WHAT ABOUT THE OTHER CORE EVALUATIONS? THE DEVELOPMENT AND VALIDATION OF A

MEASURE OF CORE-OTHER EVALUATIONS AND ITS

RELEVANCE FOR ORGANIZATIONAL

TEAMS

by

CHLOE J. LEMELLE

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2010

ACKNOWLEDGEMENTS

This dissertation was admittedly a massive undertaking and now serves as a remarkable accomplishment. To that end, there are numerous people that I would like to acknowledge. I would first and foremost like to thank God. Without guidance and direction from the righteous, I really do not know how I would be able to complete such an endeavor. I would also like to thank my family for showing love and support through the good times and the bad. If wealth were measured by family and friends, I would undoubtedly be as rich as Warren Buffett. To my wonderful husband, Kevin, for always being there to assist me and offering a shoulder to lean on. I would not be where I am today without his motivation, love, and support. To my mother, for her kind words of encouragement and her unwavering love and faith in me; she has provided me with enduring strength and confidence in my own abilities. To my father, for his continued prayers and unbridled enthusiasm for my accomplishments; he always seemed to revel in my successes more so than myself. And to my brother, for his never-ending belief in my success and constant support; it has always been my mission to make him proud. Also, a deep appreciation goes out to my extended family, in-laws, and step relatives. Family is my foundation and without them, I would surely collapse and fall. Their support, prayers, encouragement, and assistance have enabled me to rise against any obstacles attempting to thwart my progress.

I would also like to thank my wonderful friends who have shown tremendous support and assistance in this process. They are too numerous to enumerate, but I hope they know that they hold a special place in my heart. In addition, I would like to acknowledge Batrus Hollweg International (BHI) for allowing me flexibility to complete my graduate studies while working full-time and for serving as an incredible support system. Last, but not least, I would like to acknowledge the members of my committee. Without them, I would not be in the place that I am in today. Each of my committee members has been extremely flexible, helpful, and thought-provoking. I appreciate their wisdom and have learned immensely from each of them.

November 24, 2010

ABSTRACT

WHAT ABOUT THE OTHER CORE EVALUATIONS? THE DEVELOPMENT AND VALIDATION OF A

MEASURE OF CORE-OTHER EVALUATIONS AND ITS

RELEVANCE FOR ORGANIZATIONAL

TEAMS

Chloe Lemelle, Ph.D.

The University of Texas at Arlington, 2010

Supervising Professor: Shannon Scielzo

The construct of core self-evaluations has garnered a great deal of attention in the organizational psychology literature and has been hailed as one of the strongest predictors of job satisfaction. However, the construct of core self-evaluations is only one of the three core evaluations that Judge and colleagues originally described in their seminal work (Judge, Locke, & Durham, 1997). It is predicted that much can be explained about the dynamics of teams and the propensity to excel in a team environment through the study of core other-evaluations. With that objective, the current study developed and validated a comprehensive measure of core other-evaluations, referred to as the Core Other-Evaluations (COE) scale. The construction of the measure mirrored the theory of Judge and colleagues (Judge et al., 1997; Judge, Erez, Bono, & Thoresen, 2003), taking a comprehensive approach to scale development and validation. First, a series of pilot studies were conducted to develop the initial item subset, demonstrate support for the construct of core-other evaluations, examine the relations of the core-other evaluation sub-construct manifestations with the core-other evaluation items, and demonstrate support for the psychometric utility of the retained items for the core other-evaluations scale. Then a cross-validation study was conducted to further examine the retained items using IRT techniques, explore construct and criterion-related validity, and investigate the presence of group differences. The results indicated support

for the existence and efficacy of the core other-evaluations construct, the predictive power of core other-evaluations with regard to important work criteria, the incremental validity of core other-evaluations over core self-evaluations, and the lack of differential item functioning or group differences with this measure. Taken together, these findings bolster the argument that the construct of core other-evaluations is not only useful in a theoretical context, but also with regard to practical utility. Practical implications and future directions are discussed.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

How you feel about yourself has been shown to greatly affect how you feel about your job. In fact, how you feel about yourself at the core has been hailed as the single best predictor of job satisfaction by some individuals (c.f., Judge & Bono, 2001a). This concept of core self-evaluations and its relationship with job and life satisfaction was introduced by Judge, Locke, and Durham (1997) and later explicated by Judge and colleagues (c.f., Erez & Judge, 2001; Judge & Bono, 2001a; Judge & Bono, 2001b, Judge, Bono, & Locke, 2000; Judge, Erez, Bono, & Thoresen, 2003; Judge, Locke, Durham, & Kluger, 1998). Core self-evaluations are defined as "fundamental premises that individuals hold about themselves and their functioning in the world" (Judge, Erez, & Bono, 1998, p. 168).

