific industries allows organizations to better understand the cultural and environmental conditions needed to support emission of the “right” behaviors (Ricciardi, 2005). Thus, support for the development and initial validation of a managerial competency scale for transportation executives rests on the notion that competency scales are useful in improving performance.

Recent research also demonstrated that the specific core competencies essential for successful performance are partially influenced by industry context (Brownell, 2008). This conclusion comes following years of leadership and management research that has failed to agree upon one essential set of core competencies that ensure success across industry lines (Brownell, 2008). As a result, more studies are beginning to emerge that focused exclusively on applying general managerial/leadership competency research to
the development of sophisticated assessment instruments for specific industries (Calhoun et al., 2008). This is a relatively recent trend, and despite widespread use of competencies as performance criteria, certain key industries remain without psychometrically sound managerial competency scales. Many industries rely on outdated and invalid instruments (Tett et al., 2000). Transportation is one industry lacking a sophisticated and accurate managerial competency scale, while also facing increased need for services amidst workforce development shortages at all levels of management (Vogel, 2001). Few studies, if any, have focused exclusively on the development of a psychometrically sound managerial competency scale for transportation outside of corporate America. There are several reasons why transportation is an ideal industry for such a project.

As the demand for transportation continues to increase dramatically amidst the current population boom and reliance on transportation services grows, an estimated 40 to 50% of the existing local, state, and federal transportation workforce near retirement (Martin, 2001). These baby boomers lead at all levels of the transportation industry and little workforce development planning has been done to prepare for the future shortage of competent managers (CTC & Associates LLC). Sen. George Voinovich, R-Ohio estimated that by 2010 approximately 600,000 employees will retire (U.S. Dept. of Transportation). Further, the Federal Highway Administration (FHWA) reported that 45% of its workforce will be eligible to retire in 2010 (Martin, 2001). Despite workforce development challenges amidst increasing demand for services, the transportation industry remains without a valid managerial competency scale. A psychometrically sound behavioral competency instrument tailored to transportation could help the industry
establish training, recruiting, and managing practices that meet the demands of these future shortages of competent leaders at all levels of management. Establishment of a competency assessment tool for transportation managers that is based on newer approaches to scale development would greatly increase the likelihood of a smooth transition from the current aging management body by validly identifying the core competencies responsible for successful leadership in transportation. Using advanced methodology and analysis in scale development, the purpose of this study was to develop a core managerial competency scale that lists leadership competencies in terms of difficulty so that it can be used to effectively recruit, select, train and reward transportation leaders by identifying leader abilities.

Several other industries are ahead of transportation in development of a valid industry-wide leadership and/or managerial competency assessment tool (Calhoun et al., 2008). In fact, consumer product companies, financial service corporations, higher education institutions, and health care have led the way in the development of industry-specific competency-based assessment tools. In a recent study by Calhoun et al (2008), the rationale behind behavioral competency-based assessment was clarified. The study set out to develop a method of measuring the skills necessary for effective performance in all types and levels of management in health care (Calhoun et al., 2008). The goal of the study was to develop an assessment tool that could be used across various levels of management in healthcare in a variety of settings, and that would provide a common language for all managers in the healthcare industry. The study’s findings allowed the health care industry to improve business management and graduate training curricula via
their new model and accompanying scale. The current study intended to echo this movement by establishing a competency scale for managers in transportation grounded in the essential practices, traits, and behaviors of successful transportation management, which we refer to here as competencies. Pulling from existing leadership competency research, the general idea was to do the same task for transportation that other leading industries have completed; to create a managerial competency scale to improve human capital management processes.

Finally, most current scales of managerial performance developed in both academic and corporate circles utilized classical test theory (CTT) despite newer techniques that may lend additional and critical information to the development of managerial performance measures. Thus, another purpose of the current study was to develop a more psychometrically sound competency scale than currently exists. Rasch modeling, an item-response theory (IRT) approach, allows for more advanced and psychometrically sound scale development procedures. Specifically, it provides a means to examine the scale at the item and person level.

A review of the current literature examining psychometric properties of the most heavily used leadership competency assessment tools revealed that most are validated in a variety of settings with various management groups. Although numerous studies have used highly sophisticated methods of developing leadership competency models, few have specifically set out to develop a managerial competency scale for a particular industry using IRT principles. Thus, the primary purpose of this study was to develop a managerial competency scale for the transportation industry by employing tenets of both
classical test theory (CTT) and Rasch analyses. No studies were found in the academic literature on the development of a managerial competency scale for a particular industry using Rasch modeling.

**Justification for the Study**

There are several societal and economic reasons to focus on leadership performance (Levenson et al., 2006), especially in transportation. Indeed, one does not have to look too far to see how costly poor selection and training procedures can be to organizations and the nation as a whole. The press is quick to announce executive scams and failures, which have been seen regularly in recent years. In the transportation industry, the cost of safety mishaps has led to the downfall of transportation agencies and entire modes of transportation, as well as eroded the trust of the public while leading to substantial loss in revenue. Examples are the recent rail accidents in California and Washington D.C. Many immediately questioned selection and training procedures and if the rail industry is selecting the best people to do the job.

The global economic meltdown has also left companies starving for a competitive advantage to stay alive. As companies look to stay afloat during these difficult economic times, more has been and will continue to be invested in the recruiting, selecting, training, and rewarding programs in place to produce effective leaders across industry lines. Indeed, leadership matters to organizational outcomes (Holton & Lynham, 2000; Levenson et al., 2006). Establishment of reliable, valid and psychometrically sound assessment tools that can be used to recruit, train, and promote quality leaders can aid this process (Vogel, 2001). Transportation is in need of a valid instrument for this exact
reason. Thus, the primary justification for this project stemmed from the belief in the power of and need for good leadership, and the well documented research indicating that competency assessment tools is one of the best starting places to achieve this mission. Further, there are several excellent general managerial competency models and accompanying assessment tools in existence. The goal of this project was not to simply add another one. Instead, this project was charged by the premise that leadership and management in transportation require specific competencies unique from other fields. Thus, development of a tool that assesses managerial competence in transportation specifically would add to our understanding of outstanding managerial performance above and beyond that accounted for by general managerial competence assessment tools.

Finally, the majority of managerial competency assessments are based on CTT principles, despite evidence that IRT principles would lend additional useful information and most likely are more appropriate for this type of scale development project. For example, the competency-based questionnaires utilized in business today are predominantly used for HR processes such as selection, promotion, training, and rewarding leaders. These scales, the processes in which they are used, and the items that make them up are inherently hierarchical in nature. Competency in management naturally falls along a continuum. Behaviors at the lower end of managerial competency are easier to perform than behaviors and competencies at the higher end. In other words, competencies by definition are behaviors that distinguish outstanding from typical performers, they represent varying degrees of competency, and therefore are hierarchical
in nature. In CTT, items are assumed to be roughly parallel indicators of the underlying latent variable (DeVellis, 2003). In IRT, on the other hand, items are placed along a continuum and gaps in the scale can be identified (Bond & Fox, 2006). Further, by using IRT approaches items can be tuned to a particular range of ability and can then be applied to particular situations (DeVellis, 2003). Using an IRT approach for the development of a managerial competency scale for transportation would allow transportation agencies to know which items relate to certain levels of managerial competency, allowing for appropriate selection of items for particular levels of competency. Differing levels of management would not need to take the entire scale. Rather, determination of the highest item they passed would allow for clear developmental needs, as the range of performance would be clearly indicated by the scale.

**Leadership versus Management: Clarifying the Target Variable**

It is quite common to see the terms leadership and management used interchangeably as has intentionally been done in this paper to this point. Yet, the terms capture quite distinct practices and there is general consensus that management alone is unsatisfactory in measuring “what bosses do” (Stringer, 2002). In all of the literature reviewed for this project, Stringer (2002) provides the best summary explanation of the differences. He proposed that management deals with handling the “complexity” of organizational life including creating order, discipline, and structure in a way to do things right. On the other hand, leadership involves inspiring others, having a vision, or motivating. Where management involves “doing things right” leadership involves “doing the right things” (Stringer, 2002, p.105,). Further, he proposed that leaders motivate and
inspire others by “their day-to-day behaviors” (Stringer, 2002, p. 105), which he termed practices. Based on this notion that both leadership and managerial practices are critical to the success of any organization, we attempted to capture both leadership and managerial competencies in the assessment tool created to generate one assessment tool that captures the essence of successful management. The researcher did not intend to create a comprehensive list of every managerial or leadership practice or quality known to the management world, as this form of assessment can get extremely lengthy and impractical. Further, by managerial competencies the researcher is not referring to very job-specific competencies as might be seen in a job description. Instead, our target variable became simply managerial competencies as defined as any behavioral practices essential for being a successful manager of a transportation organization. The goal was to create a useful, short scale that provides a snapshot of some of the most essential managerial behavioral practices commonly needed to successfully lead in transportation. We refer to the target variable as managerial competence but include several key leadership competencies as competent management entails some leadership skill. It is always tricky to determine the appropriate balance on such a scale between leadership qualities and managerial practices. Rather than getting lost in the theoretical debate of whether a particular practice is a managerial or a leadership practice, we focused on what we define as managerial competence, which is operationally defined as the observable, behavioral and trait characteristics essential to effectively manage transportation agencies. Without a doubt, leadership practices are part of being a competent manager. Stringer (2002) suggested that there are over 350 specific leadership practices (qualities)
that have been correlated to peak performance. Thus, we followed a methodology that would provide recognition of the essential leadership practices required for competent management in transportation in particular.

**Research Questions**

Question 1: Exploratory factor analysis will identify four factors: Managing Yourself, Managing Your Team, Managing the Work (e.g., transportation specific), Managing Collaboratively (Based on the Hay Group (2001) general managerial competency model).

Question 2: The scale demonstrates acceptable internal consistency, measured by Cronbach’s Alpha.

Question 3: Confirmatory factor analysis will identify the same factors found in the EFA.

Question 4: Rasch analysis shows that items vary with increasing amounts of competency in the participant and cover the range of levels of managerial competency in the participants.

Question 5: Convergent validity is shown by strong positive correlations with scores on a measure of managerial performance.

Question 6: Discriminant validity is shown by lower correlations with scores on a measure of absenteeism.

Question 7: Concurrent validity is shown by strong correlations with peer’s ratings of the participants on a measure of managerial performance.
Question 8: The scale is an invariant measure in that managerial competency is the same thing across groups (e.g., gender).

Summary

Most industries have hired consultants, leadership associations dedicated to their particular industry, or human resource experts in their field to develop competency systems and subsequent managerial competency assessment tools for their industry. Yet, one of the largest sectors, transportation, has yet to complete this mission on a general level. Indeed, few studies, if any, have specifically focused on the development of an accurate managerial competency scale specifically designed for managers in the transportation industry. Therefore, the primary purpose of this project was to develop a managerial competency tool for transportation.

Transportation is a critical multi-national industry that maintains economic stability and allows society to carry out daily routines. Effective leadership in transportation is critical for successful, safe movement of passenger and freight. Given the lack of valid managerial competency scales, coupled with the state of the transportation workforce in modern society, this study aimed to develop and initially validate a managerial competency scale for the transportation industry by applying both CTT and Rasch Analytical methods.
Chapter 2—Literature Review

This chapter covers the relevant literature on the subject of leadership and managerial competencies. As one might expect, it is not a subject lacking in literature. Thus, attention is predominantly granted to the structure of managerial competencies that most management scales are based on. The chapter begins with a detailed description of how competencies are defined. Following the formal definition, a brief historical overview of the use of competencies is provided, a section on the prominent competency models of management follows, and then a section on leadership theory is provided. Finally, a section is dedicated to outcome research on the predictability of various individual and organizational outcomes from managerial competencies. The outcome research section discusses empirical findings on the relation between leadership competency and performance, selection, mentoring, and training.

What is a Competency?

Since psychologist David McClelland (1973) first proposed competencies as potential differentiating factors of performance beyond intelligence nearly forty years ago, a considerable amount of research on the topic and strong efforts to define what is meant by the term competency have been documented. A majority of the work on competencies has come from consulting firms and corporations who investigate competencies of leadership acumen for particular industries. Indeed, there were few studies in the academic literature until the past decade. Great efforts have been made in
recent years to clarify what is meant by the terms competency, competencies, and competency modeling. Before we progress in this paper, it seems important to clarify these concepts as the purpose of this study is to develop a general managerial competency scale for transportation based on individual competencies that together form the single latent variable of transportation managerial competency.

Like most psychological constructs, early scholars were in disagreement on the definition of competency and found it difficult to distinguish it from other similar concepts (Hollenbeck, 2009). More recently, however, researchers have focused on competencies as behavioral and observable abilities to improve the applicability of competency scales (Boyatzis, 2007). Calhoun et al. (2008), for example, defined competency as, “Those behavioral and technical characteristics (competencies) that discriminate outstanding leadership performance from typical performance” (p.377). The emphasis in the Calhoun et al (2008) definition is placed on the ability of competencies to differentiate highly effective from typical workers, which is a generally agreed upon aspect of competencies, and part of what makes them so valuable to organizational success. Further, most seem to agree that a competency refers to a skill or personal ability that is required to be effective on the job and that is critical to achieving targeted outcomes (Brownell, 2008). The definition has not always been as simple to define as one might imagine.

In Spencer and Spencer’s (1993) heavily cited text on the subject, “Competence at Work,” the definition of competency focused more on characteristics. Spencer and Spencer (1993) followed Boyatzis’s (1982) approach by suggesting that a competency is
an underlying characteristic of a person that is causally related to criterion-referenced effective and superior performance in a job. By criterion-referenced the authors meant that the competency actually predicts who does something well or poorly, based on a specific criterion (Spencer & Spencer, 1993). The criterion referenced portion of their definition still stands today, however, using the term characteristic seems to have led the field of competency systems into murky waters. Use of the term characteristic implied close relation to personality trait rather than what McClelland (1973) seemed to originally call for in a behavioral approach. Further, Spencer and Spencer (1993) suggested that underlying characteristic “means the competency is a fairly deep and enduring part of a person’s personality and can predict behavior in a wide variety of situations and job tasks” (p.9). Again, thinking of competencies in this manner, it is difficult to distinguish between competencies and personality traits, as this definition suggested that competencies are attributes that are enduring characteristics that predict highly effective versus typical or poor performers. Since this time several improved, more operational definitions of competencies have emerged.

Though the concept of competency still needs additional work to truly clarify its meaning and distinguish it from other psychological constructs, some good work has been produced to define it over the past ten years (Boyatzis, 2007; Tett et al., 2000). In a 2000 study by Tett et al. in the journal, Human Performance, the authors took a more behavioral and modern approach to define competency. The following definition was offered, “A competency is an identifiable aspect of prospective work behavior attributable to the individual that is expected to contribute positively and/or negatively to
organizational effectiveness.” In this definition, the authors acknowledged the behavioral core of competencies. Along similar lines, Ricciardi (2005) defined competencies as “distinct sets of behaviors applied to reliably complete a task that is directly linked to a critical outcome” (p.488). This behavioral, contemporary view of competencies is similar to that presented above by Calhoun et al (2008), with the emphasis on key differentiating behaviors between superior and typical performers. This is also the definition in which the scale developed in the current study will be based. Thus, competencies are considered here to be behavioral characteristics (skills) of an individual which is causally related to superior performance in a job.

Competencies have also been divided into two separate and distinct categories, “threshold” and “differentiating” (Spencer & Spencer, 1993, p.15). Threshold competencies are thought to be those characteristics that are an absolute necessity to minimally perform a certain job task (Spencer & Spencer). For example, a threshold competency for psychologists might be the ability to listen. Listening in therapy may not differentiate great and mediocre therapists, but it is an essential characteristic to be minimally effective in the job. Differentiating competencies, on the other hand, are those factors that distinguish superior and average performers. Using the same example, a differentiating competency of psychologists might be the ability to form a meaningful relationship with clients. Again, not necessarily a threshold factor, as several therapists provide counseling without the ability to establish strong interpersonal connections, but those who are great or highly effective are able to use interpersonal skills to establish
meaningful relationships. This may be a differentiating competency in that it separates superior from average psychologists.

Horton, Hondeghem, and Farnham (2002) suggested that clarification is also provided in English speaking countries between competencies and competences. They suggested that competency(ies) refer to ‘the behavioral characteristics of an individual which is causally related to effective or superior performance in a job’ (Boyatzis, 1982) while competence(s) refer ‘to the ability to perform activities within an occupation to a prescribed standard’ (Horton et al., 2002, p.4). This distinction is critical in focus in that competencies are focused on the inputs that help achieve successful performance in a job. The difference also has serious implications for the competency model that would emerge from each.

Finally, pulling from the various definitions of competency-related concepts and terms that have been presented in the literature (Garman & Johnson, 2006; Spencer & Spencer, 1993), the following definitions are used in this study.

Competencies: behavioral characteristics and skills that are causally related to effective or superior performance in a job.

Competency: distinct sets of work behaviors applied to reliably complete a task that is directly linked to a critical outcome.

Transportation Managerial Competency: a distinct set of work behaviors applied to reliably and successfully manage in transportation, and that distinguish typical from superior transportation managers.
**Competency Modeling:** a systematic process for identifying and articulating competencies at either the individual or organizational level.

Competencies are a behavioral approach to emotional, social, and cognitive intelligence (Boyatzis, 2007). Development of the transportation managerial competency scale pulls from these definitions and also takes the definition process a step further. Part of the purpose of this study is to define the specific competencies that are both threshold and distinguishing competencies to transportation management in particular. The researchers are interested in identifying and including in the assessment instrument those competencies that are both necessary to complete essential transportation management duties but that also distinguish typical transportation managers from superior transportation managers.

**Brief Historical Perspective of the Use of Competencies**

Like most movements in business, education and the social sciences, the competency movement has no single origin (Horton et al., 2002). In a way, the competency movement has been around for centuries tracing back to the mediaeval period when apprentices learned tasks associated with specific jobs by working for a master. Much later the study of jobs and skills needed to successfully perform those jobs emerged during the industrial revolution (Horton et al., 2002). Before long accreditation awards were linked to standards of literacy and knowledge in a given field, and education was greatly influenced by assessment tactics aimed at measuring one’s skills and knowledge related to the field. The competency movement really gained momentum in
the 1930s with the emergence of the Dictionary of Occupational Titles (DOT), which identified knowledge and skill sets needed for a variety of occupations (McLagan, 1997).

The early work on the apprenticeship model and the DOT gave way for the true origins of the competency movement during the changing economic and political context of the late 1960s. Intelligence was being heavily studied in psychology and scholars were investigating the means to assess job-related personality traits. With emphasis mounting on assessment at work and the study of job-related skills, the origin of competencies truly started when the concept of ‘managerial competency’ surfaced from the work of David McClelland (1973) and the McBer consultancy group in the 1970s (Bolden & Gosling, 2002). McClelland’s seminal article in the American Psychologist in 1973 is argued today as the ‘real’ beginnings of the movement when he suggested that traditional exams and tests were poor predictors of whether people could do a job well and that other means existed to look for competencies essential for success (Horton et al., 2002). McClelland’s quest to replace the study of personality traits with competencies led to exemplar methodology for developing competency models including the ‘behavioral event interview’ (BEI) and the ‘criterion-referenced assessment.’

McClelland criticized the testing movement for focusing too much on upper-class constructs in measurement and achievement. He thought that the intelligence testing movement was doing nothing more than assessing the degree of opportunities one had available. He went on to suggest that neither intelligence tests nor school grades, which he thought were based on similar non-tangible criteria that was ‘contaminated heavily by the power of those at the top of the social hierarchy’, seemed to not have much power to
predict real competence in many life outcomes (McClelland, 1973). Thus, McClelland (1973) argued for the assessment and consequential development of competencies involved in clusters of life outcomes. In 1973, McClelland proposed several new directions for the testing movement, with his central tenet focused on assessment and identification of operant as well as respondent behavior. He proposed that tests should be based more heavily on items that provide several correct responses, “among which one was better than others in terms of some criteria of efficiency that the person would have to apply” (McClelland, 1973, p.11). From this seminal article, McClelland (1973) introduced behavioral analyses to the world of education and work, and proposed a new tactic for identifying inputs associated with successful performance and positive outcomes. McClelland’s (1973) ideas were applied in the consultancy group, McBer and Company, which he founded in 1963, and went on to develop competency models for many of America’s top companies (Horton et al., 2002).

Since the McClelland (1973) article was first published in business and social sciences literature nearly forty years ago, there has been an outpouring of research on competencies at work, and how they can be utilized by organizations to improve managerial performance. In 1982 the American Management Association commissioned Richard Boyatzis, of McBer Associates, to identify which competencies distinguish superior from typical managers. A sample of approximately 1800 managers across 41 different management jobs and 12 different organizations were asked to identify the generic knowledge, motives, traits, self-image, social role or skill of a person that resulted in superior performance of a job (Horton et al., 2002). Boyatzis (1982) reported
19 generic characteristics that outstanding managers tended to possess, and he coined the plural term ‘competencies’ to refer to “the underlying characteristic of an individual that is causally related to effective or superior performance of a job” (Boyatzis, 1982). Boyatzis provided support for this study by arguing for the need to identify organization and industry-specific competency models and associated assessment tools, as he thought context does matter. Further, it was proposed that competencies can be grouped together in clusters. Boyatzis’s model grouped managerial competencies into four general clusters, each of which were thought to be related equally to the major functions of managers within any organization: achieving the goals of the organization, providing leadership, managing people, controlling and directing others. The emphasis in this model was on what managers can do and how they do things, or how they behave as opposed to what skills and knowledge they possess.

