ABSTRACT

ROSENBERG, DANIEL. Assessing Measurement Equivalence of the KEYS® Climate for Creativity Scale Across Managerial Levels. (Under the direction of S. Bartholomew Craig.)

Employee creativity has been receiving increasing attention from organizations that wish to differentiate themselves in today's competitive global marketplace (Cummings & Oldham, 1997). An important element that can serve to either enable or hinder employee creativity is employees' perceptions of how conducive their work environment is to being creative, or what has been termed the organization's climate for creativity (Amabile, Conti, Coon, Lazenby, & Herron, 1996). Because organization members' perceptions of the work environment serve as the basis for the climate for creativity, perceptual differences can have important implications for organizations that are striving to enable employee creativity. Recent research by Kwaśniewska and Nęcka (2004) found that employee perceptions of climate for creativity, as measured by the Barriers for Creativity in the Workplace Questionnaire (BCWQ), were significantly affected by managerial status. However, the BCWQ was never confirmed to display measurement equivalence, and therefore the authors' findings may not be interpretable as the group differences could be the result of measurement artifacts (Horn & McArdle, 1992; Reise, Widaman, & Pugh, 1993; Vandenberg, 2002). The purpose of this study was to investigate whether the climate for creativity, as measured by the KEYS®: Assessing the Climate for Creativity Scale (Amabile, Conti, Coon, Lazenby, & Herron, 1996), displayed measurement equivalence across three distinct managerial levels including supervisors (N = 2,100), middle managers (N = 15,829), and executives (N = 2, 960). Both confirmatory factor analyses (CFA), and the differential functioning of items and tests (DFIT; Raju, van der Linden, & Fleer, 1995), which is based on item response theory (IRT), were used to assess measurement equivalence in this study. Using the eight factor structure proposed by Rosenberg and Craig (2006), both the CFA and IRT analyses found that the KEYS® scale displayed measurement equivalence across all managerial levels. Specifically, the CFA analyses found that the full 78 item KEYS scale displayed configural, metric, and scalar equivalence across all comparison groups. Additionally, two DFIT indices were used to determine that there was no differential functioning found at either the item (NCDIF) or the test (DTF) level when using the full KEYS scale. Results are discussed in terms of implications for practitioners and researchers as well as directions for future research.

Assessing Measurement Equivalence of the KEYS® Climate for Creativity Scale Across Managerial Levels

by

Daniel Rosenberg

A thesis submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the Degree of
Master of Science

PSYCHOLOGY

Raleigh, North Carolina

February, 2007

APPROVED BY:

Lori Foster Thompson

Samuel B. Pond

S. Bartholomew Craig
Chair of Advisory Committee

BIOGRAPHY

Daniel Rosenberg was born July 8, 1974 in Howell, NJ. Upon graduating Howell High School in 1992, he completed one year of study at Pennsylvania State University and in 1996 graduated Magna Cum Laude from the University of Delaware with a Bachelor of Arts degree in Psychology. Before joining the Industrial Organizational Program at NC State, Daniel served in various HR related positions at The Colgate-Palmolive Company in New York, NY and as an HR Generalist at UNC-Chapel Hill. His work experience spans the areas of benefits, compensation, training, recruitment, position management, and employee relations. Additionally, Daniel has been involved in various systems development initiatives and has been certified as a Senior Professional in Human Resources (SPHR). While a student at NC State, Daniel's research interests included creativity and leadership in the workplace. He assisted in researching the effectiveness of Industry/University Cooperative Research Directors for a National Science Foundation grant project and presented preliminary results of this study at a symposium held during the 21st annual meeting of the Society for Industrial Organizational Psychology. Additionally, Daniel presented a poster presentation on the Factor Structure of the KEYS® Climate for Creativity Scale during the same conference.

ACKNOWLEDGEMENTS

I would like to thank the Center for Creative Leadership for providing the data used in this study. I would also like to thank my committee chair, Dr. Bart Craig, for all of his encouragement, time, efforts, and advice regarding the thesis and beyond. Additionally, I would like to thank my committee members, Drs. Bob Pond and Lori Foster Thompson for all of their useful feedback and input. Finally, I would like to thank my wife, Sharyn, for her extraordinary patience, love, and support. Your unwavering belief in me pulled me through the toughest parts of this program.

TABLE OF CONTENTS

List of Tables

Page

Section I: Literature Review

The Climate for Creativity

In the context of work, creativity has been defined as the production of novel and useful ideas produced by individuals, or groups of individuals, towards a work related goal or problem (Amabile, 1983; Shalley, 1991; Shalley, Gilson, & Blum, 2000). This construct has been recognized as a crucial factor in the success of any organization (Amabile, 1996; Amabile, Conti, Coon, Lazenby, & Herron, 1996; Oldham & Cummings, 1996; Zhou, 1998). In today's competitive global environment, creativity can serve to set any organization apart from its peers through the creation of novel products, streamlined workflows, or innovative ideas (Cummings & Oldham, 1997; Shalley, Zhou, & Oldham, 2004; Sternberg & Lubart, 1996; Zhou, 1998). As such, it is in organizations' best interest to strive to maximize their employees' creative abilities as they appear to be a necessary commodity for survival.

One important element that can serve to either enable or hinder employee creativity has been termed the organization's climate for creativity (Amabile et al., 1996). At the individual level of analysis the climate for creativity can be broadly defined as a psychological climate which takes into account organization members' perceptions of various contextual elements inherent in the work environment (Amabile et al., 1996). To the extent that there exists a shared meaning of these contextual elements across individuals, the scores associated with each organization member's psychological climate may be aggregated to measure the organizational climate for creativity (James, Joyce, & Slocum, 1988). Therefore, at the organizational level of analysis, the climate for creativity may be defined as the "…climate that promotes generation, consideration, and use of new products, services and ways of functioning" (Kwaśniewska & Nęcka, 2004, p. 188).

Because organization members' perceptions of the work environment serve as the basis for the climate for creativity, perceptual differences can have important implications for organizations that are striving to enable employee creativity. For instance, Ekvall (1989) found that employees in innovative organizations maintained significantly different perceptions of various contextual factors in the work environment when compared to employees in less innovative organizations (as cited in Ekvall, 1996). Similarly, Amabile et al. (1996) found that employees working on highly creative projects had significantly different perceptions of various aspects of their work environment than employees working on projects that did not involve as much creativity. As these studies illustrate, developing a valid and reliable measure of the climate for creativity may serve as an important indicator of whether or not an organization is truly functioning in a creative manner.

Measures of the Climate for Creativity

A myriad of instruments have been designed to assess the climate for creativity. Included in this list of instruments are the Creative Climate Questionnaire (CCQ; Ekvall, 1996) and its updated version, the Situational Outlook Questionnaire (SOQ; Isaksen, Lauer, Ekvall, & Britz, 2001), the KEYS®: Assessing the Climate for Creativity Scale (Amabile et al., 1996), the Team Climate Inventory (TCI; Anderson & West, 1996), and the Climate for Creative Productivity Index (CCPI; Witt & Beorkrem, 1989). More recent scales that have been developed include the Virtual Team Climate for Creativity (VTCC; Nemiro, 2004) and an unpublished scale named the Barriers for Creativity in the Workplace Questionnaire

(BCWQ; Kwaśniewska & Nęcka, 2004). The aforementioned instruments each propose different taxonomies of contextual stimulants and/or obstacles essential for building the climate for creativity. However, there is a great deal of overlap regarding the respective contextual factors assessed by each of these scales. As the KEYS® scale has been cited as being among the most useful for assessing these contextual factors (Mathisen & Einarsen, 2004), its dimensions will be used as the basis for discussion throughout this study.

The KEYS scale is comprised of 78 items which emerged from prior research and a critical-incidents investigation of high and low creativity events reported by 129 research and development scientists (Amabile, 1996; Amabile & Gryskiewicz, 1987; Amabile et al., 1996). Through conceptual grouping and a principal components analysis of 66 of the items, eight environmental dimensions were identified in the KEYS structure. These scales are comprised of six stimulant scales whose dimensions positively relate to creativity, and two obstacles scales whose dimensions negatively relate to creativity. The remaining 12 items were used to develop two criterion scales, whose dimensions assess broader work outcomes (Amabile, 1996; Amabile et al., 1996; Mathisen & Einarsen, 2004). A description of these scales and their corresponding dimensions can be found in Table 1.

Psychometric Properties of the KEYS Scale

Factor structure. Amabile et al. (1996) performed a confirmatory factor analysis to assess the eight factor model intended for the environment scales. Each of the 66 items in the environmental scales was allowed to load on only one factor. Specifically, the intended factors were labeled as Organizational Encouragement (15 items), Supervisory Encouragement (11 items), Workgroup Supports (8 items), Sufficient Resources (6 items), Challenging Work (5 items), Freedom (4 items), Organizational Impediments (12 items), and Workload Pressure (5 items). The analysis of fit measures revealed a “moderate fit to the data” (goodness-of-fit index = .85; adjusted-goodness-of-fit-index = .84; chi-square (2,051) = 17,305.48, $p < .001$; root mean square residual = .056” (Amabile et al., 1996, p. 1167). In their results, Amabile et al., (1996) reported that component fit measures showed that all items loaded onto their corresponding scales, but many items were found to load on more than one scale. Although marginally supportive of the proposed structure, these results are difficult to fully interpret because no factor correlation matrix was reported. Further, as noted by Mathisen and Einarsen (2004), the lack of reported results from an exploratory factor analysis makes it “difficult to evaluate the different factors in relation to the underlying theory” (p.128).

Further investigation into the factor structure of KEYS®. In response to these comments, a further investigation of the factor structure of the KEYS scale was performed by Rosenberg and Craig (2006). Using an archival dataset containing results collected using the KEYS scale from U.S. employees, Rosenberg and Craig (2006) performed an exploratory factor analysis (EFA) on all 78 items, which resulted in an eight factor solution. The eight factors were interpreted as follows: Creative and Challenging Work (12 items), Organizational Impediments (15 items), Supervisory Encouragement (11 items), Workgroup Supports (8 items), Organizational Encouragement (14 items), Sufficient Resources (7 items), Workload Pressure (4 items), and Productivity (7 items). A confirmatory factor analysis using all 78 items as indicators of their respective factors resulted in poor model fit. However, lack of fit can be due, in part, to the large number of items and factors being analyzed which can increase model complexity and prohibit good fit (Hatcher, 1994). Therefore, an additional confirmatory factor analysis was conducted using only the five highest loading items on each factor. As reported by Rosenberg and Craig (2006), the CFA conducted on the reduced 39 item scale resulted in adequate model fit. Fit indices for the full and reduced models are presented in Table 2.

Differences in factor structures. Comparing the two factor structures, seven of the eight resulting factors proposed by Rosenberg and Craig (2006) were similar to factors proposed in the original KEYS® scale. However, the separate factors of Creativity and Challenging Work proposed by Amabile et al. (1996) combined to create the Creative and Challenging Work factor in the Rosenberg and Craig study. Additionally, Rosenberg and Craig found no support for the factor of Freedom. For further discussion of the interpretation and implications of these differences see Rosenberg and Craig (2006).

Comparing the Climate for Creativity Across Organizational Levels

The analyses performed by Rosenberg and Craig (2006) provided important information about the factor structure of the KEYS scale. However, as the climate for creativity arises from individual psychological interpretations, an important question to consider was posed by Kwaśniewska and Nęcka (2004); this question is “whether organizational climates that foster creativity affect all employees to the same extent” (p. 188). To investigate this question Kwaśniewska and Nęcka used the Barriers for Creativity in the Workplace Questionnaire (BCWQ) to discover whether employees’ gender and/or position, as measured by manager/nonmanager status, would affect their rating of the climate for creativity. The results of their study indicated that managers perceived the organizational climate as "more supportive and less inhibiting" than nonmanagers (Kwaśniewska & Nęcka, 2004, p. 191). Further, managers' perceptions also significantly differed from nonmanagers in that they perceived there to be better communication, fewer complaints about stringent control, and more adequate resources.

While these findings are intriguing, it should be noted that there were several limitations in the Kwaśniewska and Nęcka (2004) study. One limitation was that those authors did not test for conceptual differences across managerial levels that may be important when measuring the climate for creativity. In essence, Kwaśniewska and Nęcka assumed that all managers maintain a common underlying conceptualization of the climate for creativity. Stated differently, Kwaśniewska and Nęcka assumed that the KEYS® scale displayed measurement equivalence across different managerial levels but never investigated the validity of this assumption.

Measurement Equivalence

Measurement equivalence, sometimes referred to as measurement invariance, occurs when "…individuals with equal standing on the trait measured by the test but sampled from different subpopulations have equal expected observed test scores" (Drasgow, 1987, p. 19). Stated differently, tests of measurement equivalence assess whether the groups being assessed have the same underlying cognitive representation of the latent construct being measured, thus allowing for scores to be meaningfully compared across the groups. Additionally, measurement equivalence tests whether the items used to obtain observed scores maintain the same relationship with the latent construct across groups. For measures where these two assumptions hold, then observed scores may be used to compare groups.

However, if group scores are obtained from a measure that is nonequivalent with regard to the groups in question, then these scores cannot be compared as they may be a function of differential cognitive representations of the latent construct or differential measurement scales underlying the relationship between any or all of the scale items and the latent construct (Bollen, 1989; Drasgow & Kanfer, 1985; Vandenberg, 2002). It follows then that comparisons based on scales that lack measurement equivalence are not interpretable as differences between observed scores may be due to nothing more than measurement artifacts (Horn & McArdle, 1992; Reise, Widaman, & Pugh, 1993; Vandenberg, 2002). This last point is troubling when one considers that most researchers assume that measurement equivalence exists and never directly test for it (Vandenberg, 2002; Vandenberg & Lance, 2000) thereby calling into question many previously reported results based on comparisons between observed scores across groups or individuals.

Despite the pervasive assumption of equivalence, however, research has not been void of investigations of the topic. For instance, Facteau and Craig (2001) found that a multisource appraisal form (MAF), which is a measure of job performance from different levels in the organization, displayed measurement equivalence across self, peer, supervisor, and subordinate rating sources. Similarly, Maurer, Raju, and Collins (1998) found equivalence across peer and subordinate rating sources when using an instrument to assess how skilled a manager is at team building. Additionally, Byrne (1991) found that the Maslach Burnout Inventory (MBI), a measure of occupational burnout, functioned differentially for intermediate educators when compared with secondary and university educators. These are just a few of the studies chosen to illustrate the recent interest in investigating measurement equivalence. Interested readers should refer to Vandenberg and Lance (2000) for a comprehensive list of studies involving investigations of equivalence across cultures, time, gender, age groups, rater groups, race and organizational levels.

Assessing Measurement Equivalence

Confirmatory factor analysis. While there are several methods for assessing measurement equivalence, one of the more common methods is confirmatory factor analysis (Meade & Lautenschlager, 2004). Vandenberg and Lance (2000) attempted to standardize the procedures and the nomenclature used to assess this issue. Their procedures follow the steps associated with Jöreskog (1971). Ultimately the recommended process for studying measurement equivalence is a sequence of steps in which one specifies successively more restrictive constraints in a series of nested model tests. These unique constraints correspond with setting specific parameters of the model to be equal across the groups being studied. Each nested model is used to test a different constraint and measurement equivalence is assessed by simultaneously analyzing the fit of the observed data to the hypothesized models. These steps begin with identifying a baseline model that adequately fits the data. Once a baseline model is established, Vandenberg and Lance (2000) recommended that researchers test for equivalence through the following sequential series of tests whose names are given, where applicable, in parentheses: equivalent covariance matrixes across groups (omnibus test), equivalent pattern of factor loadings across groups (configural equivalence), equivalent matrix of factor loadings across groups (metric equivalence), equivalent vector of item intercepts across groups (scalar equivalence), equivalent unique variances across groups, equivalent factor covariances across groups, and equivalent latent factor means across groups. It should be noted that some have argued that not all of these tests are necessary to conduct in every study (Cheung & Rensvold, 1999; Facteau & Craig, 2001; MacCallum & Tucker, 1991; Raju, Laffitte, & Byrne, 2002; Reise et al., 1993; Vandenberg & Lance, 2000).

Item response theory. Another method of analysis which can be used to investigate the issue of measurement equivalence is item response theory (IRT; Lord, 1952; Lord, 1980). In terms of measurement equivalence investigations, IRT is similar to CFA in that both procedures test whether a scale, or its associated items, functions differentially across two or more groups. However, due to some differences in approach, IRT is able to provide unique information that can prove to be useful within the context of tests for measurement equivalence (Maurer et al., 1998; Meade & Lautenschlager, 2004). While a full discussion of the similarities and differences between these two procedures is beyond the scope of this paper, interested readers can consult Raju, Laffitte, and Byrne (2002) for a more thorough discussion of the topic.

Briefly, IRT uses a nonlinear monotonic function to depict the relationship between a latent ability estimate, labeled theta (θ), and the probability of endorsing a specific item response (Hambleton, Swaminathan, & Rogers, 1991; Maurer et al., 1998). This relationship is graphically represented through an item characteristic curve (ICC), which is sometimes referred to as an item response function (IRF). Each ICC is specified by item parameters which can include one or more item difficulty parameters (*b*), an item discrimination parameter (*a*), and a pseudo-chance-level (i.e., guessing) item parameter (*c*). The item difficulty parameter (*b*) is used to locate the ICC on the ability scale at the point for which there is a 50% probability of responding to the item in a given direction, such as getting the item correct (Hambleton et al., 1991). The item discrimination parameter (*a*) is used to signify the rate at which the probability of endorsing an item increases with respect to increasing ability levels and is essentially the slope of the ICC around the item difficulty parameter (Hambleton et al., 1991). Finally, the guessing parameter (*c*) is used to represent the probability associated with an examinee of very low theta level endorsing the item (Hambleton et al., 1991).

For dichotomously scored items, IRT models that make use of only the *b* parameter to analyze data are referred to as one parameter logistic models (1PL), or Rasch models (Embretson & Reise, 2000; Hambleton et al., 1991). IRT models that make use of the *a* and *b* parameters are referred to as the 2PL models, while models that make use of all three parameters are referred to as 3PL models. For polytomous data, the graded response model (GRM; Samejima, 1969; Samejima, 1996) is an extension of the 2PL model that generates a separate operating characteristic curve (OCC) for each of the multiple ordered response categories associated with each item. Each OCC is specified by parameters that are similar to the parameters of the ICC.

Differential item or test functioning. Within the IRT framework, measurement equivalence is defined as identical (or nearly identical) item parameters across groups. Different item parameter estimates for each group result in a different shape and/or location for each group's ICC in relation to their corresponding ability estimates. Significant differences found between item parameters, or between the areas separating the ICCs, across groups indicates differential item functioning (DIF), which is the moniker for measurement inequivalence in item response theory (Embretson & Reise, 2000). It should be noted however that including the *c* parameter can lead to estimation problems unless sample size is extremely large and therefore this parameter is often set to be equal for all items under consideration in DIF studies (Embretson & Reise, 2000).

Raju, van der Linden, and Fleer (1995) developed a framework based on IRT to identify "differential functioning of items and tests" (DFIT). Essentially, DFIT uses the ICCs or OCCs estimated separately for each group to generate expected item responses for given levels of theta (called "true scores" in the DFIT framework). If significant differences between the true score estimates for each group exist, then the item or scale is said to function differentially (Flowers, Oshima, & Raju, 1999).

Investigating Measurement Equivalence Across Managerial Levels

As no studies to date have taken the aforementioned steps to investigate whether the KEYS® scale displays measurement equivalence across managerial levels, it is unknown whether scores obtained from different managerial levels are comparable. This information is important because there have been several studies that have assumed that measurement equivalence exists across organizational levels when using measures of the climate for creativity. As mentioned previously, Kwaśniewska and Nęcka (2004) assumed the BCWQ displayed measurement equivalence across managerial levels in their study comparing perceptions of the climate for creativity across managers and nonmanagers. Similarly, Ekvall (1996) assumed that the CCQ displayed measurement equivalence across managerial levels when he compared a group of managers' perceptions of the contextual factors associated with "innovative" and "stagnated" organizations. Despite the noteworthy intent of both of these studies, without having verified that their instruments displayed measurement equivalence across the managerial levels being compared, the findings are difficult to interpret.

Discriminating Among Managerial Levels

The investigation into whether all managers possess a universal conceptualization of the climate for creativity is important because, in general, managers can have a large impact on this climate through such factors as their access to resources, skill level, autonomy, and knowledge of organizational strategy (Amabile & Gryskiewicz, 1987; Ekvall, 1996; Kwaśniewska & Nęcka, 2004). Some of these same factors might also cause measures of the climate for creativity to function differentially across managerial levels. Therefore, consideration of the differences that separate managers into distinct levels is necessary to fully understand why measurement inequivalence might arise.

To date, research has supported the existence of three broadly defined, yet distinct, managerial levels (Freedman, 1998; Jacobs & Jaques, 1987; Kaiser & Craig, 2004; Katz & Kahn, 1966; Katz, 1955; Zaccaro, 2001). The first level, referred to here as the supervisory level, is comprised of managers who oversee the technical specialists (i.e., nonmanagers) within the workplace (Freedman, 1998; Kaiser & Craig, 2004). These supervisors are typically responsible for a subunit, which is a single functional area within the organization. The second level, referred to here as middle managers, tend to be employees who are responsible for "…several distinct but interdependent subunits" within the same area of business of the organization (Freedman, 1998, p. 142). These middle managers typically have supervisors, or other middle managers, as direct reports and may be responsible for areas for which they have little expertise (Freedman, 1998; Kaiser & Craig, 2004). Finally, at the highest managerial level are the executives who are responsible for managing various interdependent units of business covering different areas of the organization (Freedman, 1998). Included in this executive group are those positions that maintain the highest rank in organizations such as the CEO and, who may be held accountable to other organizational constituents, such as a board of directors or public stakeholders (Freedman, 1998; Kaiser & Craig, 2004). These executives may supervise a number of middle managers, or other executives, and may have little expertise in any areas of business for which they are responsible (Freedman, 1998; Kaiser & Craig, 2004). These three broad levels can be distinguished by differences in their functional activities, skill requirements, time horizons, and levels of abstraction in the work itself. Further, there is reason to believe that these interlevel differences might result in incumbents developing cognitive structures (e.g., implicit theories) that are specific to their own organizational levels and therefore render subjective ratings on instruments such as KEYS® to be inequivalent across levels.