Since its nascence, the construct of core self-evaluations (also referred to as CSE) has made a considerable impact in the industrial-organizational psychology realm. Not only has CSE research demonstrated strong relationships with job and life satisfaction (Judge et al., 1998; Judge, Bono, Erez, & Locke, 2005; Judge, Bono, & Locke, 2000; Piccolo, Judge, Takahashi, Watanabe & Locke, 2005), but CSE has also been linked to other important work outcomes such as job performance (Judge & Bono, 2001b), job burnout (Best, Stapleton, & Downey, 2005; Yagil, Luria, & Gal, 2008), job stress (Brunborg, 2008), work-family conflict (Boyar & Mosley, 2007), work motivation (Erez & Judge, 2001), goal self-concordance (Judge, Bono, Erez, & Locke, 2005), work success (Judge & Hurst, 2008), happiness (Piccolo et al., 2005), customer service (Salvaggio et al., 2007), salesmanship (Sager, Strutton, & Johnson, 2006), commitment to developmental goals (Bono & Colbert, 2005), coping with organizational change (Judge, Thoresen, Pucik, & Welbourne, 1999), job search persistence (Wanberg, Glomb, Song, & Sorenson, 2005), and income (Judge & Hurst, 2007). Moreover, CSE research has made a global impact with the development of a Dutch core self-evaluations scale (dePater, Schinkel, & Nijstad, 2007), a Korean version of the core self-evaluations scale (Holt & Jung, 2008), and a Japanese version (Piccolo et al., 2005). In fact, CSE researchers have claimed that conscientiousness and CSE are the two most

important personality traits in the prediction of human performance (Judge, Van Vianen, & dePater, 2004).

Judge, Locke, and Durham (1997) embarked upon the study of core self-evaluations in order to examine several potential disposition-related characteristics as predictors of job satisfaction. These variables (i.e., self-esteem, self-efficacy, locus of control, and emotional stability) comprise what is now collectively referred to as 'core self-evaluations'. However, the seminal research conducted by Judge et al. (1997) mentioned a tripartite model of core evaluations with the self being only one facet of the equation. The other areas included: core evaluations of others and core evaluations of the world. Whereas, core self-evaluations has claimed the attention of many researchers as of recent, core evaluations of others and core evaluations of the world have not ascertained the same level of academic scrutiny. In fact, the latter two have had minimal, if any, representation in the literature likely due to the fact that most attention has been captured by the study of the core self-evaluations construct, which is said to be the most fundamental of the three evaluations (Judge et al., 1997). So I then pose the question, what about the *other* core evaluations? To my knowledge, no research has been conducted regarding core evaluations of others or core evaluations of the world. Therefore, the focus of this research is to elucidate the underpinnings of core evaluations of others and to demonstrate how this construct can be useful for the study of work teams in organizations and beyond.

Teams and work groups are ubiquitous in today's organizations. In fact, there was a notable shift in the past couple of decades from individualized jobs to more team focused jobs (Kozlowski & Ilgen, 2006; Devine, Clayton, Phillips, Dunford, & Melner, 1999). Devine et al. (1999) stated that approximately half of the organizations in the U.S. utilize teams in some form or fashion. In today's society, it is seldom that a job does not include a component of teamwork or interpersonal work relations with others. Much of the research on teams has been devoted to classifying types of teams and examining team effectiveness (see Bettenhausen, 1991; Cohen & Bailey, 1997; Devine, 2002; Gully, 2000; Guzzo & Dickson, 1996; Guzzo & Shea, 1992; Hackman, 1992; Ilgen, Hollenbeck, Johnson, & Jundt, 2005; Kozlowski & Bell, 2003; and Sundstrom, McIntyre, Halfhill, & Richards, 2000 for reviews); however, little attention has been shown to the dispositional factors that individuals bring into teams that predispose them to be effective in a team environment. Helping organizations better understand an individual's propensity to be successful

in a team or interpersonal work environment can prove to be increasingly beneficial as jobs become more and more dependent on collaboration and effective group relations. Using core other-evaluations in relation to work teams can fill this void in the research. In addition, there are not many validated measures of team orientation at the individual level. Establishing a measure of an individual's likelihood to thrive in a team environment could be extremely useful for human resource practitioners and organizations looking to hire individuals based on their ability to work well within a team. The current study seeks to address both of these needs.