Since Boyatzis (1982) developed the original competency model for general managers, scholars and consultants with various backgrounds have developed endless variations of core managerial competency models. Most large companies have had competency models developed for various reasons. The use of competencies and competency models, in corporate America in particular, runs as far as imagination can reach. Indeed, most large companies have competency models in place today. Most jobs have a listing of related competencies attached by which performance appraisals are frequently made. Thus, the use of competencies and competency models is widespread across various industries (Calhoun et al., 2008).
As tools for defining and assessing performance, competencies are utilized for a variety of performance improvement purposes today (Garman & Johnson, 2006). At the individual level, competencies are frequently used to clarify job descriptions and employee duties or roles, performance expectations, and plans for strategic development. Competency modeling is considered the gold standard approach to establishing behavioral, job-related characteristics to successfully perform a job. At the organizational level, competency models can help articulate the behavioral implications of a strategic plan (Garman & Johnson, 2006). In a recent analysis of the use of competency frameworks with 31 leading North American organizations with strong involvement in executive development initiatives, Briscoe and Hall (1999) suggested that the two most common applications of competencies and accompanying competency assessment tools are executive selection and development. They went on to suggest that competencies are typically placed in an organized framework that is then used as a guide for making hiring and promotion decisions (Briscoe & Hall, 1999). The authors also suggested that competency frameworks are still relatively new to the business world, and research is certainly trying to catch up to practice. Further, Briscoe and Hall reported that organizational use of competency assessments has grown dramatically with the rise of a global competitive market. The competitive demands of today’s marketplace may be pushing companies to spend more on developing exact models of distinguishing characteristics between superior and average performance. Nevertheless, competencies, competency models, and competency assessment scales are perceived as important tools in helping organizations improve executive performance (Briscoe & Hall).
Today competency models are utilized more than ever to successfully complete various human resource and human capital management functions such as recruiting, developing, training, and assessing managers. They provide a general model to compare various managers. A primary goal of using competency assessments to evaluate individuals is to improve job performance by moving farther up the competency checklist (Levenson et al., 2006). They are also used today as the basis of performance management systems, compensation, and succession planning (Spencer & Spencer, 1993). As for performance management, the use of competency assessments is fairly common across various industries. Competency definitions can also assist companies establish human resource management practices, including recruiting, prescreening, using a balanced scorecard, identifying career ladders on which raises are based, and talent management/succession planning processes (Garman & Johnson, 2006). Identification of both general managerial and industry/organization-specific competencies and competency models pave the way for the development of valid and reliable assessment instruments based on those competencies (Spencer & Spencer, 1993). Competencies provide clarification of the necessary (threshold) and superior characteristics needed to successfully perform a certain job.

A considerable amount of competency modeling activity has been happening in various consultancy groups and within specific industries for the past twenty-five years. Most of these models have been applied through the use of assessment tools that are utilized in hiring, training, and appraisal processes in various industries. Yet, few empirically validated accompanying assessment instruments have in peer reviewed
journals. The common strategy has been to use “arm chair” methods to establish assessment instruments based on the competencies identified for leadership and management within various fields. Thus, a major purpose of this study is to use contemporary item response analytic methods to develop and initially validate a general transportation managerial competency scale that can then be used in the successful recruitment, training, and appraisal methods within transportation agencies.

**Competency Models of Management**

Examination of the literature on the topic of leadership and managerial competencies reveals a plethora of typologies, models, frameworks, taxonomies, and lists of various skills and attributes that have been generated in attempt to capture the underlying characteristics that separate highly effective from typical leadership and management practices (Brownell, 2008; Calhoun et al., 2008). Undeniably, several general managerial competency models have been produced. It is essential to recognize these models as they will provide much of the foundation for development of the transportation managerial competency scale in this study. Most of the models have great overlap in the competency clusters listed and typically “get at” the same competencies, but are then organized in unique ways. It seems the primary goal of most taxonomies has been to identify relatively few general dimensions while being as comprehensive and parsimonious as possible (Tett et al., 2000). Gentry and Leslie (2007) noted that one of the challenges organizations face in selecting a model and accompanying assessment instrument for leadership development purposes is identifying which of the 100 plus competencies are to be used. In their scan of over 100 organizations they were able to
rank the top 10 competencies used; leading employees, building relationships, risk-taking, change management, influence, communicating information and ideas, brings out the best in people, follow through, listening, and flexibility. These findings were also factored in to the development of the current scale.

Both general managerial and industry-specific competency models have surfaced over the past twenty-five years. For the purposes of this study, the competency models presented in the empirical literature and from the dominant consultancy groups in business psychology are presented as they provide the foundational theory for the development of the transportation managerial competency scale. Seven competency models or frameworks are presented here. These seven were chosen based on their scientific rigor in development, because they are heavily cited in organizational behavior, psychological, and business literature, and because they were generated by the leading scholars in the field of managerial behavior. Table 1 provides a summary of some of the reigning managerial and leadership competency models.
Table 1

*Prominent Leadership and General Managerial Competency Models*

<table>
<thead>
<tr>
<th>Source(s)</th>
<th>Sample/population</th>
<th>Methodology</th>
<th>Structure</th>
<th>Competency Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthcare Leadership Alliance (HLA; 2005a,b)</td>
<td>General health administration at all levels.</td>
<td>Collaboration of six major health administration professional associations.</td>
<td>300 competencies in 5 clusters</td>
<td>Business knowledge and skills, communication and relationship management, knowledge of healthcare environment, professionalism, and leadership.</td>
</tr>
<tr>
<td>Healthcare Leadership Competency Model (HLCM; 2008)</td>
<td>84 randomly selected managers from across the field. Also, 75 mid- and late-career leaders deemed outstanding were interviewed.</td>
<td>Initial development included behavioral event interviewing, psychometric analysis, and cross-industry sector benchmarking.</td>
<td>3 overarching domains subsuming 26 behavioral and technical competencies.</td>
<td>3 Domains that include; Transformation, Execution, and People</td>
</tr>
<tr>
<td>Hay Group (McClelland/ McBer, 1973 and updated) Manager Competency Model</td>
<td>General managerial competency model</td>
<td>Observing and interviewing outstanding performers in various industries and then grounded in solid empirical research.</td>
<td>11 competencies organized in 4 clusters</td>
<td>Managing yourself, managing your team, managing your work, and managing collaboratively</td>
</tr>
<tr>
<td>Competencies for Leadership (Weiss, 2003)</td>
<td>General leadership and managerial competency model</td>
<td>Reviewed relevant models and synthesized via empirical techniques</td>
<td>4 “SEEDS” or foundational requirements and 4 general clusters</td>
<td>4 SEEDS; Sense of purpose, Energy and optimism, Engaging, Decision-Making. 4 Clusters; Personal effectiveness, Communication, Managing others, Thinking</td>
</tr>
</tbody>
</table>
One of the most heavily cited and highly regarded general competency models comes from the Hay Group, a psychologist based consultancy group that has been developing the gold standard in competency research since the competency movement emerged in the early 1970s. The Hay Group general manager competency model stems from the original work by David McClelland (1973) and Richard E. Boyatzis (1982) with the McBer consultancy group. McBer and Company (now a part of the Hay Group) carried out the first competency study in 1973 (HayGroup, 2001). Much of the current
HayGroup model and methodology for its development stems back to the work of Boyatzis (1982). Boyatzis (1982) originally developed a competency model containing nine core competencies obtained from interview-based assessments of 253 managers of 12 Fortune 500 firms and four public agencies. The nine competencies included; efficiency, concern with impact, proactivity, self-confidence, oral presentation skills, conceptualization, diagnostic use of concepts, use of socialized power, and managing group process (Boyatzis, 1982). Ratings of the competencies were derived from coded interviews (Boyatzis, 1982). Stepwise discriminant analysis revealed that the ratings on these competencies yielded correct identification of superiors’ ordinal assessment of managers’ performance 51% of the time. However, the discriminant analysis was not tested with cross-validation, thus, classification accuracy estimates were most likely inflated due to chance associations (Russell, 2001). This suggests the need for further research on the development of and cross-validation of a competency model.

Since this original work, the HayGroup and their colleagues have conducted hundreds of competency studies on various jobs worldwide using similar methodology to that described above. Their most recent model, presented above, which contains four competency domains, is grounded in rigorous empirical methodology including behavioral event interviewing, traditional psychometric theory and analytic techniques, and in-depth analysis of the most extensive competency database in the world. In fact, they, along with Spencer and Spencer’s (1993) seminal text, “Competence at Work,” have devised the most elaborate listing of general managerial competencies, which they refer to as the competency dictionary. The HayGroup’s focus has been on the
competencies responsible for distinguishing outstanding mid- and first-level managers from typical managers. Thus, the model includes the most critical competencies for “defining excellence in a wide variety of management roles across industry” (HayGroup, 2001, p.2). The HayGroup model will serve as the foundational theory for the development of a transportation-specific managerial competency scale, as their model is constantly updated and refreshed to meet current economic and political challenges of modern day society. The model contains four clusters of competencies; Managing Yourself, Managing Your Team, Managing the Work, and Managing Collaboratively (HayGroup, 2001). Eleven competencies fall across these four domains. The eleven competencies are spread across the four clusters and include: Managing Yourself: 1) empathy (recognizes and responds to others’ feelings and concerns), 2) self-control (recognizes and manages one’s emotions and strong feelings under stress or when provoked), 3) self-confidence (possesses confidence in one’s ability to meet challenges and make the right decision); Managing Your Team: 4) developing others (helps others increase capabilities, maximize their potential, or recognize options), 5) holding people accountable (provides task focus and direction), 6) team leadership (creates an environment in which people can work together to meet organizational goals); Managing the Work: 7) results orientation (focuses on improving performance, meeting goals, and producing results), 8) initiative (sees opportunities and acts on them), 9) problem solving (identifies problems and tests alternative solutions); Managing Collaboratively: 10) influencing others (persuades, convences, or influences others to change their viewpoint), and 11) fostering teamwork (promotes cooperation and collaboration) (HayGroup, 2001).
This model is of particular importance to this study as it describes a general competency model thought to exist across industry lines. It is used here as the foundational theory to develop a transportation-specific managerial competency scale. A major charge of this study was to develop a measure that assessed these competencies plus additional competencies found to relate specifically to transportation (i.e., safety leadership), and then see if the HayGroup model could be found in the structure underlying the transportation sample.

Another recent landmark paper that must be mentioned in the discussion of empirically grounded models of competencies comes from the work of Tett et al (2000) at Wright State University in the Department of Psychology. In this eloquent study, the authors explored 12 of the most heavily cited taxonomies of managerial competence in the academic literature and then reported findings from three studies on the development and content validation of a “Hyperdimensional” Taxonomy of Managerial Competence (Tett et al., 2000). The term hyperdimensional is reportedly used to emphasize the quest for dimensions more specific than what models have previously proposed. As presented in Table 1 above, the Tett et al (2000) Hyperdimensional Taxonomy of General Managerial Competence, is a comprehensive model including 53 competencies that comprise 9 cluster or domain areas. The researchers linked each of the 53 competencies to competencies established and presented in the 12 taxonomies they reviewed. They produced this model by conducting three content validation studies. In these studies, the authors (Tett et al., 2000) mailed materials to the management participants and asked them to match 141 behavioral elements to various competency labels. The primary
research question they asked in improving specificity of their model was the degree to which behavioral element could be uniquely classified into targeted competencies. They employed the binomial test to compare observed frequencies with those expected due to chance. Thus, they were able to detect the number of people correctly classifying the element to the competency to reach statistical significance according to the binomial test. This methodology was employed to simply improve the specificity of managerial competency models. No doubt, the Tett et al (2000) model is one of the most elaborate and methodologically sound taxonomies reviewed for this study, and will be used and referenced in great detail as a comprehensive sounding board for critical competencies to include in the development of a transportation managerial competency scale.

Tett et al (2000) also dedicated a large portion of their manuscript to the current measurement issues facing the assessment instruments developed to accompany and apply the competency models reviewed. They suggested that several issues exist with most of the current assessment tools utilized today to measure general managerial competence (Tett et al.). They reported that “Psychological test developers face many challenges in creating reliable, valid, and usable measures,” (Tett et al., p.207). One of the most pressing measurement issues identified was the generality-specificity dilemma in which researchers in managerial behavior are caught between identifying general dimensions of managerial performance and also having to assume that specific exemplars within general categories are equivalent with respect to function, causes, and measurement (Tett et al.). This is a great point as it can be seen by the listing of models above that several different specific competencies of varying meanings get lumped into
certain clusters of management performance. More specifically, Tett et al. called attention
to the bandwidth and fidelity issue in personnel assessment in general. They explained
that two important concerns of the generality-specificity issue are fidelity, which denotes
the precision with which a measure captures a particular construct, and bandwidth, which
refers to the number of distinct constructs sampled by a given measure (Tett et al.). In
other words, Tett et al. suggested that researchers developing measures of competencies
are caught between either bandwidth or fidelity. That is, measuring a few things well
(high fidelity, high interpretability) or more things less well (broad bandwidth, more
comprehensive). They concluded that greater specificity should be the goal of
contemporary approaches to measurement. They accomplished this goal through the
methodology described above in their study in which they asked participants to match
behavioral elements to the appropriate and corresponding competency. They then
employed the binomial test to assess frequency versus chance responding. The Tett et al.
study raised important questions regarding the rationale behind competency measurement
tools that accompany highly complex models of managerial behavior, and their argument
will therefore be considered in the development of the current instrument.

Having just presented the reigning models of general managerial competence that
will be used as the theory backing the scale proposed for development, it is important to
also point out weaknesses of most managerial competency models prior to moving
forward. Several authors in leadership and management have alerted scholars in this area
to significant and concerning gaps in the current predominant competency models
(Bolden & Gosling, 2006; Hollenbeck, 2009). In a recent study by Bolden and Gosling
(2006), the authors compared a large review of leadership competency frameworks to an analysis of participant reports on leadership. The authors compared a total of 29 frameworks covering a vast spectrum of international organizations’ competency models to a database of 250 practicing managers’ perceptions of leadership and common challenges faced in their industry. Bolden and Gosling concluded that “a disturbing gap between attributes required of leaders as conveyed by practicing managers and popular leadership competency frameworks” (p.158) exists. The authors advanced their concern by explaining that competency frameworks tend to emphasize observable characteristics and behaviors while excluding moral and emotional concerns, which many leaders have argued are the dimensions that lie at the core of leadership (Bolden & Gosling). Further, the entire transformational leadership movement is based on moral responsibility and the ability to inspire employees at an emotional level. Thus, current leadership models may miss the boat in this area by focusing too exclusively on behavioral, observable skills (Bolden & Gosling, 2006). The current study will attempt to fill this gap by also including moral/emotional aspects in the definition and assessment of managerial competency.

**Methodology for generating the models.** Most of the reigning models of leadership and managerial competency (Calhoun et al., 2008; Hay Group, 2001; Tett et al., 2000) are based on a four phase model development process thought to be the gold standard in competency modeling (Spencer & Spencer, 1993): 1) the criteria that distinguish superior leaders from typical leaders are identified, 2) the job effectiveness criteria established in phase 1 are then used to identify a clear group of effective
transportation leaders and a comparison group of average performers, 3) data are
collected using qualitative and quantitative approaches to compare the two groups, and 4) the data is analyzed via basic descriptive statistics, qualitative analytic techniques such as content analysis, and then superior and average groups are compared on various competencies. Behavioral Event Interviews (BEI) are the most commonly employed method to identify the core leadership constructs. BEIs entail asking managers to describe times when they have reacted to challenging situations in a positive way and also times when they reacted in a negative manner (Boyatzis, 1998), and then are used to discover differences between two types of job incumbents: those who have been nominated as outstanding and those who are nominated as typical. CTT analytic techniques are commonly employed to analyze the models, validate them, and apply them to the business community.

To highlight the nature in which most competency frameworks and accompanying tools have been developed, let’s turn to a heavily cited development and initial validation study of a managerial competency scale. One study that is heavily cited in the literature for its rigor and thoroughness in developing a leadership performance scale based on a managerial competency framework dates back nearly twenty years to the work of Posner and Kouzes (1988, 1993). In this sound methodological study, two phases, “Qualitative Perspective on What Leaders Do” and then “Measuring what Leaders Do,” were conducted. These two phases follow a similar format to the four phase model (Spencer & Spencer, 1993) mentioned above as the typical competency framework development methodology used today. Managers attending leadership development seminars were first
asked to describe a “personal best as a leader.” Posner and Kouzes (1988) explained that the personal best question was really a 12 page survey that consisted of 37 open-ended questions (e.g., Who initiated the project? What made you believe you could accomplish the results you sought?) “focusing on an experience in which they got something extraordinary accomplished in an organization” (p.484). They (Posner and Kouzes, 1988) reported that over 650 surveys of the original version were completed, and then an additional 450 managers completed a shortened version of the same survey. 38 in-depth interviews with managers in various public and private sector companies were also conducted (Posner & Kouzes). The authors reported that the qualitative data gathered was content analyzed first by the authors and then by two outside raters (Posner & Kouzes). Results revealed that a “fundamental pattern of leadership behavior which emerges when people are accomplishing extraordinary things in organizations is best described by the following five practices: 1. Challenging the process, 2. Inspiring a Shared Vision, 3. Enabling other to act, 4. Modeling the way, and 5. Encouraging the heart” (Posner & Kouzes, p.485). They went on to report that approximately 80% or greater of the behaviors and strategies described in respondents’ “personal best case studies and interviews can be accounted for by these factors” (p.485).

Posner and Kouzes (1988) developed the Leadership Practices Inventory (LPI) in what they referred to as the second phase of their study. This phase of competency measurement scales is where the current study hopes to add useful scientific advancements relating to the statistical methods employed to get original psychometrics for scales. The current study is not as interested in improving model development
methodology, or the structure of competency frameworks, as that methodology has been advanced greatly in recent years (Boyatzis, 2007) and has yielded several wonderful models, such as the Hay Group (2001) model mentioned above. Nevertheless, the LPI was developed in the second phase via factor analysis, using principal factoring with iteration and varimax rotation. Posner and Kouzes reported that the factor analysis extracted five factors with eigenvalues greater than 1.0 and accounted for approximately 60% of the variance. Different subsamples were used to test the stability of the five factors and the five factors were similar to the five factors proposed (Posner & Kouzes). Internal reliabilities on the LPI ranged from .77 to .90, and the test-retest reliability was reported to be nearly .94 (Posner & Kouzes).

Though this information is useful and the scale development technique is valid, the purpose of the current study was to add more to managerial competency assessment by developing a scale with Rasch modeling techniques. Rasch modeling would allow the researcher to move beyond typical studies in managerial competence scales that were developed via similar analysis to the Posner and Kouzes (1988) study by placing the data on an interval scale and producing item-level statistics that will show that items vary with increasing amounts of competency in the participant and will cover the range of levels of managerial competency in participants. Rasch modeling methods will also move the field of competency-based performance measurement forward by ensuring unidimensionality, meeting the most basic criteria of measurement by placing the data on an interval scale, and assessing invariance in the measure (Bond & Fox, 2007), all of which is described below in greater detail in the methods section.
Leadership Theory

Clearly, one of the fundamental roles of any successful manager regardless of the industry is the ability to lead. Leadership is central to management. Many scholars have suggested that leadership is a competency or cluster area of management (Levenson et al., 2006). Indeed, it is fairly accepted today that in order to manage people one must possess some degree of leadership. Thus, in devising a managerial competency scale for transportation executives, it is important to first consult the relevant leadership theory available. Though many of the models of competency outlined above provide the necessary theory to support the proposed scale, reference to contemporary leadership theory will provide an even greater base of theory to devise a reliable, valid and comprehensive transportation managerial competency scale. In this brief review of leadership theory; leadership is defined, three general approaches to leadership (Trait, Behavioral, and Contingency approaches) that the current study rests on are presented, and two contemporary theories of leadership (Fiedler’s Contingency Theory and Charismatic and Transformational Leadership theories) are described for their contribution to the scale being developed in this study. The brief summary of the leadership theory that will guide the development of the transportation managerial competency scale ends with a discussion between the difference between leadership and management, as these two terms are often used interchangeably in a haphazard, detrimental manner (Kent, 2005). It is essential we distinguish the two prior to the development and initial validation of a general managerial competency scale, as the measure will clearly assess both aspects of management and leadership, but it is intended
to measure general transportation managerial competency as a single, overarching construct. Now, leadership is defined and then the theories that will be drawn from to develop the scale are presented.

The definitions of leadership that exist in the empirical literature alone are as diverse as the world’s population. Venturing to definitions outside of psychology’s empirical literature is actually overwhelming when attempting to synthesize what is meant by leadership. Without a doubt, leadership has become a construct that has evolved in to meaning very different things to different people (Kent, 2005). The body of literature on leadership in psychology alone is quite overwhelming. Several authors have suggested that part of the difficulty in defining it stems from the complexity of the leadership process (Yukl & Van Fleet, 1992). Leadership can and has been defined in several different ways. It has been defined as the behaviors that are enacted by a group leader, the functions they carry out while leading, or results of the behaviors (Jex, 2002). In fact, Jex (2002) reported that leadership definitions typically differ based on whether the emphasis is placed on behaviors or outcomes of those actions. Luckily, the academic quest to define the term has produced some overlap in definitions that can be drawn upon today to define it.

Yukl and Van Fleet (1992) described leadership as “a process that includes influencing the task objectives and strategies of an organization, influencing people in the organization to implement the strategies and achieve the objectives, influencing the group maintenance and identification, and influencing the culture of the organization” (p.149). Several assumptions are inherent in this definition such as the implication that leadership
involves influencing others, it is a process, and that leaders are change agents with a variety of skills (Jex, 2002). For the purpose of the current study, however, it was also important to define leadership in tandem with management, and to clarify the differences between the two.

In a 2005 article in Management Decision, Thomas Kent provided an exceptional overview and stimulating discussion of the modern view of leadership as it relates to management. Kent (2005) reported that though separate processes, managing and leading are inherent within the same individual and must be done simultaneously in most modern day positions within organizations. In this provoking article of how leadership is conceptualized in regards to management, Kent (2005) asked this important question, “From the standpoint of the purposes of the two processes, how effective is it for a leader to develop a vision for the organization and to muster support and motivation to pursue it if the manager in him/her does not procure and efficiently allocate the resources to accomplish the vision?” (p.1013). Kent (2005) eloquently highlighted the reality that though we want to separate the two to distinguish the differences, one process greatly relies upon the other. Finally, clarity on the difference is better seen in thinking of the purpose of the two processes. The purpose of leading is “to create direction and the drive to pursue it through the development of people’s thinking and valuing” while the purpose of managing, on the other hand, is “to determine and compare alternative uses and allocations of resources and to select that alternative which is most energy effective toward accomplishing or producing a product, end or goal” (Kent, 2005, p.1013). Further, various jobs in different industries will require unique degrees of each skill, leading and
managing (Kent, 2005). Thus, the position on leadership presented here is that leadership focuses on creating vision, aligning people within a group, sparking drive within employees, and other processes related to employee engagement and motivation, while management relates to administrative duties such as planning, organizing, controlling, and coordinating. Based on Kent’s (2005) argument, then, it is also assumed here that a certain degree of each is needed to carry out most managerial jobs in modern day organizations. Being able to perform both leadership and managerial competencies together would lead to what Kent (2005) referred to as a “complete leader/manager.” (p.1015). Thus, the goal of the current study was to develop a scale that assesses “complete leader/manager” ability, referred to here as general managerial competence. The researcher used Kent’s (2005) paper to assume that transportation managers will need a unique blend of the two skills sets, leading and managing. Thus, the instrument will also aim to measure a balance between both leadership and managerial ability. Greater discussion is presented following the theories of leadership below on the differences between leadership and management. This section on the differences between the two was only included here to distinguish between the two and better define leadership.