Differences in functional activities. Each of these three managerial levels maintains a distinct array of functional activities which correspond with a requisite set of knowledge, skills, and abilities that managers must possess if they are to be successful at their corresponding levels (Freedman, 1998; Jacobs & Jaques, 1987; Katz & Kahn, 1966). To distinguish between these functional activities, one can turn to the Stratified System Theory (SST) developed by Jacobs and Jaques (1987). The SST incorporates a hierarchical framework of organizational domains in which each managerial level could be placed. The first domain, located at the bottom of the hierarchy, is labeled the production domain as this is the level of the organization where production of goods or services usually takes place

(Jacobs & Jaques, 1987). Located in the production domain are the supervisors whose functional activities include assigning and supervising tasks that are consistent with the goals of the organization, motivating subordinates, and monitoring subordinate performance (Jacobs & Jaques, 1987; Kaiser & Craig, 2004; Zaccaro, 2001). As these tasks are all performed with a predefined set of directives and an assumed high level of compliance, the production domain is associated with functional tasks that involve the use, or application of, existing structure (Kaiser & Craig, 2004; Katz & Kahn, 1966).

The next highest level listed in the SST is the organizational domain. This domain distinguishes itself from the production domain in that employees in the organizational domain are not directly involved with any work on the production lines (Jacobs & Jaques, 1987). Located in this domain are middle managers whose functional activities include ensuring that all goals are effectively transmitted from the top to the bottom of the organization, creating plans to meet these goals, integrating and coordinating the work of their subunits, managing any needed changes within their subunits, generating policy to impact the company climate, and allocating resources as they see fit (Jacobs & Jaques, 1987; Kaiser & Craig, 2004). As many of these tasks include the communication, interpretation, and implementation of goals, they can be categorized as ones involved with the interpretation of structure (Kaiser & Craig, 2004; Katz & Kahn, 1966).

The highest level of the hierarchy associated with the SST is the systems domain. This domain differentiates itself from the other two lower domains in that employees in the systems domain are not fully concerned with the present production or happenings within their organization. Rather, executives, who are located in the systems domain, are also concerned with analyzing how the organization is performing in relation to the external environment, assessing the opportunities for developing or acquiring new businesses, creating strategic goals for the organization, interacting with the environment to generate acceptance of the organization and its future goals, developing or shaping an appropriate organizational climate and culture, and creating and/or securing any resources needed by the organization to accomplish its goals (Jacobs & Jaques, 1987; Kaiser & Craig, 2004; Zaccaro, 2001). As these tasks are all associated with the development or modification of the composition, goals, or culture of an organization, they can be categorized as ones involved with the creation of structure (Kaiser & Craig, 2004; Katz & Kahn, 1966).

After reviewing the differences in job responsibilities it becomes clear that despite the common notion that supervisors, middle managers, and executives may all be considered managers, their specific functional activities differ significantly. Notably apparent is that as one moves up the hierarchy of domains, specific functional activities become more complex (Jacobs & Jaques, 1987; Zaccaro, 2001). These differences in functional activities between managerial levels have important implications when considering how supervisors, middle managers, and executives may differ in their conception of the climate for creativity. For instance, because middle managers are involved with generating or refining the policies that set the tone for encouragement across the organization, they may have a distinct conceptualization of what the Organizational Encouragement factor is intended to measure when compared to supervisors who may view Organizational Encouragement from the perspective of how it is affecting their specific subunit. At the executive level, knowledge of the external environment may create differential conceptions of factors such as Productivity when compared to supervisors and middle managers. Even if supervisors and middle managers believe that productivity is at an adequate level, differences in production levels between competing organizations or changes in supply and demand, may change an executive's notion of Productivity as a whole.

Differences in skills. Just as functional responsibilities differ across managerial levels, so do the requisite skill sets. Katz (1955) developed a typology that is useful for distinguishing among the different types of skills needed at each of the three managerial levels. According to Katz (1955), the skill set that is most important for supervisors can be categorized as technical skills. Technical skills are those that correspond with the methods or techniques necessary to perform one's job (Katz, 1955). These methods or techniques may take a myriad of forms including analyzing data, skillful use of tools, or the application of specialized knowledge (Katz, 1955; Mann, 1965). Additionally, Mann (1965) added that skills involved with motivating people, labeled human-relations skills, are also important at the supervisory level. That these skills should be prevalent at the supervisory level makes sense in that the supervisors are the ones who are directly in charge of those employees who are actually performing the work. Without knowledge of how the work is done, it would be difficult for these supervisors to assess their direct reports' performance.

Although technical skills may still be useful in the organizational domain, the skill set that is thought to be most important to middle managers is that of interpersonal skills (Kaiser & Craig, 2004; Katz, 1955). Interpersonal skills refer to the ability to develop relationships with others and work effectively within a group (Kaiser & Craig, 2004; Mann, 1965). It also refers to the ability to understand the strengths, abilities, weaknesses, and limitations of oneself and others (Katz, 1955). Aside from interpersonal skills, Mann (1965) adds that middle managers should possess administrative skill, or the ability to understand, think, and act from an organizational vantage point. Because middle managers are responsible for relaying information regarding organizational objectives and goals between the executives and supervisors it should follow that interpersonal and administrative skills are requisite for their position.

Lastly, conceptual skills are thought to be most important to executives (Kaiser & Craig, 2004; Katz, 1955). These conceptual skills encompass a range of abilities including the ability to think of how each part of the organization affects the whole, as well as the ability to conceive of the symbiotic relationship between the organization and its environment (Katz, 1955). Additionally, conceptual skills involve complex reasoning, and the ability to form cognitive representations of interacting factors that one is capable of projecting and adapting to changing situations (Kaiser & Craig, 2004; Katz, 1955). That executives maintain these skills makes intuitive sense in that executives are comparatively the most concerned with the external environment and with developing policy that affects the whole organization. It should be noted that Mann (1965) asserts that at least some minimum of administrative, and human-relations skills are still necessary at this level as executives are still concerned about the organization and may need to work to obtain internal consensus to newly developed goals and ideas (Jacobs & Jaques, 1987; Mann, 1965).

Reviewing these skills demonstrates that as one moves up the managerial hierarchy the skill set shifts from working with more concrete technical skills to utilizing more abstract conceptual skills. The differences among these skill sets are important to consider as they may lead to differences in the conceptualization of the climate for creativity. For instance, the interpersonal and administrative skills required of the middle managers may force them to perceive the challenge in their jobs differently than do supervisors, in that the challenge inherent in the middle managers' job may not be related to the assigned task, but, rather, may reside in their ability to pull people together to perform that task. Therefore it is the skill required for interacting with people that may be perceived as the challenge for middle managers. As such, items measuring the Creative and Challenging Work factor may correlate highly with other factors that tend to measure team performance, such as Work Group Supports, Organizational Impediments, and Productivity, at the middle management level. This may not be true at the supervisory level where Creative and Challenging work may relate more to a specific functional activity, or task, that requires independent performance. Similarly, because executives use conceptual skills and complex reasoning when approaching their work they may tend to group factors together based on what makes overall conceptual sense rather than basing their perceptions on the everyday functional activities within their organizations. If this conceptual grouping of factors does occur uniquely at the executive level, then comparisons across managerial levels that involve the executive group might not be meaningful.

Differences in time horizons and levels of abstraction. Along with differences in functional activities and skills, time horizons also provide a point of differentiation across managerial levels. Time horizons, which are sometimes labeled time spans, generally refer to the maximum amount of time it takes to receive feedback about the accomplishment of a task once that task has been initiated by a manager (Jacobs & Jaques, 1987; Kaiser & Craig, 2004). More specifically, Jaques (1976) conceived of these time horizons as the period over which one must exercise judgment for a given task. Conceptually, time horizons are arranged in a hierarchical fashion with longer horizons corresponding to higher levels of management in the organization.

At the supervisory level, there exist short time horizons which can last anywhere from three months to two years (Jaques, 1976). At the lower end of this range, task output can be conceived of in a concrete fashion, although imagination plays a large role in this endeavor (Jaques, 1976). For instance, supervisors working with a three month time horizon must use their imagination to conceive of how a training program will look once it has been developed (Jaques, 1976). This concrete conception then guides the supervisor in determining the program's development. Above the one year time horizon though, supervisors can no longer imagine the task output as a whole entity, but rather they must be able to scan over various portions of the output domain in pieces to obtain the necessary feedback for the task as a whole (Jaques, 1976). Supervisors in charge of a number of subordinates may use this type of imaginative scanning technique when trying to account for the output of all employees in their given subunit (Jaques, 1976). It is important to note that in each of the given time horizons, the supervisors are working from a concrete and previously established conception of what the final output is supposed to look like; as such they are able to measure progress by comparing the current status of the project to the concrete mental image of the final output (Jaques, 1976).

At the middle manager level, the time horizons are longer and the levels of abstraction more complex. For middle managers, time horizons are thought to last anywhere from two to five years (Jaques, 1976). In this horizon, middle managers are forced to break from the use of concrete conceptualization of the task and make use of more abstract ways of thinking. Typically this takes the form of working with an established end product, and then using abstract thought to conceive of and produce a new product which is unlike the established predecessor (Jaques, 1976). As there is no longer reliance on previously established models, innovation is thought to begin at this level and managers must use intuition to judge task progress (Jaques, 1976).

Finally, time horizons at the executive level are typically thought to range from 10 to 20 years (Jacobs & Jaques, 1987; Kaiser & Craig, 2004). It should be noted however, that some high level middle managers and executives may perform work that maintains a time horizon between five and ten years (Jaques, 1976). In the executive time horizon, intuitive theories are used to facilitate decision making that must take place with only minimal contact with a task (Jaques, 1976). Intuitive theories are formed from generalizations based on intuitions thought to be valid for similar situations, or tasks, in the past (Jaques, 1976). Because executives' time is consumed by tasks involving long range planning for their organizations, they may not be able to attend to the details of all the shorter range projects that might be occurring. Intuitive theories, therefore, are a form of abstract reasoning which enables the executives to make quick decisions without losing sight of the long range tasks at hand. As these theories begin to emerge in a consistent manner, they may eventually lead to the establishment of policy which would guide the work of the middle managers and supervisors (Jaques, 1976).

In terms of the climate for creativity, these differences in time horizons and level of abstraction might lead to differences in the conceptualization of a number of factors measured by the KEYS® scale. For instance, supervisors, who maintain short time horizons, may feel more of a sense of urgency as the deadlines for their work come at a quicker rate than those higher in the managerial hierarchy. This may lead to differences in how they conceive of items measuring sufficient time and time pressure in the Workload Pressure factor when compared to middle managers and executives. Similarly, the time required to process information through abstract thought may cause middle managers to conceive of the Productivity factor differently than supervisors do. Whereas supervisors may think of Productivity as tangible output produced over a short time period, middle managers may be more willing to consider time spent in thought as being productive. Finally, as executives make use of intuitive theories rather than relying on first hand observation, they may maintain differential interpretations of any of the factors in the KEYS scale compared to other managerial levels.

Resulting Differences in Cognitive Map Complexity and Implicit Theories

As a result of the differences that have been discussed to this point, it seems possible that managers at each level might possess their own unique implicit representations, or cognitive maps (Hunt, 1983; Sternberg, 2003), of their environments (Jacobs & Jaques, 1987; Zaccaro, 2001). As internal representations of the environment, cognitive maps enable one to understand the structural relationships among variables of interest to the organization, and create strategies to manipulate these relationships (Hunt, 1983). Therefore, from an organizational perspective, cognitive maps emerge from a manager's ability to use information to form implicit representations of the functioning of the organization. These cognitive maps can then be used to understand the relative importance of different factors within the organization, and how those factors are related to one another (Jaques, 1976). Additionally, these cognitive maps can serve as the basis for strategy formulation (Hunt, 1983; Jaques, 1976; Nadler & Tushman, 1983).

With the changing nature of skills, levels of abstraction, and time horizons across managerial levels, it follows that the implicit representations of a manager's environment becomes more complex as one ascends the hierarchy of managerial levels (Zaccaro, 2001). One reason why the complexity increases with each organizational level is because cognitive maps depend on the amount and type of information received from the environment; therefore, the greater the amount of input relative to one's position, the more complex the cognitive map will become (Jacobs & Jaques, 1987). In light of this, supervisors' cognitive maps might be less complex because they are not required to be as knowledgeable of the variables which affect the whole organization compared with middle managers and executives (Jacobs & Jaques, 1987). Further, because supervisors maintain the shortest time horizons and work with a concrete understanding of their tasks, they may have a firmer understanding of the variables inherent in their cognitive maps compared with higher level managers (Jacobs & Jaques, 1987).

Alternatively, because middle managers are more involved with the broader organizational domain, their cognitive maps might be expected to require more information to form an appropriate representation of the functioning of the organization compared to supervisors (Jacobs & Jaques, 1987). As there is more information needed at this level, there may be more periods of uncertainty in which middle managers must work with insufficient data (Jacobs & Jaques, 1987). This uncertainty may last until such a time as the necessary information is located. These periods of uncertainty lend credence to the idea that middle managers must be able to utilize abstract thought in work that takes place over longer periods of time, compared with supervisors.

Finally, cognitive maps at the executive level might well be more complex than those at any of the lower levels because executives are engaged in activities that provide external as well as internal input (Jacobs & Jaques, 1987). This external input may stem from a myriad of sources including multinational constituencies that contain foreign laws, unique political landscapes, and diverse cultural viewpoints that all need to be integrated into the executive's cognitive map of the organization (Zaccaro, 2001). In light of this, the executives' cognitive map must be able to consistently adapt to changes in the external environment (Zaccaro, 2001), whereas the cognitive maps of supervisors and middle managers may only require fundamental changes as a function of the executives' decisions. Additionally, as their position requires the longest period of time to receive task-related feedback, executives must utilize their cognitive maps to develop strategies that reach further into the future than any other level in the managerial hierarchy (Zaccaro, 2001).

If cognitive maps do vary in complexity across managerial levels, one might expect these differences to affect the underlying conceptualizations of the climate for creativity. For instance, supervisors may only be aware of those aspects of the organizational environment that affect their subunit. This could create differences in how supervisors conceive of the Organizational Encouragement and Organizational Impediments factors when compared with middle managers and executives who have a broader understanding of the variables inherent in the organization. Additionally, as executives take the future of the organization into consideration when utilizing their cognitive maps, they may approach many factors on the KEYS® scale using a completely different conceptualization of the organization compared to lower managerial levels. As such, cross-level comparisons on any of the dimensions on the KEYS scale would be suspect as ratings may have been obtained from different temporal orientations which could affect each level's underlying conceptualizations of the scale.

Implicit theories. It should be noted that, conceptually, cognitive maps seem similar to implicit theories which are commonly held beliefs about a construct of interest that do not have to be supported by scientific evidence (Sternberg, 1985). Research has supported the existence of implicit theories of creativity across different cultures (Niu & Sternberg, 2002; Sternberg, 1985), and has found that people actively use these theories when rating such factors as job performance (Borman, 1987) or another individual's level of creativity (Puccio & Chimento, 1987). From these results the argument could be made that the differences in functional activities, skill sets, time horizons, and levels of abstraction may result in distinct implicit theories about what the climate for creativity is across managerial levels which, in turn, can lead to a lack of measurement equivalence when using measures such as the KEYS instrument.

Section II: Present Study

Based on the differences in requisite functional activities, skill sets, time horizons, level of abstraction, and cognitive map complexity, it is evident that distinctions can be made across managerial levels. For reference purposes, these differences are summarized in Table

3. The primary question, however, is whether these differences result in each managerial level having a distinct underlying conceptualization of the climate for creativity. Therefore, the purpose of the present study is to examine whether the KEYS® scale displays measurement equivalence across supervisors, middle managers, and executives. This would be important information for any organization to possess if it is interested in investigating its climate for creativity. Additionally, a study of this nature will further enhance the validity of the KEYS instrument and, therefore, serve as a response to the recommendation by Mathisen and Einarsen (2004) that these types of studies take place. The research questions are as follows:

Measurement Equivalence Across Supervisors and Middle Managers

1) Will the KEYS instrument demonstrate configural measurement equivalence across supervisors and middle managers?
    a. If configural measurement equivalence is found, then will the KEYS instrument demonstrate metric measurement equivalence across supervisors and middle managers?
    b. Where metric measurement equivalence is found, will the KEYS instrument demonstrate scalar measurement equivalence across supervisors and middle managers?

Measurement Equivalence Across Supervisors and Executives

2) Will the KEYS instrument demonstrate configural measurement equivalence across supervisors and executives?

a. If configural measurement equivalence is found, then will the KEYS® instrument demonstrate metric measurement equivalence across supervisors and executives?

b. Where metric measurement equivalence is found, will the KEYS instrument demonstrate scalar measurement equivalence across supervisors and executives?

Measurement Equivalence Across Middle Managers and Executives

3) Will the KEYS instrument demonstrate configural measurement equivalence across middle managers and executives?

    a. If configural measurement equivalence is found, then will the KEYS instrument demonstrate metric measurement equivalence across middle managers and executives?

    b. Where metric measurement equivalence is found, will the KEYS instrument demonstrate scalar measurement equivalence across middle managers and executives?

Section III: Method

Participants

An archival dataset containing ratings collected using the KEYS scale (n = 74,415) was obtained for this study. Participants completed the scale either as part of an organizational assessment or as part of an individual leader development program contracted to an international nonprofit leadership development organization. As part of the background information section of KEYS, respondents were asked to endorse one of seven organizational levels including *Top*, *Executive*, *Upper Middle*, *Middle*, *First Level*, *Hourly Employees*, and *Not relevant in my situation*. Each of these response options was anchored by a set of generic job titles that served to exemplify the types of positions found at that level. The dataset was reduced to include only employees from the United States who identified themselves as holding a managerial position at any organizational level (*n* = 20,889). Therefore, all employees who endorsed the "Hourly Employees" or "Not relevant in my situation" options were excluded from the analyses. Additionally, due to algorithm convergence issues associated with missing item values, analyses were only performed on employees with complete data. The employees in this dataset (12,604 male, 8,081 female, and 204 unidentified) were grouped according to the three established managerial levels previously discussed with supervisors (*n* = 2,100) represented by those participants who indicated they were forepersons, crew chiefs, or section supervisors. Middle managers (*n* = 15,829) were represented by those participants who indicated they were either middle office managers, professional staff, mid-level administrators, department executives, plant managers, or senior professional staff. Executives (*n* = 2,960) were represented by those participants who indicated they were executive vice presidents, directors, board-level professionals, chief executives, operating officers, or presidents. Participants also represented a range of functions including: administration/management (36.0%), education/training/human resources (13.1%), engineering/product development (8.6%), law (1.8%), manufacturing (3.2%), marketing/advertising (4.8%), medicine (0.4%), research and development (6.3%), sales (8.6%), support services (4.8%), and unidentified employees (12.1%). Approximately

26% of the employees had 5 years of service or fewer and about 27% had over 20 years of service.

Analyses

Both CFA and IRT were used to investigate the measurement equivalence of the KEYS® scale across managerial levels. The reason why both methods were used in this analysis is that each method can provide unique information when investigating measurement equivalence (Maurer et al., 1998; Meade & Lautenschlager, 2004). For instance, as outlined in Raju et al. (2002), CFA can be used to assess measurement equivalence across groups for all factors inherent in a scale simultaneously. Conversely, one is required to analyze one factor at a time when performing measurement equivalence tests using a unidimensional model of IRT. However, IRT provides more item-specific information than CFA, such as the probability of endorsing, or not endorsing, any specific response category when using polytomous data. As the KEYS scale uses polytomous items, the results from both the CFA and IRT analyses were useful in the determination of measurement equivalence. Although a full discussion of the similarities and differences between these two procedures is beyond the scope of this paper, the interested reader is referred to Raju et al. (2002) for a more thorough discussion of the topic.

Confirmatory factor analysis. The procedure proposed by Vandenberg and Lance (2000) to investigate measurement equivalence was used for conducting the CFAs in this study. All CFAs were performed using the Mplus computer program version 3.13 (Muthén & Muthén, 1998-2005). As mentioned previously, the recommended steps for investigating measurement equivalence in a CFA framework began with identifying a baseline model that adequately fit each group being compared. As the dataset used by Rosenberg and Craig (2006) was a subset of the dataset used in this study, the eight factor model specified by their initial EFA of the 78 item KEYS® scale was analyzed for fit as the initial baseline for each managerial level. Model fit was assessed by analyzing several fit indices. These fit indices included the chi-square test, which is a test of the null hypothesis that the hypothesized model perfectly fits the model implied by the observed data (Bollen, 1989). As such, when attempting to establish a baseline model one is looking to accept the null hypothesis and therefore is seeking a nonsignificant chi-square value (Bollen, 1989). In many cases though, this nonsignificant value is hard to obtain because the chi-square statistic has been shown to be very sensitive to large sample sizes (Hatcher, 1994).