More specifically, this research aspires to: 1) provide a thorough review of core evaluations research, 2) expound upon a theoretical framework of core other-evaluations (guided by the research of Judge et al., 1997), 3) develop and validate a comprehensive measure of core evaluations of others using item response theory, 4) establish criterion-related validity with core other-evaluations and relevant constructs, and 5) determine the predictive ability of core other-evaluations over and beyond that of core self-evaluations. I will begin with a review of the relevant, extant literature on core evaluations.

<u>1.1 Core Evaluations</u>

Packer, a clinical psychologist, claimed that at the base of every appraisal is a fundamental evaluation that influences situation-specific appraisals (Packer, 1985; Judge et al. 1997). She called this deep-seated appraisal a "core evaluation" (Packer, 1985). She also mentioned that everyone holds these core evaluations subconsciously, and they are divided into three fundamental groups: core self-evaluations, core evaluations of others, and core evaluations of the world. This proposition is rooted in appraisal theory, which contends that "emotions are the form in which one experiences subconscious appraisals of objects, people, or events in relation to one's perceived values, needs, or commitments" (Judge et al., 1997, p. 157). She asserts that because these evaluations are core and all-encompassing, they affect other situationally-specific appraisals. Therefore, an individual who makes an appraisal of a certain situation as bad subconsciously holds negative views about the world, other people, or themselves that impacts their assessment of their current situation. Contrarily, a person who enters into the same situation and appraises it to be good likely does so because they hold fundamental positive views about the world, other people, or themselves. A more concrete, real-world example would be a person who gets confronted by someone they have never met when out in a social setting. Someone who

holds deep-seated negative views about the world, other people, or themselves may view this situation as frightening or unwelcomed. They may feel threatened or annoyed at the prospect of someone invading their personal space. Whereas, someone who holds positive core evaluations about themselves, the world, and other people may view this situation as innocent, non-threatening, and acceptable. They may feel eager and excited at the opportunity to potentially help someone or encounter a new acquaintance.

Rokeach (1972), a social psychologist, made a similar assertion by stating that individuals have "belief systems" that are fundamental values that people hold about themselves, the physical world, and the social world. This represents the same trichotomous breakdown that Packer (1985) used to describe core evaluations. Rokeach (1972) also contends that these belief systems are less easily changed than more peripheral traits and, thus, beliefs are more dispositional than situational. Unanimously proclaimed by researchers in two different disciplines is this notion of deeper, fundamental beliefs or evaluations that individuals hold about themselves, the world, and others. It is also deemed to be deep-seated and central and, thereby, affecting more peripheral traits. So, in other words, how one reacts to a specific situation may be due in part to the evaluations that this individual holds at their core. Both Packer (1985) and Rokeach (1972) suggest that these beliefs or appraisals are fundamental, evaluative, and broad in nature. They also contend that the evaluation of the self is the most central and, thus, affects all other evaluations. Because the core evaluation of the self is said to be the most central, this may explain the heightened attention that this area of core evaluations has garnered in academic research in comparison to the other two.

Using the core evaluations research as a guide, Judge et al. (1997) focused primarily on core self-evaluations and determined that the three important inclusion components of core self-evaluations were fundamentality, evaluation-focus, and scope. Following the same logic, these inclusion criteria were also applied to core other-evaluations. Judge et al. (1997) also focused on the relationship between core evaluations and job satisfaction and purported that there is both a direct and indirect relationship between these two variables at play. Each of the three components said to drive these central evaluations and why they would be related to job satisfaction will be discussed in turn.

1.2 Inclusion Criteria

Fundamentality assesses the centrality of a trait to one's self-concept. This further explains how core self-evaluations would be considered the most fundamental of the three core evaluations, seeing as how you feel about yourself is the most central evaluation you can make regarding your self-concept. Cattell (1965) made a clear distinction between peripheral or surface-level traits and source traits, where he indicated that source traits are more basic in nature. Because these traits are more basic and less specific, they are more able to get to the core of an individual's judgment. Judge et al. (1997) claimed that more fundamental traits have more connections to other traits than more peripheral traits. With this proposition, Judge et al. (1997) predicted that fundamental traits would significantly impact job satisfaction to a greater degree than that of less fundamental traits.