Another recent and pressing article that relates strongly to the current review and added tremendous insight to the conceptualization of leadership came from Kaiser, Hogan, and Craig (2008). In this study published in the American Psychologist last year, the authors proposed the idea of conceptualizing leadership and evaluating leaders in terms of organizational and team performance (Kaiser et al., 2008). The authors reported
that the 100-year old body of leadership theory reveals methodological diversity which could point to “a robust literature but may also reflect a lack of definitional clarity” (Kaiser et al., 2008, p.97). In this study, the authors looked to how leadership has been measured for greater clarity on the topic. They reviewed 10 meta-analytic studies to determine how research has been measured over the past century. The meta-analyses included over 280,000 leaders from 1,124 samples and 1,695 statistical tests of the relation between leader predictors and criteria (Kaiser et al., 2008). The authors then content analyzed the criterion variable and identified two categories of leadership measures, each with two subcategories; 1) measures focusing on individual leaders and 2) measures focused on groups, teams, and organizations (Kaiser et al., 2008). The study yielded support for the notion that leadership effectiveness should be defined by the performance of the organization, that measures of leadership should be aware of the difference between perceptions of leaders and actual leadership effectiveness, and that more research should focus on organizational outcomes to enhance applicability and “real-world relevance” (Kaiser et al., 2008). The Kaiser et al. (2008) article is reviewed here for its contribution to the conceptualization of leadership.

There are several distinct approaches to leadership that provide additional clarity to this conversation. As a result, the following theories will greatly shape the development of this scale. The first three theories presented are classic leadership theories that relate to the development of a managerial competency scale, and the last two are more contemporary approaches on the topic.
The Trait approach. The Trait approach is one of the older theories of leadership behavior within organizational psychology. The initial general premise was that effective leaders possess traits that are different from those who are less effective leaders (Zaccaro, 2007). Thus, scholars under the trait approach were primarily concerned with identifying the traits that distinguish superior from typical leaders. After a series of studies found equivocal support for the exact traits responsible for successful leadership, research in this area moved to the study of leader emergence (Zaccaro, 2007). Researchers in this area have identified traits that predict who will emerge as a leader in a group when no formally designated leader was identified. Some of the factors found to predict leader emergence include intelligence, higher needs for dominance, and high self-monitoring (Jex, 2002). Yukl and Van Fleet (1992) suggested that several traits have also been identified under this approach to predict managerial effectiveness including high energy, stress tolerance, emotional maturity, and self-confidence. The trait approach also offers a means to measure leadership as it has identified the traits that are most critical and should be measured. Items can easily be written to reflect specific leadership traits. A great deal of work has been done on this in psychology already that can be used for the current project as many of the leader and managerial competencies essential for successful management rely on key traits. One critical difference between the trait and competency approach, however, is that traits refer to more longstanding, pervasive aspects of personality while competencies refer to specific, observable behaviors.

The behavioral approach. The competency movement, as described earlier in this chapter, was primarily based on the notion that aptitude testing needed to be focused
more on observable, behavioral characteristics as opposed to less measurable, socially constructed concepts, as seen in intelligence testing. Thus, the behavioral approach to leadership may resonate and better qualify as the theory behind this instrument than any other approach presented in this text. The basic idea of the behavioral approach is similar to the trait approach only behaviors are the focus of distinguishing characteristics of successful leadership (Ricciardi, 2005). The behavioral approach posits that leadership behavior is divided into two general domains; (1) initiating structure and (2) consideration. Initiating structure behaviors include those that are geared toward facilitating the task performance of groups (e.g., communicating performance expectations, keeping subordinates focused, or organizing work for employee units) (Jex, 2002). Further, consideration behaviors are aimed at showing employees that they are valued and recognized. Controversy over the division of all leader behaviors into two broad categories is still heavily debated (Jex, 2002). The behavioral approach has been marked with criticism about the inability of researchers in this area to identify a core set of behaviors that are consistently related to effectiveness. This may suggest that leader behaviors may be based on context, or dependent upon the unique leadership position and situation. This issue led to the contingency approach to leadership.

**The contingency approach.** The central tenet to the contingency approach to leadership is that effective leadership is based on the relationship between leader behaviors and traits and characteristics of the situation the leader is in. Thus, contingency theorists believe that the job of the leader is to assess which behaviors and traits are necessary to successfully perform various duties within unique contexts (Fiedler &
Mahar, 1979). Nearly all leadership theories that have emerged over the past thirty years have been contingency theories. There is general consensus in the field that the basic premise of contingency theories is true in most leadership domains. Thus, it may be important in any leadership or managerial competency scale to assess the flexibility in ability to apply the various skills needed to perform specific duties.

**Fiedler’s Contingency Theory.** One of the more established and well regarded contingency theories to emerge over the past three decades is Fiedler’s Contingency Theory. Like all theories under the contingency umbrella, the general idea is that the success of a leader is dependent upon the interaction between characteristics of the leader and the situation. Fiedler took the theory a step further by looking at aspects of the situation that are more or less favorable toward a positive outcome. This theory identifies three factors that lead to what they refer to as situation favorability, or the likelihood that a situation is favorable to a leader. The most favorable situations for leaders, based on Fiedler’s theory, are those in which leader-member relations are good, task structure is high, and position power (the amount of formal authority over subordinates) is high. The second portion of Fiedler’s theory relates to the characteristics of the leader, which is more relevant to the task of developing a managerial competency scale for transportation managers. This theory states that leaders can be reliably distinguished based on two dimensions, or whether they are “task-oriented” or “relationship-oriented.” Fiedler then proposed that leaders who are task-oriented fare best in either highly unfavorable or highly favorable situations, while relationship-oriented leaders do best in moderately favorable situations (Fiedler & Garcia, 1987). In other words, when the situation is
clearly favorable or not for the leader, the best outcome will most likely result with a
task-oriented leader. On the other hand, if the situation is one in which it is only
moderately favorable, or there is ambiguity and division, relationship-oriented leaders
will fare the best. The logic stems from the idea that when situations are not clearly
defined managers must rely on interpersonal skill to navigate the group to the best
possible outcome. Despite numerous outcome studies on Fiedler’s theory, equivocal
support has been found (Jex, 2002).

Charismatic and transformational leadership. The charismatic and
transformational leadership theories are two of the most recent approaches to leadership.
The general idea with transformational leadership is that there are “certain traits and
behaviors that not only influence subordinates but may also inspire them to perform well
beyond their capabilities” (Jex, 2002, p. 285). It is also posited that both have the
capability to influence meaningful change in organizations. Under this theory, the term
transactional leadership is used to describe the opposite style of leadership than
transformational and charismatic. A transactional leader is thought to be one who
enforces rules to ensure that workers get the job done rather than inspiring subordinates
or facilitating meaningful change. One central concept to this theory is called vision
implementation. This concept relates to the ability of a leader to persuade others to
implement their vision. The general premise is that it is not very helpful for a leader or
manager to possess a vision and have drive if they cannot implement the vision and spark
drive among the organization’s members. Finally, this theory suggests that a core
attribute of a transformational and charismatic leader is one who possesses a charismatic
communication style. A charismatic communication style is based on a captivating voice, direct eye contact, and other specified characteristics of communicating. These attributes of transformational and charismatic leaders have been empirically linked to employee job performance, satisfaction, and positive perceptions of leaders (Bass & Avolio, 1994). This theory is applied via selection and assessment procedures. For instance, based on this theory, superior managers most likely possess attributes thought to be associated with transformational and charismatic leadership.

One of the greatest contributions from a transformational theorist to date came from the work of Bernard Bass and Bruce Avolio (1994) in their seminal text, “Improving Organizational Effectiveness.” In this book, Bass and Avolio (1994) show how the concepts under transformational and transactional leadership apply to specific areas of leadership, management, and organizational development. Bass and Avolio (1994) provided an excellent clarification of the term transformational leadership by suggesting that it is seen when leaders, “stimulate interest among colleagues and followers to view their work from new perspectives, generate awareness of the mission or vision of the team and organization, develop colleagues and followers to higher levels of ability and potential, and motivate colleagues and followers to look beyond their own interests toward those that will benefit the group” (p.2). Much of their work in this text explained how transformational leaders develop their people to higher levels of potential. They provided several competencies essential for successful management at all levels of an organization by proposing a “full-range model” of transformational, transactional, and non-transactional leadership (Bass & Avolio, 1994). For instance, the authors suggested
that transformational leaders “behave in ways to achieve superior results by employing one or more of the “Four I’s”” (Bass & Avolio, 1994, p.3). The four I’s include: 1. Idealized influence (e.g., being a role model), 2. Inspirational motivation (e.g., behave in ways that motivate those around them by providing meaning), 3. Intellectual stimulation (e.g., question assumptions of followers), and 4. Individualized consideration (e.g., pay special attention to subordinate’s needs for achievement and growth). The authors (Bass & Avolio, 1994) go on to present ways that transformational leadership theory can be unpackaged to boost individual leader and organizational potential. In developing the transportation managerial competency scale, then, this theory will be referenced to ensure that the attributes of transformational leadership that have been linked to employee job performance are included in the scale.

**Managing versus leading.** Managers and leaders are distinct groups of people within organizations yet the two terms are commonly used interchangeably (Kent, 2005). Like other constructs in psychology, operationally defining leadership has not always been the easiest task. One of the challenges is the relation and overlap it shares with the similar term of management. Jex (2002) reported that this issue is especially important to practitioners or administrators in the business world. Similarly, the distinction is equally important in devising a valid assessment tool that truly measures one construct, not leadership and management, but managerial competence. A manager is typically defined as one who engages in traditional administrative behaviors and duties such as planning, overseeing the work of employees, and providing disciplinary actions to ensure all work gets done. Leaders, conversely, are not only asked to fulfill these essential administrative
duties but are asked to inspire and motivate people to new levels of being and working. Thus, leadership requires going beyond simply carrying out managerial tasks to perform these tasks in a way that moves the organization forward with charisma and drive. For the purposes of the current study, leadership was thought to be a characteristic of a competent manager, or subsumed in management as presented by several of the reigning managerial competency frameworks listed in Table 1. Therefore, we used the term managerial competence as the variable of measurement interest but made sure to include several competencies that clearly relate more to the leadership side of great management.

**Managerial Competency Outcome Research**

The concept of management competency has become omnipresent within the field of performance assessment and organizational development (Bolden & Gosling, 2006). Despite widespread use of competencies to improve leadership performance and billions of dollars spent on leadership competency systems, the academic and applied research literature is only in its infancy and has lagged behind (Boyatzis, 2007). Steady research on the topic has come from consulting firms and American corporation’s human resource departments, yet little has been published in the academic literature until the past decade (Boyatzis, 2007). It is important we review this literature prior to the development of an assessment tool of managerial competency to determine which factors or competencies underlying general managerial competence have been empirically linked to positive outcomes at the individual and organizational level.

The following research questions guided the review of the empirical literature in order to assess the need and usefulness of a general managerial competency scale for the
transportation industry. Do competencies predict performance? If yes, which competencies specifically discriminate between superior and typical managers? Can managerial competency assessment tools assist in the selection, training, and development process? Several studies were found that yielded empirical evidence to support the development of a transportation managerial competency scale.

Much of the outcome research on the efficacy of using competency systems focuses on performance. At their core, competencies are “an employee’s ability to perform the skills required for a specific job” (Levenson et al., 2006, p.361). For the past fifty years organizations have placed major stock on the notion that assessments of employee’s competencies can yield an effective means of predicting job performance (McClelland, 1973; Spencer & Spencer, 1993). Though competency systems are heavily used today across various industries to select, reward, and promote managers, limited empirical evidence exists for the effectiveness of managerial competency systems (Hollenbeck, 2009; Levenson et al., 2006). More specifically, little data exists that shows that managerial competency systems increase managerial effectiveness. Despite large gaps, competency assessments have been shown to predict individual managerial success as measured by 360-degree ratings (Goldstein, Yusko, & Nicolopoulos, 2001), and other promising findings have emerged relating competency measurement to performance outcomes.

In a recent study by Levenson et al. (2006), the researchers examined the relationship between managerial competencies and performance at both the individual and organizational unit levels. The researchers were interested in determining whether
competency-based assessment tools, as the one proposed for development in this study, are better than traditional human resource tools. The sample was composed of 1,279 first and mid-level managers of a Fortune 500 consumer products company. Levenson et al. (2006) found that higher level competency managers had higher individual performance ratings, which was hypothesized. Further, they suggested that years of experience was positively correlated with individual performance (Levenson et al.). Individual performance was determined by both business (e.g., manufacturing line uptime, defects, shipment accuracy) and people results (injury rates and other objective personnel metrics), both rated from 1 (worst) to 5 (best). The site performance ratings were generated by archival data and are a weighted average of the cost reduction, productivity, and injury goals (Levenson et al.). In general, the authors concluded that a positive relationship existed for this sample between higher competency levels and individual-level performance, and a weaker relationship was found with site-level performance. A positive relationship was also detected between mentoring based on the competency system and individual performance, suggesting that indeed, competencies can be used for training and development as well. Levenson et al. (2006) demonstrated that competencies are more strongly related to performance than traditional human capital is. Human capital in this study was defined by the typical or standard human capital variables such as education and experience (Levenson et al.). Finally, the authors concluded that aggregated managerial competencies were only related to site performance for the medium and large organizations and not the smaller ones. This suggests a “contingency interpretation” of the relation between competencies and organization-wide performance.
In other words, competencies may be more conducive and appropriate for larger companies, and may predict performance better the larger the organization. Having said that, there were a few limitations to this study that should also be mentioned. For instance, the definition of human capital appears limiting and slightly outdated (e.g., educational level). Further, they explained the competency system but did not report on the development of the system and the methods employed. This limits the usefulness of the findings in that methodological concerns cannot be addressed or answered.

Another relevant study from the past decade that explored the relation between managerial competencies and job performance came from Goldstein et al. (2001). In this study of 633 employees from various public and private organizations representing numerous industries, the investigators examined black and white subgroup differences of managerial competencies and the moderating effect of ethnicity (black versus white) on the relation of competencies and job performance. The general premise of this research was that validity and diversity have been competing factors in personnel selection as validity of predicting individual job performance and diversity of the selected individuals on characteristics such as race have been at odds. Participants in this study were asked to complete a cognitive ability test in conjunction with the competency inventory. Thus, the study investigated the extent to which job-relevant managerial competencies vary by ethnicity (white versus black) and if managerial competencies with less cognitive load had smaller black-white subgroup mean differences than managerial competencies with higher levels of cognitive load. Cognitive load was defined as the degree of cognitive content involved with the competency. The idea with the cognitive load component of the
study is to measure noncognitive capabilities that are critical to performance and tend to have fewer subgroup differences than cognitive ability (Goldstein et al., 2001). The results demonstrated that managerial competencies do vary in the size of their black-white subgroup differences (Goldstein et al., 2001). Further, abilities with less cognitive load (e.g., effectiveness under stress) tended to lead to nonsignificant Black-White subgroup differences while still contributing incremental validity above cognitive ability to job performance. Based on these important findings, the current project of developing a managerial competency scale should be based on the notion that ethnicity and cognitive load do affect the development, administration and interpretation of the assessment tool.

In other words, race matters on the various competencies that make up the primary construct of general managerial competence, and both cognitive and noncognitive items should be included in the scale proposed. In regards to performance, the results suggested that competencies with greater cognitive load more strongly predicted cognitive aspects of job performance as compared to noncognitive aspects of job performance. Though this study contributed significantly to the competency literature, the limitations also must be mentioned. A major concern was how representative the sample was of the management population to which the researchers were attempting to generalize. The sample size of African-Americans was relatively small at 88 for the statistical analysis that was conducted. Nevertheless, the study produced some intriguing findings that will certainly shape future interpretations of competency findings across racial lines.

Another recent study (Dreyfus, 2008) investigated the competencies that predict highly effective performance in science and engineering managers. All 35 participants
were working as first level managers at a major US government research center in the Mid-West. Nine factors were found to discriminate highly effective from typical managers. The researchers reported that the nine variables grouped respectively into two competencies: managing groups and interpersonal sensitivity. It was also concluded that the highly effective managers demonstrated more interpersonal competency than their average peers, therefore demonstrating that interpersonal skills make a difference.

Further, the researchers hypothesized and demonstrated the importance of including social and emotional intelligence competencies in management competency models, as these were found to be discriminating competency areas between superior and typical performers.

In a recent Guest Editorial by Boyatzis (2007) in the *Journal of Management Development*, the legendary competency scholar expanded on the notion of emotional, social, and cognitive intelligence competencies while providing an update on competencies in the 21st century. This manuscript ties in nicely with the discussion on what we know about competencies that relate strongly to performance and discriminate typical from superior performers. Boyatzis reported that “An integrated concept of emotional, social, and cognitive intelligence competencies offers more than a convenient framework for describing human dispositions. It offers a theoretical structure for the organization of personality and linking it to a theory of action and job performance” (p.21). He then defined emotional intelligence competency as an ability to recognize and use emotional information about oneself that leads to or causes effective or superior performance (Boyatzis). Boyatzis made the point that emotional intelligence
competencies very much relate to performance and cited numerous studies to empirically support this point. Thus, in the current study, incorporation of emotional competencies was a core aspect of the development of a general managerial competency scale for managers in transportation.

Another highly regarded article on competencies surfaced in *Human Resource Development International* that explored if different leadership competencies were important in different countries, and if so, which ones were important to which countries (Kowske & Anthony, 2007). Kowske and Anthony administered Personnel Decisions International’s (PDI) PROFILER®, which is an instrument that assesses leadership competence importance. The sample was composed of 18,425 participants representing 561 companies in 12 countries. The authors concluded that there are differences in leadership competency importance ratings across countries, yet some countries share ideas about what’s most important. Further, Kowske and Anthony suggested that Analyse Issues (e.g., gathers relevant information systematically, considers broad range of issues or factors, seeks input from others) and Foster Teamwork (e.g., builds effective teams committed to organizational goals, uses teams to address relevant issues) were identified as two competencies that are significantly viewed differently across countries. However, the authors concluded that these two competencies are rated consistently within the top six competencies on importance. Thus, Kowske and Anthony reported that these two competencies may be considered universal or generic managerial/leadership duties. This article also proposed countries archetypes of leadership style. Thus, it is a great resource
for the development of a culturally-sensitive measure of managerial competence for transportation executives in the US.

Finally, Hopkins and Bilimoria (2007) added to the literature on the relation of competencies to performance when they analyzed data from a sample composed of 130 upper-level executives (90 males and 40 females). In this recent examination, the authors looked at the gender differences in the demonstration and predictability of emotional and social competencies to performance. Three primary research questions were explored: (1) Are there differences between male and female leaders in their overall level of emotional and social intelligence competencies? (2) What is the relationship between the overall level of emotional and social intelligence competencies and success, and does gender moderate that relationship? (3) Are there any significant differences between the most successful male and female leaders in their overall level of emotional and social intelligence competencies? The authors reported no significant differences between male and female leaders in emotional and social competencies. However, they concluded that there were significant differences found between the most successful male and female leaders. Hopkins and Bilimoria (2007) reported that gender does moderate the relation between emotional and social competence and success. Further, only male leaders were more successful when they demonstrated higher competencies (Hopkins & Bulimoria).

The findings from this study suggest that gender is an important factor when identifying the core competencies that distinguish superior from typical leaders.

As one can see from the studies presented above, there is mounting evidence that suggests that competencies can and do predict performance. The research appears to
suggest that ethnic, racial and gender differences most likely moderate the relation between competencies and performance. These will be central questions to address in future research with the scale created in this study. Further, the current literature provides sufficient guidance on the key competencies to include in the development of a transportation managerial competency scale. Several aspects of the current state of transportation warrant additional research toward managerial performance improvement.

**State of Transportation Leadership**

Several glaring concerns and current factors in the transportation industry make this an ideal and possibly critical time to develop a valid and psychometrically sound instrument to assess leadership competency in the transportation industry. Transportation agencies in the public and private sector move people and products. Thus, safety and efficiency is key to success (Marti, 2007). Several studies in various industries have documented the relation between leadership competency assessment systems and individual and organizational performance outcomes (Levenson et al., 2008). It is alarming that transportation operates at the responsibility level it does without such competency systems in place to ensure quality leadership that can guide various transportation agencies to safe and efficient practices. The transportation workforce is facing a potential crisis if more planning and attention is not garnered.

A 2007 article in the Minnesota Local Technical Assistance Program (LTAP) newsletter provided one of the best summaries of the causes of the workforce development crisis in transportation. Marti (2007) reported that the mix of baby boomers nearing retirement at alarming rates, whom lead at all levels of various private and public
transportation sectors, coupled with increasing demands for transportation amidst worldwide efforts for environmentally friendly modes of moving people and products, have left the transportation workforce with the ingredients for the perfect storm. Further, the workforce is changing rapidly in socio-economic status, age, and values, and leaders are needed that can motivate diverse employees to safely compete in a global market.

A separate report by the Wisconsin Department of Transportation (WisDOT; 2005) reported workforce development issues a top priority. The report, prepared by CTC & Associates, LLC for the WisDot Research Administrator, suggested that transportation agencies will be asked to do more with less in the 21st century. The report also noted that the pool of qualified applicants has shrunk due to public image issues (CTC & Associates, 2005). Further, the report acknowledged the expected complexities of leading in transportation in the years to come by highlighting the strong possibility that transportation agencies will face downsizing, outsourcing, and greater public interaction, all of which require keen leadership and effective, flexible manager training programs. The need for a reliable and valid instrument for transportation to compete is clear.

In 2001 the Federal Transit Administration (FTA) sponsored the American Public Transportation Association’s (APTA) Workforce Development Initiative. With the Transit Cooperative Research Board (TCRP), the study identified the most important challenges that the transit industry currently faces in workforce development. Brian Vogel (2001), principal investigator of the study, reported that it was universally agreed upon that workforce issues are critical to the success of the industry. Vogel (2001) suggested that the industry faces severe recruitment and retention of quality worker issues
in the face of increasing demand for services. Further, Vogel (2001) specified the need for industry-developed models that could be used to measure the return on training or other human resources development. Additional issues identified included an aging workforce, problems with succession planning, recruitment difficulty, training issues, and planning issues. Nearly every issue identified was also related to leadership and managerial concerns. The system in place for recruiting, identifying, selecting, and training talented and highly effective leaders in the transportation industry appears to be out dated (Vogel, 2001). The culture of the transportation industry has been slow to respond to the vast societal changes and subsequent transit needs abroad. Thus, development of a managerial competency scale for leaders in transportation is warranted. Further, the literature supports the use of Rasch modeling methods as transportation manager’s competencies can be best explored with item-level analytic methods using interval data.