In light of the sensitive nature of the chi-square statistic to sample size, other fit indices provided by Mplus (Muthén & Muthén, 1998-2005), such as the comparative fit index (CFI; Bentler, 1990), Tucker-Lewis index (TLI; Tucker & Lewis, 1973), standardized root mean square residual (SRMR; Jöreskog & Sörbom, 1996), and root mean square error of approximation (RMSEA; Browne & Cudeck, 1992; Steiger & Lind, 1980) were used to analyze model fit throughout this study. These indices assess model fit in different ways. For instance, the SRMR is an absolute fit index that summarizes the average covariance residuals resulting from the differences between the observed and model implied covariances (Jöreskog & Sörbom, 1996; Kline, 1998). This model is termed absolute because it is only comparing the hypothesized model to the model implied by the observed data (Kline, 1998). The RMSEA is another absolute fit index in that it indicates how well the hypothesized model fits the model implied by the data per degree of freedom (MacCallum, Browne, &

Sugawara, 1996). In this way, the RMSEA will indicate a better overall fit for a model which contains fewer parameters than for one that contains more parameters even if both models fit the data equally well (Facteau & Craig, 2001; MacCallum et al., 1996). As opposed to the absolute fit indices, incremental fit indices assess the overall fit of the hypothesized model compared to that of the null model in which all variables are uncorrelated and simultaneously assess the fit of the hypothesized model to the observed data (Browne, MacCallum, Kim, Andersen, & Glaser, 2002; Kline, 1998). The CFI and TLI are incremental fit indices which include corrections for sample size and model complexity respectively (Kline, 1998).

Hu and Bentler (1999) proposed empirically based cutoff criteria for each of these indices. For the CFI or TLI, values close to .95 indicate a good model fit, while for the RMSEA and SRMR, cutoff values close to .06 and .08 respectively indicate adequate model fit. To decrease the chance of Type I and Type II errors, the authors also ultimately recommended using combinations of these fit indices which are as follows: a TLI (or CFI) index greater than .95 with an SRMR index less than .06; or a RMSEA index less than .06 with a SRMR index less than .09. These cutoff values and combinational rules were all applied in the current study to assess whether a baseline model was established.

Once a baseline model was established for each managerial level, tests for measurement equivalence were conducted by simultaneously analyzing model fit in a series of tests where specific parameters are constrained to be equal across groups being compared. The first test recommended by Vandenberg and Lance (2000) is the omnibus test which investigates whether the covariance matrices (Σ) are equivalent across groups. Specifically, this is a test of the null hypothesis that H_o: $\Sigma_1 = \Sigma_2 = \ldots \Sigma_G$, (where G is the number of groups). In this test a model in which all parameters are constrained to be equal across groups is compared to one in which all parameters are free to vary. If one is unable to reject the null hypothesis then the groups in question can be considered equivalent except for their factor mean structures (Byrne, Shavelson, & Muthén, 1989). In the event that the null hypothesis is rejected, one may then move onto the next in the sequence of nested model equivalence tests in order to find the source of inequivalence. However, Cheung and Rensvold (1999) have argued that one would not expect to find support for the notion that all parameters are equal across groups unless random samples were taken from the same population. In light of the fact that the null hypothesis is expected to be rejected on most occasions, this test has been omitted by most researchers (Cheung & Rensvold, 1999; Raju et al., 2002) and was not conducted in this study.

As the omnibus test was not performed, the test for configural equivalence was conducted which involved simultaneously constraining the pattern of factor loadings to be equivalent across the groups being compared (Horn & McArdle, 1992; Vandenberg & Lance, 2000). This constraint ensured that the underlying conceptualizations of the construct were equivalent across comparison groups. Similar to testing the fit of the baseline models for each specific managerial level, the overall fit of this model was assessed by analyzing the value of the fit indices according to Hu and Bentler's (1999) recommendations mentioned previously.

Once configural equivalence was established, then the test for metric measurement equivalence was conducted by constraining the matrix of factor loading magnitudes (Λ) to be equivalent across groups (Vandenberg & Lance, 2000). Essentially, this model tested whether the relationships between the observed variables and the latent constructs were similar across groups. Conceptually, this test assessed whether the slope of the regression lines relating the observed variables to the latent constructs was the same across groups (Bollen, 1989). As a nested model, this test for metric measurement equivalence was even more restrictive than the previous test for configural equivalence because the data across all groups under study were forced to maintain a high degree of consistency (Horn & McArdle, 1992). The overall fit of this model was assessed relative to the model which implied configural equivalence through a chi-square difference test. Significant values on this test would indicate that the model which implied metric measurement equivalence fit the observed data significantly worse than the previous, less constrained, model.

Once metric measurement equivalence was established then the test for scalar equivalence was conducted by investigating whether the vector of item intercepts (τ) was equivalent across groups (Vandenberg & Lance, 2000). This was a test of the null hypothesis H_o: $\tau_1 = \tau_2 = \ldots \tau_G$. Essentially, this test examined whether one group tended to consistently give higher ratings on the construct of interest for a particular item than another group despite having equivalent slopes (Λ) (Bollen, 1989); as such this test may be interpreted as a test for systematic response bias or leniency, rather than a test of measurement equivalence *per se* (Bollen, 1989; Vandenberg & Lance, 2000). To test for scalar measurement invariance the item intercepts were constrained to be equivalent across groups, and the chi-square difference test was conducted to assess whether the model which implied scalar measurement equivalence fit the data worse than the previous, less constrained, model.

It should be noted that there are three other tests which could be performed when examining equivalence across groups, but that were not used in the current study. The first two tests would have assessed whether the factor variances and covariances were equivalent across groups ($\varphi_1 = \varphi_G$), and whether the unique (error) variances were equivalent across groups ($\theta_1 = \theta_G$). However, there has been contention as to whether testing the equivalence of covariance matrices across groups is a necessary step as factor variances and covariances are expected to differ across samples taken from different populations (MacCallum & Tucker, 1991). In addition to these parameters, unique variances are also said to be sample specific (Reise et al., 1993). Because these data represented responses taken from different populations, these additional equivalence tests were not performed in this study. Similarly, the final test of equivalent latent factor means across groups ($\kappa_1 = \kappa_G$), was not conducted as this test does not assess the underlying properties of the scale (Vandenberg & Lance, 2000) which was the main focus of this study.

Item response theory. As mentioned, item response theory (IRT; Lord, 1952; Lord, 1980) can also be used to obtain additional information when conducting studies of measurement equivalence. It should be noted that certain common assumptions about the data need to be confirmed before conducting IRT analyses. The first assumption, termed local independence, is achieved when the probability of endorsing an item is independent of responses to all other items on the test when item and person parameters are held constant (Embretson & Reise, 2000). Inherent in the local independence assumption is the concept of unidimensionality, which specifies that the items being analyzed all correspond to a single latent construct (Embretson & Reise, 2000). Finally, a second common assumption is that the form of the relationship between the latent ability estimates and the probability of endorsing an item is depicted by a nonlinear monotonic relationship (Embretson & Reise, 2000). Each of these assumptions was tested before analyses began.

To test for unidimensionality, an exploratory factor analysis, using an oblique rotation, was performed on each of the eight subscales of the KEYS® instrument for each managerial level. This yielded a total of 24 factor analyses. Scree plots were analyzed for one clearly dominant factor as well as the amount of variance accounted for by the first factor in each scale. As per Reckase (1979), the assumption of unidimensionality was considered to have been met if the first factor accounted for 20% or more of the common variance on each scale. As reported in Table 4, each of the analyzed scales met the requirements for unidimensionality posed by Reckase (1979). Additionally, all of the scree plots indicated the presence of a clearly dominant factor on all scales for the responding supervisors, middle managers, and executives.

Because the response scale from the KEYS instrument is polytomous in nature, the DFIT framework proposed by Raju et al. (1995) was used to assess measurement equivalence in this study. The DFIT framework provides some useful measures when trying to assess significant levels of differential functioning either at the item or scale level. Specifically, noncompensatory differential item functioning (NCDIF) is an index of the magnitude of differences in the expected item responses given theta ("item true scores" in DFIT terminology) across groups which does not take into account the DIF levels of other items on the scale (Facteau & Craig, 2001; Flowers et al., 1999; Raju et al., 1995). Conversely, the compensatory differential item functioning index (CDIF) yields similar information to

NCDIF but accounts for the functioning of other items on the test. Specifically, the CDIF index indicates the net change in the scale's differential functioning that would occur if the focal item were removed from the scale. Differential test functioning (DTF) is the last index which reflects the difference between the two groups' true scores on the test as a whole (Flowers et al., 1999; Raju et al., 1995).

CDIF can be considered important in a practical sense because sometimes items can function differentially but in opposite directions (Flowers et al., 1999; Raju et al., 1995). In such a case one item may display DIF that favors the reference group while another item displays DIF that favors the focal group, thereby canceling each item's contribution to the differential functioning of the scale (Flowers et al., 1999; Raju et al., 1995). If this cancellation does occur, one might find significant NCDIF levels but nonsignificant DTF levels for the given items (Flowers et al., 1999; Raju et al., 1995); CDIF may be able to point to the items that contribute to this situation.

As each of the indices within the DFIT framework yields different information, one should consider when it would be appropriate to use each index. In terms of item level indices, CDIF is useful if a test developer is forced to include some items that display DIF in favor of the reference group and some items that display DIF in favor of the focal group (Raju et al., 1995). NCDIF is appropriate if some items on the test are thought to be biased towards one group in particular or if one wishes to compare the differences in magnitude of DIF across the item set (Raju et al., 1995). Finally, DTF is useful whenever differences in scale level scores are of interest (Raju et al., 1995). Because this study was equally concerned with both item and scale level scores, NCDIF and DTF were considered for investigation. If significant DTF was found for a given scale, then CDIF was investigated.

Item parameter estimation. As mentioned, DIF is assessed in the DFIT framework by looking at the difference between true score estimates across two groups. Typically, these groups are separated into a majority, or reference group (R), and a minority, or focal group (F) (Raju et al., 1995). As such, for each comparison group in the DFIT analyses, the managerial level with a greater number of respondents was designated the reference group while the managerial level with fewer respondents was designated the focal group. It is important to remember that true scores are based on item and person parameters which are estimated separately for each group (Embretson & Reise, 2000). For this study, item and person parameters were estimated using the MULTILOG computer program (Thissen, Chen, & Bock, 2003). Initially, item parameters were estimated for each subscale of the KEYS® instrument within each managerial level. Therefore a total of 24 separate MULTILOG analyses were performed to obtain item parameter estimates. As all items in the KEYS scale contained four response categories, MULTILOG estimated three category threshold (*b)* parameters and one item discrimination (*a*) parameter for each item.

Person parameter estimation. The estimated item parameters for each managerial level were used by MULTILOG to produce theta estimates for all respondents in the dataset at their respective managerial level. Using maximum likelihood estimation procedures, MULTILOG yielded a theta estimate for each person in the dataset such that the probability of each response pattern was maximized.

Equating. In raw form, item parameters estimated by separate MULTILOG analyses are on different metrics and cannot be directly compared. In order to allow comparisons of true score estimates, item parameters must be placed on a common metric which is typically accomplished through a linear transformation. For this study, the EQUATE 2.1 computer program (Baker, 1995) was used to estimate a set of linear transformation coefficients according to the procedure developed by Stocking and Lord (1983). These linear transformation coefficients were used to place the reference group item parameters onto the metric of the focal group to enable cross group comparisons (Flowers et al., 1999). Following the procedures for linking outlined by Candell and Drasgow (1988), a DIF analysis was performed after initially linking the item parameter metrics. Any items displaying significant DIF, as indicated by NCDIF, were removed and a new linking equation was produced using the reduced item set. This procedure was continued until no items displaying DIF had been used in estimating the transformation coefficients. Through this iterative process, item and person parameter estimates from both groups were on the same metric. Group comparisons then commenced by obtaining a true score estimate for an examinee from the focal group at a given level of theta. This true score estimate was compared to a true score estimate based on reference group parameters for the same examinee at an equivalent level of theta (Flowers et al., 1999). If significant differences between the estimates were found, then the item was said to function differentially (Flowers et al., 1999).

Determination of DIF. In this study, DIF was assessed with the DFITPS6 computer program (Raju, 1999). DFITPS6 utilizes the linear transformation coefficients produced in the equating procedure to transform the item parameters onto a common scale so that the reference and focal groups may be compared and DIF may be assessed. Within the DFIT framework, DIF is considered to be significant if the DFIT indices exceed predetermined critical values and if the associated chi-square statistic is significant (Facteau & Craig, 2001). Critical values for the NCDIF index are dependent on the number of response options available in the items (Facteau & Craig, 2001). For the KEYS® scale there were four response options per item and therefore the critical value for NCDIF was 0.054 (Raju, 1999). The critical value for DTF of a scale composed of items with four response options was 0.054 multiplied by the number of items on the scale (Facteau & Craig, 2001).

Section IV: Results

CFA Analyses

Establishing a baseline model. To establish the baseline model, each managerial level was randomly split into two groups consisting of a calibration sample and a hold-out sample. As mentioned previously, the eight factor model specified by Rosenberg and Craig (2006) was used as the a priori model to establish a baseline across all three calibration samples. All eight factors were free to correlate with each other and no items were permitted to cross load. Additionally, the variance of each factor was set to one to set the measurement scale of the latent variables; this also enabled the factor loadings of all manifest variables to be freely estimated. Specifically, 12 items were loaded onto Factor 1 measuring Creative and Challenging Work (item numbers 1, 2, 5, 7, 23, 36, 47, 52, 53, 55, 69, and 76), 15 items were loaded onto Factor 2 measuring Organizational Impediments (item numbers 4, 10, 12, 16, 17, 20, 24, 30, 31, 34, 39, 43, 66, 77, and 78), 11 items were loaded onto Factor 3 which measured Supervisory Encouragement (item numbers 9, 21, 27, 33, 37, 51, 59, 60, 68, 72, and 73), 8 items were loaded onto Factor 4 which measured Workgroup Support (item numbers 6, 15, 19, 25, 29, 41, 58, and 67), 14 items were loaded onto Factor 5 which measured Organizational Encouragement (item numbers 8, 18, 22, 28, 35, 40, 42, 45, 49, 50, 56, 61, 62, and 64), 7 items were loaded onto Factor 6 which measured Sufficient Resources (item numbers 26, 32, 44, 46, 57, 63, and 75), 4 items were loaded onto Factor 7 which measured Workload Pressure (item numbers 3, 11, 38 and 70), and 7 items were loaded onto Factor 8 which measured Productivity (item numbers 13, 14, 48, 54, 65, 71, and 74).

As shown in Table 5, contradictory results were obtained upon trying to fit the baseline model to the calibration samples of each managerial level. Specifically, each of the resulting chi-square values indicated a lack of fit, although this was to be expected given the large sample sizes (Hatcher, 1994). Additionally, the incremental fit indices (i.e. the CFI, and TLI), indicated a lack of fit as well across all managerial levels. In contrast to this, the absolute fit indices (i.e. the RMSEA and SRMR) independently indicated an adequate model fit for each of the calibration samples and also met one of Hu and Bentler's (1999) combinational rules for reducing Type I and Type II errors when viewed together. Although there was some evidence of adequate model fit, additional exploratory analyses were performed on the calibration samples in order to assess whether an adequate model fit could be obtained by each of the incremental and absolute fit indices.

Revisions to the model were suggested by the modification indices produced by Mplus. Modification indices were given for a number of specific parameters that were constrained in the a priori model. The modification indices provided by Mplus were based on Sörbom's (1989) index labeled MI, and represented the expected drop in chi-square if the corresponding parameter were allowed to be freely estimated (Muthén & Muthén, 1998-2005). In order to ensure that the resulting baseline models were equivalent across groups, those modifications that represented the greatest drop in chi-square and which affected all three calibration samples were taken into consideration. Additionally, only those modifications that made theoretical sense were added to the model. For instance, the error covariance was estimated between Item 53 and Item 7 because they both measured the Creative and Challenging Work construct and had nearly identical wording. In all, a total of 12 modifications were added to the original a priori model, which included estimating eight error covariances, allowing two items to cross load on more than one factor, and deleting two items from the model. Although the modifications improved the previously acceptable values for the absolute fit indices, the resulting values for the incremental fit indices were still well below the acceptable fit level of .95 given by Hu and Bentler (1999). The best values obtained for the fit indices across all three groups were .867, .862, .044, and .050 for the CFI, TLI, RMSEA and SRMR respectively.

The contradictory results provided by the various fit indices are not unprecedented; there are several recent examples in the literature of instances where the RMSEA and/or SRMR indicate adequate fit but the incremental fit indices (e.g., CFI, TLI) do not (Browne et al., 2002; Cheung, Leung, & Au, 2006; Cheung & Rensvold, 2001, 2002; Rigdon, 1996). As discussed in the literature, there are several reasons such discrepancies might occur. First, Cheung and Rensvold (2001) discuss the role of parsimony error in model specification and evaluation. Parsimony error occurs when the aggregation of many small secondary relationships (i.e., secondary factor loadings and error term correlations lacking theoretical bases) result in a lack of model fit, even when the hypothesized model adequately explains the primary relationships among the data (Cheung & Rensvold, 2001). In a large scale study involving both simulated and real data, Cheung and Rensvold (2001) found that the only significant predictor of the discrepancy function was the degrees of freedom, which are a function of the number of variables and the number of estimated parameters in the model. They concluded that using traditional cutoff values for the incremental fit indices (e.g., CFI, TLI) in the presence of parsimony error may lead to the rejection of acceptable models (Cheung & Rensvold, 2001).

In a follow-up study, Cheung and Rensvold (2002) investigated the behavior of 20 different goodness of fit indices (of those available in MPLUS, the CFI, TLI, and RMSEA were included) in the presence of high model complexity. Although many goodness of fit indices claim to adjust for model complexity, Cheung and Rensvold (2002) found that only the RMSEA was completely unaffected by the number of items, number of factors, and item by factor interactions. Therefore, they recommend that the RMSEA be used to establish overall model fit and configural equivalence, particularly with models that have many items and/or factors (Cheung & Rensvold, 2002).

In addition to issues related to parsimony and model complexity, there is also evidence that incremental fit indices such as the CFI can be affected by sample size. Rigdon (1996) found that the use of a null model in calculations of model fit for incremental fit indices such as the TLI and CFI resulted in less stable estimation across different methods. Conversely, the RMSEA led to more stable results, particularly for large sample sizes.

Rigdon (1996) concluded that "CFI [is] better suited to more exploratory, small sample cases, and RMSEA [is] better suited to more confirmatory, large sample situations" (p. 376).

The third instance of discrepant results among the various fit indices occurs when the unique variances are small (Browne et al., 2002). Browne et al. (2002) found that in such situations, incremental fit indices such as the CFI and TLI may indicate lack of model fit even when good fit is plausible. The authors recommended using the RMR (root-mean-square residual) index, or the equivalent SRMR, in such situations because it is based only on the residual matrix; that is, the difference between the model implied correlation matrix and the correlation matrix based on actual data. Browne et al. (2002) noted that, "Large values of RMR will occur only if there are large residuals" (p. 408). They concluded that low values of the CFI and TLI may not necessarily indicate poor model fit, particularly if the residuals are small (Browne et al., 2002). The explanation provided is that, "the RMR estimates a measure of misfit, and the chi-square-based fit indices estimate measures of misfit detectability" (Browne et al., 2002, p. 417). Because the SRMR is a direct measure of the model residuals, it is not affected by other characteristics of the model (Browne et al, 2002).

Given the large number of items and large samples used in the present study, the contradictory results among fit indices were most likely due to parsimony error and sample size; consequently, the RMSEA was likely to provide more accurate results about model fit than the CFI or TLI (Cheung & Rensvold, 2001, 2002; Cheung et al., 2006; Rigdon, 1996). The use of a 78-item scale was likely to result in significant parsimony error due to the accumulation of many secondary factor loadings and error correlations that lack theoretical meaning. The very large overall sample size of 20,889 provided further justification for the use of the RMSEA because this fit index is more stable in such situations. Based on Hu and Bentler's (1999) recommendation that more than one fit index be used to minimize type I and type II errors, the SRMR was chosen in addition to the RMSEA. Although the present study did not appear to contain very small unique variances, the study conducted by Browne et al. (2002) indicated that the SRMR is a direct measure of model fit and is unlikely to be adversely affected by other model characteristics. Consequently, the combinational rule of .06 or less for RMSEA and .09 or less for SRMR was used for model evaluation of each of the baseline models and for establishing configural equivalence (Hu & Bentler, 1999). As shown in Table 5, the values for the RMSEA and SRMR resulting from the original unmodified eight factor model, met this combinational rule cutoff criteria posed by Hu and Bentler (1999) for the calibration samples of each managerial level. Additionally, as shown in Table 5, when the same baseline model was tested on the holdout samples for each managerial level, similar values for all fit indices were produced; therefore, an acceptable baseline model for each group had been established.

Testing measurement equivalence. To establish configural measurement equivalence, the established baseline model was simultaneously analyzed across each of the three pairs of comparison groups. As no modifications were included in the previously established baseline models, all 78 items were included in this analysis and no items were permitted to crossload or covary with other items. The variances of all latent variables were set to one in order to establish the measurement scale. Additionally, the factor loadings of all manifest variables were freely estimated.

Table 6 displays the results of the configural equivalence analyses for each of the three comparison groups. Similar to the results obtained when establishing the baseline models, the configural equivalence tests resulted in contradictory information for each comparison group in terms of the model fit indicators. As before, both the chi-square value and the incremental fit indices indicated a lack of model fit while the absolute fit indices resulted in a fair to adequate model fit both independently and in combination with one another. These results are not surprising given that a similar pattern of results was obtained for the baseline model of each managerial level. Similar to evaluating the fit of the baseline models, configural equivalence was considered to have been established across all comparison groups because the RMSEA was close to, or at, .05 for each comparison group. Additionally, Hu and Bentler's (1999) combinational rules concerning the SRMR and RMSEA were met under all three simultaneous comparisons. As part of the initial CFA analyses, these results affirmatively answered research questions 1, 2, and 3, and allowed further investigations of measurement equivalence when using the KEYS® scale.