The second criteria, evaluation-focus, deals with whether the trait involves an evaluation or a description. It was suggested that evaluation-focused traits involve some form of self-evaluation as opposed to merely being descriptive. Judge et al. (1997) cited an example involving self-esteem and assertiveness. Whereas self-esteem requires an evaluative process, assertiveness is more of a description of an individual's behavior. For example, a typical question designed to assess self-esteem is, "Overall, I am satisfied with myself" (Judge, Erez, Bono & Thoresen, 2003), and a typical question designed to assess assertiveness is, "You speak out in meetings to oppose those whom you feel sure are wrong" (Guilford, Guilford, & Zimmerman, 1978). In the case of self-esteem, answering this question requires an introspective evaluation of thoughts about oneself. Whereas in the case of assertiveness, the question is centered more around the behaviors a person may or may not exhibit in a given situation. Judge et al. (1997) proposed that because job satisfaction is more evaluative in nature, then traits that are also more evaluative should influence job satisfaction to a greater extent than traits that are more descriptive in nature.

Lastly, the criterion of scope refers to the breadth of a trait. It was proposed that the broader a trait is, the more likely it is to relate to other traits. The example that Judge et al. (1997) used was the difference between a person's overall self-evaluation versus an instance of one's specific ability (e.g., athletic ability). In this example, overall self-evaluation is far broader in nature than athletic ability. Another example that is more related to evaluations of others is the difference between a trait like agreeableness,

which measures the extent to which an individual is able to get along with others, and a construct like team orientation, which measures the degree to which a person is inclined to work well in a team environment. The latter specifically refers to the team setting and is thus more limiting in nature, whereas, the former refers to a more all-encompassing view about other people. The proposition here was that because job satisfaction is a broader concept, it should correlate with other such broad concepts.

Using these inclusion criteria as a guide, Judge et al. (1997) focused primarily on which traits would exemplify core self-evaluations and included self-esteem, self-efficacy, emotional stability, and locus of control. Because the same process employed by Judge et al. (1997) to include relevant traits under the construct of core self-evaluations was used in the current study, it may be beneficial to briefly review core self-evaluations and the rationale behind including the traits that comprise the construct.

<div align="center">1.3 Core Self-Evaluation Traits</div>

1.3.1 Self-Esteem

Self-esteem describes a basic evaluation of an individual's self worth (Harter, 1990). Due to the fact that self-esteem is somewhat synonymous with self-evaluation, self-esteem is believed to be the most fundamental trait of core self-evaluations (Judge et al., 1997; Judge et al., 1998). Some have argued that self-esteem depends upon situational influences and results as an outcome of the situation an individual faces (e.g., personal success; Rosenberg, Schooler, & Schoenbach, 1989). However, others have claimed that self-esteem is fairly stable and shows little variability beyond adolescence (Costa & McCrae, 1994; Tharenou, 1979). Hence, self-esteem is essentially a dispositional trait that influences individuals to behave in a certain way but can also be affected by situational variables.

Pertaining to the inclusion criteria, self-esteem meets the objectives of fundamentality, scope, and evaluation-focus. Self-esteem is evaluative in the sense that it requires an appraisal of the self rather than description. Additionally, it has broad associations with behavioral, cognitive, and affective processes (Blascovich & Tomaka, 1991) and can, therefore, be related to many other traits. Furthermore, there is a great deal of empirical support for the relationship between self-esteem and job satisfaction (Locke, McClear, & Knight, 1996).

<div align="center">6</div>

1.3.2 Generalized Self-Efficacy

Bandura (1997) originally described self-efficacy as being task specific, but Judge et al. (1997) sought to broaden the scope of the concept. These researchers looked only at the generality dimension of Bandura's (1986) conceptualization and described generalized self-efficacy as an individual's ability to show control over their life events (Judge, Locke, Durham, & Kluger, 1998). Self-efficacy is actually a sub-component of self-esteem (Judge et al., 1998). It was stated that each of the core self-evaluation traits are not theoretically distinct, but yet are indicative a single, higher-order construct (Judge et al., 1998). Thus, one would expect these traits to be highly related. In relation to the inclusion criteria, generalized self-efficacy sufficiently meets the three criteria. It requires an evaluation of one's own capabilities. Also, being that it represents the competence aspect of self-esteem, it is fundamental to the self. Finally, it is broad in the sense that it is associated to many different constructs (e.g., self-control, problem solving, task performance, effort, etc.; Judge et al., 1997). In reference to job satisfaction, because generalized self-efficacy pertains to an individual's coping ability, it was predicted that those who are better able to cope through stressful situations are more likely to experience satisfaction at work (Johnson, Rosen, & Levy, 2007).