**Benefits of Rasch Analysis**

Though some researchers have turned to Rasch analysis to develop and validate leadership measures used in business, most researchers continue to rely on Classical Test Theory (CTT) techniques. One of the unique aspects and purposes of the current study is the application of Item-Response Theory (IRT) principles and procedures to the development of an assessment tool for leadership competency. Few if any studies to date have used Rasch analyses methods to develop a managerial competency scale. IRT is an alternative approach to CTT (CTT; DeVallis, 2003), in that it focuses primarily on individual items and their characteristics, where CTT is more concerned with composites
and scales. Thus, in IRT researchers assess each item’s relationship to the variable of interest (DeVallis, 2003). One of the great advantages of IRT’s analysis at the item level is that the measure’s invariance can be assessed. Thus, if the managerial competency scale is developed with IRT techniques, it should be a useful measurement of managerial competency regardless of the sample or transportation agency. Further, several well-documented limitations of CTT are avoided by using IRT procedures. Some of the limitations of scale development with CTT are that the item difficulty (or item position) is dependent upon the distribution of person scores and therefore sample dependent, it does not measure on an interval scale and therefore is difficult to detect difference in position, and finally, CTT attends only to item fit and ignores person fit or ability. These are merely a few of the limitations of CTT that IRT attempts to overcome.

Rasch analysis and other IRT models are rapidly gaining popularity over CTT approaches. Rasch examines item-level psychometrics by first converting an instrument’s ordinal data into interval data, and as a result, meeting the most basic criteria of true measurement (Pomeranz, Byers, Moorhouse, Velozo, & Spitznagel, 2008). Pomeranz et al. (2008) explained that Rasch techniques can provide psychometric information that CTT techniques have not been able to provide. For example, by applying Rasch analytical techniques to managerial competency measurement, we will be able to account for the unidimensionality of the measure, or the degree to which it assesses one latent variable, in this case, managerial competency. Unidimensionality is determined by assessing item “fit” statistics, which indicate which items need to be removed to ensure that the instrument is unidimensional (DeVallis, 2003).
By applying Rasch analysis to the development of the managerial competency scale, item hierarchy will also be determined. Rasch modeling was used more for item hierarchy and targeting in this study than for any other reasons. Item hierarchy identifies items from least to most difficult, and also provides the researcher with knowledge of any gaps in the measurement tool that would reduce the efficiency and accuracy (Bond & Fox, 2007). It tells the researcher what types of items (i.e., more difficult or less difficult to endorse) are needed to be written to cover the full range of managerial competency abilities. Needless to say, there are several advantages to using Rasch analysis as the primary development and validation technique in this study, and as a result, application of this contemporary technique is one of the most exciting aspects of this study.

Competency assessments are also commonly used with two groups (e.g., upper and lower level managers, typical versus superior performers, men and women, management groups in different agencies, etc.). IRT is particularly appropriate for this approach to measurement by use of differential item functioning (DIF). DIF allows the researcher to identify whether responses differ based on groups, that is, are the items reliable across groups? In order to compare two groups, as is frequently done in practice with competency assessments, it must be assumed that the competency scale performs identically with both groups and any differences detected are the result of the attribute of interest (DeVellis, 2003). DeVellis (2003) suggested that “Classical methods may tell part of the story (e.g., by noting markedly different factor patterns across groups) but may not detect more subtle processes” (p.151). The current models that most competency scales are based on, including the Hay Group model (2001), stem from work done with
CTT methods. Though these scales are useful and provide a great jumping off place for managerial competency measurement, the Rasch approach from the IRT framework moves managerial competency assessment to an entire new level, with more advanced data at the item level, and more possibilities for leadership development application emerging as a result.

**Absenteeism**

Absenteeism was observed in this study as another means to assess the validity of the new scale. The idea was that absenteeism should correlate negatively with any measure of performance or competence. Absenteeism as an organizational outcome metric is not uncommon in the industrial-organizational literature (Johns, 2001). It is often studied as an outcome associated with job satisfaction (Farell & Stamm, 1988). Most agree that absenteeism is a sign of either job dissatisfaction (Johns, 2001) or psychological distress (Johns & Xie, 1998). It is therefore assumed that it would inversely relate to performance, or in this case, managerial competence, as both psychological distress and job dissatisfaction have been found to predict important job outcomes. Absenteeism has long been thought of as a consequence of burnout as well (Maslach & Leiter, 1999). Burnout has also been linked to performance (Petitta & Vecchione, 2011). Thus, absenteeism was used in this study to assess initial discriminant validity of the new scale.

**Summary**

The contemporary concept of competencies dates back to the work of psychologist David McClelland (1973). McClelland was concerned with the widespread
use of intelligence and related aptitude tests, which he deemed too removed from practical outcomes at work and in education (Garman & Johnson, 2006). He proposed that competencies, thought here to describe underlying characteristics related to successful performance of a job (Boyatzis, 1982), should garner greater attention as a more useful means to measuring aptitude. Since then, an outpouring of sophisticated research studies on the development of competency models and outcome research on the efficacy of competencies to key performance outcomes has taken place. Today competencies are being used to clarify key job descriptions and to recruit, train, and assess employee performance based on behaviorally defined characteristics.

Several general managerial competency models are in use today as well. One of the most heavily cited and recognized models came from the McBer consultancy group, who proposed four clusters of general managerial competency: Managing yourself, managing your team, managing your work, and managing collaboratively. This model is just one of several general managerial competency frameworks that was used in this study as a foundation to generate the transportation managerial competency scale. Most of these general models are used in specific industries without thought of industry-specific tasks. However, there is a push to develop industry-specific managerial competency scales. The research on competencies also demonstrates soft or people skill and knowledge is as important as technical expertise across industry lines. Fostering teamwork, social competencies, and emotional competencies are just a few of the “people skills” thought to be critical characteristics that distinguish superior from average leaders.
Finally, several studies have explored the essential competencies of managerial competence but few have done so specifically for transportation.
Chapter 3—Methodology

The purpose of the current study was to develop a scale for assessing transportation managerial competency that could be used in the recruitment, selection, and training of transportation leaders. More specifically, the aim was to obtain preliminary psychometric data on the scale and conduct an initial validation study. This chapter describes the two studies conducted to initially develop and evaluate the Managerial Success Factors Inventory: Transportation Version (MSFI: TV). The scale was designed to specifically assess four factors found to underlie general managerial competence: Managing yourself, managing your team, managing your work, and managing collaboratively (HayGroup, 2001). An 8-step, three study approach (DeVellis, 2003) to constructing a sound measurement tool was followed, and both factor analytic and Rasch modeling techniques were conducted to develop an initial version of the MSFI: TV. The specific procedural and analytical steps taken are explained in detail in this chapter.

Sample

Three separate samples were obtained in this study: one sample of managers who participated in one of three focus groups held prior to data collection, one self-assessment sample composed of managers who assessed their own leadership competencies, and then one sample of peer raters (i.e., supervisors) who assessed the individual who requested their feedback to provide a validation step.
Fifteen managers at a state Department of Transportation located in the Western half of the United States participated in one of three focus groups. Each focus group was composed of five participants who had been identified by a senior human resources manager at the DOT as outstanding in his or her position and expert in transportation leadership. These fifteen managers were diverse in ethnicity, age, rank, years with the DOT, and position or job title. Of the fifteen focus group participants, two were female.

The self-assessment sample consisted of 287 managers (227 males and 60 females) at a state DOT located in the Western half of the United States. The ethnic breakdown of the self-assessment sample was: 215 Caucasians (74.9%), 34 Hispanics, Latino/as (11.8%), 7 African-Americans (2.4%), 7 Asian or Pacific Islanders (2.4%), 6 Bi-racials (2.1%), 5 Multiracials (1.7%), 2 American Indian or Native Alaskan (.7%), and 11 who indicated Other (3.8%). The average age of the self-assessment sample was 48.6 years old and the mean annual salary was $85,710.67. On average, self-assessment participants supervised 17 people at the time of data collection. 285 of the self-assessment managers indicated that they were full-time while only 2 managers indicated that they were part-time employees. There was variance in job classification or title as well: 10 Entry-level manager (i.e., Team Leaders) (3.5%), 173 Supervisors (60.3%), 86 Managers (30%), 7 Directors (2.4%), 3 Executives (1%), and 8 indicated Other (2.8%). The average length of time in this position for this sample was 5.38 years, while the average length of time in the field for this sample was 20.23 years. This self-assessment sample was randomly split in two equal halves to allow the researcher to follow the conventional three study scale development research design. The file was split randomly.
to ensure equal representation across groups. The first half \((n = 144)\) was utilized as the exploratory factor analysis (developmental) sample and the second half became the confirmatory sample \((n = 143)\).

The peer rater sample consisted of 287 managers, supervisors, directors, and executives at a DOT in the Western half of the United States. These individuals were only asked to indicate their relationship to the ratee who provided a self-assessment and then rate that individual requesting feedback on a five item measure of performance. The peer feedback sample indicated their relationship to the ratee requesting feedback: 118 Managers/Supervisors (41.1%), 81 Colleague/Peers (28.2%), 67 Direct Reports (22.3%), 9 Customers (3.1%), and 12 Unspecified (5.3%).

**Measures**

**Demographic questionnaire.** A short demographic questionnaire (Appendix A) was included in the MSFI: TV. The demographic section assessed each participant’s age, ethnicity, gender, length of time in the field, length of time in the current position, marital status and salary.

**Absenteeism.** Absenteeism was assessed by self-report of a single item that read, “In the past 30 days, how many full days of work have you missed?”. Alice and Gregory (2010) demonstrated support for self-report, single-item measures of absenteeism when compared to organizational records. They found a strong positive correlation, a strong intra-class correlation and strong Cronbach’s alpha for both measures of absenteeism; organizational records and self-report time missed. The literature supports use of self-
report absenteeism measures (Alice & Gregory, 2010; Short, M., Goetzel, R., Pei, X., Tabrizi, M., Ozminkowski, R., Gibson, T., Dejoy, D., & Wilson).

**Performance questionnaire.** A short five item performance questionnaire (Appendix B) was also included in the survey. The performance items assessed the participant’s beliefs about their performance over the past year. Participants were asked to respond to each question based on how much they agree with the 5 statements (Very Strongly Disagree = 1, Very Strongly Agree = 6). Two internal consistency estimates of reliability were computed for the Performance scale: a split-half coefficient expressed as a Spearman-Brown corrected correlation and coefficient alpha. For the split-half coefficient, the scale was split into two halves such that the two halves would be equivalent as possible. In splitting the items, we took into account the sequencing of the items as well as whether items assessed the same aspect of performance. The first half included items 1, 3, and 5 from the performance scale, while the second half included items 2 and 4. The Spearman-Brown coefficient for unequal length was .91. A coefficient alpha was also computed to assess the reliability of the performance scale. For the coefficient alpha, the greater the consistency in responses among items, the higher the coefficient alpha will be. The coefficient alpha was found to be .89, which suggests that the scale scores are reasonably reliable for respondents like those in the study.

**Supervisor’s ratings.** Managers who completed the MSFI: TV were asked to list a supervisor who could rate their performance as a manager. Using a six-point scale, where 1 is Very Strongly Disagree and 6 is Very Strongly Agree, each manager’s supervisor rated the manager by responding to a 5-item performance scale described
above. The scale is included in Appendix B. A composite score was then generated. A coefficient alpha was also computed to assess the reliability of the performance scale for supervisors. The coefficient alpha was found to be .93, indicating good consistency in responses among items.

**Managerial Success Factors Inventory: Transportation Version (MSFI:TV)-in development.** The MSFI: TV is a web-based or paper-and-pencil 47-item questionnaire that was constructed by the researcher in this study to measure transportation managerial competence. It consists of behavioral and technical leadership and management characteristics to which an individual responds to each characteristic with respect to how much they agree with each statement based on how they feel they can perform that characteristic of management (e.g., “accurately identifies his or her own strengths and weaknesses, responds well to criticism”). The questionnaire assesses performance of various leadership competencies in transportation and can be used as a 360 degree assessment tool. Managers are asked to rate their own performance of each competency on a six-point Likert-type rating scale (1=Very Strongly Disagree to 6=Very Strongly Agree).

**Procedures**

The procedures for this project followed the 8 step scale development guidelines presented by Devellis (2003) and the conventional three study scale development research design (EFA, CFA, and validity/reliability analysis). The following 8 steps have been identified to direct the development of scales: Step 1- Determine clearly what is to be measured; Step 2- Generate an item pool; Step 3- Determine the format for
measurement; Step 4- Have the initial item pool reviewed by experts; Step 5- Consider inclusion of validation items; Step 6- Administer items to a development sample; Step 7- Evaluate the items, and Step 8- Optimize scale length (DeVellis, 2003).

**Item development for the MSFI:TV.** An initial item pool of 33 items were generated in previous research as a part of a grant from the National Center for Intermodal Transportation (NCIT) to identify essential managerial competencies relating to outstanding transportation management performance. These initial 33 items were developed from two sources of information; 1) qualitative data gathered from focus groups comprised of human resource managers in various transportation agencies, and 2) a thorough literature review in business management and psychology academic databases. The researchers utilized existing theory in leadership covered in Chapter 2 of this paper to devise the original 33 items. They were written to measure several identified and agreed upon success factors of leadership and management (i.e., trust, fostering teamwork, coaching, integrity, etc.). This initial item pool formed the basis of the MSFI:TV. These original items were pilot tested on a sample of 62 managers in transportation. The researchers found strong reliability for the scale with a Cronbach’s Alpha of .98. However, factor analytic techniques indicated we were measuring one factor, leadership, and that the scale was not very useful for leadership development practices. Durr and Sherry (2009) presented these initial findings at the American Psychological Association’s national convention in 2009.

As a part of the current investigation, the researcher revisited the literature and previous findings to expand the original 33 item pool to 60 items. The MSFI:TV was
developed rationally to represent the managerial and leadership competencies found to underlie the structure of the reigning general managerial competency models. More specifically, 60 items were written to measure each of the 11 competencies identified in the Hay Group (2001) general managerial competency model that fall across four clusters; empathy, self-control, self-confidence, developing others, holding people accountable, team leadership, results orientation, initiative, problem solving, influencing others, and fostering teamwork (HayGroup, 2001). At least one item was written to assess each competency.

The researcher then held three focus groups at a state DOT that included 15 experts in transportation management and leadership currently holding managerial and/or leadership positions within the DOT. The participants were notified that the DOT was conducting a research study to identify a leadership competency model and accompanying leadership survey which could be used for future leadership development initiatives. Three focus groups were then held on site in a DOT meeting room. Each focus group participant was asked at the beginning of the group to look at the preliminary version of the MSFI: TV on one of the laptops set-up in the meeting room and rank each item in terms of relevance/importance, and to make note of reactions regarding what the item may be intending to measure and how it relates to their work specifically as leaders in transportation. Upon completion of reviewing the initial item pool, the researcher facilitated a discussion about essential characteristics and qualities of leadership within transportation. The focus group participants were asked to tell personal accounts of successful leadership experiences at the DOT. Participants were also asked to comment
on the survey they had reviewed and to discuss specific concerning items or areas they
did not think were well covered. Each participant was asked to review each of the 60
items on relevance and importance to managerial competence in transportation. Each
member of this expert panel had been identified by a key human resource manager within
a state DOT as a star or peak performer. The group provided feedback about specific
items and assisted the researchers in tailoring the item language to transportation. The
focus group members were notified that the researcher constructed the scale to reflect the
four clusters of managerial competence found to underlie general management (i.e.,
managing your team, managing yourself, managing your work, and managing
collaboratively), and that we were interested in creating a scale that mirrored the
preeminent general managerial competence model in business but that was tailored
specifically to transportation management competence. The experts were asked to
comment on missing competencies, the organization of the competencies across the four
clusters, and to describe specific managerial and leadership practices they have utilized as
a transportation leader. The focus group participants evaluated the item pool in terms of
clarity, conciseness, and fit specific to transportation. They were also asked to discuss
how items link to the competencies they were written to measure. Finally, the experts
were asked to brainstorm about essential leadership functions not included in the initial
item pool. More specifically, they were asked to share stories of successful and
challenging leadership situations, and to attempt to identify the core behaviors and skills
required to successfully navigate those experiences.
In addition to the three expert focus groups, two cognitive interviews were conducted to assess the 60 items. One participant was an Industrial-Organizational Psychologist (I/O Psychologist) at the DOT and the second was a high ranking human resource manager at the DOT. The I/O Psychologist has over 25 years of executive experience in both public and private transportation, and the executive in human resources holds a doctoral degree in cognitive psychology, possesses over 20 years of experience selecting, promoting, and training transportation managers at various stages of professional development, and has written job descriptions and competency models for all levels of the organizational chart for an entire state DOT. Both individuals were asked about leadership/managerial competencies essential to successful performance as a transportation manager, and each reviewed the items for clarity and appropriateness to transportation management.

Both in the cognitive interviews and in all three focus groups a great need for a competency cluster related to creating a culture of safety was identified. Further, the concepts of “safety leadership” and “emotional competence” surfaced repeatedly as essential competencies in transportation management. Items that were deemed either unrelated or unclear were edited or removed.

Following these procedures, the researcher took the original 60 items and wrote new items to cover the missing competencies, changed the language of several items to reflect transportation specific terms, and removed several items marked as both redundant and irrelevant. Specifically, the researcher wrote 5 items relating to safety leadership and included more items related specifically to interpersonal, emotional, and social
competence, as these were skill areas specifically highlighted as essential to success in transportation leadership positions, and as deficit areas in management in public transportation in particular.

Following the focus groups, cognitive interviews, and editing of the original items, the researcher arrived at the version of the MSFI: TV administered in this study, which consisted of 47 items. The scale was written with intent to measure the four competency clusters found to underlie general managerial competence: Managing yourself, managing your team, managing your work, and managing collaboratively (HayGroup, 2001). Items were also added to measure an additional 3 competencies found to be essential to successful management in transportation by experts in the field. The competencies of Integrity, Emotional Competence, and Safety Leadership were added to the Managing Yourself and Managing the Work clusters in an effort to tailor the instrument. Even with the addition of these three competencies, we predicted the four cluster model to emerge in an exploratory factor analysis as we thought the new competencies fit nicely in to the existing model.

Study procedures. Prior to data collection, approval for this project was granted by the Institutional Review Board (IRB) at the University of Denver. The self-assessment sample (n = 287) was obtained by a series of email solicitations that were sent out via the DOT’s employee email list from a senior human resource manager. Only managers who were at the Team Leaders position or above in a managerial structure received the email. An initial email was sent out from the Executive Director introducing the study and providing a rationale for the data collection, as well as explicitly indicating voluntary
participation. The second email contained the informed consent form and a link to the on-line version of the MSFI: TV. Participants who were interested in participating in the study completed the survey on-line. Approximately 400 managers were invited to participate. Of those invited, 287 completed the on-line assessment in its entirety (71.8%). As mentioned, one half of this sample of 287 managers served as the development sample in this study and the other half as the confirmatory sample. Each participant was asked to list one supervisor who could rate their performance on a 5-item measure of managerial performance. The email address of the supervisor was obtained from the participant and then a system-generated email was sent to their email with a link to the survey on line. Supervisor’s responses were confidential and not shared directly with the ratee. Self-assessment participants were only provided their self-assessment results and aggregate supervisor results.

Informed consent was obtained prior to participation in the study by including the informed consent form in the email sent and notifying participants that by completing the survey they are consenting. Further, confidentiality was ensured to all participants. This was accomplished by requiring security passwords for the on-line version of the survey, and by storing data and identifying information in separate data files. Individual participants were assigned a participant number and there name and identifying information was maintained in a separate file corresponding to their assigned participant identification number. Only aggregate results were presented to the participating organization at the conclusion of the study.
Analysis

The self-assessment sample of managers in transportation were asked to complete the items of the MSFI:TV and then the sample was randomly split into two halves to allow the researcher to follow a three study format: an exploratory factor analysis (EFA) was conducted on the first half of the sample and then a confirmatory factor analysis (CFA) was conducted on the second half of the sample. Analysis conducted on the first half of the sample is referred to in this paper as phase 1 of the study while the analysis conducted on the second half of the sample is referred to as phase 2. This study included the first two studies (EFA and CFA) of the conventional three study scale development research design.

The dimensionality of the 47 items from the MSFI:TV was analyzed using the maximum likelihood factor analysis with IBM SPSS 19.0. A principal factoring method with varimax rotation was employed to examine the factor structure in phase 1. The coefficient alphas for each subscale were estimated to assess initial internal consistency. Correlations among the subscales were also generated.

In phase 2, the second half of the data sample was used as the confirmatory sample to further examine the reliability and validity of the scale, and assess the stability of the factor structure using CFA. To examine the structure underlying the MSFI:TV in greater depth, the researcher conducted a CFA of the 34 items of the MSFI:TV retained from the first study using LISREL 8.80 (Joreskog & Sorbom, 2006). The linear structural relationship (LISREL) model, a special case of the structural equation model (SEM), was assessed as part of the CFA to assess the goodness-of-fit. The structural stability of the
phase 1 factor solution was assessed. CFA allowed the researcher to generate path diagrams to describe how the latent variables relate to the observed constructs. The five factor solution identified in the EFA was then examined by observing the strength of each path as well as various indices of model fit. Using LISREL, model fit indices were used to determine how well the model obtained in the EFA fit the data. Chi-square, the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), the Non-normed fit index (NNFI), and the Goodness of Fit index (GFI) were all observed to assess different aspects of the fit of the model to the data.

The study also involved an initial validation procedure in which the manager subscale scores on the MSFI:TV were correlated with a performance measure, supervisor ratings, and absenteeism. Composite scores were generated for each subscale and the total scale, and then correlated with the five-item performance measure to assess convergent validity. Further, the second quantitative sample collected was made of supervisor’s composite supervisor (peer) ratings of the managers were used as an additional validation step. Concurrent validity was assessed in a 3rd phase comparing the self-assessment sample’s responses to their supervisor’s ratings of their performance.

**Rasch Modeling.** Analyses were then performed using the Rasch model with the WINSTEPS software version 3.72.0 (Linacre, 2009). The Rasch model is a mathematical framework that declares the relationship between the persons and the items that operationalize a single trait (Chiang, Green, & Cox, 2009). The model incorporates a method for ordering persons based on their ability and items according to their difficulty (Bond & Fox, 2007). Thus, the likelihood of higher scores increases as people possess
more of the trait, in this case, managerial competency in the transportation industry.