The next analyses performed concerned testing for metric measurement equivalence across all three comparison groups as indicated by research questions 1a, 2a, and 3a. For these analyses, the factor loadings were set to be equal within each comparison group. The variance of each latent variable was set to one and no item was permitted to cross load or covary with any other item. Simultaneous analyses were performed across each pair of managerial levels being analyzed.

Table 6 displays the results of the metric measurement equivalence tests for each comparison group. As displayed in Table 6 each of the chi-square difference tests was significant. This indicated a lack of metric measurement equivalence as there was a significant drop in fit across all three comparison groups when the factor loadings were constrained to be equal. However, it is well documented that the chi-square difference test is similar to the chi-square statistic in that it is sensitive to large sample sizes and is likely to result in significant values (Brannick, 1995; Cheung & Rensvold, 2002; Facteau & Craig, 2001; Kelloway, 1995). As an alternative, Cheung and Rensvold (2002) empirically determined that the change in magnitude of the CFI is a superior indicator of the fit of a nested model compared to the chi-square difference test. This is because the change in CFI is neither affected by sample size nor correlated with CFI values obtained in prior steps of establishing equivalence. In terms of guidelines, Cheung and Rensvold (2002) recommended that changes in CFI which are smaller than, or equal to, -0.01 indicate that equivalence is displayed across nested models. As seen in Table 6, the change in value for the CFI across all three comparison groups was never larger than -0.01. Additionally, the changes in values for the remainder of the incremental and absolute fit indices were negligible in size. Across all comparisons, the change in SRMR never exceeded .008, the change in CFI or TLI was never larger than .003, and there was no change in the RMSEA for any of the comparison groups. This indicated that the metric measurement equivalence model provided a similar degree of fit to the data when compared to the less constrained model. Thus, it was concluded that the KEYS® scale displayed metric measurement equivalence across all three comparison groups and affirmatively answered research questions 1a, 2a, and 3a.

The test for scalar measurement equivalence, which was the final test in the CFA framework, began once metric measurement equivalence was established for all managerial levels. The test for scalar measurement equivalence was identical to the test for metric measurement equivalence except all item intercepts were constrained to be equal. As before, simultaneous analyses were performed across each pair of managerial levels being analyzed.

The results for the scalar measurement equivalence analyses are displayed in Table 6. While the chi-square difference tests were significant across all three comparison groups, the change in CFI once again was smaller than the -0.01 guideline used to assess equivalence (Cheung & Rensvold, 2002). Additionally, as with the metric equivalence tests, the changes in values for the remainder of the incremental and absolute fit indices were negligible in size. Across all comparisons, the change in SRMR never exceeded -.001, the change in CFI or TLI was never larger that -.008, and there was no change in the RMSEA for any of the comparison groups. This indicated that the scalar measurement equivalence model provided a similar degree of fit to the data when compared to the metric measurement equivalence model. Consequently, in answer to research questions 1b, 2b, and 3b, it was concluded that there was no significant difference between the metric and the scalar measurement equivalence models across any of the groups in our comparison. Given these results, the CFA analyses suggested that the 78 item KEYS® scale functions equivalently across supervisors, middle managers, and executives.

IRT Analyses

DFIT analyses. For each comparison group, the DFIT analyses generated 86 indices to be analyzed for differential item and test functioning. This included eight scale level, or DTF, indices as well as 78 item level, or NCDIF, indices. Across all comparison groups a total of 258 indices were analyzed. Table 7 displays the cutoff and observed DTF and NCDIF values, as well as the equated item parameters, for each item in the comparison of the supervisors to middle managers. Table 8 displays similar information for the DFIT analyses comparing the middle managers to executives, while Table 9 displays these same results comparing the supervisors to executives. As shown in Tables 7 - 9, none of the items displayed a significant amount of differential functioning across managerial levels. While almost all of the items displayed significant chi-square values, none of the items exceeded the NCDIF cutoff value of 0.054, which would be necessary to designate them as DIF items. At the item level, Item 16, an indicator of the Organizational Impediments factor, displayed the highest NCDIF index of .012 for the analysis comparing supervisors to middle managers. For the analysis comparing middle managers to executives Item 38, an indicator of the Workload Pressure factor, displayed the highest NCDIF index of .021. Finally, Item 1, an indicator of the Creative and Challenging Work factor, displayed the highest NCDIF index of .038 for the analyses comparing supervisors to executives. Each of these values was still well below the NCDIF cutoff value established by Raju (1999) for a polytomous item with four response options and therefore no items were found to function differentially across any of the comparison groups.

The IRT analyses also found that no differential functioning occurred at the scale level across any of the comparison groups. The Workgroup Supports factor displayed the highest observed DTF relative to its cutoff value of .432 for the analyses comparing supervisors to middle managers (DTF = .009) and supervisors to executives (DTF = .025). For the analysis comparing middle managers to executives, the Organizational Encouragement factor emerged as having the highest DTF (.016) relative to its cutoff value

(.756). However, none of the DTF indexes were large enough to signify differential test functioning. Given these results, the IRT analyses concluded that the KEYS® scale displays measurement equivalence across all managerial levels at both the item and scale level.

Interrater Agreement Analyses

An additional analysis was performed in order to investigate James et al.'s (1988) position that the psychological climate can be aggregated to create an organizational climate if there exists a shared perception of the climate among individuals within the organization. The analysis was performed by calculating $r_{wg(J)}$, an index of interrater agreement (James, Demaree, & Wolf, 1984, 1993), for each of the 55 U.S. organizations included in the analyses. The index $r_{wg(J)}$ utilizes both the observed and expected error variance of raters' responses to assess the amount of agreement found for scales with multiple items measuring the same construct (James et al., 1984, 1993). James et al. (1984) defined $r_{wg(J)}$ as follows:

$r_{wg(J)} = \dfrac{J[1-(\overline{s_{xj}^{2}}/\sigma_{EU}^{2})]}{J[1-\overline{s_{xj}^{2}}/\sigma_{EU}^{2})]+(\overline{s_{xj}^{2}}/\sigma_{EU}^{2})}$, where J is the number of items measuring the same construct, $\overline{s_{xj}^{2}}$ is the average observed variance over the J items, and σ_{EU}^{2} is the expected error variance for a uniform distribution. In most cases, $r_{wg(J)}$ is expected to fall within the range of 0 to 1, with higher values interpreted as greater agreement (James et al., 1993). A value of .70 has been used as a criterion for $r_{wg(J)}$ to establish significance in past studies (Castro, 2002), and therefore was used to judge significance for these analyses as well. James et al. (1993) suggested that "…$r_{wg(j)}$ is best used in situations in which judges are believed to be interpreting the rating scale in a similar manner" (p. 308). As the CFA and IRT analyses concluded that the KEYS scale displayed measurement equivalence across managerial levels, this study previously established a similar level of interpretation existed among managers.

For the organizations analyzed, the average $r_{wg(j)}$ ranged from .82 to .94 across all scales while the standard deviation ranged from .01 to .07 across all scales. In light of the high level of agreement among managers it was concluded that the organizational climate for creativity can be derived by aggregating all the individual managerial scores for each factor intended for the KEYS® scale.

Section V: Discussion

The results of this study indicated that the KEYS scale displays measurement equivalence across supervisors, middle managers, and executives. Through both CFA and IRT analyses, it was established that managers of all levels maintain a similar cognitive representation of the factors that comprise the KEYS instrument. Additionally it was found that supervisors, middle managers, and executives use a similar measurement scale with equivalent item intercepts when responding to all of the 78 items on the KEYS instrument. These findings are critical because they allow for unambiguous interpretation of comparisons made between groups or individuals that differ in managerial levels (Drasgow & Kanfer, 1985; Horn & McArdle, 1992; Reise et al., 1993; Vandenberg, 2002). Such comparisons may include analyzing group or individual differences in observed or latent means for any of the items or factors on the KEYS scale, analyzing differences in the unique variances for any of items, or other more specific comparisons of interest (Vandenberg, 2002; Vandenberg & Lance, 2000). Consequently, studies which involve investigating the effects of managerial level on the climate for creativity, as measured by KEYS®, can be more confidently performed and interpreted.

It is interesting to note that none of the analyses performed in this study found measurement inequivalence across managerial levels. This included all analyses performed at the scale and the item levels. These results imply that, for U.S. organizations, the differences in functional activities, skill sets, time horizons, levels of abstraction, or cognitive map complexity across managerial levels are not salient enough to result in differential conceptualization of factors in the KEYS scale. This supports a claim by Katz (1955), who stated that a great deal of overlap exists in the skill sets found at each managerial level. He further stated that, "…in practice, these skills are so closely interrelated that it is difficult to tell where one ends and another begins" (p. 37). Perhaps the same type of practical overlap exists across managerial levels in their corresponding functional activities, time horizons, levels of abstraction, and cognitive map complexity thereby explaining the resulting measurement equivalence found in this study.

Alternatively, it is possible that managerial levels differ in functional activities, skill sets, time horizons, levels of abstraction, and cognitive map complexity but the KEYS scale does not assess these dimensions extensively enough to be affected by the differences. For instance, there are no items on the KEYS scale that specifically reference the mix of activities, skill sets, or time required to perform one's job. Similarly there are no items on the KEYS scale that reference the cognitive processes used to make decisions while performing one's job, nor are there any items that reference how employees use their time horizons. In order to study the differences in managerial levels, Katz (1955) recommended that emphasis on these differences should be varied by managerial level. Currently, there is no such emphasis found on any of the items on the KEYS® scale. Essentially Amabile et al. (1996) designed the instrument so that respondents answered in reference to their “current work environment” (p. 1170), and not to any one particular activity within that environment. It would be interesting to see if items that served to differentiate between managerial levels across the discussed variables resulted in differential conceptualizations of the climate for creativity for supervisors, middle managers, and executives. However, given that the purpose of the KEYS scale is to assess the organizational climate for creativity, the fact that the scale does not address all of the differences that may exist across managerial levels is not problematic.

A final possible explanation for the finding of measurement equivalence in this study is that organizational level was not measured precisely enough by the demographic items on the KEYS scale. As indicated earlier, subjects were given 15 generic job titles that served to differentiate their managerial levels in the organization. However, the functions of these generic anchors may not be generalizable across organizations and therefore may have inadvertently confused subjects into endorsing the wrong managerial level. To the extent that this happened, differences across managerial levels in functional activities, skill sets, time horizons, levels of abstraction and cognitive map complexity may have been masked resulting in the finding of measurement equivalence across managerial levels. Further studies should investigate whether the generic job titles included on the KEYS scale are appropriate to use as anchors for their corresponding organizational levels.

Aside from finding measurement equivalence across managerial levels, another important finding from the current study is that there was sufficient agreement among managers within organizations to justify aggregating individual KEYS® scores to create an index of the organizational climate for creativity. With the exception of Amabile et al. (1996), no other studies could be located that tested the assumption that individual scores on the KEYS scale can be aggregated to the organizational level. This may be because Amabile et al. (1996) calculated an intraclass correlation (ICC) among different raters of a project and found modest support for the aggregation of climate for creativity perceptions to the organizational level when using different respondents from the same project team (ICC = .58). However, the authors only used a sample of individuals who were responding to KEYS based on a specific project they were involved with and therefore were not rating the entire "current work environment" the instrument was designed to assess (Amabile et al., 1996, p. 1170). Additionally, as the ICC assesses the ratio of between item variance to the within item variance, there may be cases where ICC values are low in situations where there is a high level of agreement across parallel items (James et al., 1984; Lahey, Downey, & Saal, 1983). As the current study contained KEYS ratings from managers who were assessing their general work environment and analyzed the level of interrater agreement using $r_{wg(J)}$, the findings add further support to the claim that individual scores on KEYS can be aggregated to the organizational level.

Implications

The results of this study have several important implications for researchers and practitioners. First, as no studies have investigated this issue to date, these findings justify using the KEYS® scale to make comparisons across supervisors', middle managers', and executives' perceptions of the climate for creativity. Although Kwaśniewska and Nęcka (2004) investigated this issue using the BCWQ, no studies have investigated this issue using a scale that has been shown to display measurement equivalence.

Additionally, the findings have implications for some established theoretical frameworks of creativity. For instance, in their theory of organizational creativity, Woodman, Sawyer, and Griffin (1993) postulate that individual, group, and organizational characteristics interact with each other to produce creative behavior and creative situations. Similar to the climate for creativity, the creative situation includes all of the environmental influences that interact with individual and group creativity; as such the creative situation affects, and is affected by, the creative behavior that is present in the organization. This interaction between the creative situation and creative behavior is postulated to ultimately result in the construct of organizational creativity (Woodman et al., 1993). Similarly, Amabile's (1996) update to the componential framework of creativity, which was originally proposed by Amabile (1983), posits that creative performance is influenced by interactions that occur between an individual's level of task motivation and contextual factors as measured by the KEYS scale. Given the present finding of measurement equivalence across managerial levels, it seems likely that the underlying conceptualization of the climate for creativity is common to all managers despite interactions between task motivation and the climate for creativity. Given these theories, future research could investigate whether the KEYS scale displays measurement equivalence across a range of individual, group, and organizational characteristics.

Finally, the implications of these findings are positive for practitioners who wish to communicate about the organizational climate for creativity, as any underlying differences in conceptualization of the climate for creativity seem not to lead to differential interpretation of the message across managerial levels. Given the importance of interpersonal and human-relations skills across managerial levels, this finding is particularly important as it implies that communications regarding the climate for creativity could be simultaneously distributed across a broad array of supervisors, middle managers, and executives.

Limitations

Despite the important implications of this study, there are some limitations that should also be noted. In the CFA analyses, no baseline model could be found that resulted in acceptable values for the incremental fit indices across all groups. Even after accounting for all theoretically sound adjustments to the model, as specified by the modification indices provided by Mplus, it was concluded that acceptable values for the CFI and TLI could not be achieved. Although this finding was attributed to such factors as parsimony error resulting from the number of items and factors in the model (Cheung & Rensvold, 2001, 2002), as well as the large sample size used in this study (Cheung et al., 2006; Rigdon, 1996), it may actually have been indicative of a truly misspecified baseline model. Despite this limitation however, the RMSEA resulted in an acceptable value and the combinational rule for the RMSEA and SRMR established by Hu and Bentler (1999) was met when using the full 78 item KEYS® scale as the baseline model for each managerial level. Given this finding it was concluded that an acceptable baseline model was achieved for each managerial level. Identical reasoning was used to establish configural equivalence across all comparison groups. Additionally, it should be noted that the central question of measurement equivalence is not whether a given model is "correct" in an absolute sense (although that is certainly a desirable state of affairs), but instead whether the model fits equivalently to data from different groups. The baseline model used here did fit equivalently in each group, regardless of its absolute goodness of fit in any one group.

Another limitation of the CFA analyses concerned the chi-square difference tests used to judge whether the KEYS® scale displayed metric and scalar equivalence across comparison groups. For all comparison groups, the chi-square difference test resulted in significant values indicating a lack of metric and scalar equivalence. However, as the chi-square difference test is sensitive to sample size (Brannick, 1995; Cheung & Rensvold, 2002; Facteau & Craig, 2001; Kelloway, 1995) an alternative test developed by Cheung and Rensvold (2002) was used to assess whether the KEYS scale displayed metric and scalar equivalence across supervisors, middle managers, and executives. By analyzing the change in CFI across nested models it was concluded that the KEYS scale displayed both metric and scalar equivalence as the value for CFI never changed by more than .01 when further restrictions were placed on the models (Cheung & Rensvold, 2002).

Aside from the limitations with the CFA analyses, there were also some limitations that should be noted regarding the IRT analyses performed in this study. For instance, the number of items used to investigate whether each unidimensional scale in KEYS displayed measurement equivalence was suboptimal. This was because the marginal maximum likelihood estimation procedure employed by MULTILOG to generate theta estimates for the focal group is dependent upon the number of items in a scale in order to gain accurate results.

Analyzing scales with fewer than 20 items can result in inaccurate theta estimates (Embretson & Riese, 2000). As noted earlier, the number of items in each scale of KEYS® used in these analyses ranged from a minimum of four items in the Workload Pressure scale to a maximum of 15 items in the Organizational Encouragement Scale. As such the theta estimates used by DFIT to assess if items or scales displayed DIF may have been inaccurate. However, because DFIT addresses the question of whether theta estimates differ depending on which group's item parameters are used, the absolute accuracy of the theta estimates is less important. Consequently, this limitation seems unlikely to be a significant source of error in the results.

It should be noted that several item parameter estimates were found to have high standard errors across all three comparison groups. For all three managerial groups, Item 38, which served as an indicator of the Workload Pressure scale, revealed the highest standard errors for the $b_1$category threshold parameters. Item 38 is worded as "The organization has an urgent need for successful completion of the work I am now doing." The high standard errors associated with this item are also evidenced by the extraordinarily low values for $b_1$category threshold parameters across all three DIF analyses. For the supervisors and executives, these errors in estimation were most likely due to the low number of respondents to the first option of this item. In these analyses only 2.8% (*n* = 58) of the supervisors and 0.7% (*n* = 22) of the executives responded to the first option of this item. As the middle managers served as the reference group for both DIF analyses for which they were included, the extremely low values of their b_1 category threshold parameter were most likely the result of being placed on the metric of a focal group. However, it should be noted that only 2.1% (*n* = 325) of the middle managers responded to the first response option and therefore this parameter is not free from estimation errors for this group, as is noted by the standard error of .35 across both analyses in which they were included. More accurate estimates of these parameters for all three managerial levels would probably have been obtained had more respondents endorsed this first response option (Embretson & Reise, 2000). Additionally, the first response option could have been collapsed with the second response option. Response options were not collapsed, however for various reasons. Firstly, Andrich (1995) argued against the practice of collapsing response options for Likert scales when using IRT models such as the graded response model. The central point of this argument was that one cannot sum the probabilities associated with responses across collapsed response categories because changing the number of response alternatives affects the probabilistic calculations for all of the item's operating characteristic curves thereby making theta estimates less precise. Because the probabilities associated with the operating characteristic curves for all response categories are affected, summing the new probabilities of only the collapsed response categories would be inappropriate. Overall, by collapsing the response alternative, the item is fundamentally changed due to the resulting differences in the distribution of theta estimates. This statement had previous empirical support in that Andrich (1993) found that, given the same respondents, adding an extra category into a graded response scale affected the probabilities of all the response categories, and not just the ones that were collapsed (as cited in Andrich, 1995).

Additionally, given that the wording of the anchors associated with the KEYS® scale, collapsing response categories did not make intuitive sense. Amabile et al. (1996) developed the following response alternatives that currently serve as anchors for the KEYS® scale: *Never*, *Sometimes*, *Often*, and *Always*. Despite clarification in the KEYS legend that the anchors of *Never* and *Always* could be interpreted as “Almost never” and “Almost always”(Amabile et al., 1996, p. 1165), subjects were likely to be influenced by the literal meanings of the actual item anchors. It was therefore decided not to collapse any response alternatives in order to maintain the integrity, and consistency, of the original response scale across all items.

Finally, although the results generated by the DFIT software program indicated that there were no significant differences in item or scale functioning, it should be noted that the cutoff values used to assess the significance of both the NCDIF and DTF indices are based on past research that may not be fully generalizable to the present study (Oshima, Raju, & Nanda, 2006). Currently, Oshima et al., have indicated that approaches to identify more appropriate cutoff values for NCDIF have been developed for dichotomous data, and advances are being made to identify more appropriate DTF values that will generalize to polytomous data as well. It is recommended that research be conducted to replicate the current study with more applicable cutoff values once these methods become available.

Future Research

In addition to the directions for future research already discussed some other areas for research should be mentioned as well. While the current study found that the KEYS scale functions equivalently across managerial levels, it is important to note that non-managerial employees were not included in the analyses. This is critical because, as discussed previously, Kwaśniewska and Nęcka (2004) found that managers perceive the climate for creativity differently than non-managers. The authors explained their findings by asserting that managers have more access to resources, and face fewer obstacles to implement creative solutions. However, without performing a formal study of measurement equivalence, it is impossible to tell if the perceptual disparities between managers and non-managers are due to true perceptual differences or are just measurement artifacts (Horn & McArdle, 1992; Reise et al., 1993; Vandenberg, 2002). As the current study found that the KEYS® scale functions equivalently across managerial levels, it is foundational for this research.

Additionally, as all respondents in this study were located in the United States it would be interesting to investigate whether the KEYS scale displays measurement equivalence across cultures. As mentioned previously, research has supported the idea that implicit theories of creativity differ across cultures (Niu & Sternberg, 2002; Sternberg, 1985) and has found that people actively use these theories when rating such factors as job performance (Borman, 1987) or another individual's level of creativity (Puccio & Chimento, 1987). As noted by Niu and Sternberg (2002), while some aspects of implicit theories of creativity are universal, there are differences that emerge between Eastern and Western cultures. In their study, Niu and Sternberg (2002) found that "…people in Eastern societies seem to emphasize the social and moral components of creativity, whereas people in Western societies appear to emphasize some individual characteristics such as humor and aesthetic tastes" (p. 277). Given these findings, perhaps measurement inequivalence would have been found if different cultures were represented in the present study. Mikdashi (1999) found that when the KEYS scale was given to a sample of 300 Lebanese managers, a different factor structure emerged than what was originally proposed by Amabile et al. (1996). It is important to note, however, that one of the factors that emerged in Mikdashi's (1999) study was the Creative and Challenging Work factor which was also found by Rosenberg and Craig (2006) and used in the current study. Also, Mikdashi (1999) only used exploratory factor analysis to assess the factor structure of KEYS® and never assessed the fit of the proposed six factor structure with confirmatory factor analyses. Finally, in addition to other methodological limitations, Mikdashi's (1999) study might have benefited from more rigorous tests of measurement equivalence before concluding that Lebanese managers conceptualize the KEYS scale in a different manner than other respondents. However, Mikdashi's (1999) study emphasizes the question of whether or not the KEYS scale displays measurement equivalence across cultures. Future research is needed in this area to assess this question more explicitly.