1.3.3 Emotional Stability

Emotional stability is one of the Big Five dimensions of personality and assesses an individual's likelihood to become emotionally unbalanced in stressful or challenging situations (Costa & McCrae, 1988; Dilchert, Ones, Van Rooy, & Viswesvaran, 2006). Individuals who score low on emotional stability are more prone to feelings of worry, vulnerability, and insecurity (Judge et al., 1998) than those who score high on emotional stability. It is often referred to as neuroticism, seeing as neuroticism is on the opposite side of the continuum. Judge et al. (1997) described neuroticism as the opposite of self-esteem. Because of this connection to one of the other CSE traits, Judge and colleagues further contend that core self-evaluation traits are really facets of a single, underlying construct. It was also proposed that those lower in emotional stability are more likely to jump to negative conclusions, therefore, leading to lower job satisfaction (Judge et al. 1997).

1.3.4 Locus of Control

Rotter (1966) distinguished locus of control into two components: internal locus of control and external locus of control. Internals believe that self-guided actions lead to their fate, whereas externals attribute peripheral or contextual factors to their successes or failures. Locus of control has been called into question regarding its adherence to all the inclusion criteria. It is said to be less evaluation-focused than the other traits. However, Judge and Bono (2001a) note that many of the items included in measures of locus of control involve self-evaluations (e.g., "My life is determined by my own actions"; Levenson, 1981). Additionally, Locus of control has been theoretically linked to other core self-evaluation traits (i.e., generalized self-efficacy). In an effort to differentiate the two traits, Judge et al., (1998) state that, "self-efficacy pertains to confidence with respect to actions or behaviors, whereas locus is more concerned with confidence in being able to control outcomes" (p. 19). It is believed that an individual's perceived ability to control one's destiny (i.e., internal locus of control) leads to higher levels of job satisfaction.

<u>1.4 Core Other-Evaluation Traits</u>

Shifting focus now to the traits that underlie core other-evaluations, the same inclusion criteria were employed to determine which traits would be most appropriate. Because the focus of the proposed research is on individuals' fundamental feelings about other people, the traits that fall within the inclusion criteria are trust, agreeableness, collectivism, and other orientation. These traits were chosen due to their adherence to the inclusion criteria and their overall representation of feelings towards others. Each area and its reason for inclusion will be discussed in turn.

1.4.1 Trust

Judge et al. (1997) included trust vs. cynicism as the main dispositional trait measuring core other-evaluations. They made reference to Erikson's (1950) research on child development and trust. In his theory of psychosocial development, Erikson (1963) described a series of eight stages that individuals face throughout development, beginning with Trust vs. Mistrust. In fact, one of the earliest conclusions a child makes in their life is whether or not someone can be trusted. Inherent in this theory of psychosocial development is the role that the mother or parental figure plays in developing the child's ability to trust or distrust others. However, several researchers assert that trust goes beyond infancy and childhood (Katz & Rotter, 1969; Judge et al., 1997). For instance, Judge et al. (1997) remark that trust not only refers to

the stage an individual reaches in infancy but also encompasses an individual's ability to trust whether adults can be rational in their interactions, "for example, explain things, give valid reasons for their actions, give valid explanations of the meaning of events, provide accurate knowledge of what is needed to succeed in the world, make the world predictable, offer valid moral principles to live by, and explain how to deal with other people" (p. 166).