Rasch modeling first converts the instrument’s data into interval data by providing two parameter estimates: a person location and item location, typically known as person and item logits. A logit is defined as a nonlinear translation of the raw score (Bond & Fox, 2007). Once persons and items are placed on a common scale, equal-interval measures can be constructed (Chiang et al., 2009). After the data have been converted into this format, analysis of the psychometric properties of the scale can be completed. Analyses included unidimensionality, reliability, invariance, targeting, scale continuity, item hierarchy, differential item functioning, and person separation. Unidimensionality refers to how well each item measures or “fits” a construct (Bond & Fox, 2007). Dimensionality has been assessed in previous studies using the dimensionality coefficient, the overall fit of the data to a one-dimensional model, a Rasch principal components analysis of residuals, and individual item fit (Chiang et al., 2009). Reliability of the scale was then assessed again using person and item separation, which measures the spread of both items and persons in standard error units (Chiang et al.). For the instrument to be reliable and useful, the separation statistic must be higher than 2.0. Cronbach’s alphas were also calculated.
Chapter 4—Results

This chapter describes the results found in the current investigation. Following the scale development process, the chapter begins with the key findings that emerged from the initial phases of the scale development process including the core themes from the focus groups and cognitive interviews. Following the presentation of the focus group findings, the results of the quantitative data analysis obtained from administration of the MSFI: TV to the developmental sample are provided. Phase 1 results include the exploratory factor analysis, subscale correlations, and initial internal consistencies. Research questions 1 and 2 are addressed under phase 1 findings. Results related to questions 5, 6, and 7, which all relate to validity of the measure, are then presented. The second phase of the study results are presented next including the confirmatory factor analysis and Rasch analysis results. This section addresses question 3. Results of the Rasch analysis related to research questions 4 and 8 are presented next. In this section the subscale and total scale scores of the MSFI: TV are correlated with scores on a performance measure, absenteeism scores, and compared to composite scores of the supervisor’s ratings.

Focus Group Content Analysis

Content analysis of qualitative responses obtained in the focus groups revealed several ways to tailor the instrument specifically to transportation leadership. Given a primary task of this study was to tailor reigning models of leadership and management to
transportation management specifically, the results of the qualitative analysis are presented here. The assessment tool was revised prior to administration to the DOT after reviewing the focus group data. The focus groups made the following recommendations: remove inappropriate wording and items that do not relate (e.g., “knowledge of global market”, “transportation management strategies”); change the wording to reflect transportation culture (e.g., “replace company with organization”, “change consumer to customer/citizen”); and be sure to measure important aspects of leadership not measured in the preliminary version (e.g., safety leadership, more items tapping developing others, emotional competence, integrity). Focus group participants were encouraged to reflect on their experience as a manager in transportation and tell both success and failure stories. Primary themes identified in these exercises yielded keys to managerial competence in transportation that greatly informed the revision process following the focus groups. These themes allowed the researcher to tailor the general managerial competency model underlying the MSFI: TV to transportation. The themes are presented in Table 2.
Table 2

*Primary Managerial Competence Themes Identified in Focus Groups*

<table>
<thead>
<tr>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Get involved/make a difference</td>
</tr>
<tr>
<td>2. Clear minded, predictable decision making</td>
</tr>
<tr>
<td>3. Empower others to make decisions and take responsibility – the ability to inspire others</td>
</tr>
<tr>
<td>4. Listen and respect input from diverse views</td>
</tr>
<tr>
<td>5. Create opportunities for others to shine – ability to foster growth and development</td>
</tr>
<tr>
<td>6. Honesty and commitment to organizations initiative and values</td>
</tr>
<tr>
<td>7. Self-awareness (e.g., “Ability to identify own strengths and weaknesses”)</td>
</tr>
<tr>
<td>8. Understanding of one’s emotional reaction and how they impact others emotions.</td>
</tr>
<tr>
<td>9. Champions safety over productivity</td>
</tr>
</tbody>
</table>

**Phase 1**

**Exploratory Factor Analysis.** Required assumptions for a principal components factor analysis were assessed. The Kaiser-Meyer-Olkin measure of sampling adequacy was .91 and Bartlett’s test of sphericity approximated the chi-square statistic to be significant at the .001 level, indicating that a factor analysis could be conducted. The dimensionality of the 47 items from the MSFI: TV was analyzed using maximum likelihood factor analysis. Four criteria were used to determine the number of factors to rotate: the *a priori* hypothesis that the measure was unidimensional, the scree test, the number of eigenvalues over 1 and the interpretability of the factor solution. The eigenvalues and scree plot for the first sample indicated that our initial hypothesis of unidimensionality was incorrect.
Eleven factors had eigenvalues higher than 1.0. The first five factors had eigenvalues ranging from 1.42 to 17.31. The scree test indicated that a four or possibly five factor solution might be most interpretable. Based on existing theory (HayGroup model, 2001), the plot, the Eigenvalues, and the interpretability of the factor solutions; One-, four-, five-, six-, seven-, eight- and eleven-factor solutions were rotated using the varimax rotation procedure. Surprisingly, the four factor (cluster) general leadership/managerial competency model that the instrument was written to reflect was not interpretable. Though the initial analysis revealed eleven factors based on eigenvalues over 1.0, the researcher found the five-factor solution to be the most interpretable. The five factors accounted for 47.72% of the variance. Items were then selected for the measure based on the factor pattern matrix using the following criteria: (1) a factor loading above .40 on the factor and (2) cross-loadings on other factors of less than .40. In other words, for an item to be retained, it had to load higher than .40 on the factor it loaded highest on and not load higher than .40 on other factors. If either of these criteria were not met, it was eliminated. Based on these criteria, 34 items out of the original 47 were retained. The rotated solution yielded five interpretable factors; 1) Problem Solving (10 items that assess problem solving competencies), 2) Emotional Competence (12 items assessing emotional intelligence and the ability to develop others), 3) Results Orientation (4 items assessing initiative and drive for results), 4) Safety Leadership (5 items assessing safety orientation), and 5) Integrity (3 items assessing ethics and honesty). The factor loadings, means, and standard deviations of the 34 retained items from the maximum likelihood factor analysis and the eigenvalue associated with each of the five factors are
presented in Table 3. The Problem Solving factor accounted for 13.42% of the item variance, the Emotional Competence factor accounted for 11.02% of the item variance, the Results Orientation factor accounted for 7.88% of the item variance, the Safety Leadership factor accounted for 7.85% of the item variance, and the Integrity factor accounted for 7.54% of the item variance.
### Table 3

**Items, Factor Loadings, and Standard Deviations**

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comes up with new ideas and solutions to recurring problems.</td>
<td>.63</td>
<td>.07</td>
<td>.17</td>
<td>.17</td>
<td>.27</td>
<td>4.51</td>
<td>.79</td>
</tr>
<tr>
<td>2. Is self-confident.</td>
<td>.59</td>
<td>.09</td>
<td>.12</td>
<td>.20</td>
<td>.11</td>
<td>4.63</td>
<td>.80</td>
</tr>
<tr>
<td>3. Negotiates effectively with other units in the organization to get things done</td>
<td>.59</td>
<td>.34</td>
<td>.23</td>
<td>.10</td>
<td>.15</td>
<td>4.58</td>
<td>.71</td>
</tr>
<tr>
<td>4. Is able to persuade others to adopt new points of view.</td>
<td>.56</td>
<td>.39</td>
<td>.12</td>
<td>.15</td>
<td>.07</td>
<td>4.33</td>
<td>.66</td>
</tr>
<tr>
<td>5. Takes action before being forced to do so by events.</td>
<td>.53</td>
<td>.17</td>
<td>.29</td>
<td>.08</td>
<td>.20</td>
<td>4.53</td>
<td>.77</td>
</tr>
<tr>
<td>6. Is able to incorporate information from various sources to make the correct decision.</td>
<td>.53</td>
<td>.23</td>
<td>.27</td>
<td>.10</td>
<td>.37</td>
<td>4.72</td>
<td>.72</td>
</tr>
<tr>
<td>7. Is able to use conflict in groups effectively.</td>
<td>.53</td>
<td>.28</td>
<td>.17</td>
<td>.37</td>
<td>.07</td>
<td>4.05</td>
<td>.82</td>
</tr>
<tr>
<td>8. Acts as a catalyst to a committee’s decision-making process.</td>
<td>.47</td>
<td>.17</td>
<td>.14</td>
<td>.31</td>
<td>.03</td>
<td>4.21</td>
<td>.71</td>
</tr>
<tr>
<td>9. Identifies high potential candidates.</td>
<td>.47</td>
<td>.31</td>
<td>.27</td>
<td>.10</td>
<td>.15</td>
<td>4.67</td>
<td>.82</td>
</tr>
<tr>
<td>10. Is able to rally support for important initiatives.</td>
<td>.46</td>
<td>.37</td>
<td>.25</td>
<td>.12</td>
<td>.08</td>
<td>4.42</td>
<td>.71</td>
</tr>
<tr>
<td>11. Connects with people at all levels of the organization.</td>
<td>.45</td>
<td>.29</td>
<td>.22</td>
<td>.09</td>
<td>.16</td>
<td>4.82</td>
<td>.88</td>
</tr>
<tr>
<td>12. Knows when to become involved in a conflict and when not to.</td>
<td>.44</td>
<td>.34</td>
<td>.29</td>
<td>.27</td>
<td>.20</td>
<td>4.48</td>
<td>.86</td>
</tr>
<tr>
<td>13. Communicates goals or a vision that motivates others.</td>
<td>.43</td>
<td>.39</td>
<td>.25</td>
<td>.25</td>
<td>.15</td>
<td>3.36</td>
<td>.71</td>
</tr>
<tr>
<td>14. Responds well to emotions of others.</td>
<td>.32</td>
<td>.72</td>
<td>.06</td>
<td>.32</td>
<td>.04</td>
<td>4.28</td>
<td>.87</td>
</tr>
<tr>
<td>15. Recognizes where others are in the professional development process.</td>
<td>.19</td>
<td>.59</td>
<td>.08</td>
<td>.01</td>
<td>.13</td>
<td>4.46</td>
<td>.75</td>
</tr>
<tr>
<td>16. Able to recognize his/her own emotional reaction to people, events, and situations.</td>
<td>.08</td>
<td>.53</td>
<td>.21</td>
<td>.24</td>
<td>.14</td>
<td>4.33</td>
<td>.75</td>
</tr>
<tr>
<td>17. Coaches employees to help them achieve success.</td>
<td>.33</td>
<td>.52</td>
<td>.11</td>
<td>.19</td>
<td>.23</td>
<td>4.67</td>
<td>.73</td>
</tr>
<tr>
<td>18. Demonstrates awareness of how his/her own actions affect the emotions of others.</td>
<td>.25</td>
<td>.50</td>
<td>.32</td>
<td>.24</td>
<td>.11</td>
<td>4.40</td>
<td>.76</td>
</tr>
<tr>
<td>19. Is able to understand how others feel about a decision or situation.</td>
<td>.26</td>
<td>.46</td>
<td>.04</td>
<td>.25</td>
<td>.14</td>
<td>4.35</td>
<td>.71</td>
</tr>
<tr>
<td>20. Is clear when delegating responsibilities.</td>
<td>.25</td>
<td>.44</td>
<td>.33</td>
<td>.22</td>
<td>.20</td>
<td>4.50</td>
<td>.76</td>
</tr>
<tr>
<td>21. Provides feedback that is accurate and helpful for improving performance.</td>
<td>.36</td>
<td>.42</td>
<td>.30</td>
<td>.23</td>
<td>.18</td>
<td>4.49</td>
<td>.70</td>
</tr>
<tr>
<td>22. Asks for feedback.</td>
<td>.10</td>
<td>.41</td>
<td>.06</td>
<td>.22</td>
<td>.13</td>
<td>4.48</td>
<td>.80</td>
</tr>
<tr>
<td>23. Delivers what is promised.</td>
<td>.23</td>
<td>.23</td>
<td>.67</td>
<td>.06</td>
<td>.19</td>
<td>4.79</td>
<td>.80</td>
</tr>
<tr>
<td>24. Manages self effectively.</td>
<td>.27</td>
<td>.19</td>
<td>.62</td>
<td>.13</td>
<td>.16</td>
<td>4.60</td>
<td>.86</td>
</tr>
<tr>
<td>25. Anticipates what needs to be done and does it.</td>
<td>.26</td>
<td>.08</td>
<td>.55</td>
<td>.14</td>
<td>.24</td>
<td>4.95</td>
<td>.75</td>
</tr>
<tr>
<td>26. Knows when to use informal versus formal corrective actions.</td>
<td>.25</td>
<td>.36</td>
<td>.43</td>
<td>.19</td>
<td>.13</td>
<td>4.43</td>
<td>.82</td>
</tr>
<tr>
<td>27. Holds meaningful safety meetings.</td>
<td>.13</td>
<td>.23</td>
<td>.09</td>
<td>.83</td>
<td>.03</td>
<td>4.17</td>
<td>.85</td>
</tr>
<tr>
<td>28. Is a model for the use of safe work practices.</td>
<td>.15</td>
<td>.17</td>
<td>.15</td>
<td>.79</td>
<td>.14</td>
<td>4.33</td>
<td>.75</td>
</tr>
<tr>
<td>29. Provides necessary training or tools to safely perform the job.</td>
<td>.38</td>
<td>.03</td>
<td>.06</td>
<td>.67</td>
<td>.15</td>
<td>4.64</td>
<td>.84</td>
</tr>
<tr>
<td>30. Focuses on identifying the root cause of safety incidents and not placing blame.</td>
<td>.23</td>
<td>.25</td>
<td>.15</td>
<td>.54</td>
<td>.21</td>
<td>4.47</td>
<td>.73</td>
</tr>
<tr>
<td>31. Ensures the safe performance of the job first.</td>
<td>.16</td>
<td>.15</td>
<td>.18</td>
<td>.42</td>
<td>.37</td>
<td>5.09</td>
<td>.79</td>
</tr>
<tr>
<td>32. Is honest and trustworthy.</td>
<td>.12</td>
<td>.24</td>
<td>.32</td>
<td>.12</td>
<td>.69</td>
<td>5.18</td>
<td>.82</td>
</tr>
<tr>
<td>33. Makes ethical decisions when faced with conflicting choices.</td>
<td>.22</td>
<td>.09</td>
<td>.39</td>
<td>.04</td>
<td>.65</td>
<td>4.87</td>
<td>.78</td>
</tr>
<tr>
<td>34. Appreciates different ideas and perspectives.</td>
<td>.24</td>
<td>.17</td>
<td>.05</td>
<td>.21</td>
<td>.56</td>
<td>4.85</td>
<td>.71</td>
</tr>
</tbody>
</table>

Factor Loadings > .40 are in bold. Factor 1 = Problem Solving; Factor 2 = Emotional Competence; Factor 3 = Results Orientation; Factor 4 = Safety Leadership; Factor 5 = Integrity.
**Internal consistency and correlations among the subscales.** A composite subscale score was generated for each of the five factors (subscales) and then internal consistency statistics (i.e., Cronbach’s Alpha) were estimated for each subscale and the total scale. The coefficient alpha assesses the consistency in scores for each of the scale and then subscale items (Green & Salkind, 2005). The total scale showed good internal consistency reliability (Cronbach’s Alpha = .95). Thus, support was found for question 2. The Cronbach’s Alpha, eigenvalue, and percentage of variance accounted for by each factor are listed in Table 4. This was an important finding given the need for the scale to produce a reliable total managerial competence in transportation score in addition to the five distinct subscale scores. Practical applications of the scale warrant analysis at both the total scale and subscale level as consulting endeavors frequently require both total managerial competence and more specific managerial competency scores. The results indicated adequate levels of reliability for all five subscales as well: Problem Solving had a Cronbach’s Alpha of .91; Emotional Competence had a Cronbach’s Alpha of .84, Results Orientation had a Cronbach’s Alpha of .78, Safety Leadership had a Cronbach’s Alpha of .85, and the Integrity had a Cronbach’s Alpha of .78. The Results Orientation and Integrity subscales appear to be the least reliable based on the coefficient alpha, though they also contain the fewest items. The Cronbach’s Alpha’s for each of the subscales support the scales internal consistency and confirm hypothesis 3.
Table 4

*Leadership Scales and Associated Cronbach’s Alpha, Eigenvalue, and Percentage of Variance Accounted for by the Factors*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Coefficient Alpha</th>
<th>Eigenvalue</th>
<th>Percentage of Variance</th>
<th># of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Problem-Solving</td>
<td>.91</td>
<td>6.31</td>
<td>13.42</td>
<td>13</td>
</tr>
<tr>
<td>2. Emotional Competence</td>
<td>.84</td>
<td>5.18</td>
<td>11.02</td>
<td>9</td>
</tr>
<tr>
<td>3. Results Orientation</td>
<td>.78</td>
<td>3.71</td>
<td>7.88</td>
<td>4</td>
</tr>
<tr>
<td>4. Safety Leadership</td>
<td>.85</td>
<td>3.69</td>
<td>7.85</td>
<td>5</td>
</tr>
<tr>
<td>5. Integrity</td>
<td>.78</td>
<td>3.55</td>
<td>7.54</td>
<td>3</td>
</tr>
</tbody>
</table>

**Validity.** Each manager who completed the survey was also asked to list a supervisor who could assess their performance on the same five item measure of managerial performance they completed. A composite score was generated for this five item measure. Supervisor ratings were then used to assess the concurrent validity of the MSFI: TV. A measure of absenteeism and the self-assessment scores on the performance measure allowed for assessment of discriminant and convergent validity. Pearson product-moment correlation coefficients were generated between the MSFI: TV total scale score, the subscales, a measure of performance, and a measure of absenteeism. A correlation was also computed between each subscale and supervisor’s ratings of each participant on a measure of performance.

A Pearson product-moment correlation of .75 (p < .001) was found between the total MSFI: TV score and the total score on the performance measure, indicating a strong correlation between the scales. This finding offers initial convergent validity for the scale.
and confirms question five. The correlations among the subscales were all positive and relatively high. The correlations were also all statistically significant at $p < .01$ and are listed in Table 5. The highest correlation was found between the Problem-Solving factor and Emotional Competence at .78. This finding may support the notion that effective leaders who are emotionally competent managers are seen as successful at solving the complex problems faced within public transportation. More specifically, managers who can competently read other’s emotions, understand their own emotional reactions, and predict how behaviors impact the emotions of others are more likely to be able to successfully rally needed support for key initiatives (i.e., overcome complex organizational challenges via social and interpersonal skills). Safety Leadership correlated .65 with Problem-Solving, .60 with Emotional Competence, .55 with Results Orientation, and .51 with Integrity. Results orientation and Problem-Solving also correlate highly at .70. The positive, significant correlations between the subscales indicates that higher levels of competency in one managerial competency domain are associated with higher levels of competency in other domains of managerial competency (i.e., stronger problem solving and collaboration skills are associated with stronger safety leadership skills). This also supports the notion that the subscales relate closely to one another and potentially to the construct of managerial competence. This may indicate dimensionality of the scale as well.
Table 5

*Correlations Among the Leadership Subscales, Performance Measure, and Demographic Variables*

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Problem Solving</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Emotional Competence</td>
<td>.78**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Results Orientation</td>
<td>.70**</td>
<td>.69**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Safety Leadership</td>
<td>.65**</td>
<td>.60**</td>
<td>.55**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Integrity</td>
<td>.66*</td>
<td>.62**</td>
<td>.63**</td>
<td>.51**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>6. MSFI Total</td>
<td>.94**</td>
<td>.90**</td>
<td>.81**</td>
<td>.77**</td>
<td>.76**</td>
<td>1.00</td>
</tr>
<tr>
<td>7. Performance (self-assessment)</td>
<td>.77**</td>
<td>.59**</td>
<td>.64**</td>
<td>.54**</td>
<td>.59**</td>
<td>.75**</td>
</tr>
<tr>
<td>8. Absenteeism</td>
<td>-.11</td>
<td>-.05</td>
<td>-.11</td>
<td>.03</td>
<td>.01</td>
<td>-.07</td>
</tr>
<tr>
<td>9. Supervisor’s Rating of Performance</td>
<td>.02</td>
<td>-.05</td>
<td>-.01</td>
<td>-.01</td>
<td>-.06</td>
<td>-.01</td>
</tr>
<tr>
<td>Gender</td>
<td>-.02</td>
<td>.04</td>
<td>.16</td>
<td>-.22</td>
<td>.07</td>
<td>-.02</td>
</tr>
<tr>
<td>Years in the Position</td>
<td>.16</td>
<td>.10</td>
<td>.01</td>
<td>.12</td>
<td>.06</td>
<td>.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50.00</td>
<td>50.00</td>
<td>50.01</td>
<td>91.66</td>
<td>25.58</td>
<td>50.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Standard Deviation</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>17.05</td>
<td>5.86</td>
<td>10.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Skewness</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.32</td>
<td>.33</td>
<td>.06</td>
<td>.36</td>
<td>-.09</td>
<td>.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Kurtosis</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-.21</td>
<td>-.20</td>
<td>-.71</td>
<td>-.36</td>
<td>-.99</td>
<td>-.23</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01

Significant, strong, positive correlations were also found between the subscales of the MSFI: TV and a five-item measure of performance. The strongest correlation was between Performance and Problem-Solving (*r* = .77). These strong and positive
correlations support the scales convergent validity. It was predicted that higher scores on the MSFI:TV would indicate higher performance ratings.

Surprisingly, the correlations between absenteeism (i.e., full days of work missed in the past 30 days) and the subscales were non-significant at $p < .05$. Thus, question 6 was not supported. Correlations between the subscales and gender, and the subscales and years in the position, were also non-significant, as listed in Table 5. Negative correlations between gender and the subscales indicate that females are more likely to score higher on the subscale where positive correlations suggest the opposite. Finally, results indicate a slightly inverse, non-significant relation between supervisor’s ratings on a performance measure and each of the subscales of managerial competence. None of the bivariate correlations between the supervisor’s ratings of the self-assessment participant on a performance measure and the self-assessment participant’s ratings of themselves on each of the subscales correlated significantly. The total MSFI: TV score and supervisor’s ratings were also not significantly correlated. Thus, support was not found for question 7.