Another direction for future research is to investigate whether the KEYS scale displays measurement equivalence across gender. Kwaśniewska and Nęcka (2004) found that there were significant differences in perceptions of the climate for creativity across gender as measured by the BCWQ. Specifically, the results of their study indicated that women perceived the factor of Excessive Control as more inhibiting to the climate for creativity than men. While these findings are consistent with literature documenting gender differences in the workplace, (Barnett, Biener, & Baruch, 1987; Greenglass & Burke, 1988; Pretty & McCarthy, 1991) they must still be questioned as no formal study of measurement equivalence across gender for a well established climate for creativity scale has been performed. Given the findings of the current study, future research could investigate whether the KEYS® scale displays measurement equivalence across male and female managerial employees.

Finally, additional research should be performed to develop new fit indices, or alternative ways to interpret existing fit indices such as the CFI and TLI, to account for parsimony error. This research is particularly relevant to the study of the climate for creativity as many of the scales used to assess this construct contain a very large number of items and factors. For instance, the CCQ (Ekvall, 1996) and its updated version, the SOQ (Isaksen et al., 2001) each contain 50 items that load onto anywhere from 7 to 10 factors (Mathisen & Einarsen, 2004). Similarly, the TCI (Anderson & West, 1996) contains anywhere from 38 items, on a shortened version of the scale, to 61 items on the original version that load onto four factors (Mathisen & Einarsen, 2004). Given the large number of items and factors on these scales, there is an increased likelihood that parsimony error could affect evaluation of model fit when attempting to establish baseline models or investigating configural equivalence across comparison groups. Research into alternative ways to interpret existing fit indices, or developing new fit indices that accommodate for parsimony error, is imperative for this field of study.

Section VI: References

Amabile, T. (1983). *The social psychology of creativity*. New York: Springer-Verlag.

Amabile, T. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, Colo.: Westview Press.

Amabile, T., & Gryskiewicz, S. S. (1987). *Creativity in the R&D laboratory*. Greensboro, N.C.: Center for Creative Leadership.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *Academy of Management Journal, 39*(5), 1154-1184.

Anderson, N., & West, M. A. (1996). The team climate inventory: Development of the TCI and its applications in teambuilding for innovativeness. *European Journal of Work and Organizational Psychology, 5*(1), 53-66.

Andrich, D. (1995). Distinctive and incompatible properties of two common classes of IRT models for graded responses. *Applied Psychological Measurement, 19*(1), 101-119.

Baker, F. (1995). *Equate 2.1: Computer program for equating two metrics in item response theory*. Madison: University of Wisconsin, Laboratory of Experimental Design.

Barnett, R. C., Biener, L., & Baruch, G. K. (1987). *Gender and stress*. New York London: Free Press; Collier Macmillan Publishers.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Borman, W.C. (1987). Personal constructs, performance schemata, and “folk theories” of subordinate effectiveness: Explorations in an army officer sample. *Organizational Behavior and Human Decision Processes*, *40*, 307-322.

Brannick, M. T. (1995). Critical comments on applying covariance structural modeling. *Journal of Organizational Behavior, 16*(3), 201-213.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230-258.

Browne, M. W., MacCallum, R. C., Kim, C.-T., Andersen, B. L., & Glaser, R. (2002). When fit indices and residuals are incompatible. *Psychological Methods, 7*(4), 403-421.

Byrne, B. M. (1991). The Maslach burnout inventory: Validating factorial structure and invariance across intermediate, secondary, and university educators. *Multivariate Behavioral Research, 26*(4), 583-605.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466.

Candell, G. L., & Drasgow, F. (1988). An iterative procedure for linking metrics and assessing item bias in item response theory. *Applied Psychological Measurement, 12*(3), 253-260.

Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlation coefficients, $r_{wg(j)}$, hierarchical linear modeling, within- and between-analysis, and random group resampling. *Leadership Quarterly, 13*(1), 69-93.

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25*(1), 1-27.

Cheung, G. W., & Rensvold, R. B. (2001). The effects of model parsimony and sampling error on the fit of structural equation models. *Organizational Research Methods, 4*(3), 236-264.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.

Cheung, M. W. L., Leung, K., & Au, K. (2006). Evaluating multilevel models in cross-cultural research: An illustration with social axioms. *Journal of Cross-Cultural Psychology, 37*(5), 522-541.

Cummings, A., & Oldham, G. R. (1997). Enhancing creativity: Managing work contexts for the high potential employee. *California Management Review, 40*(1), 22-28.

Drasgow, F. (1987). Study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology, 72*(1), 19-29.

Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology, 70*(4), 662-680.

Ekvall, G. (1996). Organizational climate for creativity and innovation. *European Journal of Work and Organizational Psychology, 5*(1), 105-123.

Embretson, S. E., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers.

Facteau, J. D., & Craig, S. B. (2001). Are performance appraisal ratings from different rating sources comparable? *Journal of Applied Psychology, 86*(2), 215-227.

Flowers, C. P., Oshima, T. C., & Raju, N. S. (1999). A description and demonstration of the polytomous-DFIT framework. *Applied Psychological Measurement, 23*(4), 309-326.

Freedman, A. M. (1998). Pathways and crossroads to institutional leadership. *Consulting Psychology Journal: Practice and Research, 50*(3), 131-151.

Greenglass, E. R., & Burke, R. J. (1988). Work and family precursors of burnout in teachers: Sex differences. *Sex Roles, 18*(3), 215-229.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, Calif.: Sage Publications.

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117-144.

Hu, L.T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus. *Structural Equation Modeling, 6*(1), 1-55.

Hunt, E. (1983). On the nature of intelligence. *Science, 219*(4581), 141-146.

Isaksen, S. G., Lauer, K. J., Ekvall, G., & Britz, A. (2001). Perceptions of the best and worst climates for creativity: Preliminary validation evidence for the situational outlook questionnaire. *Creativity Research Journal, 13*(2), 171-184.

Jacobs, T. O., & Jaques, E. (1987). Leadership in complex systems. In J. Zeidner (Ed.), *Human productivity enhancement* (Vol. 2, pp. 7-65). New York: Praeger.

James, L. R., Joyce, W. F., & Slocum, J. W. (1988). Organizations do not cognize. *Academy of Management Review, 13*(1), 129-132.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*(1), 85-98.

James, L. R., Demaree, R. G., & Wolf, G. (1993). R_{wg}: An assessment of within-group interrater agreement. *Journal of Applied Psychology, 78*(2), 306-309.

Jaques, E. (1976). *A general theory of bureaucracy*. London: Heinemann.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, Vol. 36*(4), 409-426.

Jöreskog, K. G., & Sörbom, D. (1996). *Lisrel 8: User's reference guide*. Chicago: Scientific Software International.

Kaiser, R. B., & Craig, S. B. (2004). *What gets you there won't keep you there: Managerial behaviors related to effectiveness at the bottom, middle, and top.* Symposium presented at the 19th Annual Conference of the Society for Industrial Organizational Psychology, Chicago, IL.

Katz, D., & Kahn, R. L. (1966). *The social psychology of organizations*. New York: Wiley.

Katz, R. L. (1955). Skills of an effective administrator. *Harvard Business Review, 33*(1), 33-42.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Kwaśniewska, J., & Nęcka, E. (2004). Perception of the climate for creativity in the workplace: The role of the level in the organization and gender. *Creativity and Innovation Management, 13*(3), 187-196.

Lahey, M. A., Downey, R. G., & Saal, F. (1983). Intraclass correlations: There's more there than meets the eye. *Psychological Bulletin, 93*(3), 586-595.

Lord, F. (1952). A theory of test scores. *Psychometric Monographs*(7).

Lord, F. M. (1980). *Applications of item response to theory to practical testing problems*. Hillsdale, N.J.: L. Erlbaum Associates.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130-149.

MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in the common-factor model: Implications for theory and practice. *Psychological Bulletin, 109*(3), 502-511.

Mann, F. C. (1965). Toward an understanding of the leadership role in formal organization. In L. Broom (Ed.), *Leadership and productivity; some facts of industrial life* (pp. 68-103). San Francisco: Chandler Pub. Co.

Mathisen, G. E., & Einarsen, S. (2004). A review of instruments assessing creative and innovative environments within organizations. *Creativity Research Journal, 16*(1), 119-140.

Maurer, T. J., Raju, N. S., & Collins, W. C. (1998). Peer and subordinate performance appraisal measurement equivalence. *Journal of Applied Psychology, 83*(5), 693-702.

Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods, 7*(4), 361-388.

Mikdashi, T. (1999). Constitutive meaning and aspects of work environment affecting creativity in Lebanon. *Participation & Empowerment: An International Journal, 7*(3), 47-55.

Muthén, L., & Muthén, B. (1998-2005). *Mplus users guide*. Los Angeles: Muthén & Muthén.

Nadler, D. A., & Tushman, M. L. (1983). A general diagnostic model for organizational behavior: Applying a congruence perspective. In J. R. Hackman, E. E. Lawler III & L. W. Porter (Eds.), *Perspectives on behavior in organizations* (pp. 112-124). New York: McGraw-Hill.

Nemiro, J. E. (2004). *Creativity in virtual teams: Key components for success*. San Francisco: Pfeiffer.

Niu, W., & Sternberg, R. (2002). Contemporary studies on the concept of creativity: The east and the west. *Journal of Creative Behavior, 36*(4), 269-288.

Oldham, G. R., & Cummings, A. (1996). Employee creativity: Personal and contextual factors at work. *Academy of Management Journal, 39*(3), 607-634.

Oshima, T. C., Raju, N. S., & Nanda, A. O. (2006). A new method for assessing the statistical significance in the differential functioning of items and tests (DFIT) framework. *Journal of Educational Measurement, 43*(1), 1-17.

Pretty, G. M. H., & McCarthy, M. (1991). Exploring psychological sense of community among women and men of the corporation. *Journal of Community Psychology, 19*(4), 351-361.

Puccio, G. J., & Chimento, M. D. (2001). Implicit theories of creativity: Laypersons' perceptions of the creativity of adaptors and innovators. *Perceptual and Motor Skills, 92*(3,Pt1), 675-681.

Raju, N. S. (1999). *DFITPS6: Differential functioning of items and tests for the polytomous case* (Samejima's GRM) [computer program].Chicago: Illinois Institute of Technology.

Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology, 87*(3), 517-529.

Raju, N. S., van der Linden, W. J., & Fleer, P. F. (1995). IRT-based internal measures of differential functioning of items and tests. *Applied Psychological Measurement, 19*(4), 353-368.

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*(3), 207-230.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*(3), 552-556.

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling, 3*(4), 369-379.

Rosenberg, D., & Craig, S. B. (2006). *Factor structure of the KEYS® Climate for Creativity scale.* Poster presented at the 21st Annual Meeting of the Society for Industrial and Organizational Psychology., Dallas, TX.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph, 17*.

Samejima, F. (1996). Handbook of modern item response theory. In W. J. Van der Linden & R. K. Hambleton (Eds.), (pp. Xv, 510 p.). New York: Springer.

Shalley, C.E. (1991). Effects of productivity goals, creativity goals, and personal discretion on individual creativity. *Journal of Applied Psychology, 76*(2), 179-185.

Shalley, C. E., Gilson, L. L., & Blum, T. C. (2000). Matching creativity requirements and the work environment: Effects on satisfaction and intentions to leave. *Academy of Management Journal, 43*(2), 215-223.

Shalley, C. E., Zhou, J., & Oldham, G. R. (2004). The effects of personal and contextual characteristics on creativity: Where should we go from here? *Journal of Management, 30*(6), 933-958.

Sörbom, D. (1989). Model Modification. *Psychometrika, 54*, 371-384.

Steiger, J. H., & Lind, J. M. (1980, June). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology, 49*(3), 607-627.

Sternberg, R. J. (2003). *Cognitive psychology* (3rd ed.). Belmont, CA: Thomson Wadsworth.

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist, 51*(7), 677-688.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7*(2), 201-210.

Thissen, D., Chen, W.H., & Bock, D. (2003). *MULTILOG for Windows* (Version 7.0) [computer program]. Lincolnwood, IL: Scientific Software International.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, Vol. 38*(1), 1-10.

Vandenberg, R. J. (2002). Toward a further understanding of an improvement in measurement invariance methods and procedures. *Organizational Research Methods, 5*(2), 139-158.

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

Witt, L., & Beorkrem, M. N. (1989). Climate for creative productivity as a predictor of research usefulness and organizational effectiveness in an R&D organization. *Creativity Research Journal, 2*(1-2), 30-40.

Woodman, R. W., Sawyer, J. E., & Griffin, R. W. (1993). Toward a theory of organizational creativity. *Academy of Management Review, 18*(2), 293-321.

Zaccaro, S. J. (2001). *The nature of executive leadership: A conceptual and empirical analysis of success* (1st ed.). Washington, DC: American Psychological Association.

Zhou, J. (1998). Feedback valence, feedback style, task autonomy, and achievement orientation: Interactive effects on creative performance. *Journal of Applied Psychology, 83*(2), 261-276.

Table 1.

Dimensions of the KEYS®: Assessing the Climate for Creativity Scale

Dimension	Description
Stimulant Scales	
Organizational Encouragement	Measures the extent to which the organization values, rewards, and includes employees in efforts related to creativity.
Supervisory Encouragement	Measures the extent to which managers can effectively set goals, create an environment of open communication, and support their team members' ideas.
Work Group Supports	Measures employee perceptions of the amount of diversity inherent in the team, the degree to which ideas and processes are challenged, the team's openness to novel ideas, and the amount of shared communication and collaboration inherent in the team.
Sufficient Resources	Measures the extent to which employees perceive that all of the appropriate resources for being creative are available. This would include access to such things as "funds, facilities, materials, and information" (Amabile et al., 1996, p.1166).
Freedom	Measures the extent to which employees have autonomy in their tasks and can choose how to conduct their daily work.
Challenging Work	Measures the extent to which employees perceive their work to be meaningful and challenging.
Obstacle Scales	
Workload Pressure	Measures the extent to which employees perceive time pressures to be too constricting and feel that there are unreasonable expectations for performing their work.
Organizational Impediments	Measures the extent to which employees perceive any of the following: they will receive harsh negative feedback on ideas, there will be destructive competition, or the organization does not want to change or take risks.
Criterion Scales	
Creativity	Measures the extent to which employees feel that being creative is an integral part of their job, the functioning of their department, and the functioning of their organization as a whole.
Productivity	Measures the extent to which employees perceive their department, unit or organization to be efficient, effective, and/or productive.

Note: All information in this table is based on Amabile et al. (1996)

Table 2.

Fit Indexes for Full and Reduced KEYS® Scale Models

Model	χ^2	*df*	GFI	CFI	TLI	RMSEA	χ^2_{diff}	Δ*df*	ΔGFI	ΔCFI	ΔTLI	ΔRMSEA
Full Scale	10,889.234**	2897	.759	.826	.819	.050						
Reduced Scale	2,546.300**	674	.895	.916	.907	.049	8342.934	2223	.136	.090	.088	.001

Note. Full scale includes all 78 items. Reduced scale includes the 5 highest loading items for each factor for a total of 39 items (only 4 items loaded onto the Workload Pressure factor). ** $p < .0001$.

Table 3.

Summary of Differences Across Managerial Levels

Level	Domain	Functional Activities	Skill Set	Time Horizon	Level of Abstraction	Cognitive Map Complexity
Supervisor	Production Domain	**Application of Existing Structure:** Assign and supervise tasks that are consistent with the goals of the organization, motivate subordinates, and monitor subordinate performance.	Technical & Human-Relation	3 Months to 2 years	Use of imagination to conceptualize end product. Task progress is judged by comparison to concrete conceptualization of end product.	Least complex. Cognitive map is formed in relation to a specific subunit. Map accounts for present and short term future activities.
Middle Manager	Organizational Domain	**Interpolation of Structure:** Ensure that goals are effectively transmitted from the top to the bottom of the organization, create plans to meet these goals, integrate and coordinate the work of their subunits, manage any changes within their subunits, generate policy to impact the company climate, and allocate resources as they see fit.	Interpersonal & Administrative	2 to 5 years	Use of abstract thinking techniques to innovate new products based on pre-established models. Task progress is judged by intuition.	Moderately complex. Cognitive map is formed in relation to the organization as a whole. Map accounts for present and short term future activities.

Table 3 (continued).

Level	Domain	Functional Activities	Skill Set	Time Horizon	Level of Abstraction	Cognitive Map Complexity
Executive	Systems Domain	**Creation of Structure:** Analyze how the organization is performing in relation to the external environment, assess the opportunities for developing or acquiring new businesses, create strategic goals for the organization, interact with the environment to generate acceptance of the organization and its future goals, develop or shape an appropriate organizational climate and culture, and create and/or secure any resources needed by the organization to accomplish its goals.	Conceptual & Administrative & Human-Relation	10 to 20+ years	Use of intuitive theories to make quick decisions regarding tasks that cannot be fully attended to.	Very complex. Cognitive map is formed in relation to internal and external organizational environment. Map accounts for present and long term future activities.

Table 4.

Percent of Variance Accounted for by Each KEYS® Factor

Scale and Managerial Level	Factor	Percent of Variance	Cumulative Percent of Variance
Creative and Challenging Work			
Supervisors	1	43.493	43.493
	2	10.470	53.963
	3	8.802	62.765
Middle Managers	1	46.106	46.106
	2	9.039	55.145
	3	8.752	63.897
Executives	1	42.426	42.426
	2	10.180	52.606
	3	9.629	62.235
Organizational Impediments			
Supervisors	1	34.993	34.993
	2	7.217	42.210
	3	6.953	49.163
Middle Managers	1	37.095	37.095
	2	6.905	44.000
	3	6.550	50.550
Executives	1	35.194	35.194
	2	6.961	42.155
	3	6.565	48.719
Supervisory Encouragement			
Supervisors	1	55.933	55.933
	2	8.052	63.985
	3	7.095	71.080
Middle Managers	1	55.254	55.254
	2	8.183	63.437
	3	7.292	70.730

Table 4 (continued).

Scale and Managerial Level	Factor	Percent of Variance	Cumulative Percent of Variance
Executives	1	50.937	50.937
	2	8.679	59.617
	3	7.973	67.590
Work Group Supports			
Supervisors	1	54.002	54.002
	2	7.906	61.907
	3	7.624	69.532
Middle Managers	1	53.680	53.680
	2	8.668	62.348
	3	7.659	70.008
Executives	1	51.009	51.009
	2	9.022	60.031
	3	8.558	68.589
Organizational Encouragement			
Supervisors	1	47.202	47.202
	2	6.808	54.010
	3	6.340	60.350
Middle Managers	1	50.609	50.609
	2	6.631	57.240
	3	5.597	62.838
Executives	1	50.692	50.692
	2	6.274	56.967
	3	5.723	62.690
Sufficient Resources			
Supervisors	1	47.762	47.762
	2	12.773	60.535
	3	11.228	71.764

Table 4 (continued).

Scale and Managerial Level	Factor	Percent of Variance	Cumulative Percent of Variance
Middle Managers	1	48.875	48.875
	2	12.353	61.229
	3	11.141	72.370
Executives	1	47.971	47.971
	2	12.689	60.660
	3	11.727	72.386
Workload Pressure			
Supervisors	1	53.748	53.748
	2	23.642	77.390
	3	11.758	89.148
Middle Managers	1	54.510	54.510
	2	23.924	78.434
	3	11.662	90.096
Executives	1	52.427	52.427
	2	23.982	76.409
	3	12.634	89.043
Productivity			
Supervisors	1	56.948	56.948
	2	13.868	70.816
	3	7.806	78.622
Middle Managers	1	59.507	59.507
	2	13.608	73.115
	3	7.113	80.228
Executives	1	56.303	56.303
	2	14.494	70.797
	3	7.710	78.507

Note. Only the first three extracted factors are reported for each scale across managerial Levels. Maximum likelihood extraction was used across all analyses.

Table 5

Fit Indexes for Establishing Baseline Models across Managerial Levels

Sample & Managerial Level	χ^2	*df*	CFI	TLI	RMSEA	SRMR
Calibration Sample						
Supervisors	11,860.798*	2,897	.784	.776	.054	.070
Middle Managers	59,622.258*	2,897	.824	.818	.050	.062
Executives	14,082.352*	2,897	.802	.795	.051	.065
Holdout Sample						
Supervisors	12,024.880*	2,897	.783	.775	.055	.070
Middle Managers	59,532.022*	2,897	.823	.817	.050	.063
Executives	13,691.637*	2,897	.805	.798	.050	.066

* $p < .05$.

Table 6

Fit Indexes for Tests of Measurement Equivalence

Comparison Group & Equivalence Test	χ^2	*df*	CFI	TLI	RMSEA	SRMR	χ^2_{diff}	Δ*df*	ΔCFI	ΔTLI	ΔRMSEA	ΔSRMR
Supervisors & Middle Managers												
Configural	136,487.390*	5,794	.820	.813	.050	.063						
Metric	136,685.887*	5,872	.820	.816	.050	.063	198.497*	78	.000	.003	.000	.000
Scalar	137187.932*	5,942	.819	.817	.050	.062	502.045*	70	-.001	.001	.000	-.001
Middle Managers & Executives												
Configural	140445.897*	5794	.822	0.815	.050	.063						
Metric	140877.634*	5872	.821	0.817	.049	.069	431.737*	78	-.001	.002	-.001	.006
Scalar	142393.458*	5942	.819	.817	.049	.068	1515.824*	70	-.002	.000	.000	-.001

Table 6 (continued).