Rotter (1971) distinguished between Erikson's (1950) idea of basic trust and what he deems interpersonal trust. He defines interpersonal trust as "an expectancy held by an individual or a group that the word, promise, or written statement of another individual or group can be relied on" (Rotter, 1971, p. 444). The idea behind interpersonal trust is based upon social learning theory principles. According to this theory, individuals will engage in certain behaviors based on the anticipation that they will experience a preferred outcome or reinforcement (Katz & Rotter, 1969). Therefore, if an individual trusts someone who turns out to be untrustworthy, then the conclusion that the individual may reach is not to trust people so easily. Rotter's (1967) conceptualization of trust is utilized in the context of the current study, because it is the most fundamental and broad representation of trust found in the extant literature, and it refers to trust as a dispositional trait rather than an aspect of relationships (Schoorman, Mayer, & Davis, 2007). Judge et al. (1997) described the converse of trust as being cynicism. Cynical individuals are highly wary about the motives of others and jump to negative conclusions regarding others' intentions more readily than someone who is trusting. In reference to the inclusion criteria, because trust is developed in childhood and well ingrained throughout adult life, it ostensibly meets the criteria of fundamentality. Additionally, trust involves the assessment of whether or not someone can be reliable in their words and actions and, thus, involves an evaluation. Furthermore, fundamental trust is broad in scope, because it can refer to numerous different social referents such as parents, friends, teachers, and co-workers. Trust not only affects personal relationships but also affects work relationships as well.

Trust is a valuable component in the workplace and can lead to many important work outcomes. Bennis and Nanus (1985) called for increased attention in the area of trust in the workplace. Trust shows strong relationships with job satisfaction (Chiaburu & Byrne, 2009; Dirks & Ferrin, 2002; Mulki, Jaramilla, & Locander, 2006; Requena, 2002; Aryee, Budhwar, & Chen, 2002; Willenbrock & Kauffeld, 2010), increased cooperation between co-workers and supervisors (Tyler, 2003), organizational commitment

(Albrecht & Travaglione, 2003; Tan & Tan, 2000), organizational citizenship behaviors (Chiaburu & Byrne, 2009; Robinson & Morrison, 1995), and team cohesion (Hansen, Morrow, & Batista, 2002; Mayer, Davis, & Schoorman, 1995; Ricketta & van Dick, 2005). It is expected that those variables indicative of core other-evaluations will be related to job satisfaction and team relationships and cohesion.

1.4.2 Agreeableness

The Big Five or Five-Factor Model of personality has been generally accepted by researchers as a parsimonious framework for classifying differences in personality (Costa & McCrae, 1992; Digman, 1990; Goldberg, 1990; Hough & Oswald, 2000; Tett, Jackson, & Rothstein, 1991). This model of personality includes the following traits: conscientiousness, agreeableness, extraversion, emotional stability, and openness to experience. Agreeableness refers to an individual's ability to show cooperation, kindness, courteousness, caring, help, and trust for others (Costa & McCrae, 1992), thereby standing out as a trait indicator of core other-evaluations. Individuals high in agreeableness are described as having a caring orientation and easy to get along with (Zhao & Seibert, 2006). Contrarily, individuals low in agreeableness are said to be unconcerned with the well-being of others, less likely to cooperate with other people, and more likely to put self interests ahead of the interests of others (Costa & McCrae, 1992). Agreeableness has also been linked to helping behaviors (Kamdar & Van Dyne, 2007). In fact, some researchers claim that agreeableness is the best personality predictor of helping behavior (Barrick, Stewart, Neubert, & Mount, 1998).

Of the five personality dimensions, agreeableness is also said to be the most reflective of an individual's interpersonal orientation (Graziano, Jensen-Campbell, & Hair, 1996). In other words, agreeableness describes a person's fundamental feelings toward others, which further necessitates its need for inclusion in the core other-evaluations construct. Moreover, Graziano et al. (1996) mention the link that agreeableness has to social evaluation, which fulfills the evaluation-focus inclusion criteria. Agreeableness also represents a broad dimension of personality, encompassing many different forms of interpersonal relationships from cooperation to trust to altruism. As stated previously, the traits that underlie core evaluations are not expected to be theoretically distinct, therefore, overlap between the traits is permissible (Judge, Locke, Durham, & Kluger, 1998).