Phase 2

Confirmatory Factor Analysis. The five-factor solution obtained from the EFA was examined further by conducting a CFA on the 34 items retained from phase 1 of the study. The CFA was conducted using LISREL 8.80 Student Edition for Windows (Joreskog & Sorbom, 2006). The intercorrelation matrix served as the starting point for the subsequent LISREL maximum-likelihood parameter estimation. In addition, the resulting goodness of fit indices including Chi-Square ($X^2$), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Comparative Fit Index (CFI), and Root Mean
Square Error of Approximation (RMSEA) suggested by Joreskog & Sorbom (1986) were observed. Each of the fit statistics offer unique information about the fit of the data to the model thus each are reported. This large class of fit measures exists to assess overall model fit to the data. A large $X^2$ and rejection of the null hypothesis indicates that the model does not fit the data well. Thus, the smaller the $X^2$ and rejection of the null indicates acceptable model fit for the data sample. Due to sampling size sensitivity with $X^2$, the other fit statistics listed above were computed to determine if the model fits the data in an acceptable fashion. The RMSEA, a measure of fit that is less sensitive to sample size, was one statistic observed to assess model fit. It is suggested that RMSEA values of .05 or less suggest acceptable fit. However, some suggest RMSEA scores in the .05-.08 range can indicate acceptable fit. The CFI compares model fit to a baseline model and assumes that there is no relationship among the variables. Scores on CFI range from 0 to 1 with closer scores to 1 indicating better fit. Though there is no single agreed upon rule of thumb for cut-off values for declaring fit of a model to the data, Hu and Bentler (1999) provided some examples of acceptable fit scenarios. For example, they suggested that if RMSEA values are close to the .06 range or lower and CFI is greater than .95, the model most likely fits the data well. Thus, these fit statistics are presented here in addition to the $X^2$ statistic. Comparisons were also made between the five factor model identified in the EFA and a higher-order model including the managerial competence variable as a sixth latent or $2^{nd}$ order variable. The hypothesized Hay Group model (2001) was not found in the EFA in phase 1 of this study so it did not make sense to test it against the five-factor solution obtained in the EFA.
Table 6 presents the fit statistics for each of the competing models using the 34 item MSFI: TV. The five-factor model obtained in the EFA consisted of five latent variables; Problem Solving (13 items), Emotional Intelligence (9 items), Results Orientation (4 items), Safety Leadership (5 items), and Integrity (3 items), as demonstrated in Figure 1.
Figure 1. Five-Factor Model
The results of the CFA indicate that this model represented an acceptable fit to the data, however, only certain fit indices observed were considered good ($X^2 = 902.72$, CFI = .97). The CFI of .97 indicates a moderate fit. The EFA derived five-factor model gave a RMSEA of .07, a GFI of .73 and an AGFI of .69. These values do not meet the recommended cutoff values for acceptable model fit. The GFI and AGFI are lower than recommended but this may be attributable to the relatively low sample size compared to the relatively high degrees of freedom. However, taking all of the fit statistics together ($X^2$, CFI, RMSEA, GFI), the model is acceptable. However, to attempt to improve model fit, the five-factor model was assessed again but this time the model included a second-order factor of managerial competence (termed leadership in Figure 2).
Figure 2. Second-order Model
The higher-order model produced very similar fit statistics to that of the five factor model. The primary difference was the increase in $X^2$ from 902.72 to 917.49. The RMSEA, GFI, AGFI, and CFI all stayed the same with the second-order factor. The second-order model indicated acceptable fit to the data as well. Though there is minimal difference in fit between the five-factor model and the second-order model, the second-order model is more practical in terms of usefulness of the scale as it allows for a total managerial competence in transportation score in which leaders can be compared. Despite the slightly higher second-order model fit results, acceptable fit of the second-order model allows for use of a total scale score. This is a desirable feature of the instrument in that leaders can be given total managerial competence scores. This finding indicates that the scale has five related factors or subcomponents that also relate to this one overarching score. Given the acceptable fit to the data of the second-order model, the 34-item higher order model was analyzed using Rasch analysis.

### Table 6

<table>
<thead>
<tr>
<th></th>
<th>$X^2$</th>
<th>df</th>
<th>p Value</th>
<th>RMSEA</th>
<th>AGFI</th>
<th>GFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Five-Factor Model (EFA)</td>
<td>902.72</td>
<td>517</td>
<td>.001</td>
<td>.07</td>
<td>.69</td>
<td>.73</td>
<td>.97</td>
</tr>
<tr>
<td>Higher-Order Model</td>
<td>917.49</td>
<td>522</td>
<td>.001</td>
<td>.07</td>
<td>.69</td>
<td>.73</td>
<td>.97</td>
</tr>
</tbody>
</table>

**Rasch modeling results.** Rasch analysis was performed using WINSTEPS, version 3.72.0 single-user license (Wright & Linacre, 2009). The WINSTEPS program provides detailed statistics for each item as well as for the overall instrument. Since the hypothesized four cluster-eleven competency model was not found in the EFA, but the
five-factor solution obtained in the EFA was confirmed with the CFA with a higher-order factor. Rasch analysis was used to examine the entire MSFI: TV scale. All 34 items of the MSFI: TV obtained from the EFA and confirmed in the CFA were examined. Results of the first Rasch analysis on all 34 items suggested that the data fit the model well. Initial results also supported unidimensionality of the scale. However, there were three misfitting items detected in the first analysis. For items to be deemed misfitting, the item must exceed both the mean squares (MS) and standardized Z score (ZSTD) criteria (MS = 1.30 and ZSTD = 2.0). Items 6, 30 and 34 fell outside the acceptable criterion for fitting items. These items seemed to underfit the measure. Though they assessed critical areas of managerial competence (i.e., Item 6 read “Responds well to emotions of others”, Item 30 read “Asks for feedback”, and Item 34 read “Connects with people at all levels of the organization”), the researcher assumed the scale is unidimensional from the CFA and initial Rasch analysis and therefore removed any misfitting items that were threatening unidimensionality. These items were removed and then the Rasch analysis was rerun on the remaining 31 items. The results that follow are from the analysis conducted with those items removed.

**Dimensionality.** In Rasch analysis, unidimensionality refers to whether the instrument measures a single construct (Pomeranz et al, 2008). It is assumed that items assess a single or unidimensional construct by the Rasch model (Bond & Fox, 2007). Rasch analysts use two statistics to determine an instrument’s unidimensionality: 1) mean square standardized residuals (MS) and 2) standardized Z scores (ZSTD). According to Linacre and Wright (1994), a reasonable MS value is less than or equal to 1.30 (30%
variance), with a desired value of 1.0. Items with SD scores greater than 2.0 are too unpredictable (Pomeranz et al., 2008). Thus, for it to be “misfitting,” it must both exceed an MS of 1.3 and a ZSTD of 2.0. Dimensionality was assessed here by observing overall fit of the data to a one-dimensional model, a principal components analysis of residuals by the Rasch model, and individual item fit.

Fit was assessed by examining overall fit of the data to a unidimensional construct. Rasch analysis provides two chi-square ratios as fit statistics: infit and outfit mean square statistics (Bond & Fox, 2007). Both are produced as mean squares (MNSQ) with an expected value of +1.0, which indicates perfect fit, and possible values ranging from zero up. The fit statistics estimate average fit of persons and items. A MNSQ of less than 1.0 indicates less variability than expected, while an MNSQ greater than 1.0 indicates more variability than expected by the model. It is generally agreed upon that MNSQ fit values between .5 and 1.5 are acceptable (Bond & Fox, 2007). These were the values considered for overall fit.

Based on the Rasch analysis conducted on the 31 items retained in the EFA, it appears that the data fit the model well because the MNSQ for both the infit and outfit was less than 1.3 and the ZSTD was less than 2.0. The ZSTD is essentially zero, which indicates the data fit the unidimensional model quite well. These scores are also known as person fit statistics as they indicate how well the sample aligned with the predicted model. In other words, participants with higher levels of managerial competence in transportation should produce higher scores than those with lower managerial competence in transportation. The total MSFI: TV scale produced a MNSQ of 1.0 for
both infit and outfit. Thus, person fit was in the expected range for the scale. This also indicates that individuals are responding consistently to the scale.

A Rasch principal components analysis of residuals was also conducted to assess dimensionality. Unlike principal components analysis in EFA, when it is used to construct variables, it was used here to determine if a second dimension exists. Linacre (2011) reported that the variance explained by the measure should be 40% or greater for the scale to be unidimensional. Additionally, the unexplained variance should be less than 5% (Linacre, 2011). The resulting explained variance of 45.9% indicates that the instrument is probably unidimensional as the variance accounted for by the first dimension is greater than 40%. Further, the unexplained variance for the first construct was less than 5%. These findings suggest that the measure is unidimensional, or that it is most likely transportation managerial competence. This supports what was found in the CFA with the second-order model. The Rasch results indicate that a unidimensional structure best fits the data. Unidimensionality is assessed further at the item level.

Scaling. Rasch analysis also provides useful information regarding the rating scale. The rating category can be assessed to determine if people are using the scale as it is intended to be used. In other words, a researcher can look at the category fit statistics to determine if the steps up the scale worked appropriately for this sample. Similar to the use of a ruler, the thresholds that indicate successive categories need to be ordered to be interpretable. The MSFI: TV utilizes a 6-point rating scale (1=Very Strongly Disagree to 6=Very Strongly Agree). The results indicate that the estimates were not properly ordered, as listed in Table 7.
Table 7

Summary of Category Structure

<table>
<thead>
<tr>
<th>Category Label</th>
<th>Observed Average</th>
<th>Structure Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.26</td>
<td>-.93</td>
</tr>
<tr>
<td>2</td>
<td>.67</td>
<td>-2.28</td>
</tr>
<tr>
<td>3</td>
<td>.72</td>
<td>-3.26</td>
</tr>
<tr>
<td>4</td>
<td>1.77</td>
<td>-.82</td>
</tr>
<tr>
<td>5</td>
<td>3.12</td>
<td>2.59</td>
</tr>
<tr>
<td>6</td>
<td>4.54</td>
<td>4.79</td>
</tr>
</tbody>
</table>

These results suggest that the people in this sample did not use the scale as it was intended to be used. The rating scale categories should increase in difficulty from 1 to 6. However, the results indicate that the rating scale categories were disordered. The higher observed average for category 1 than category 2-5 implies disorder in the category definitions (Linacre, 2011). These results also indicate disordering of the average ability of the people observed in each category. The step calibration statistics listed in Table 7 were also disordered. Categories 3-6 were ordered appropriately, indicating that this level of the scale was used as intended by the people in this sample. Nevertheless, Linacre (2011) noted that disordered of the step calibration may imply that the category may narrowly correspond to intervals on the latent variable, managerial competence in transportation. Thus, such findings indicate a need for future versions of the MSFI: TV to use alternative rating scales, such as a 5-point rating scale so that categories are collapsed.
to correspond appropriately to managerial competence in transportation intervals.

Analysis at the item and person level were also needed.

**Item Fit.** Though overall fit of the instrument was discussed above, Rasch analysis also determined how well each of the 31 items contributed to the managerial competence construct. Identification of items that “misfit” allows the researcher to determine if certain items are threatening the unidimensionality of the measure. Table 8 presents the 31 items retained after the first Rasch analysis and their respective infit statistics. A complete listing of the items can also be found in Appendix C. These statistics indicate whether or not any of the items are threatening the unidimensionality of the measure. Item fit describes how well responses to items match predictions of the model. As mentioned in the discussion of unidimensionality and overall fit, MNSQ is modeled to take a value of 1.0 and ZSTD score is modeled to take a value of zero. However, Smith (1991) reported that at the item level a negative bias in the standardized fit statistic exists and said that criteria of <-3 or >+2 for identifying misfitting items might be best. In order to maintain reliability and usefulness of the scale in terms of retaining items, the Smith (1991) criteria of <-3 to >+2 was followed at the item level for the ZSTD. Linacre (2011) suggested that MNSQ values above 1.50 contribute little value to the measure. However, given MNSQ tends to shrink with sample size and the current sample is less than 300, a 1.30 MNSQ cutoff value was established. Rasch analysts commonly examine infit scores because they yield information about how well the observations fit the Rasch modeled expectations (Bond & Fox, 2007; Pomeranz et al, 2008), and are typically more sensitive to unexpected behavior regarding trait level and
item endorsement. Using the established cutoff MNSQ infit value of 1.30, none of the 31 retained items fell outside the acceptable criteria for a unidimensional construct. Table 8 shows the fit and logit position for each of the items on the MSFI: TV scale. It appears that items 29, “Is able to persuade others to adopt new points of view” and 21, “Is able to take in information from various sources and make the correct decision (e.g., close a road),” both overfit the measure based on a ZSTD value exceeding -3.0. However, the researcher is less concerned about overfitting items as this merely indicates the item fits too well. Given neither of the overfitting items are overly extreme, these items were retained. As seen in Table 8, the very low logit difficulty for most of the items indicates that the items do not adequately capture the range of logit ability of the sample. Items 5, “Is honest and trustworthy,” 22, “Puts the safe performance of the job first,” and 18, “Anticipates what needs to be done and does it,” appear to be particularly low logit positions relative to logit ability of the sample. Overall the logit position of most of the items is low compared to the managerial competence in transportation observed in the sample. Implications from these findings for further research development initiatives with this scale are discussed in Chapter 5.
Table 8

*Item Statistics for the MSFI: TV Total Scale*

<table>
<thead>
<tr>
<th>Item #</th>
<th>Logit Position</th>
<th>MNSQ Infit</th>
<th>ZSTD Infit</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>1.40</td>
<td>1.18</td>
<td>1.6</td>
</tr>
<tr>
<td>24</td>
<td>1.05</td>
<td>1.23</td>
<td>2.1</td>
</tr>
<tr>
<td>26</td>
<td>-.10</td>
<td>1.21</td>
<td>2.4</td>
</tr>
<tr>
<td>22</td>
<td>-1.45</td>
<td>1.19</td>
<td>2.4</td>
</tr>
<tr>
<td>3</td>
<td>-.21</td>
<td>1.17</td>
<td>2.0</td>
</tr>
<tr>
<td>28</td>
<td>1.01</td>
<td>1.11</td>
<td>1.1</td>
</tr>
<tr>
<td>10</td>
<td>-.29</td>
<td>1.14</td>
<td>1.6</td>
</tr>
<tr>
<td>2</td>
<td>-.18</td>
<td>1.13</td>
<td>1.6</td>
</tr>
<tr>
<td>14</td>
<td>.25</td>
<td>1.13</td>
<td>1.5</td>
</tr>
<tr>
<td>15</td>
<td>.34</td>
<td>1.10</td>
<td>1.1</td>
</tr>
<tr>
<td>20</td>
<td>-.02</td>
<td>1.10</td>
<td>1.1</td>
</tr>
<tr>
<td>18</td>
<td>-1.10</td>
<td>1.03</td>
<td>.4</td>
</tr>
<tr>
<td>11</td>
<td>.23</td>
<td>.98</td>
<td>-.2</td>
</tr>
<tr>
<td>25</td>
<td>.56</td>
<td>1.06</td>
<td>.7</td>
</tr>
<tr>
<td>8</td>
<td>.46</td>
<td>1.05</td>
<td>.6</td>
</tr>
<tr>
<td>17</td>
<td>-.63</td>
<td>1.00</td>
<td>.0</td>
</tr>
<tr>
<td>7</td>
<td>.59</td>
<td>1.03</td>
<td>.3</td>
</tr>
<tr>
<td>1</td>
<td>.43</td>
<td>1.05</td>
<td>.5</td>
</tr>
<tr>
<td>4</td>
<td>-.89</td>
<td>1.03</td>
<td>.5</td>
</tr>
<tr>
<td>32</td>
<td>-.71</td>
<td>.93</td>
<td>-.9</td>
</tr>
<tr>
<td>12</td>
<td>-.41</td>
<td>.95</td>
<td>-.6</td>
</tr>
<tr>
<td>5</td>
<td>-1.76</td>
<td>.94</td>
<td>-.7</td>
</tr>
<tr>
<td>27</td>
<td>.31</td>
<td>.89</td>
<td>-1.3</td>
</tr>
<tr>
<td>13</td>
<td>.14</td>
<td>.90</td>
<td>-1.2</td>
</tr>
<tr>
<td>19</td>
<td>.11</td>
<td>.86</td>
<td>-1.7</td>
</tr>
<tr>
<td>31</td>
<td>-.08</td>
<td>.85</td>
<td>-1.9</td>
</tr>
<tr>
<td>23</td>
<td>.24</td>
<td>.80</td>
<td>-2.4</td>
</tr>
<tr>
<td>9</td>
<td>.15</td>
<td>.77</td>
<td>-3.0</td>
</tr>
<tr>
<td>16</td>
<td>.45</td>
<td>.77</td>
<td>-2.8</td>
</tr>
<tr>
<td>29</td>
<td>.57</td>
<td>.73</td>
<td>-3.2</td>
</tr>
<tr>
<td>21</td>
<td>-.49</td>
<td>.68</td>
<td>-4.4</td>
</tr>
</tbody>
</table>

**Person Reliability and Separation.** Rasch analysis also provides two unique statistics to measure the reliability of an instrument; person reliability and separation. Person reliability refers to the replicability of a person’s placement along the ability
continuum, while person separation refers to how well the sample separates into distinct performance levels of strata (Pomeranz et. al., 2008). In other words, person separation shows the spread of persons across the item difficulty levels. Most researchers conclude that values over 1.0 are acceptable while values over 2.0 are ideal (Pomeranz et al.). Based on the REAL RMSE, the separation was 4.00 and Reliability of Person Separation was .94. Based on the Model RMSE, which removes the misfit, this sample produced a person reliability coefficient of .95 and person separation of 4.41. Pomeranz et al. (2008) proposed that a person reliability coefficient of .70 is an acceptable level of reliability, while .80 represents a good level. Thus, it appears that together the items are working well to reproduce a participant’s transportation managerial competence. Pomeranz et al. (2008) also suggested that the separation statistic can be placed in the formula: \[ HP = \frac{4GP + 1}{3} \], where \( HP \) = the number of strata and \( GP \) = the person separation statistic, to determine the number of distinct managerial competence strata. The number of strata equaled 6.21, suggesting that the sample can be separated into six distinct transportation management competence groups. These results indicate that persons were spread fairly well across the continuum of item difficulty.

**Targeting and Item Hierarchy.** The points made above about item difficulty being lower than person ability are further demonstrated in the Person-Item Map (Figure 3). The Person-Item Map lends information about targeting of the measure, or in other words, the ability of the measure to assess the entirety of the construct. Targeting is one the greatest contributions of Rasch analysis over CTT methods. Targeting is made possible in Rasch analysis by placing persons and items on the same metric (logit
position) to assess how ability and difficulty relate. In other words, this approach yields information about the extent to which the items cover the range of managerial competence in the sample, and also if the variance in ability in the sample covers the range of difficulty of the items. As demonstrated in Figure 3, the Rasch model software graphs person and item locations to simplify comparison of person ability and item difficulty. This approach also allows for identification of potential gaps in the measure and to determine the success of the scale to create a continuum of managerial competence.
TABLE 1.2 Rasch data set
INPUT: 287 PERSON 31 ITEM REPORTED: 287 PERSON 31 ITEM 7 CATS WINSTEPS
3.72.2

---

PERSON - MAP - ITEM
<more>|<rare>

8     # +
     |
     |

7     .  +
     |
     |

6     # +
     .## |
     .# T|

5     # +
     ## |
     ### |

4     .## S+
     ####### |
     ###### |

3      +
     ############# |
     ####### M|

2      +
     .########## |
     .###### |

1      .#####  +  ITEM24 ITEM28
     .########## |S  ITEM1 ITEM16 ITEM25 ITEM29 ITEM7 ITEM8
     .### |  ITEM11 ITEM13 ITEM14 ITEM15 ITEM23 ITEM27 ITEM9
0      ## +M ITEM19 ITEM20 ITEM26 ITEM31
     . T ITEM10 ITEM2 ITEM3 |
     . | ITEM12 ITEM21
     S ITEM17 ITEM32
-1     + ITEM18 ITEM4
     |
     T ITEM22

-2     +
     <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

Figure 3. Person-Item Map
From the Person-Item Map it is clear that the items cover a relatively narrow range of the construct of managerial competence. The participants are fairly well spread but the items appear easy for this sample. There are not enough items covering more difficult competencies in management in transportation. This is discussed in greater depth in Chapter 5. The mean logit position for items is set at 0.0 while the average logit position for people is 2.5. Thus, the samples managerial competence in transportation exceeds the managerial competence in transportation measured by this set of 31 items by 2.5 logits. Further, the distribution of items along scale had items generally appearing between logits -1.76 and 1.40, which can be seen on the right side of Figure 1. This indicates that the MSFI: TV items currently provide a good measurement of managerial competence in transportation on the lower end of the continuum. However, there is a noticeable gap in items covering greater levels of competence. Further, there appears to be redundancy in items assessing lower levels of competency. Several of the items appear redundant. Items 16, 29, and 9, in particular, appear redundant, and had poorer fit statistics. Closer analysis of what the items assess was warranted. These items measure areas of competence critical to the measure and their fit statistics indicate they do not threaten the unidimensional structure. Further, elimination of these items may weaken the reliability of the subscales. This is discussed in greater detail in Chapter 5. On a general level, the people in this sample do not line up with the items because the items do not tap the full range of competence of the sample.

**Differential Item Functioning.** Differential item functioning (DIF) relates to whether or not items vary in meaning based on different time points or groups. DIF, or
invariance, assesses if the meaning prescribed to a variable remains constant from one occasion (or in this case group) to the next. The idea is that the items should remain constant across groups (i.e., men and women; Bond & Fox, 2007). The researcher was interested in knowing if the MSFI: TV items have the same difficulty level for men and women. Assessment of DIF also indicates if the variable values are independent of the instrument and consistently measuring the same way regardless of the sample. To assess the DIF of the items that comprise the MSFI: TV, the t-test was observed for significance comparing items between males (n = 226) and females (n = 59), a Pearson correlation was generated between gender based on item difficulty, and the DIF plot was observed. This was an important step in the analysis of the instrument as one of the intended uses of the scale is to be able to meaningfully compare leaders in the workplace, and it is essential that the items mean the same thing across key groups, like gender. In other words, measurement equivalence must be maintained across groups for this to be a possibility (Chiang et al., 2009).

Based on the significance of the t-test, 8 of the 31 items showed DIF between gender. Item 2, “Manages self effectively” (p value < .01), Item 9, “Provides feedback that is accurate and helpful to improving performance” (p value = .03), Item 15, “Knows when to become involved in a conflict and when not to”, (p value < .01), Item 17, “Delivers what is promised“ (p value = .03), Item 23, “Focuses on identifying the root cause of safety incidents and not on placing blame” (p value = .02), Item 24, “Holds meaningful and constructive safety meetings” (p value < .01), Item 25, “Is a model for the use of safe work practices” (p value < .01), and Item 26, “Provides necessary training
and/or tools to safely perform the job” \( (p \text{ value} < .01) \). Interestingly, items 23 – 26 make up the safety leadership subscale, indicating that it is a measure that varies by gender. This subscale could potentially measure safety leadership differently for men than it does women. However, Linacre (2011) reported that significance tests should be interpreted conservatively because differences can be statistically significant, but very small in terms of having any impact on the meaning of the measure. Thus, these items were retained as they assess a key area of managerial competence in transportation. Nevertheless, these are important results that need further attention in future studies as they may indicate a strong need for further assessment of the validity of the instrument across ethnicity/race, age and gender.