Comparison Group & Equivalence Test	χ^2	*df*	CFI	TLI	RMSEA	SRMR	χ^2_{diff}	Δ*df*	ΔCFI	ΔTLI	ΔRMSEA	ΔSRMR
Supervisors & Executives												
Configural	45150.918*	5794	.798	.790	.052	.066						
Metric	45521.917*	5872	.796	.792	.052	.074	370.999*	78	-.002	.002	.000	.008
Scalar	47158.452*	5942	.788	.786	.052	.074	1636.535*	70	-.008	-.006	.000	.000

* $p < .05$.

Table 7

Results of DFIT Analyses across Supervisors and Middle Managers

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Creative and Challenging Work					0.648	0.005	2100.19
Item1					0.054	0. 009	434,632.40
Supervisors	0.985 (.07)	-5.089 (.41)	-1.368 (.11)	1.596 (.12)			
Middle Managers	1.003 (.02)	-5.402 (.17)	-1.769 (.05)	1.330 (.03)			
Item2					0.054	0.002	88,383.07
Supervisors	1.563 (.08)	-3.778 (.25)	-1.038 (.06)	1.017 (.06)			
Middle Managers	1.643 (.03)	-3.771 (.10)	-1.21 (.02)	0.941 (.02)			
Item5					0.054	0.001	5,038.45
Supervisors	1.358 (.07)	-2.800 (.15)	-0.079 (.05)	2.099 (.11)			
Middle Managers	1.631 (.03)	-2.608 (.05)	-0.116 (.02)	1.944 (.03)			
Item7					0.054	0.000	2,501.75
Supervisors	1.662 (.08)	-3.326 (.19)	-1.011 (.06)	1.063 (.06)			
Middle Managers	1.752 (.03)	-3.576 (.09)	-1.039 (.02)	1.102 (.02)			
Item23					0.054	0.003	56,214.02
Supervisors	0.616 (.05)	-4.421 (.42)	-1.579 (.17)	2.345 (.23)			
Middle Managers	0.746 (.02)	-4.159 (.13)	-1.629 (.06)	2.107 (.06)			
Item36					0.054	0.002	21,517.45
Supervisors	2.062 (.09)	-2.762 (.13)	-0.411 (.04)	1.457 (.06)			
Middle Managers	2.239 (.04)	-2.611 (.05)	-0.354 (.01)	1.611 (.02)			
Item47					0.054	0.000	7,293.18
Supervisors	1.831 (.08)	-2.559 (.12)	-0.226 (.04)	1.399 (.07)			
Middle Managers	1.98 (.03)	-2.523 (.05)	-0.238 (.01)	1.497 (.02)			

Table 7 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,099**
Item52					0.054	0.001	40,673.58
Supervisors	2.296 (.09)	-2.227 (.09)	-0.394 (.04)	1.349 (.06)			
Middle Managers	2.260 (.04)	-2.370 (.04)	-0.612 (.01)	1.030 (.02)			
Item53					0.054	0.000	6,293.74
Supervisors	1.903 (.08)	-2.722 (.13)	-0.701 (.05)	0.903 (.05)			
Middle Managers	2.108 (.03)	-2.565 (.05)	-0.665 (.02)	0.935 (.02)			
Item55					0.054	0.002	<u>2,155.67</u>
Supervisors	2.312 (.10)	-1.867 (.07)	0.127 (.03)	1.434 (.06)			
Middle Managers	2.206 (.03)	-2.008 (.03)	0.076 (.01)	1.627 (.02)			
Item69					0.054	0.002	28,730.85
Supervisors	1.962 (.08)	-2.609 (.12)	-0.414 (.04)	1.376 (.06)			
Middle Managers	1.98 (.03)	-2.565 (.05)	-0.368 (.02)	1.52 (.02)			
Item76					0.054	0.004	17,671.48
Supervisors	2.246 (.09)	-2.248 (.09)	-0.173 (.04)	1.524 (.06)			
Middle Managers	2.12 (.03)	-2.337 (.04)	-0.095 (.01)	1.746 (.02)			
Organizational Impediments					**0.810**	**0.002**	**<u>2,171.43</u>**
Item4					0.054	0.001	2,309.83
Supervisors	1.143 (.06)	-2.54 (.15)	0.202 (.06)	1.829 (.11)			
Middle Managers	1.052 (.02)	-2.836 (.06)	0.235 (.02)	2.043 (.05)			
Item10					0.054	0.001	5,710.29
Supervisors	0.928 (.06)	-1.962 (.14)	1.209 (.10)	3.783 (.26)			
Middle Managers	1.088 (.02)	-1.740 (.04)	1.128 (.03)	3.415 (.08)			

Table 7 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,099**
Item12					0.054	0.003	14,302.06
Supervisors	0.991 (.06)	-1.83 (.12)	0.811 (.08)	2.824 (.18)			
Middle Managers	0.928 (.02)	-1.837 (.04)	1.066 (.03)	3.252 (.08)			
Item16					0.054	0.012	261,867.00
Supervisors	1.187 (.06)	-2.943 (.17)	-0.144 (.06)	1.398 (.09)			
Middle Managers	1.422 (.03)	-2.636 (.04)	0.023 (.02)	1.620 (.03)			
Item17					0.054	0.002	8,712.60
Supervisors	1.386 (.07)	-2.153 (.11)	0.546 (.05)	2.371 (.13)			
Middle Managers	1.229 (.03)	-2.464 (.04)	0.516 (.02)	2.641 (.06)			
Item20					0.054	0.000	81,275.98
Supervisors	1.927 (.09)	-0.691 (.04)	1.223 (.06)	2.428 (.11)			
Middle Managers	1.948 (.03)	-0.653 (.01)	1.266 (.02)	2.503 (.04)			
Item24					0.054	0.002	3,228.65
Supervisors	2.115 (.09)	-1.397 (.06)	0.302 (.04)	1.397 (.06)			
Middle Managers	2.081 (.03)	-1.588 (.02)	0.336 (.01)	1.438 (.02)			
Item30					0.054	0.000	2,117.48
Supervisors	1.058 (.06)	-1.702 (.12)	1.300 (.09)	3.128 (.19)			
Middle Managers	1.186 (.02)	-1.561 (.03)	1.125 (.03)	2.879 (.06)			
Item31					0.054	0.005	26,712.88
Supervisors	1.434 (.07)	-1.235 (.07)	1.072 (.07)	2.672 (.14)			
Middle Managers	1.379 (.03)	-1.526 (.03)	1.034 (.02)	2.702 (.05)			
Item34					0.054	0.001	11,637.29
Supervisors	1.173 (.06)	-2.453 (.14)	0.147 (.06)	1.996 (.12)			
Middle Managers	1.316 (.03)	-2.375 (.04)	0.178 (.02)	1.931 (.04)			

Table 7 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square df = 2,099
Item39					0.054	0.000	2,287.22
Supervisors	1.353 (.07)	-0.948 (.07)	1.035 (.07)	2.637 (.14)			
Middle Managers	1.351 (.03)	-1.002 (.02)	1.113 (.03)	2.879 (.06)			
Item43					0.054	0.001	2,554.32
Supervisors	1.064 (.06)	-2.219 (.14)	0.869 (.08)	2.673 (.16)			
Middle Managers	1.240 (.02)	-2.004 (.04)	0.763 (.02)	2.390 (.05)			
Item66					0.054	0.001	29,225.84
Supervisors	1.802 (.09)	-1.168 (.06)	1.343 (.06)	2.768 (.14)			
Middle Managers	1.886 (.03)	-1.220 (.02)	1.202 (.02)	2.652 (.05)			
Item77					0.054	0.002	32,273.80
Supervisors	1.479 (.07)	-1.552 (.08)	1.203 (.07)	2.559 (.13)			
Middle Managers	1.391 (.03)	-1.715 (.03)	1.111 (.03)	2.817 (.06)			
Item78					0.054	0.002	13,859.22
Supervisors	2.009 (.09)	-0.500 (.04)	1.304 (.06)	2.348 (.10)			
Middle Managers	2.228 (.04)	-0.588 (.01)	1.169 (.02)	2.233 (.03)			
Supervisory Encouragement					**0.594**	**0.003**	**2,229.25**
Item9					0.054	0.001	8,947.31
Supervisors	1.564 (.08)	-2.17 (.12)	-0.115 (.05)	1.507 (.07)			
Middle Managers	1.548 (.03)	-2.076 (.04)	-0.026 (.02)	1.496 (.03)			
Item21					0.054	0.001	7,389.86
Supervisors	2.895 (.13)	-2.188 (.09)	-1.320 (.05)	0.005 (.03)			
Middle Managers	2.516 (.04)	-2.258 (.04)	-1.334 (.02)	0.051 (.01)			

Table 7 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item27					0.054	0.001	11,510.26
Supervisors	2.708 (.12)	-1.965 (.07)	-0.587 (.04)	0.573 (.03)			
Middle Managers	2.929 (.04)	-1.928 (.03)	-0.608 (.01)	0.540 (.01)			
Item33					0.054	0.001	10,346.87
Supervisors	2.246 (.11)	-2.514 (.13)	-1.381 (.06)	0.294 (.04)			
Middle Managers	2.228 (.04)	-2.508 (.05)	-1.360 (.02)	0.350 (.01)			
Item37					0.054	0.000	2,217.80
Supervisors	2.549 (.12)	-2.263 (.10)	-1.297 (.05)	0.283 (.03)			
Middle Managers	2.405 (.04)	-2.341 (.04)	-1.419 (.02)	0.269 (.01)			
Item51					0.054	0.000	9,651.73
Supervisors	2.592 (.12)	-2.578 (.13)	-1.063 (.05)	0.281 (.03)			
Middle Managers	2.679 (.04)	-2.564 (.05)	-1.089 (.02)	0.244 (.01)			
Item59					0.054	0.000	2,207.48
Supervisors	2.562 (.11)	-2.085 (.09)	-1.136 (.05)	0.304 (.03)			
Middle Managers	2.81 (.04)	-2.052 (.03)	-1.106 (.02)	0.337 (.01)			
Item60					0.054	0.001	2,216.49
Supervisors	1.738 (.08)	-2.103 (.10)	0.044 (.04)	1.757 (.07)			
Middle Managers	1.823 (.03)	-1.984 (.04)	0.118 (.01)	1.635 (.03)			
Item68					0.054	0.000	9,688.16
Supervisors	2.652 (.12)	-2.484 (.12)	-1.011 (.04)	0.454 (.03)			
Middle Managers	2.613 (.04)	-2.513 (.05)	-0.981 (.02)	0.463 (.01)			
Item72					0.054	0. 001	10,823.92
Supervisors	2.446 (.11)	-2.529 (.13)	-0.866 (.04)	0.622 (.04)			
Middle Managers	2.738 (.04)	-2.529 (.05)	-0.867 (.02)	0.568 (.01)			

Table 7 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item73					0.054	0.000	5,692.53
Supervisors	2.726 (.13)	-2.471 (.12)	-0.734 (.04)	0.537 (.03)			
Middle Managers	2.747 (.04)	-2.560 (.05)	-0.791 (.02)	0.525 (.01)			
WorkGroup Supports					**0.432**	**0.009**	**3,923.94**
Item6					0.054	0.004	29,832.09
Supervisors	2.193 (.10)	-3.296 (.21)	-1.316 (.05)	0.620 (.04)			
Middle Managers	2.191 (.04)	-3.167 (.08)	-1.206 (.02)	0.757 (.01)			
Item15					0.054	0.001	2,494.60
Supervisors	2.275 (.10)	-2.256 (.09)	-0.729 (.04)	0.747 (.04)			
Middle Managers	2.192 (.04)	-2.410 (.04)	-0.822 (.02)	0.802 (.01)			
Item19					0.054	0.001	4,447.73
Supervisors	1.843 (.09)	-2.352 (.11)	-0.354 (.04)	1.567 (.07)			
Middle Managers	1.737 (.03)	-2.684 (.05)	-0.487 (.02)	1.593 (.02)			
Item25					0.054	0.000	2,172.67
Supervisors	2.060 (.10)	-3.285 (.20)	-0.693 (.04)	1.040 (.05)			
Middle Managers	2.115 (.04)	-3.149 (.07)	-0.780 (.02)	1.096 (.02)			
Item29					0.054	0.000	2,670.14
Supervisors	2.439 (.10)	-3.243 (.20)	-1.159 (.05)	0.422 (.03)			
Middle Managers	2.659 (.04)	-3.020 (.07)	-1.139 (.02)	0.451 (.01)			
Item41					0.054	0.000	4,518.94
Supervisors	1.712 (.15)	-3.456 (.22)	-1.314 (.07)	0.884 (.05)			
Middle Managers	1.870 (.03)	-3.213 (.07)	-1.197 (.02)	0.913 (.02)			

Table 7 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item58					0.054	0.001	9,199.49
Supervisors	1.938 (.08)	-3.845 (.29)	-1.348 (.06)	0.525 (.04)			
Middle Managers	1.877 (.03)	-3.894 (.13)	-1.501 (.03)	0.475 (.02)			
Item67					0.054	0.001	4,476.77
Supervisors	2.310 (.09)	-2.393 (.10)	-0.872 (.04)	0.676 (.04)			
Middle Managers	2.592 (.04)	-2.538 (.05)	-0.886 (.02)	0.751 (.01)			
Organizational Encouragement					**0.756**	**0.010**	**3,989.42**
Item8					0.054	0.000	9,624.18
Supervisors	1.814 (.08)	-2.505 (.12)	-0.145 (.04)	1.585 (.07)			
Middle Managers	1.805 (.03)	-2.686 (.05)	-0.188 (.02)	1.600 (.02)			
Item18					0.054	0.002	30,196.63
Supervisors	2.440 (.10)	-2.467 (.11)	-0.409 (.03)	0.866 (.04)			
Middle Managers	2.602 (.04)	-2.413 (.04)	-0.368 (.01)	0.979 (.01)			
Item22					0.054	0.001	210,619.00
Supervisors	1.251 (.07)	-2.564 (.14)	-0.368 (.06)	1.730 (.10)			
Middle Managers	1.350 (.03)	-2.626 (.05)	-0.447 (.02)	1.650 (.03)			
Item28					0.054	0.001	55,033.61
Supervisors	1.605 (.07)	-2.435 (.12)	-0.591 (.05)	1.057 (.06)			
Middle Managers	1.705 (.03)	-2.612 (.05)	-0.660 (.02)	1.037 (.02)			
Item35					0.054	0.001	5,812.56
Supervisors	2.391 (.11)	-1.925 (.07)	0.078 (.03)	1.779 (.06)			
Middle Managers	2.390 (.04)	-2.052 (.03)	0.152 (.01)	1.754 (.02)			

Table 7 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item40					0.054	0. 002	15,867.05
Supervisors	1.928 (.09)	-2.080 (.09)	-0.128 (.04)	1.477 (.06)			
Middle Managers	2.215 (.04)	-1.906 (.03)	-0.052 (.01)	1.515 (.02)			
Item42					0.054	0.000	13,204.90
Supervisors	2.450 (.11)	-2.539 (.11)	-0.311 (.03)	1.538 (.06)			
Middle Managers	2.569 (.04)	-2.479 (.04)	-0.339 (.01)	1.486 (.02)			
Item45					0.054	0.000	2,103.60
Supervisors	0.629 (.05)	-2.295 (.23)	1.639 (.16)	4.701 (.39)			
Middle Managers	0.669 (.02)	-2.266 (.08)	1.577 (.05)	4.652 (.13)			
Item49					0.054	0.001	3,569.58
Supervisors	2.667 (.11)	-2.343 (.10)	-0.411 (.03)	0.969 (.04)			
Middle Managers	2.659 (.04)	-2.373 (.04)	-0.424 (.01)	1.101 (.01)			
Item50					0.054	0.001	4,467.93
Supervisors	2.253 (.10)	-1.704 (.07)	0.270 (.04)	1.805 (.07)			
Middle Managers	2.311 (.04)	-1.825 (.03)	0.334 (.01)	1.811 (.02)			
Item56					0.054	0.002	7,130.18
Supervisors	1.601 (.07)	-2.515 (.12)	-0.408 (.05)	1.160 (.06)			
Middle Managers	1.856 (.03)	-2.241 (.04)	-0.335 (.01)	1.189 (.02)			
Item61					0.054	0.004	10,575.83
Supervisors	2.853 (.12)	-1.321 (.04)	0.430 (.03)	1.783 (.06)			
Middle Managers	2.728 (.04)	-1.305 (.02)	0.551 (.01)	1.933 (.02)			

Table 7 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,099**
Item62					0.054	0.000	6,010.64
Supervisors	1.866	-1.354	0.703	2.311			
	(.08)	(.06)	(.05)	(.10)			
Middle Managers	2.060	-1.362	0.735	2.305			
	(.03)	(.02)	(.01)	(.03)			
Item64					0.054	0.009	188,291.00
Supervisors	1.529	-1.770	0.207	1.862			
	(.07)	(.09)	(.05)	(.09)			
Middle Managers	1.582	-2.056	0.022	1.766			
	(.03)	(.04)	(.02)	(.03)			
Sufficient Resources					**0.378**	**0.001**	**3,875.77**
Item26					0.054	0.002	31,395.71
Supervisors	1.886	-2.978	-1.024	0.906			
	(.07)	(.15)	(.05)	(.05)			
Middle Managers	1.736	-3.335	-1.115	0.817			
	(.02)	(.06)	(.02)	(.02)			
Item32					0.054	0.000	2,183.82
Supervisors	2.482	-3.049	-0.892	1.069			
	(.10)	(.16)	(.04)	(.04)			
Middle Managers	2.294	-2.797	-0.898	1.067			
	(.04)	(.04)	(.01)	(.02)			
Item44					0.054	0.004	50,214.23
Supervisors	1.040	-3.833	-1.219	1.371			
	(.06)	(.26)	(.09)	(.10)			
Middle Managers	0.989	-4.356	-1.506	1.248			
	(.02)	(.11)	(.04)	(.03)			
Item46					0.054	0.001	144,694.50
Supervisors	1.375	-2.408	-0.557	1.753			
	(.06)	(.13)	(.05)	(.09)			
Middle Managers	1.261	-2.411	-0.506	1.850			
	(.02)	(.04)	(.02)	(.03)			
Item57						0.002	90,572.25
Supervisors	2.075	-2.816	-0.487	1.433	0.054		
	(.08)	(.14)	(.04)	(.06)			
Middle Managers	2.080	-2.708	-0.425	1.536			
	(.03)	(.04)	(.01)	(.02)			

Table 7 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item63					0.054	0.000	2,289.76
Supervisors	1.708 (.08)	-3.445 (.21)	-2.034 (.10)	0.498 (.05)			
Middle Managers	1.923 (.03)	-3.085 (.06)	-2.015 (.03)	0.495 (.02)			
Item75					0.054	0.000	25,219.14
Supervisors	2.107 (.08)	-2.845 (.14)	-0.325 (.04)	1.569 (.06)			
Middle Managers	2.178 (.03)	-2.679 (.04)	-0.309 (.01)	1.596 (.02)			
Workload Pressure					**0.216**	**0.000**	**5997.63**
Item3					0.054	0.001	8,783.21
Supervisors	2.643 (.11)	-2.313 (.08)	-0.127 (.03)	0.982 (.04)			
Middle Managers	3.229 (.05)	-2.177 (.03)	-0.186 (.01)	0.963 (.01)			
Item11					0.054	0.001	17,330.88
Supervisors	2.377 (.10)	-2.037 (.07)	-0.211 (.03)	1.751 (.06)			
Middle Managers	2.226 (.03)	-2.090 (.03)	-0.174 (.01)	1.921 (.02)			
Item38					0.054	0.002	228,063.00
Supervisors	0.489 (.05)	-7.506 (.85)	-2.188 (.26)	2.217 (.26)			
Middle Managers	0.484 (.02)	-8.097 (.35)	-2.547 (.11)	2.102 (.08)			
Item70					0.054	0.001	175,272.60
Supervisors	2.154 (.09)	-2.583 (.11)	-0.528 (.04)	0.756 (.04)			
Middle Managers	2.178 (.03)	-2.50 (.04)	-0.490 (.01)	0.806 (.01)			
Productivity					**0.378**	**0.000**	**2,347.15**
Item13					0.054	0.000	2,357.13
Supervisors	2.570 (.11)	-2.745 (.12)	-0.651 (.04)	1.363 (.05)			
Middle Managers	3.044 (.05)	-2.578 (.04)	-0.610 (.01)	1.370 (.02)			

Table 7 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,099**
Item14					0.054	0.001	9,275.69
Supervisors	1.700 (.07)	-2.084 (.09)	-0.057 (.04)	1.648 (.08)			
Middle Managers	1.728 (.03)	-2.173 (.04)	-0.160 (.02)	1.691 (.03)			
Item48					0.054	0.001	24,053.86
Supervisors	1.667 (.08)	-4.210 (.36)	-1.636 (.08)	0.565 (.05)			
Middle Managers	1.672 (.03)	-4.127 (.12)	-1.685 (.03)	0.657 (.02)			
Item54					0.054	0.000	<u>2,144.32</u>
Supervisors	2.222 (.09)	-3.367 (.24)	-1.238 (.05)	0.981 (.05)			
Middle Managers	2.089 (.03)	-3.562 (.09)	-1.246 (.02)	0.986 (.02)			
Item65					0.054	0.001	2,490.19
Supervisors	2.981 (.12)	-2.710 (.14)	-0.881 (.04)	1.008 (.04)			
Middle Managers	3.611 (.06)	-2.702 (.04)	-0.849 (.01)	0.951 (.01)			
Item71					0.054	0.001	4,968.07
Supervisors	2.926 (.11)	-2.158 (.08)	-0.438 (.03)	1.350 (.05)			
Middle Managers	3.243 (.05)	-2.057 (.02)	-0.346 (.01)	1.352 (.02)			
Item74					0.054	0.000	30,437.69
Supervisors	2.358 (.10)	-2.798 (.14)	-0.906 (.04)	1.194 (.05)			
Middle Managers	2.349 (.03)	-2.986 (.06)	-0.935 (.02)	1.176 (.02)			

Note. For all analyses, the middle managers served as the reference group and the supervisors served as the focal group. For all scales, the DFIT cutoff values and indices refer to the DTF index. For all items, the DFIT cutoff values and indices refer to the NCDIF index. All chi-square values are significant at the $p < .05$ level, except for those that are underlined.

Table 8

Results of DFIT Analyses across Executives and Middle Managers

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Creative and Challenging Work					**0.648**	**0.001**	**2,962.34**
Item1					0.054	0.006	57,749.08
Executives	0.838 (.05)	-7.579 (.74)	-3.197 (.21)	0.461 (.07)			
Middle Managers	0.908 (.02)	-6.667 (.17)	-2.652 (.05)	0.773 (.03)			
Item2					0.054	0.006	64,530.77
Executives	1.600 (.07)	-5.770 (.88)	-2.171 (.09)	0.150 (.04)			
Middle Managers	1.487 (.03)	-4.864 (.10)	-2.034 (.02)	0.344 (.02)			
Item5					0.054	0.002	64,438.96
Executives	1.647 (.07)	-3.746 (.22)	-0.714 (.04)	1.558 (.06)			
Middle Managers	1.476 (.03)	-3.579 (.05)	-0.825 (.02)	1.452 (.03)			
Item7					0.054	0.001	8,233.63
Executives	1.474 (.07)	-4.882 (.41)	-2.172 (.10)	0.459 (.04)			
Middle Managers	1.586 (.03)	-4.649 (.09)	-1.845 (.02)	0.521 (.02)			
Item23					0.054	0.005	122,004.40
Executives	0.652 (.05)	-6.713 (.61)	-3.022 (.26)	1.400 (.13)			
Middle Managers	0.675 (.02)	-5.293 (.13)	-2.496 (.06)	1.632 (.06)			
Item36					0.054	0.001	31,163.33
Executives	2.119 (.09)	-3.482 (.20)	-1.041 (.04)	1.136 (.04)			
Middle Managers	2.026 (.04)	-3.582 (.05)	-1.088 (.01)	1.084 (.02)			
Item47					0.054	0.008	90,083.87
Executives	1.876 (.07)	-3.276 (.16)	-0.781 (.04)	1.167 (.05)			
Middle Managers	1.792 (.03)	-3.485 (.05)	-0.959 (.01)	0.958 (.02)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item52					0.054	0.001	3,237.71
Executives	2.433 (.09)	-2.963 (.14)	-0.923 (.04)	0.844 (.03)			
Middle Managers	2.103 (.04)	-3.120 (.04)	-1.052 (.01)	0.839 (.02)			
Item53					0.054	0.003	37,958.98
Executives	2.045 (.08)	-3.465 (.20)	-1.397 (.05)	0.221 (.03)			
Middle Managers	1.908 (.03)	-3.532 (.05)	-1.432 (.02)	0.336 (.02)			
Item55					0.054	0.002	24,316.02
Executives	2.257 (.08)	-2.712 (.11)	-0.521 (.03)	1.193 (.04)			
Middle Managers	1.996 (.03)	-2.916 (.03)	-0.613 (.01)	1.102 (.02)			
Item69					0.054	0.002	18,753.06
Executives	1.939 (.07)	-3.058 (.14)	-0.914 (.04)	1.053 (.04)			
Middle Managers	1.792 (.03)	-3.531 (.05)	-1.103 (.02)	0.983 (.02)			
Item76					0.054	0.003	31,692.17
Executives	2.227 (.08)	-3.180 (.17)	-0.676 (.03)	1.357 (.05)			
Middle Managers	1.919 (.03)	-3.280 (.04)	-0.802 (.01)	1.233 (.02)			
Organizational Impediments					**0.810**	**0.016**	**9,132.44**
Item4					0.054	0.003	†
Executives	1.134 (.05)	-2.389 (.12)	0.683 (.06)	2.580 (.03)			
Middle Managers	1.004 (.02)	-2.616 (.06)	0.604 (.02)	2.499 (.05)			
Item10					0.054	0.000	26,257.27
Executives	1.148 (.06)	-1.445 (.08)	1.549 (.08)	3.897 (.21)			
Middle Managers	1.038 (.02)	-1.467 (.04)	1.540 (.03)	3.937 (.08)			

Table 8 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,959**
Item12					0.054	0.000	8,243.61
Executives	0.882 (.05)	-1.762 (.11)	1.555 (.10)	4.190 (.26)			
Middle Managers	0.885 (.02)	-1.568 (.04)	1.475 (.03)	3.767 (.08)			
Item16					0.054	0.001	5,197.77
Executives	1.556 (.07)	-2.581 (.11)	0.451 (.04)	2.177 (.09)			
Middle Managers	1.357 (.03)	-2.406 (.04)	0.381 (.02)	2.056 (.03)			
Item17					0.054	0.002	107,187.00
Executives	1.344 (.06)	-2.428 (.11)	0.728 (.05)	2.780 (.14)			
Middle Managers	1.173 (.03)	-2.225 (.04)	0.899 (.02)	3.126 (.06)			
Item20					0.054	0.004	38,169.58
Executives	1.829 (.07)	-0.497 (.04)	1.792 (.07)	3.135 (.15)			
Middle Managers	1.858 (.03)	-0.327 (.01)	1.685 (.02)	2.981 (.04)			
Item24					0.054	0.002	3,358.04
Executives	1.895 (.07)	-1.464 (.05)	0.840 (.04)	2.161 (.08)			
Middle Managers	1.985 (.03)	-1.307 (.02)	0.710 (.01)	1.864 (.02)			
Item30					0.054	0.000	5,432.55
Executives	1.251 (.06)	-1.278 (.07)	1.529 (.07)	3.224 (.16)			
Middle Managers	1.131 (.02)	-1.279 (.03)	1.537 (.03)	3.375 (.06)			
Item31					0.054	0.000	120,374.20
Executives	1.335 (.06)	-1.281 (.07)	1.391 (.07)	3.036 (.15)			
Middle Managers	1.316 (.03)	-1.243 (.03)	1.441 (.02)	3.189 (.05)			
Item34					0.054	0.000	3,487.53
Executives	1.420 (.06)	-2.057 (.08)	0.454 (.04)	2.157 (.10)			
Middle Managers	1.255 (.03)	-2.132 (.04)	0.544 (.02)	2.381 (.04)			

Table 8 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,959**
Item39					0.054	0.002	209,847.00
Executives	1.385 (.06)	-0.645 (.05)	1.682 (.07)	3.565 (.19)			
Middle Managers	1.288 (.03)	-0.693 (.02)	1.524 (.03)	3.375 (.06)			
Item43					0.054	0.009	213,042.20
Executives	1.178 (.06)	-1.498 (.08)	1.475 (.08)	3.543 (.19)			
Middle Managers	1.183 (.02)	-1.743 (.04)	1.158 (.02)	2.863 (.05)			
Item66					0.054	0.002	142,621.00
Executives	1.793 (.07)	-1.009 (.05)	1.483 (.06)	2.865 (.12)			
Middle Managers	1.799 (.03)	-0.922 (.02)	1.617 (.02)	3.138 (.05)			
Item77					0.054	0.002	55,849.71
Executives	1.430 (.06)	-1.625 (.07)	1.434 (.07)	3.260 (.17)			
Middle Managers	1.327 (.03)	-1.441 (.03)	1.522 (.03)	3.311 (.06)			
Item78					0.054	0.008	35,356.70
Executives	2.205 (.09)	-0.449 (.03)	1.470 (.05)	2.559 (.10)			
Middle Managers	2.125 (.04)	-0.259 (.01)	1.582 (.02)	2.698 (.03)			
Supervisory Encouragement					**0.594**	**0.003**	**3,533.35**
Item9					0.054	0.001	3,983.90
Executives	1.169 (.06)	-2.906 (.16)	-0.244 (.05)	1.676 (.08)			
Middle Managers	1.328 (.03)	-2.582 (.04)	-0.192 (.02)	1.582 (.03)			
Item21					0.054	0.000	2,982.65
Executives	2.162 (.09)	-2.648 (.11)	-1.644 (.06)	-0.111 (.03)			
Middle Managers	2.158 (.04)	-2.794 (.04)	-1.717 (.02)	-0.102 (.01)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item27					0.054	0.004	15,366.38
Executives	2.749 (.10)	-2.375 (.09)	-0.909 (.03)	0.354 (.03)			
Middle Managers	2.513 (.04)	-2.409 (.03)	-0.871 (.01)	0.467 (.01)			
Item33					0.054	0.000	3,514.84
Executives	2.094 (.09)	-3.216 (.17)	-1.803 (.07)	0.311 (.03)			
Middle Managers	1.911 (.04)	-3.086 (.05)	-1.748 (.02)	0.246 (.01)			
Item37					0.054	0.002	22,220.97
Executives	1.918 (.08)	-2.901 (.14)	-1.733 (.07)	0.237 (.03)			
Middle Managers	2.063 (.04)	-2.891 (.04)	-1.816 (.02)	0.152 (.01)			
Item51					0.054	0.002	11,691.53
Executives	2.628 (.10)	-2.949 (.13)	-1.353 (.05)	0.061 (.03)			
Middle Managers	2.298 (.04)	-3.150 (.05)	-1.432 (.02)	0.122 (.01)			
Item59					0.054	0.002	10,338.09
Executives	2.436 (.10)	-2.603 (.11)	-1.400 (.05)	0.331 (.03)			
Middle Managers	2.410 (.04)	-2.553 (.03)	-1.451 (.02)	0.231 (.01)			
Item60					0.054	0.001	31,393.69
Executives	1.535 (.07)	-2.465 (.11)	-0.005 (.04)	1.814 (.07)			
Middle Managers	1.564 (.03)	-2.475 (.04)	-0.024 (.01)	1.744 (.03)			
Item68					0.054	0.000	2,972.46
Executives	2.276 (.09)	-3.168 (.16)	-1.294 (.05)	0.383 (.03)			
Middle Managers	2.241 (.04)	-3.092 (.05)	-1.305 (.02)	0.378 (.01)			
Item72					0.054	0.000	15,233.17
Executives	2.378 (.09)	-3.018 (.15)	-1.166 (.04)	0.461 (.03)			
Middle Managers	2.349 (.04)	-3.110 (.05)	-1.173 (.02)	0.501 (.01)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item73					0.054	0.001	17,921.54
Executives	2.482 (.10)	-3.284 (.19)	-1.092 (.04)	0.385 (.03)			
Middle Managers	2.356 (.04)	-3.146 (.05)	-1.084 (.02)	0.450 (.01)			
WorkGroup Supports					**0.432**	**0.009**	**3,233.66**
Item6					0.054	0.001	18,009.93
Executives	1.984 (.08)	-4.098 (.13)	-1.843 (.07)	0.459 (.03)			
Middle Managers	1.952 (.04)	-4.041 (.08)	-1.840 (.02)	0.365 (.01)			
Item15					0.054	0.002	33,591.62
Executives	1.894 (.08)	-3.399 (.18)	-1.339 (.05)	0.500 (.04)			
Middle Managers	1.953 (.04)	-3.190 (.04)	-1.408 (.02)	0.415 (.01)			
Item19					0.054	0.005	217,238.60
Executives	1.704 (.07)	-3.675 (.21)	-1.193 (.05)	1.127 (.05)			
Middle Managers	1.548 (.03)	-3.499 (.05)	-1.033 (.02)	1.303 (.02)			
Item25					0.054	0.000	4,228.15
Executives	2.116 (.09)	-3.811 (.27)	-1.360 (.05)	0.744 (.04)			
Middle Managers	1.884 (.04)	-4.020 (.07)	-1.361 (.02)	0.745 (.02)			
Item29					0.054	0.004	8,465.78
Executives	2.520 (.10)	-4.009 (.38)	-1.687 (.05)	0.191 (.03)			
Middle Managers	2.369 (.04)	-3.875 (.07)	-1.764 (.02)	0.020 (.01)			
Item41					0.054	0.002	57,609.79
Executives	1.830 (.08)	-3.920 (.27)	-1.834 (.07)	0.445 (.04)			
Middle Managers	1.665 (.03)	-4.092 (.07)	-1.829 (.02)	0.539 (.02)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item58					0.054	0.002	32,976.04
Executives	1.869 (.08)	-4.872 (.54)	-2.231 (.09)	-0.040 (.03)			
Middle Managers	1.672 (.03)	-4.856 (.13)	-2.170 (.03)	0.047 (.02)			
Item67					0.054	0.001	5,970.62
Executives	2.572 (.10)	-3.258 (.19)	-1.353 (.04)	0.441 (.03)			
Middle Managers	2.309 (.04)	-3.334 (.05)	-1.480 (.02)	0.357 (.01)			
Organizational Encouragement					**0.756**	**0.016**	**5,058.06**
Item8					0.054	0.000	3,322.38
Executives	1.801 (.07)	-3.642 (.22)	-0.775 (.04)	1.171 (.05)			
Middle Managers	1.799 (.03)	-3.212 (.05)	-0.705 (.02)	1.089 (.02)			
Item18					0.054	0.005	24,290.01
Executives	2.659 (.09)	-2.969 (.16)	-0.819 (.03)	0.603 (.03)			
Middle Managers	2.593 (.04)	-2.938 (.04)	-0.886 (.01)	0.465 (.01)			
Item22					0.054	0.000	3,754.61
Executives	1.381 (.06)	-3.319 (.18)	-1.050 (.06)	1.189 (.06)			
Middle Managers	1.345 (.03)	-3.152 (.05)	-0.965 (.02)	1.139 (.03)			
Item28					0.054	0.000	15,714.54
Executives	1.790 (.07)	-3.230 (.17)	-1.158 (.05)	0.585 (.04)			
Middle Managers	1.699 (.03)	-3.137 (.05)	-1.180 (.02)	0.524 (.02)			
Item35					0.054	0.003	26,847.96
Executives	2.504 (.09)	-2.485 (.10)	-0.310 (.03)	1.401 (.04)			
Middle Managers	2.382 (.04)	-2.575 (.03)	-0.365 (.01)	1.243 (.02)			

Table 8 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,959**
Item40					0.054	0.001	24,661.85
Executives	2.326 (.09)	-2.456 (.10)	-0.542 (.03)	1.076 (.04)			
Middle Managers	2.207 (.04)	-2.429 (.03)	-0.569 (.01)	1.004 (.02)			
Item42					0.054	0.001	21,918.12
Executives	2.578 (.09)	-2.984 (.16)	-0.826 (.03)	1.052 (.03)			
Middle Managers	2.560 (.04)	-3.004 (.04)	-0.857 (.01)	0.974 (.02)			
Item45					0.054	0.017	24,243.24
Executives	0.849 (.05)	-2.637 (.17)	0.609 (.07)	3.090 (.18)			
Middle Managers	0.666 (.02)	-2.790 (.08)	1.066 (.05)	4.151 (.13)			
Item49					0.054	0.002	27,393.15
Executives	2.713 (.10)	-2.878 (.13)	-0.873 (.03)	0.678 (.03)			
Middle Managers	2.650 (.04)	-2.898 (.04)	-0.942 (.01)	0.588 (.01)			
Item50					0.054	0.001	27,461.52
Executives	2.540 (.09)	-2.189 (.08)	-0.140 (.03)	1.364 (.04)			
Middle Managers	2.303 (.04)	-2.348 (.03)	-0.181 (.01)	1.300 (.02)			
Item56					0.054	0.000	4,476.79
Executives	2.042 (.08)	-2.716 (.12)	-0.783 (.04)	0.703 (.04)			
Middle Managers	1.850 (.03)	-2.766 (.04)	-0.853 (.01)	0.677 (.02)			
Item61					0.054	0.011	44,506.65
Executives	2.742 (.10)	-1.780 (.05)	0.163 (.03)	1.652 (.05)			
Middle Managers	2.719 (.04)	-1.826 (.02)	0.036 (.01)	1.423 (.02)			
Item62					0.054	0.004	8,563.39
Executives	2.507 (.09)	-1.813 (.06)	0.148 (.03)	1.606 (.05)			
Middle Managers	2.053 (.03)	-1.883 (.02)	0.220 (.01)	1.796 (.03)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item64					0.054	0.017	63,595.88
Executives	1.613 (.07)	-2.859 (.13)	-0.852 (.05)	1.081 (.05)			
Middle Managers	1.577 (.03)	-2.580 (.04)	-0.495 (.02)	1.255 (.03)			
Sufficient Resources					**0.378**	**0.000**	**3,592.62**
Item26					0.054	0.004	38,919.41
Executives	1.706 (.07)	-3.450 (.19)	-1.397 (.06)	0.409 (.04)			
Middle Managers	1.798 (.02)	-3.432 (.06)	-1.290 (.02)	0.575 (.02)			
Item32					0.054	0.000	7,213.56
Executives	2.197 (.08)	-2.941 (.13)	-1.093 (.04)	0.845 (.04)			
Middle Managers	2.377 (.04)	-2.913 (.04)	-1.080 (.01)	0.816 (.02)			
Item44					0.054	0.011	118,215.00
Executives	0.985 (.06)	-4.920 (.37)	-2.127 (.12)	0.693 (.07)			
Middle Managers	1.025 (.02)	-4.418 (.11)	-1.667 (.04)	0.991 (.03)			
Item46					0.054	0.001	27,500.52
Executives	1.443 (.06)	-2.475 (.11)	-0.581 (.04)	1.616 (.07)			
Middle Managers	1.306 (.02)	-2.541 (.04)	-0.702 (.02)	1.572 (.03)			
Item57						0.001	136,876.70
Executives	2.170 (.08)	-2.687 (.11)	-0.546 (.03)	1.325 (.05)	0.054		
Middle Managers	2.155 (.03)	-2.827 (.04)	-0.624 (.01)	1.269 (.02)			
Item63					0.054	0.001	11,649.82
Executives	2.127 (.07)	-3.235 (.18)	-1.998 (.07)	0.336 (.03)			
Middle Managers	1.992 (.03)	-3.191 (.06)	-2.159 (.03)	0.264 (.02)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item75					0.054	0.003	107,310.40
Executives	2.338 (.08)	-2.608 (.10)	-0.412 (.03)	1.460 (.05)			
Middle Managers	2.256 (.03)	-2.799 (.04)	-0.512 (.01)	1.327 (.02)			
Workload Pressure					**0.216**	**0.002**	**10,023.19**
Item3					0.054	0.004	104,849.10
Executives	2.905 (.10)	-2.674 (.10)	-0.468 (.03)	0.806 (.03)			
Middle Managers	2.987 (.05)	-2.684 (.03)	-0.532 (.01)	0.711 (.01)			
Item11					0.054	0.009	53,089.93
Executives	2.169 (.08)	-2.447 (.09)	-0.261 (.03)	1.889 (.06)			
Middle Managers	2.059 (.03)	-2.591 (.03)	-0.518 (.01)	1.747 (.02)			
Item38					0.054	0.021	†
Executives	0.517 (.08)	-9.413 (1.31)	-3.763 (.59)	1.180 (.37)			
Middle Managers	0.447 (.02)	-9.085 (.35)	-3.084 (.11)	1.943 (.08)			
Item70					0.054	0.001	83,100.09
Executives	2.042 (.08)	-2.985 (.13)	-0.817 (.04)	0.594 (.03)			
Middle Managers	2.014 (.03)	-3.033 (.04)	-0.860 (.01)	0.541 (.01)			
Productivity					**0.378**	**0.002**	**3,430.89**
Item13					0.054	0.000	7,060.35
Executives	2.762 (.10)	-2.947 (.13)	-1.027 (.03)	1.107 (.04)			
Middle Managers	2.822 (.05)	-3.163 (.04)	-1.040 (.01)	1.096 (.02)			
Item14					0.054	0.002	90,465.48
Executives	1.707 (.06)	-2.850 (.12)	-0.706 (.04)	1.382 (.06)			
Middle Managers	1.601 (.03)	-2.726 (.04)	-0.554 (.02)	1.443 (.03)			

Table 8 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,959
Item48					0.054	0.001	43,972.75
Executives	1.524 (.06)	-5.952 (.92)	-2.375 (.10)	0.406 (.04)			
Middle Managers	1.550 (.03)	-4.835 (.12)	-2.200 (.03)	0.327 (.02)			
Item54					0.054	0.000	18,019.98
Executives	2.039 (.08)	-4.321 (.40)	-1.856 (.07)	0.666 (.04)			
Middle Managers	1.936 (.03)	-4.225 (.09)	-1.726 (.02)	0.682 (.02)			
Item65					0.054	0.000	18,360.71
Executives	3.387 (.12)	-3.052 (.16)	-1.227 (.03)	0.668 (.03)			
Middle Managers	3.347 (.06)	-3.297 (.04)	-1.298 (.01)	0.644 (.01)			
Item71					0.054	0.001	3,626.51
Executives	3.611 (.13)	-2.459 (.08)	-0.702 (.03)	1.078 (.03)			
Middle Managers	3.006 (.05)	-2.601 (.02)	-0.755 (.01)	1.077 (.02)			
Item74					0.054	0.001	15,462.38
Executives	2.340 (.08)	-3.867 (.32)	-1.320 (.04)	0.934 (.04)			
Middle Managers	2.177 (.03)	-3.603 (.06)	-1.391 (.02)	0.887 (.02)			

Note. For all analyses, the middle managers served as the reference group and the executives served as the focal group. For all scales, the DFIT cutoff values and indices refer to the DTF index. For all items, the DFIT cutoff values and indices refer to the NCDIF index. All chi-square values are significant at the $p < .05$ level, except for those that are underlined. † =Chi-square values greater than 999,999.99.