A correlation was also generated between values of item logit positions across calibrations of the two groups, men and women. This provided a second way to assess DIF by gender. The correlation between gender by logit position was high, \( r = .85 \ (p \text{ value} < .01) \). This is contrary to that found in the t-test significance tests presented above. This high correlation indicates the items most likely work in a generally stable fashion across gender.

A third step in detecting DIF of the items between gender was to examine the DIF plot (Figure 4). This plot demonstrates the DIF of the MSFI: TV items between gender. It appears that items 2, 9, 15, 24, 25, and 26 all may exceed a difference between calibrations of .50 logits or greater, indicating likelihood of variance between groups. The plot appears to identify the same items as identified above by the t-test significance values. However, significance levels may be more accurate. Given none of these items
threaten unidimensionality of the scale and predominantly measure an entire subscale, the items were retained. These findings warrant additional assessment of DIF of the safety leadership subscale items in particular.
Summary

This chapter presented the results of the analysis of the Managerial Success Factors Inventory: Transportation Version (MSFI: TV). In sum, the hypothesized general managerial competence model thought to adequately capture managerial competence across industry was not found in the EFA. Instead, a five-factor solution was found to be most interpretable. The five-factor solution differed from the general competency model indicating that managerial competence in transportation may require unique competencies not required in other fields. The original 47 items were reduced to 34 items based on cross-loadings in the EFA. The model obtained in the EFA was confirmed in the CFA with a second-order factor, though some of the fit statistics were not strong. The CFA also indicated that the items load significantly well on each factor. The higher-order factor model indicated unidimensionality of the MSFI: TV by showing that all subscales relate significantly to one another and to the overarching construct of interest, managerial competence in transportation. Thus, Rasch analysis was conducted on the entire scale.

Initial findings from the Rasch analysis indicated that overall the data for the MSFI: TV fit the Rasch Model quite well. However, three items were detected as misfitting and threatened the unidimensionality of the scale. Thus, they were removed and Rasch analysis was rerun on the remaining 31 items. In regards to the relation of item difficulty and person ability, overall the scale items are far too easy for the sample. There are far too few items measuring higher levels of managerial competence. This indicates that more items are needed that tap greater levels of managerial competence in transportation. The Rasch analysis also revealed that certain key items may function
differently for men and women. The safety leadership subscale items seem to lack invariance in particular. In terms of initial validation for the scale, the results are not as hypothesized. Evidence for convergent validity was obtained by finding high correlations between total MSFI: TV scores and subscale scores with a measure of overall performance. However, initial evidence of concurrent and discriminant validity was not obtained by assessing correlations between the MSFI: TV scores and supervisor’s ratings and absenteeism. In terms of the research questions, questions 2, 3, and 5 were supported whereas 1, 4, 6, 7 and 8 were not. Implications of these findings are presented in greater detail along with additional research needs to enhance the psychometrics of this scale in Chapter 5.
Chapter 5—Discussion

This discussion chapter begins with an overview of the study including the purpose and rationale. From there, a brief summary of the literature reviewed for this study is provided. Then a discussion of the results that pertain to each of the eight identified research questions is provided. Recommended next steps for development of the scale based on the findings of this study are then provided. Theoretical implications related to general management and management specific to transportation from the study are also summarized, as well as potential uses of the scale. Finally, a section on the limitations of the study precedes concluding remarks.

The purpose of this project was to begin the development and initial validation of a managerial competency measure that could be used in the recruitment, retention, and rewarding practices of transportation agencies. Development procedures and evidence of initial reliability and validity for this new instrument were presented. This study used both classical test theory and item response theory, as well as structural equation modeling to obtain initial psychometric data on the Managerial Success Factors Inventory: Transportation Version (MSFI: TV). The results of each of the eight research questions are presented in the Results Chapter 4 and then expanded on below. The key findings are discussed as well as recommended uses of the scale in business in this chapter.
Brief Literature Review

Leadership plays an integral role in organizational outcomes. Thus, the means by which we select, promote and reward leaders is a critical aspect of achieving operational excellence. Leadership competency assessment tools and accompanying competency models have grown from a new concept to common practice over the last three decades. Since first introduced by David McClelland as key differentiating factors of successful leadership, competencies have been established by several industries (Calhoun et al., 2008). Despite widespread practical growth and use among various industries, the academic literature has lagged behind. Further, certain key industries, such as transportation, remain without a tailored managerial competency instrument.

Extensive review of the reigning models of business management and leadership in various academic databases revealed several core competencies thought to cut across industry lines as key differentiating factors of peak performance. The most heavily cited and well established structures of leadership include various competencies organized in a myriad of ways. However, agreement amongst scholars on a single core group of competencies has been difficult to achieve for the over 100 years of research on the topic. Thus, recent research has suggested the need to explore industry-specific models of management. The current investigation was charged by this premise.

Finally, most current scales of managerial performance developed in both academic and corporate circles utilized classical test theory (CTT) despite newer techniques that may lend additional critical information to the development of managerial performance scales.
A review of the current literature examining psychometric properties of the most heavily used managerial competency assessment tools revealed that most are validated in a variety of settings with various management groups. Few have specifically set out to develop a managerial competency scale for a particular industry using IRT principles. No studies were found in the academic literature on the development of a managerial competency scale for a particular industry using a mix of Rasch modeling and CTT tenets. Thus, a major contribution of this study was the development of a managerial competency scale for the transportation industry that was developed by employing Rasch analyses in addition to traditional CTT approaches to allow for assessment of the scale at the item and person level.

Key Findings

Research Question 1. Research question one stated that EFA would identify an interpretable four-factor solution based on the Hay Groups general managerial competency model; Managing Yourself, Managing Your Team, Managing the Work, and Managing Collaboratively. Instead, a five-factor model solution was found using exploratory factor analysis methods (the scree plot, interpretability of the solutions, and unidimensionality of the measure). This five-factor model contained items that were originally intended to load on different factors, leading to a completely different model than that hypothesized. The researcher rotated 1 factor, 4 factors (to represent the four clusters of the Hay Group Model) and 11 factors (to represent the Hay Group’s 11 competencies) and could not find an interpretable solution that mimicked this general managerial competency model. Based on the scree, Eigenvalues, and interpretability after
maximum likelihood rotation, a five-factor solution was most interpretable. The small sample size could have contributed to the inability to detect the model from EFA, but nevertheless, the managerial competency model thought to predict peak performance regardless of industry was not found. Instead, a five factor model was obtained from the EFA; Problem Solving, Emotional Competence, Results Orientation, Safety Leadership, and Integrity. Items that were written to assess different clusters of the Hay Group model ended up grouping together in a different fashion with this sample. Thus, new subscale names were given to reflect the items that grouped. Rather than general cluster titles, each of the five subscales (or subcomponents) of the measure were named a more specific competency (i.e., integrity). After setting criteria to retain items, the original 47 items were reduced to a total of 34 items that were organized across five unique competencies. Though this made the five-factor solution work, it reduced the number of items for the scale and eliminated some of the competencies intended to be measured. This relates to the scale development conflict of length (reliability and comprehensiveness versus time to complete or practicality). The researcher was not able to find the Hay Group model in the EFA with a transportation sample, thus, question 1 was not supported. However, five interpretable, practical subscales and a total scale that can be used in consulting and coaching endeavors emerged.

**Research Question 2.** It was predicted that the scale would demonstrate acceptable internal consistency by producing high Cronbach Alpha values for both the total scale and subscales. Internal consistency for the composite managerial competence in transportation score was found to be strong, indicating the items are measuring a
unitary construct and relate to one another or the latent variable of managerial competence. Acceptable levels of internal consistency reliability were also found for each of the subscales, though some higher than others. Thus, question two was supported. It is important to mention that the subscales of problem solving, emotional competence, and safety leadership each produced a Cronbach Alpha well above acceptable levels. However, the results orientation and integrity subscales produced a Cronbach Alpha just below the .80 level, indicating the items may not be strongly related to each other or their respective variables intended to measure (i.e., results orientation and integrity). Most likely, the lower internal consistency values for these subscales are the result of the relatively few items retained to measure that aspect of managerial competence. The results orientation factor, for example, was left with only 4 items and the integrity factor only 3 items following the EFA. Conversely, the problem solving factor, which produced the highest Coefficient Alpha, had the most items. These results indicate a need to write more items in future studies that address results orientation and integrity, should the researcher decide these are important factors to measure in assessing managerial competence. Ultimately, reliability estimates decrease with fewer items. This is an issue for the entire scale as well. The researcher sought to develop a measure that is both reliable (has plenty of items that accurately assess the construct in a comprehensive and consistent nature) and at the same time is useful and practical (few enough items that managers in a busy transportation agency actually complete it and respond in an honest fashion). This issue will need to be a focus of future studies.
**Research Question 3.** Research question three predicted that a confirmatory factor analysis would confirm the structure of managerial competence in transportation found in the exploratory phase of the study. A five-factor structure was obtained from the EFA: problem-solving, emotional competence, results orientation, safety leadership, and integrity. This model demonstrated adequate model fit and all of the items were significantly loaded on the factors. A higher-order factor model was tested as well to assess if adding the construct of managerial competence would improve model fit. The higher-order model slightly dropped in model fit yet it too reached acceptable fit statistics. Each of the latent variables (i.e., subscales) loaded significantly well on the second-order factor and correlated very highly. The items were all significantly loaded on their respective factors. The correlations between the latent variables were actually too high, indicating that it might be best to collapse two of the latent variables into one. However, when a four factor solution was rotated in the EFA it was not interpretable. Plus, a four- rather than five-factor solution further limits the number of subscale scores that would be available for use in business application (i.e., leadership development initiatives). Often times more subscales are desirable in coaching relationships to provide more specific scores related to key competency areas. Thus, the researcher maintained the five-factor model.

Most importantly, the CFA results suggested the presence of a unidimensional structure underlying the MSFI: TV. These results are at odds with what was found in the EFA which found a five-factor structure. However, as Chiang et al. (2009) noted, that is probably due to factor analysis being more likely to find multiple factors while other
methods, such as Rasch analysis, may be more likely to find unidimensionality. Based on the results of the CFA and Rasch analysis, the researcher concluded that the current version of the MSFI: TV is unidimensional. Therefore, the scale is most likely measuring just managerial competence but also has five subcomponents (or subscales) that relate to each other and each relates closely to a total managerial competence in transportation score. This is quite favorable in terms of usefulness of the scale as it is often necessary to compare leaders based on a total composite score. The higher-order model, though fit slightly decreased, is favorable as it is more practical to have five related factors that also relate to an overarching total managerial competency in transportation score. The unidimensionality of the scale lends itself nicely to leadership development initiatives.

Though it was decided that both the five-factor and second-order models achieved acceptable fit, some of the fit statistics for both models, such as the Goodness-of-Fit statistic, was not in an acceptable range. This is problematic in that it indicates lack of fit of the model to the data and should be addressed. However, this could be the result that the degrees of freedom (df = 517 and 522) were large relative to the sample size (n = 143) in this particular study. Nevertheless, the lack of completely acceptable fit statistics indicates additional work on the model in future research. Future research should include conducting a CFA on a larger sample of transportation managers to compare the Hay Group model and the model obtained in this study. This would make a significant contribution to the literature. This was beyond the scope of this study, especially given the Hay Group model was not found in the EFA. Since the Hay Group model (2001) was not found in the EFA it was not compared to the five-factor model in the CFA. Given the
overall acceptable fit of the two models tested in this study, research question 3 was supported.

**Research Question 4.** One of the most interesting findings in this study stemmed from the Rasch analysis conducted on the scale to assess if item difficulty covers the range of managerial competency ability found in the sample. It was predicted that this analysis would show that items vary with increasing amounts of competency in the participant and cover the range of levels of managerial competency in the participants. The Rasch analysis offered significant findings about the item difficulty of the scale. This was potentially the single greatest contribution of the Rasch analysis to the study. The logit position values provided in Table 8 offer a common value between item difficulty and person ability in which the two can be compared. Items with higher logit positions need more of the latent trait to endorse. In other words, we were able to see which items required more managerial competence in transportation to endorse, and therefore have a better idea of the difficulty of the pool of items on the scale. For example, item 33, which reads, “Is able to use conflict in groups effectively”, had the highest logit position indicating more managerial competence in transportation would be required to endorse it. Item 5 (“Is honest and trustworthy”) was found to be the least difficult item to endorse. Further, item 28 (“Acts as a catalyst to a committee’s decision-making process”) has a higher logit difficulty than item 18 (“Anticipates what needs to be done and does it”). Conceptually, this makes sense. The ability to move a team toward a decision is considered to be much more difficult than simply taking initiative.
Results from the Rasch analysis revealed that the items cover a relatively narrow range of managerial competence. Overall, the item difficulty is much lower than ability of the sample in this study. This indicates that more items are greatly needed that tap higher levels of managerial competence. Rasch modeling allowed for this finding by converting the data to an interval scale to create a “yardstick” in which item difficulty can be examined simultaneously to logit ability of the persons. The scale is currently not very well targeted to the sample and the potential ability of managers in transportation at the top end of competency. This finding indicates a dire need for future studies to write more difficult items. The results indicate a significant gap in item difficulty at the high end of managerial competence. It appears the scale was not difficult enough for the sample. There could be several explanations for this. One may be that the sample was concerned about the confidentiality of the results given the study was conducted at a place of employment. Even knowing that only aggregate data were to be presented to management, respondents may have attempted to paint a favorable impression of their competencies to avoid more trainings or having their department appear to be lacking in key competency areas. This would have made the items appear less difficult due to a high endorsement rate. Nevertheless, the results of the targeting analyses indicate the need to address the lack of difficult items. The current version of the MSFI: TV does not adequately measure greater levels of managerial competence in transportation. Options include rewriting items to be more difficult, adding more items that assess high levels of competence, or administering the scale to samples that may be less concerned about
potential negative consequences of low scores on the measure to retest this finding. Support for research question 4 was not garnered in this study.

**Research Questions 5, 6, and 7.** The total MSFI:TV score and all five subscale scores were correlated with three different measures of performance to assess validity of the scale. Questions 5-7 predicted the scale would show initial convergent (by correlating positively with a measure of managerial performance), discriminant (by correlating inversely with a measure of absenteeism), and concurrent (by correlating positively with supervisor’s ratings on a measure of managerial performance) validity. Both the subscale and scale scores correlated significantly and positively with the 5-item performance measure, offering initial support for convergent validity of the scale. However, the subscale and scale correlations with both the supervisor ratings and absenteeism scores were non-significant and very small. All of the correlations were close to zero, indicating no relationship. Obviously, this is a surprising finding that does not bode well for the scale, as it indicates it could be measuring something different than competence in managing a transportation agency. However, there are several reasons this could have happened. One explanation of this could be the small sample size. Another may be the nature of the variables of absenteeism and supervisor’s ratings. Absenteeism, for example, could have been higher than what was self-reported. Future studies that use absentee rates may want to try to get the company to provide absenteeism data rather than relying on self-report. As for the supervisor’s ratings, it is perplexing as to why there would be virtually no relationship detected between how one rates him/herself on performance and how their supervisor rates them. This was one of the more surprising
findings in the study. One possible reason for the lack of correlation could be the issue of the self-assessment sample response style leaning heavily toward a favorable impression. Supervisors knew that their scores of the participants were not going to be reported so they were in a position of rating more freely. Though the self-assessment sample was notified that their responses would not be given to management, participants still could have attempted to paint an overly favorable picture of themselves, thereby lowering the correlation. Regardless, research question 5 was supported while questions 6 and 7 were not. Initial validity outcomes of the scale are not strong, indicating a need to refine the items and add more to better assess competence in management in transportation more completely.

**Research Question 8.** It was predicted that the scale would be an invariant measure in that managerial competency would be the same thing across groups (e.g., gender). Invariance is important to the integrity of a measure as it holds that when two persons in different groups have the same expected raw item responses they also have the same level of the latent trait. The researcher was interested in determining if the difficulty of the MSFI: TV items were constant across groups, especially between gender, as it is a common group identity of key leaders in business. This hypothesis was assessed with Rasch analysis software by examining differential item functioning (DIF) of the MSFI: TV items. Only partial support for this question was found. The following items were found to DIF between gender, Item 2, “Manages self effectively”, Item 9, “Provides feedback that is accurate and helpful to improving performance”, Item 15, “Knows when to become involved in a conflict and when not to”, Item 17, “Delivers what is promised“,
Item 24, “Holds meaningful and constructive safety meetings”, Item 25, “Is a model for the use of safe work practices”, and Item 26, “Provides necessary training and/or tools to safely perform the job”. This indicates that men and women with the same levels of the latent trait, managerial competence in transportation, may produce different responses to those particular items. The difficulty level of these items varied significantly between men and women in this sample. It is important to point out that items 24-26 make up 3 of the 5-item safety leadership subscale. Thus, this subscale is particularly lacking invariance. Despite DIF detected from significance levels of t-tests between groups, a correlation conducted between the items logit value between groups was high, indicating stability across groups. This plus the importance of the areas each of these items assess led the researcher to retain each of the items. Nevertheless, it is important to use the instrument with caution given these findings, especially when comparing men and women on the total or subscale scores. Further work is clearly required on the safety leadership subscale in particular and the entire measure to ensure invariance.

**General Conclusions from Rasch Modeling**

Examination of scale fit, separation statistics, item fit, logit difficulty, targeting of the scale, and differential item functioning yielded several interesting findings that both provided initial support for the scale and gaps in ability coverage that must be addressed in future versions of the MSFI: TV. Based on the fit statistics, the data fit the model quite well. Both person and item fit for the scale are within acceptable ranges. The Rasch model predictions fit that observed in both person and item behavior. These findings suggest that as intended, participants higher in managerial competence in transportation
produce higher scores on the measure than those with lower levels of managerial competence. The Rasch analysis results also strongly indicated that the scale is unidimensional. However, in the first Rasch analysis all 34 items were analyzed and 3 items were found to be misfitting, indicating they may be measuring a different construct and therefore threatening the unidimensionality of the measure. The three underfitting items were removed and the global fit statistics improved. All of the remaining 31 items fit the measure well in the second Rasch analysis. Based on the acceptable fit of the higher-order model and the findings from the Rasch analysis, it appears the MSFI: TV can be treated as unidimensional. The unidimensional model has five subcomponents that relate strongly to one another and the total managerial competence in transportation score. Such findings have strong implications for the uses of the scale, which are discussed later in this chapter.

In terms of reliability of the scale based on the Rasch analysis, the scale showed excellent separation indicating that the sample was diverse enough to order the items. The items appear to be working well together to reproduce a person’s transportation managerial competence. The results indicated that the participants can be distinguished from one another quite well. More specifically, we found that the sample can be separated into six distinct transportation management competence groups. The very high person reliability statistic obtained from the analysis indicates that the scale replicates a person’s placement along the competence continuum reliably. This is a very important finding in terms of usefulness of the scale. Based on this finding, consultants using the measure with transportation leaders in a coaching relationship could assume that
increases in scores on the measure over time were the result of improvement in the leader’s competence and not unreliable measurement.

**Theoretical Implications**

As reported in Chapter 2 of this paper, numerous taxonomies, models, and theories of leadership and management exist. Most people can name a few qualities of manager’s they feel are critical to success. Yet, there is little agreement on the structure of management or organization of competencies that must be included in a comprehensive theory of managerial and leadership competence. Thus, it is unclear which qualities of management must be included in comprehensive assessment of managerial competence in a particular field. Indeed, despite 100 years of research on the topic, there is little agreement on a comprehensive set of competencies that make up managerial or leadership competence.

The Hay Group general manager competency model is one of the best. It stems from the original work by David McClelland (1973) and Richard E. Boyatzis (1982) with the McBer consultancy group. Boyatzis (1982) originally developed a nine competency model from interview-based assessments of 253 managers of 12 Fortune 500 firms and four public agencies. The nine competencies included; efficiency, concern with impact, proactivity, self-confidence, oral presentation skills, conceptualization, diagnostic use of concepts, use of socialized power, and managing group process (Boyatzis, 1982). Since this original work, the HayGroup and their colleagues have conducted hundreds of competency studies on various jobs worldwide that utilized leading research methodology and statistics including structural equation modeling, CTT and IRT. Their
The most recent model, presented above, which contains four competency domains, includes the most critical competencies for “defining excellence in a wide variety of management roles across industry” (HayGroup, 2001, p.2). The HayGroup model served as the foundation for the writing phase of a transportation-specific managerial competency scale in this study. The model contains four clusters of competencies; Managing Yourself, Managing Your Team, Managing the Work, and Managing Collaboratively (HayGroup, 2001) and eleven competencies fall across these four domains (empathy, self-control, self-confidence, developing others, holding people accountable, team leadership, results orientation, initiative, problem solving, influencing others, and fostering teamwork (HayGroup, 2001)). A major task of this study was to see if the Hay Group model could be found in a sample of transportation managers by developing a measure of managerial competence in transportation based on the Hay Group model and then factor analyzing it using EFA procedures.

Based on the results of the EFA, there were no interpretable solutions that closely resembled the Hay Group’s (2001) model. The researcher rotated a 1 factor, 4 factor (to explore if the items hung together in a similar fashion to the 4 cluster model), and an 11 factor model (to explore if the items hung together based on the 11 competencies in which the instrument was written). Instead, a five factor solution was most interpretable. Each of the five factors was composed of items that were originally written to tap different clusters from the Hay Group model. This led the researcher to devise a new model for management in transportation. Though five factors led to five useful scales that can be used in coaching and consulting endeavors, the validity of each scale is weak at
this point as there are certain items that do not fit well conceptually. Further, certain competencies measuring very important aspects of transportation management were removed due to cross loading. From the EFA results, it seems safe to say that the five competencies currently measured are key to transportation management: problem solving, emotional competence, safety leadership, results orientation, and integrity, but there are probably many others that exist that need to be incorporated in future versions of the MSFI: TV. For instance, the current version of the MSFI: TV does not assess self-control, technical abilities, analytical thinking or financial skills well, to name a few. A major contribution of this study is that the model underlying managerial competence in transportation may differ slightly from other industries (i.e., more weight put on emotional/social competence, the organization of items to make competencies and competencies to make clusters may be unique, and entirely unique competencies (i.e., safety leadership) are needed). Thus, the derived model most likely accounts for additional managerial competence in transportation above and beyond that previously accounted for by general models.