Table 9

Results of DFIT Analyses across Supervisors and Executives

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Creative and Challenging Work					**0.648**	**0.006**	**2,869.45**
Item1					0.054	0.038	229,286.20
Supervisors	0.985 (.07)	-5.089 (.41)	-1.368 (.11)	1.596 (.12)			
Executives	0.937 (.05)	-6.156 (.74)	-2.233 (.21)	1.041 (.07)			
Item2					0.054	0.015	333,646.70
Supervisors	1.563 (.08)	-3.778 (.25)	-1.038 (.06)	1.017 (.06)			
Executives	1.788 (.07)	-4.536 (.88)	-1.315 (.09)	0.762 (.04)			
Item5					0.054	0.001	3,091.74
Supervisors	1.358 (.07)	-2.800 (.15)	-0.079 (.05)	2.099 (.11)			
Executives	1.840 (.07)	-2.725 (.22)	-0.011 (.04)	2.022 (.06)			
Item7					0.054	0.003	9,718.54
Supervisors	1.662 (.08)	-3.326 (.19)	-1.011 (.06)	1.063 (.06)			
Executives	1.646 (.07)	-3.741 (.41)	-1.316 (.10)	1.039 (.04)			
Item23					0.054	0.021	†
Supervisors	0.616 (.05)	-4.421 (.42)	-1.579 (.17)	2.345 (.23)			
Executives	0.729 (.05)	-5.380 (.61)	-2.076 (.26)	1.881 (.13)			
Item36					0.054	0.004	35,424.03
Supervisors	2.062 (.09)	-2.762 (.13)	-0.411 (.04)	1.457 (.06)			
Executives	2.368 (.09)	-2.488 (.20)	-0.303 (.04)	1.645 (.04)			
Item47					0.054	0.012	196,982.10
Supervisors	1.831 (.08)	-2.559 (.12)	-0.226 (.04)	1.399 (.07)			
Executives	2.096 (.07)	-2.304 (.16)	-0.071 (.04)	1.672 (.05)			

Table 9 (continued).

Scale and Item	***a* (SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,099**
Item52					0.054	0.007	10,885.41
Supervisors	2.296 (.09)	-2.227 (.09)	-0.394 (.04)	1.349 (.06)			
Executives	2.719 (.09)	-2.024 (.14)	-0.198 (.04)	1.384 (.03)			
Item53					0.054	0.001	2,511.47
Supervisors	1.903 (.08)	-2.722 (.13)	-0.701 (.05)	0.903 (.05)			
Executives	2.285 (.08)	-2.473 (.20)	-0.622 (.05)	0.826 (.03)			
Item55					0.054	0.005	11,463.10
Supervisors	2.312 (.10)	-1.867 (.07)	0.127 (.03)	1.434 (.06)			
Executives	2.521 (.08)	-1.799 (.11)	0.162 (.03)	1.696 (.04)			
Item69					0.054	0.014	85,024.13
Supervisors	1.962 (.08)	-2.609 (.12)	-0.414 (.04)	1.376 (.06)			
Executives	2.166 (.07)	-2.109 (.14)	-0.190 (.04)	1.570 (.04)			
Item76					0.054	0.016	67,812.86
Supervisors	2.246 (.09)	-2.248 (.09)	-0.173 (.04)	1.524 (.06)			
Executives	2.489 (.08)	-2.218 (.17)	0.023 (.03)	1.842 (.05)			
Organizational Impediments					**0.810**	**0.025**	**2,242.33**
Item4					0.054	0.003	13,689.19
Supervisors	1.143 (.06)	-2.54 (.15)	0.202 (.06)	1.829 (.11)			
Executives	1.187 (.05)	-2.625 (.12)	0.310 (.06)	2.121 (.13)			
Item10					0.054	0.002	8,547.45
Supervisors	0.928 (.06)	-1.962 (.14)	1.209 (.10)	3.783 (.26)			
Executives	1.203 (.06)	-1.723 (.08)	1.136 (.08)	3.379 (.21)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item12					0.054	0.004	5,312.17
Supervisors	0.991 (.06)	-1.83 (.12)	0.811 (.08)	2.824 (.18)			
Executives	0.924 (.05)	-2.026 (.11)	1.142 (.10)	3.659 (.26)			
Item16					0.054	0.022	76,297.77
Supervisors	1.187 (.06)	-2.943 (.17)	-0.144 (.06)	1.398 (.09)			
Executives	1.629 (.07)	-2.808 (.11)	0.088 (.04)	1.736 (.09)			
Item17					0.054	0.009	75,967.16
Supervisors	1.386 (.07)	-2.153 (.11)	0.546 (.05)	2.371 (.13)			
Executives	1.407 (.06)	-2.662 (.11)	0.352 (.05)	2.312 (.14)			
Item20					0.054	0.002	3,479.06
Supervisors	1.927 (.09)	-0.691 (.04)	1.223 (.06)	2.428 (.11)			
Executives	1.915 (.07)	-0.818 (.04)	1.368 (.07)	2.651 (.15)			
Item24					0.054	0.007	<u>2,101.75</u>
Supervisors	2.115 (.09)	-1.397 (.06)	0.302 (.04)	1.397 (.06)			
Executives	1.985 (.07)	-1.741 (.05)	0.459 (.04)	1.720 (.08)			
Item30					0.054	0.001	<u>2,171.90</u>
Supervisors	1.058 (.06)	-1.702 (.12)	1.300 (.09)	3.128 (.19)			
Executives	1.310 (.06)	-1.563 (.07)	1.117 (.07)	2.736 (.16)			
Item31					0.054	0.008	57,749.16
Supervisors	1.434 (.07)	-1.235 (.07)	1.072 (.07)	2.672 (.14)			
Executives	1.398 (.06)	-1.566 (.07)	0.985 (.07)	2.557 (.15)			
Item34					0.054	0.001	<u>2,108.58</u>
Supervisors	1.173 (.06)	-2.453 (.14)	0.147 (.06)	1.996 (.12)			
Executives	1.487 (.06)	-2.307 (.08)	0.091 (.04)	1.717 (.10)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item39					0.054	0.003	10,073.98
Supervisors	1.353 (.07)	-0.948 (.07)	1.035 (.07)	2.637 (.14)			
Executives	1.451 (.06)	-0.959 (.05)	1.263 (.07)	3.061 (.19)			
Item43					0.054	0.015	747,117.10
Supervisors	1.064 (.06)	-2.219 (.14)	0.869 (.08)	2.673 (.16)			
Executives	1.233 (.06)	-1.774 (.08)	1.065 (.08)	3.041 (.19)			
Item66					0.054	0.006	30,367.61
Supervisors	1.802 (.09)	-1.168 (.06)	1.343 (.06)	2.768 (.14)			
Executives	1.877 (.07)	-1.307 (.05)	1.073 (.06)	2.393 (.12)			
Item77					0.054	0.006	31,825.68
Supervisors	1.479 (.07)	-1.552 (.08)	1.203 (.07)	2.559 (.13)			
Executives	1.498 (.06)	-1.895 (.07)	1.026 (.07)	2.770 (.17)			
Item78					0.054	0.018	41,631.53
Supervisors	2.009 (.09)	-0.500 (.04)	1.304 (.06)	2.348 (.10)			
Executives	2.309 (.09)	-0.772 (.03)	1.061 (.05)	2.101 (.10)			
Supervisory Encouragement					**0.594**	**0.005**	**2,651.86**
Item9					0.054	0.002	6,170.83
Supervisors	1.564 (.08)	-2.17 (.12)	-0.115 (.05)	1.507 (.07)			
Executives	1.363 (.06)	-2.353 (.16)	-0.071 (.05)	1.576 (.08)			
Item21					0.054	0.001	8,957.41
Supervisors	2.895 (.13)	-2.188 (.09)	-1.320 (.05)	0.005 (.03)			
Executives	2.522 (.09)	-2.132 (.11)	-1.272 (.06)	0.043 (.03)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item27					0.054	0.008	10,020.74
Supervisors	2.708 (.12)	-1.965 (.07)	-0.587 (.04)	0.573 (.03)			
Executives	3.206 (.10)	-1.898 (.09)	-0.641 (.03)	0.442 (.03)			
Item33					0.054	0.001	4,701.64
Supervisors	2.246 (.11)	-2.514 (.13)	-1.381 (.06)	0.294 (.04)			
Executives	2.442 (.09)	-2.619 (.17)	-1.408 (.07)	0.405 (.03)			
Item37					0.054	0.002	8,388.11
Supervisors	2.549 (.12)	-2.263 (.10)	-1.297 (.05)	0.283 (.03)			
Executives	2.237 (.08)	-2.349 (.14)	-1.347 (.07)	0.342 (.03)			
Item51					0.054	0.004	6,770.67
Supervisors	2.592 (.12)	-2.578 (.13)	-1.063 (.05)	0.281 (.03)			
Executives	3.065 (.10)	-2.390 (.13)	-1.022 (.05)	0.190 (.03)			
Item59					0.054	0.002	4,980.78
Supervisors	2.562 (.11)	-2.085 (.09)	-1.136 (.05)	0.304 (.03)			
Executives	2.841 (.10)	-2.093 (.11)	-1.062 (.05)	0.422 (.03)			
Item60					0.054	0.001	2,556.70
Supervisors	1.738 (.08)	-2.103 (.10)	0.044 (.04)	1.757 (.07)			
Executives	1.790 (.07)	-1.976 (.11)	0.134 (.04)	1.694 (.07)			
Item68					0.054	0.000	5,484.79
Supervisors	2.652 (.12)	-2.484 (.12)	-1.011 (.04)	0.454 (.03)			
Executives	2.654 (.09)	-2.578 (.16)	-0.971 (.05)	0.467 (.03)			
Item72					0.054	0.003	10,129.56
Supervisors	2.446 (.11)	-2.529 (.13)	-0.866 (.04)	0.622 (.04)			
Executives	2.773 (.09)	-2.449 (.15)	-0.862 (.04)	0.533 (.03)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item73					0.054	0.002	11,866.89
Supervisors	2.726 (.13)	-2.471 (.12)	-0.734 (.04)	0.537 (.03)			
Executives	2.895 (.10)	-2.677 (.19)	-0.798 (.04)	0.469 (.03)			
WorkGroup Supports					**0.432**	**0.025**	**3,422.09**
Item6					0.054	0.010	24,613.59
Supervisors	2.193 (.10)	-3.296 (.21)	-1.316 (.05)	0.620 (.04)			
Executives	2.222 (.08)	-3.226 (.34)	-1.212 (.07)	0.844 (.03)			
Item15					0.054	0.003	9,156.97
Supervisors	2.275 (.10)	-2.256 (.09)	-0.729 (.04)	0.747 (.04)			
Executives	2.121 (.08)	-2.602 (.18)	-0.762 (.05)	0.880 (.04)			
Item19					0.054	0.010	32,699.25
Supervisors	1.843 (.09)	-2.352 (.11)	-0.354 (.04)	1.567 (.07)			
Executives	1.909 (.07)	-2.849 (.21)	-0.632 (.05)	1.440 (.05)			
Item25					0.054	0.000	2,103.64
Supervisors	2.060 (.10)	-3.285 (.20)	-0.693 (.04)	1.040 (.05)			
Executives	2.369 (.09)	-2.970 (.27)	-0.780 (.05)	1.098 (.04)			
Item29					0.054	0.008	9,073.62
Supervisors	2.439 (.10)	-3.243 (.20)	-1.159 (.05)	0.422 (.03)			
Executives	2.822 (.10)	-3.147 (.38)	-1.073 (.05)	0.604 (.03)			
Item41					0.054	0.001	3,065.00
Supervisors	1.712 (.15)	-3.456 (.22)	-1.314 (.07)	0.884 (.05)			
Executives	2.049 (.08)	-3.067 (.27)	-1.204 (.07)	0.831 (.04)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	B_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item58					0.054	0.005	39,379.87
Supervisors	1.938 (.08)	-3.845 (.29)	-1.348 (.06)	0.525 (.04)			
Executives	2.092 (.08)	-3.917 (.54)	-1.559 (.09)	0.398 (.03)			
Item67					0.054	0.005	7,681.95
Supervisors	2.310 (.09)	-2.393 (.10)	-0.872 (.04)	0.676 (.04)			
Executives	2.880 (.10)	-2.476 (.19)	-0.774 (.04)	0.828 (.03)			
Organizational Encouragement					**0.756**	**0.030**	**4,941.91**
Item8					0.054	0.002	5,384.64
Supervisors	1.814 (.08)	-2.505 (.12)	-0.145 (.04)	1.585 (.07)			
Executives	1.807 (.07)	-3.114 (.22)	-0.257 (.04)	1.682 (.05)			
Item18					0.054	0.012	29,502.06
Supervisors	2.440 (.10)	-2.467 (.11)	-0.409 (.03)	0.866 (.04)			
Executives	2.668 (.09)	-2.444 (.16)	-0.301 (.03)	1.116 (.03)			
Item22					0.054	0.004	77,687.05
Supervisors	1.251 (.07)	-2.564 (.14)	-0.368 (.06)	1.730 (.10)			
Executives	1.386 (.06)	-2.792 (.18)	-0.531 (.06)	1.700 (.06)			
Item28					0.054	0.000	6,408.14
Supervisors	1.605 (.07)	-2.435 (.12)	-0.591 (.05)	1.057 (.06)			
Executives	1.796 (.07)	-2.703 (.17)	-0.639 (.05)	1.099 (.04)			
Item35					0.054	0.005	23,889.29
Supervisors	2.391 (.11)	-1.925 (.07)	0.078 (.03)	1.779 (.06)			
Executives	2.512 (.09)	-1.961 (.10)	0.206 (.03)	1.911 (.04)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item40					0.054	0.005	40,095.75
Supervisors	1.928 (.09)	-2.080 (.09)	-0.128 (.04)	1.477 (.06)			
Executives	2.334 (.09)	-1.932 (.10)	-0.025 (.03)	1.588 (.04)			
Item42					0.054	0.000	3,003.40
Supervisors	2.450 (.11)	-2.539 (.11)	-0.311 (.03)	1.538 (.06)			
Executives	2.587 (.09)	-2.458 (.16)	-0.308 (.03)	1.564 (.03)			
Item45					0.054	0.006	7,044.25
Supervisors	0.629 (.05)	-2.295 (.23)	1.639 (.16)	4.701 (.39)			
Executives	0.852 (.05)	-2.112 (.17)	1.122 (.07)	3.595 (.18)			
Item49					0.054	0.006	11,399.44
Supervisors	2.667 (.11)	-2.343 (.10)	-0.411 (.03)	0.969 (.04)			
Executives	2.722 (.10)	-2.353 (.13)	-0.355 (.03)	1.191 (.03)			
Item50					0.054	0.004	22,688.59
Supervisors	2.253 (.10)	-1.704 (.07)	0.270 (.04)	1.805 (.07)			
Executives	2.549 (.09)	-1.666 (.08)	0.376 (.03)	1.874 (.04)			
Item56					0.054	0.006	9,340.91
Supervisors	1.601 (.07)	-2.515 (.12)	-0.408 (.05)	1.160 (.06)			
Executives	2.049 (.08)	-2.191 (.12)	-0.265 (.04)	1.216 (.04)			
Item61					0.054	0.018	10,571.23
Supervisors	2.853 (.12)	-1.321 (.04)	0.430 (.03)	1.783 (.06)			
Executives	2.752 (.10)	-1.258 (.05)	0.678 (.03)	2.162 (.05)			
Item62					0.054	0.001	2,114.19
Supervisors	1.866 (.08)	-1.354 (.06)	0.703 (.05)	2.311 (.10)			
Executives	2.515 (.09)	-1.291 (.06)	0.663 (.03)	2.116 (.05)			

Table 9 (continued).

Scale and Item	***a*** **(SE)**	**b_1 (SE)**	**b_2 (SE)**	**b_3 (SE)**	**DFIT Cutoff Value**	**Index**	**Chi Square *df* = 2,099**
Item64					0.054	0.062	163,162.80
Supervisors	1.529 (.07)	-1.770 (.09)	0.207 (.05)	1.862 (.09)			
Executives	1.619 (.07)	-2.334 (.13)	-0.334 (.05)	1.592 (.05)			
Sufficient Resources					**0.378**	**0.004**	**3,704.92**
Item26					0.054	0.012	51,936.05
Supervisors	1.886 (.07)	-2.978 (.15)	-1.024 (.05)	0.906 (.05)			
Executives	1.647 (.07)	-3.353 (.19)	-1.226 (.06)	0.644 (.04)			
Item32					0.054	0.000	2,476.18
Supervisors	2.482 (.10)	-3.049 (.16)	-0.892 (.04)	1.069 (.04)			
Executives	2.121 (.08)	-2.826 (.13)	-0.912 (.04)	1.096 (.04)			
Item44					0.054	0.031	93,935.02
Supervisors	1.040 (.06)	-3.833 (.26)	-1.219 (.09)	1.371 (.10)			
Executives	0.951 (.06)	-4.876 (.37)	-1.983 (.12)	0.939 (.07)			
Item46					0.054	0.004	500,608.70
Supervisors	1.375 (.06)	-2.408 (.13)	-0.557 (.05)	1.753 (.09)			
Executives	1.393 (.06)	-2.343 (.11)	-0.381 (.04)	1.895 (.07)			
Item57						0.005	244,960.60
Supervisors	2.075 (.08)	-2.816 (.14)	-0.487 (.04)	1.433 (.06)	0.054		
Executives	2.095 (.08)	-2.562 (.11)	-0.345 (.03)	1.593 (.05)			
Item63					0.054	0.001	4,956.96
Supervisors	1.708 (.08)	-3.445 (.21)	-2.034 (.10)	0.498 (.05)			
Executives	2.053 (.07)	-3.131 (.18)	-1.849 (.07)	0.569 (.03)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item75					0.054	0.004	97,916.55
Supervisors	2.107 (.08)	-2.845 (.14)	-0.325 (.04)	1.569 (.06)			
Executives	2.257 (.08)	-2.480 (.10)	-0.206 (.03)	1.734 (.05)			
Workload Pressure					**0.216**	**0.002**	**9,662.95**
Item3					0.054	0.001	12,147.23
Supervisors	2.643 (.11)	-2.313 (.08)	-0.127 (.03)	0.982 (.04)			
Executives	3.145 (.10)	-2.164 (.10)	-0.126 (.03)	1.050 (.03)			
Item11					0.054	0.021	90,868.76
Supervisors	2.377 (.10)	-2.037 (.07)	-0.211 (.03)	1.751 (.06)			
Executives	2.349 (.08)	-1.954 (.09)	0.064 (.03)	2.050 (.06)			
Item38					0.054	0.035	†
Supervisors	0.489 (.05)	-7.506 (.85)	-2.188 (.26)	2.217 (.26)			
Executives	0.560 (.08)	-8.388 (1.31)	-3.169 (.59)	1.396 (.37)			
Item70					0.054	0.004	173,967.40
Supervisors	2.154 (.09)	-2.583 (.11)	-0.528 (.04)	0.756 (.04)			
Executives	2.210 (.08)	-2.451 (.13)	-0.449 (.04)	0.855 (.03)			
Productivity					**0.378**	**0.003**	**2,100.08**
Item13					0.054	0.000	3,132.39
Supervisors	2.570 (.11)	-2.745 (.12)	-0.651 (.04)	1.363 (.05)			
Executives	2.997 (.10)	-2.363 (.13)	-0.594 (.03)	1.373 (.04)			
Item14					0.054	0.008	29,255.91
Supervisors	1.700 (.07)	-2.084 (.09)	-0.057 (.04)	1.648 (.08)			
Executives	1.852 (.06)	-2.274 (.12)	-0.298 (.04)	1.627 (.06)			

Table 9 (continued).

Scale and Item	*a* (SE)	b_1 (SE)	b_2 (SE)	b_3 (SE)	DFIT Cutoff Value	Index	Chi Square *df* = 2,099
Item48					0.054	0.002	11,399.12
Supervisors	1.667 (.08)	-4.210 (.36)	-1.636 (.08)	0.565 (.05)			
Executives	1.654 (.06)	-5.133 (.92)	-1.836 (.10)	0.727 (.04)			
Item54					0.054	0.000	8,531.37
Supervisors	2.222 (.09)	-3.367 (.24)	-1.238 (.05)	0.981 (.05)			
Executives	2.212 (.08)	-3.630 (.40)	-1.357 (.07)	0.967 (.04)			
Item65					0.054	0.001	2,195.39
Supervisors	2.981 (.12)	-2.710 (.14)	-0.881 (.04)	1.008 (.04)			
Executives	3.675 (.12)	-2.460 (.16)	-0.778 (.03)	0.969 (.03)			
Item71					0.054	0.003	3,363.21
Supervisors	2.926 (.11)	-2.158 (.08)	-0.438 (.03)	1.350 (.05)			
Executives	3.918 (.13)	-1.914 (.08)	-0.294 (.03)	1.347 (.03)			
Item74					0.054	0.000	6,481.61
Supervisors	2.358 (.10)	-2.798 (.14)	-0.906 (.04)	1.194 (.05)			
Executives	2.539 (.08)	-3.211 (.32)	-0.864 (.04)	1.213 (.04)			

Note. For all analyses, the executives served as the reference group and the supervisors served as the focal group. For all scales, the DFIT cutoff values and indices refer to the DTF index. For all items, the DFIT cutoff values and indices refer to the NCDIF index. All chi-square values are significant at the $p < .05$ level, except for those that are underlined. † =Chi-square values greater than 999,999.99.