The findings from this study offered great value to the differences between general managerial competence and managerial competence specific to transportation. The age old question of whether general managerial competency models hold up across industry lines was a central question for this study. The researcher was interested in if the structure of managerial competence assessments on a general level would be found in a sample of manager’s in transportation. Based on an extensive literature review, cognitive interviews, and three focus groups held at a DOT focusing on the essential differentiating
factors of successful transportation leadership, safety leadership and emotional competence were two frequently identified competency areas that are thought to be key to this industry. Safety leadership in particular was mentioned numerous times as an industry-specific competency essential for successful management in transportation. Thus, the current version of the MSFI: TV focuses heavily on these two aspects of transportation leadership. These two areas or subscales were found to have two of the three highest item loadings, and the items loaded less on other factors than other items did. Factor two was labeled the emotional competence factor and was made up of items such as, “Responds well to emotions of others” and “Able to recognize his/her own emotional reaction to people, events, and situations.” The fourth factor, termed Safety Leadership, reflected the leader’s ability to promote and establish a value of safety first. Without a doubt, the DOT is interested in identifying the specific leadership practices that promotes a culture of safety.

There was consistent emphasis placed on inclusion of items that measure integrity, credibility, ethical behavior, and trustworthiness as well. Most telling, of the 47 items (leadership competencies) listed, the top two items identified as most important by the focus groups related to being trustworthy and the ability to build trust. We began to realize that transportation wants, and needs, leaders who build a sense of camaraderie and trust among their team. Integrity matters and valid assessments that can be used to recruit, train, and promote leaders need to measure this competency. It is clear that transportation managers think that integrity plays a critical role in achieving great leadership practices. This is considered an important finding from this study. Thus, the fifth competency on
the MSFI: TV was found to be most related to trust and integrity, and future endeavors will focus even more exclusively on how to measure integrity as a leadership competency. More items are needed that assesses the full range of integrity as a managerial competency.

**Usefulness of the Scale**

Accurate leadership assessment is at the core of organizational performance and success. Competency-based scales have become the gold standard in management and leadership assessment instruments. Most industries have started the process of identifying the essential competencies at the heart of outstanding management in their field (Calhoun et al., 2008). Transportation, a key industry in societal and economic success, has lagged behind other industries in developing an agreed upon managerial competency model and an accompanying assessment tool that can be used in recruitment, training, and selection procedures. Vogel (2001) suggested that the industry faces severe recruitment and retention of quality worker issues in the face of increasing demand for services. Thus, one of the key contributions of the current study was the development and initial validation efforts of a scale intended to fill this gap in one of the world’s most critical industries. Initial psychometric findings on a measure intended to assess managerial competence in transportation in particular are presented. It appears the scale has good internal consistency, as do each of the five subscales. Further, reliability estimates obtained in the Rasch analysis at the item and person level indicate the measure produces consistent assessments of the person’s managerial competence level. Both the MSFI: TV total scale score and each of the five subscales provide useful information that can greatly enhance
executive coaching, organizational development, team development, and professional
development services that are already being offered at numerous transportation agencies
with measures that have absolutely no empirical evidence or with no measures at all.

A closer look at each of the five subscales reveals that the items of the MSFI: TV
actually assess a variety of necessary skills to successfully manage in transportation.
There are a number of uses for both the total scale and subscale scores of the MSFI: TV.
For instance, the problem solving subscale alone is composed of items that really get at
how a manager works with people to solve problems (i.e., “Negotiates effectively with
other units in the organization to get things done”, “Is able to persuade others to adopt
new points of view”, “Is able to use conflict in groups effectively”, “Is able to rally
support for important initiatives”, “Connects with people at all levels of the
organization”, and others). This first scale, though at its core assesses one’s ability to be
creative and problem solve, equally could have been labeled “people”, as it really
measures human resource potential, interpersonal and social understanding, relationship
building, self-confidence, and team leadership. These are essential qualities for any
manager to possess, and the MSFI: TV Problem Solving scale score will reliably assess
these competencies. Transportation agencies can use this scale alone to identify where
key individual executives, teams and even entire departments fall in terms of each of
these competencies, and then devise leadership development programs that will provide
growth in deficit areas identified by the measure.

The safety leadership and emotional competence scales each produced high
internal consistency estimates and the items all have excellent face validity. The
emotional competence scale seems to also tap coaching behaviors with items that read, “Coaches employees to help them achieve success” and “Recognizes where others are in the professional development process”. The total scale score of the Emotional Competence scale provides a very useful measure of one’s EQ, and will highlight professional development needs in terms of internal coaching competencies.

Unlike many of the commonly used leadership and managerial competency scales, the MSFI: TV is unidimensional and can easily be used to generate a composite managerial competence in transportation score. The results from this study suggest that the items assess one overarching construct of managerial competence in transportation. This is a nice feature of the scale as it will allow for comparison between job applicants and current leaders on one total scale score by producing a single value. Often times companies look to make decisions regarding hiring and promotion based on a leader’s overall talent. The MSFI: TV total scale score will be a nice summary score to factor in these types of decisions.

The analysis of the MSFI: TV revealed several initial issues that will clearly need to be addressed in future research. The results of the Rasch analysis proved to highlight the significant contribution of Item Response Theory to the scale development process. No doubt, use of Rasch analysis added key information above and beyond that obtained by the CTT methods employed. In terms of item hierarchy and targeting, it is clear that more items are severely needed that measure higher levels of managerial competence. Unless the current sample presented themselves in an extremely favorable light, the current items do not assess upper bands of competency in transportation management and
this will need to be addressed. There are also issues with invariance and the current rating scale that will need to be further examined with larger samples. Overall, results indicate the MSFI: TV is unidimensional and reliable but lacks validity. The results of the CFA and Rasch analysis supported use of the total scale score in coaching endeavors, but also brought several glaring issues to light that can be tackled in future studies. These findings expose lingering questions in managerial competence theory and the way it is assessed.

**Factor analysis.** The exploratory analysis of the structure underlying managerial competence in transportation did not lead to similar item groupings of the model proposed by the Hay Group (2001). In fact, items that were written to reflect the four cluster areas specifically ended up correlating with very different clusters and there were no interpretable solutions that even closely resembled the Hay Group (2001) model. Future work on this scale will entail reworking the clusters to better fit the findings in this study and transportation’s needs.

**Implications for future research from Rasch Modeling.** The outcomes of this study clearly highlight the offerings of IRT methods over CTT. Much was learned about the MSFI: TV in terms of specific item and person behavior that is not offered in traditional CTT approaches. Ultimately, the items appear to be working well together but seem to be far too easy for the sample. There is a large gap in item difficulty at the top end of the managerial competence ability. There were also obvious items that appear redundant. Based on the Person-Item Map (Figure 3), it appears that items 9 and 15 are two of the redundant items that could be reworded or one removed in future revisions. The two items read very similarly (“Identifies high potential candidates” and
“Recognizes where others are in the professional development process”). Due to concerns about the instrument getting too short to maintain reliability and usefulness, both items were retained in this analysis. Future studies may want to slightly reword one of these items to ensure they measure unique information. It is probably going to be most useful to simply add more items that measure higher logit difficulties to ensure the scale is long enough to assess all key areas of managerial competence in transportation. It is also recommended that the number of items for the results orientation and integrity subscales be increased as well, as they currently have 4 and 3 items respectively, and additional items would improve reliability of the scale and validity of the measurement.

**Limitations and Future Research**

Much was gained from this investigation about the structure of leadership in public transportation, however, more work is clearly needed on the study of the essential success factors underlying outstanding leadership in this industry. The researcher hopes to continue to validate this measure and assess the 5-factor leadership solution found in the EFA in future research studies with other DOTs. There are several limitations of this study that warrant additional research and attention. One key limitation was the lack of diversity of the sample regarding transportation agency, ethnicity/racial identity, gender, and geographical location. Most of the sample was composed of middle aged, white, heterosexual males. This greatly limits generalizability of the study and poses serious issues if used for coaching engagements with diverse groups. Practitioners who utilize this instrument with racially/ethnically diverse leaders are strongly encouraged to consider how one’s racial/ethnic make-up may alter the meaning prescribed to the results
of this assessment. This managerial sample was a good one but it lacked diversity in terms of representing the myriad of agencies one might want to use a “transportation management competency scale”.

Further, the sample is small for the research question, methodology employed, and statistical analyses conducted. This is a major limitation and could have potentially impacted the findings of the factor analysis. The size of the sample led to clear limitations of what was possible regarding EFA. Future studies are needed that follow similar steps in terms of research methodology but with a much larger sample.

Another limitation of the study was the relatively few items incorporated on the measure to assess specific aspects of leadership and management that are thought to be important. Key competencies are not assessed by the current version and there is a good reason for this limitation. Managers who participated in this study were asked to complete the measure during their regular work shift. It was also available on line to be completed at home but many reported wanting to complete it during regular business hours. Though they were given adequate time to complete the measure, they were not relinquished of their regular work duties for the day so this added some restriction to the amount of time managers would willingly take to complete the survey. Thus, this limited the number of items that the researchers could incorporate on the scale. As a result, key managerial competency areas beyond the 11 identified by the Hay Group (2001) were not incorporated on the instrument (i.e., technical expertise, organizational awareness, professionalism, financial skills, etc.). In other words, to comprehensively assess all of the behaviors, qualities, and skills required to successfully function as a manager in
transportation, a scale may need well over 100 items. This is not practical or useful. Instead, this study contributed a measure that appears to assess a few very critical competencies well. Future research on the measure should look to include more competencies than the five identified in the current version. Further, the scale has been completely developed on self-assessment data to this point. Future work is needed that incorporates more 360 data.

An additional limitation of the study was the measures used to assess performance and absenteeism. As mentioned, the company that provided the participants put survey length restrictions in place that greatly limited the researcher’s ability to incorporate additional measures. Restrictions were placed on survey length that restricted the researcher from incorporating performance measures with extensive psychometric support. Thus, the researcher included the 5-item performance measure used in this study to maintain adequate survey length. Though the performance measure used in this study demonstrated adequate initial psychometrics, it does not have substantial psychometric data to support its use. This is a limitation. Additionally, absenteeism was measured by self-report as opposed to being obtained from the company due to the company being unable to provide such information. Self-reported absenteeism rates may be less reliable than obtaining absenteeism rates from the company. Thus, this is another limitation that should be addressed in future studies.

Another limitation of the current investigation was the high likelihood of response bias in the self-assessment sample. The self-assessment and supervisor ratings demonstrated no correlation indicating a high likelihood of response bias in the self-
assessment sample. It appears the self-assessment sample inflated their scores in effort to paint a favorable image of their skill level. The inflated or overly favorable responses could be the result of issues pertaining to fear of layoff or additional training programs if results identified deficit areas.

Future studies should also look to bring findings from this study together with other leadership projects underway in transportation. There are very important studies taking place within the airline and train industries that could merge well with the contributions from this study to inform future projects. Incorporating previous studies of a similar mission could help the industry arrive at a universal transportation leadership measurement tool based on an industry-wide competency model. Further, it will be essential that future investigations of the structure of managerial competency in transportation obtain samples from private as well as public companies.

**Summary**

The present study reported (1) development of a measure of managerial competency in transportation and (2) evidence of initial reliability and validity for this new instrument. The Managerial Success Factors Inventory: Transportation Version (MSFI: TV) was written to reflect a well-regarded general managerial competency model in business known as the Hay Group model (2001), as well as three additional competencies identified in the scale development process; integrity, emotional competence, and safety leadership. The study described the development of the scale using CTT, and then provided initial evidence of reliability and validity using CTT, EFA, CFA, and Rash modeling. The results of the EFA led to a five-factor solution; problem
solving, emotional competence, results orientation, safety leadership, and integrity. The results of the CFA indicated acceptable model fit for both the five-factor model obtained in the EFA and a higher-order factor model including the construct of managerial competence in transportation. This higher-order model was of particular importance as it improves application of the scale by use of a total managerial competence scale score. The CFA also offered strong support for the item loadings on each of the factors. The Rasch analysis supported the unidimensional structure.

Many of the findings offered strong initial support for the measure including convergent validity obtained from total and subscale correlations with a measure of performance, strong person and item reliability estimates obtained from Rasch analysis, and strong Cronbach Alpha values for both the subscales and scale score. The results of the Rasch modeling, however, suggest that the items were not well targeted for the sample. The MSFI: TV would benefit greatly from more items assessing higher levels of managerial competence. More items are also needed for the results orientation and integrity competencies. Further, inclusion of items that measure competencies thought to be central to successful leadership and management in transportation but not measured by the current version of the MSFI: TV are also needed. The use of both CTT and Rasch modeling procedures in this study provided information that can be used to further develop this scale and enhance our understanding of management as it relates to key industries.

In closing, leadership matters a great deal to organizational performance. Thus, we need accurate, reliable, and valid instruments to select, train and promote great
leaders. Especially in key industries, like transportation, as this industry is being asked to do more with less including safely moving people and freight across the world. Accurate assessment instruments can be used in coaching engagements, leadership programs, team development initiatives, culture change projects, engagement practices, employee reviews, training programs, and many other ways for transportation companies to flourish. Though assessment is only one piece in the complex process of leadership and organizational development, it is an important one. The primary reason for this project was to apply advanced statistical analyses to the development and initial validation of an instrument to assess managerial competency in transportation, and fill a need for an accurate, reliable assessment in transportation at the managerial level. From this exploration, transportation is one step closer to having a tool that can be used to secure leaders who are a good fit.
References


### DEMOGRAPHIC INFORMATION
Please answer the following biographical questions about yourself. These will be used for statistical analysis only. For the following questions, please write in, circle or shade in the option that best applies to you.

1. What is your age?  

2. What is your gender?  

3. Which of the following categories do you feel best describes your race or ethnicity?  

<table>
<thead>
<tr>
<th>Category</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian</td>
<td>O</td>
</tr>
<tr>
<td>Hispanic, Latino/a</td>
<td>O</td>
</tr>
<tr>
<td>African-American</td>
<td>O</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>O</td>
</tr>
<tr>
<td>American Indian or Alaskan Native</td>
<td>O</td>
</tr>
<tr>
<td>Multiracial</td>
<td>O</td>
</tr>
<tr>
<td>Other (Please Specify)</td>
<td></td>
</tr>
</tbody>
</table>

4. What is your marital status?  

<table>
<thead>
<tr>
<th>Status</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single (never married)</td>
<td>O</td>
</tr>
<tr>
<td>Married/Remarried</td>
<td>O</td>
</tr>
<tr>
<td>Divorced/Separated</td>
<td>O</td>
</tr>
<tr>
<td>Widowed</td>
<td>O</td>
</tr>
</tbody>
</table>

5. Which of the following best describes your present position?  

<table>
<thead>
<tr>
<th>Position</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry-level manager</td>
<td>O</td>
</tr>
<tr>
<td>Mid-level manager</td>
<td>O</td>
</tr>
<tr>
<td>Senior or Upper-level manager</td>
<td>O</td>
</tr>
<tr>
<td>Executive</td>
<td>O</td>
</tr>
<tr>
<td>Other (Please Specify)</td>
<td></td>
</tr>
</tbody>
</table>

6. How long have you been in this position?  

<table>
<thead>
<tr>
<th>Duration</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than one year</td>
<td>O</td>
</tr>
<tr>
<td>1-3 years</td>
<td>O</td>
</tr>
<tr>
<td>4-6 years</td>
<td>O</td>
</tr>
<tr>
<td>7-10 years</td>
<td>O</td>
</tr>
<tr>
<td>Over 10 years</td>
<td>O</td>
</tr>
</tbody>
</table>

7. How long have you been working in this field?  

<table>
<thead>
<tr>
<th>Duration</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than one year</td>
<td>O</td>
</tr>
<tr>
<td>1-3 years</td>
<td>O</td>
</tr>
<tr>
<td>4-6 years</td>
<td>O</td>
</tr>
<tr>
<td>7-10 years</td>
<td>O</td>
</tr>
<tr>
<td>Over 10 years</td>
<td>O</td>
</tr>
</tbody>
</table>

8. Are you working full or part-time?  

<table>
<thead>
<tr>
<th>Option</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-Time</td>
<td>O</td>
</tr>
<tr>
<td>Part-Time</td>
<td>O</td>
</tr>
</tbody>
</table>

9. In the past 30 days, how many full days of work have you missed?  

144
Appendix B

Performance Questionnaire

The following questions relate to how you feel about the quality of your work. Please check or shade in the most appropriate response for how well you think you have been performing. If you are a supervisor, please respond to each question in terms of how well you think your supervisee has been performing.

<table>
<thead>
<tr>
<th></th>
<th>Very Strongly Disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Very Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Overall, is an effective leader.</td>
<td>0</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>2. Overall, is an effective communicator.</td>
<td>0</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>3. Overall, has a high level of technical skill or expertise (e.g., transportation, logistics, etc.).</td>
<td>0</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4. Overall, is an effective performer within the organization.</td>
<td>0</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>5. Overall, is an exceptional performer when compared to others.</td>
<td>0</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
Appendix C

Managerial Success Factors Inventory: Transportation Version (MSFI: TV) – Developmental Scale Items

47 Items ran for EFA

71. Is able to understand how others feel about a decision or situation.
19. Responds well to criticism.
47. Makes ethical decisions when faced with conflicting choices.
7. Behaves in a way that builds trust.
39. Is honest and trustworthy.
9. Responds well to emotions of others.
31. Able to recognize his/her own emotional reaction to people, events, and situations.
95. Demonstrates awareness of how his/her own actions effect the emotions of others.
53. Provides feedback that is accurate and helpful to improving performance.
65. Provides opportunities for others to grow and contribute.
84. Identifies high potential candidates.
10. Recognizes where others are in the professional development process.
11. Coaches employees to help them achieve success.
38. Is clear when delegating responsibilities.
28. Knows when to use informal versus formal corrective actions.
77. Knows when to become involved in a conflict and when not to.
14. Communicates goals or a vision that motivates others.
56. Recognizes people for good performance.
37. Is able to communicate a clear vision for the team.
35. Inspires employees to achieve exceptional results.
36. Conveys a positive outlook.
33. Encourages innovation & risk taking.
60. Strives for excellence by setting challenging goals.
87. Consistently meets performance goals.
23. Delivers what is promised.
5. Anticipates what needs to be done and does it.
61. Takes action before being forced to do so by events.
24. Is able to clearly describe a problem.
66. Comes up with new ideas and solutions to recurring problems.
79. Is able to take in information from various sources and make correct decision (e.g., close a road).
55. Puts the safe performance of the job first.
96. Focuses on identifying the root cause of safety incidents and not on placing blame.
97. Holds meaningful and constructive safety meetings.
98. Is a model for the use of safe work practices.
99. Provides necessary training and/or tools to safely perform the job.
21. Is able to rally support for important initiatives.
91. Acts as a catalyst to a committee’s decision making process.
57. Is able to persuade others to adopt new points of view.
6. Asks for feedback.
69. Negotiates effectively with other units in the organization to get things done.
76. Listens to peers and subordinates ideas to improve overall processes.
90. Collaborates with diverse co-workers.
104. Appreciates different ideas and perspectives.
93. Is able to use conflict in groups effectively.
17. Connects with people at all levels of the organization.

**34 Items Retained after EFA**

71. Is able to understand how others feel about a decision or situation.
47. Makes ethical decisions when faced with conflicting choices.
39. Is honest and trustworthy.
9. Responds well to emotions of others.
31. Able to recognize his/her own emotional reaction to people, events, and situations.
95. Demonstrates awareness of how his/her own actions effect the emotions of others.
53. Provides feedback that is accurate and helpful to improving performance.
84. Identifies high potential candidates.
10. Recognizes where others are in the professional development process.
11. Coaches employees to help them achieve success.
38. Is clear when delegating responsibilities.
28. Knows when to use informal versus formal corrective actions.
77. Knows when to become involved in a conflict and when not to.
14. Communicates goals or a vision that motivates others.
23. Delivers what is promised.
5. Anticipates what needs to be done and does it.
61. Takes action before being forced to do so by events.
66. Comes up with new ideas and solutions to recurring problems.
79. Is able to take in information from various sources and make correct decision (e.g., close a road).
55. Puts the safe performance of the job first.
96. Focuses on identifying the root cause of safety incidents and not on placing blame.
97. Holds meaningful and constructive safety meetings.
98. Is a model for the use of safe work practices.
99. Provides necessary training and/or tools to safely perform the job.
21. Is able to rally support for important initiatives.
91. Acts as a catalyst to a committee’s decision making process.
57. Is able to persuade others to adopt new points of view.
6. Asks for feedback.
69. Negotiates effectively with other units in the organization to get things done.
104. Appreciates different ideas and perspectives.
93. Is able to use conflict in groups effectively.
17. Connects with people at all levels of the organization.

SAME 34 items retained after EFA just numbers removed and replaced with 1-34

1. Is able to understand how others feel about a decision or situation.
3. Is self-confident.
4. Makes ethical decisions when faced with conflicting choices.
5. Is honest and trustworthy.
6. Responds well to emotions of others.
7. Able to recognize his/her own emotional reaction to people, events, and situations.
8. Demonstrates awareness of how his/her own actions effect the emotions of others.
9. Provides feedback that is accurate and helpful to improving performance.
10. Identifies high potential candidates.
11. Recognizes where others are in the professional development process.
12. Coaches employees to help them achieve success.
13. Is clear when delegating responsibilities.
14. Knows when to use informal versus formal corrective actions.
15. Knows when to become involved in a conflict and when not to.
16. Communicates goals or a vision that motivates others.
17. Delivers what is promised.
18. Anticipates what needs to be done and does it.
19. Takes action before being forced to do so by events.
20. Comes up with new ideas and solutions to recurring problems.
21. Is able to take in information from various sources and make correct decision (e.g., close a road).
22. Puts the safe performance of the job first.
23. Focuses on identifying the root cause of safety incidents and not on placing blame.
24. Holds meaningful and constructive safety meetings.
25. Is a model for the use of safe work practices.
26. Provides necessary training and/or tools to safely perform the job.
27. Is able to rally support for important initiatives.
28. Acts as a catalyst to a committee’s decision making process.
29. Is able to persuade others to adopt new points of view.
30. Asks for feedback.
31. Negotiates effectively with other units in the organization to get things done.
32. Appreciates different ideas and perspectives.
33. Is able to use conflict in groups effectively.
34. Connects with people at all levels of the organization.