Like trust, agreeableness also shows relevance to the workplace and, more specifically, to work teams. O'Neill and Kline (2008) report that agreeableness has shown numerous relationships with team outcomes. For example, Taggar (2000) found that agreeableness was positively related to conflict resolution. Similarly, Graziano et al. (1996) found that agreeable individuals were more prone to reducing interpersonal conflict. Additionally, Neuman and Wright (1999) reported that agreeableness was more effective than cognitive ability and individual skills at predicting teamwork performance and team communication. In regard to organizational teamwork, many studies have found positive relationships between agreeableness and teamwork (Halfhill, Sundstrom, Lahner, Calderone, & Nielsen, 2005; Kamdar & Van Dyne, 2007; Kline & O'Grady, 2009; O'Neill & Kline, 2008; Mount, Barrick, & Stewart, 1998). Those high in agreeableness are more adept at interpersonal facilitation (Hurtz & Donovan, 2000) and less likely to succumb to within-group competition (Graziano, Hair, & Finch, 1997). Furthermore, agreeableness has been shown to positively relate to individual-targeted citizenship behaviors, which are activities such as "altruism, helping, courtesy, cooperative behavior, and interpersonal facilitation" that are geared toward benefitting others on the team (Ilies, Fulmer, Spitzmuller, & Johnson, 2009).

Agreeableness has also been shown to predict job performance, but only in specific types of jobs (Barrick, Mount, & Judge, 2001). Zhao and Seibert (2006) suggest that high levels of agreeableness may be detrimental for managerial positions, because it decreases an individual's ability to make difficult decisions regarding their employees. Moreover, Seibert and Kramer (2001) found a negative relationship between agreeableness and salary level in a managerial population. In regard to job satisfaction, however, agreeableness has been shown to yield a positive relationship through meta-analytic research (Judge, Heller, & Mount, 2002). Additionally, because agreeableness measures an individual's interpersonal orientation (Graziano et al., 1996), it is expected that facets of job satisfaction that involve social interaction will be rated more highly for agreeable individuals than those that do not.

1.4.3 Individualism vs. Collectivism

Hofstede (2001) studied cultural variations and defined five dimensions by which different cultures can be classified. Of these five dimensions, individualism vs. collectivism refers to how much individuals define themselves by their group membership. Hofstede (2001) goes on to describe that individualistic cultures prefer autonomy, self-reliance, and independence. In contrast, collectivistic

societies prefer camaraderie, togetherness, and interdependence (Hofstede, 2001). Individualists often put their personal needs and desires ahead of the group and are less willing to sacrifice personal gain for the greater good of the group. Contrarily, collectivists value group interests over self-enhancing interests, typically look out for the best interests of the group, and are willing to sacrifice their own personal gain for the betterment of the group (Wagner & Moch, 1986).

Hofstede's (1980) original proposition referred to differences at the national level, and this conceptualization received a great deal of criticism. His critics contend that inherent in his supposition is the assumption that national cultures are homogenous, which is often not the case in modern society. He later clarifies that the cultural dimensions can be viewed within cultures rather than solely at the national level. Moreover, individualism vs. collectivism should not apply only to cultural differences and can be beneficial for other purposes as well (Gundlach, Zivnuska, & Stoner, 2006). Individualism vs. collectivism can be examined at the individual level (Colquitt, 2004; Colquitt, Noe, & Jackson, 2002; Cox, Lobel, & McLeod, 1991; Dickson & Weaver, 1997; Eby & Dobbins, 1997; Gundlach et al., 2006; Moorman & Blakely, 1995; Pillai & Meindl, 1998; Triandis, 1995; Triandis, Leung, Villareal, & Clack, 1985; Van Dyne, Vandewalle, Kostova, Latham, & Cummings, 2000; Wagner & Moch, 1996) or across teams (Bell, 2007; Chen, Chen, & Meindl, 1998; Cox, Lobel, & McLeod, 1991; Earley & Gibson, 1998; Kirkman & Shapiro, 2001; Koslowski, Gully, Nason, & Smith, 1999; Murphy, 1999; Perrewe & Spector, 2002; Wagner, 1995). In an effort to distinguish between the individual-level version of individualism-collectivism and the cultural-level version, Triandis et al. (1985) suggested using the terms idiocentrism (which refers to individualism) and allocentrism (which refers to collectivism). However, other researchers claim that Triandis et al.'s (1985) labels may be confusing and that the construct could be better described as "psychological collectivism" (Jackson, Colquitt, Wesson, & Zapata-Phelan, 2006). As far as the inclusion criteria go, Gundlach et al. (2006) remark that "individual differences in individualism-collectivism exist in the very essence of people's self-concept," thereby, fulfilling the requirement of fundamentality (p. 1609). Additionally, viewing individualism-collectivism as an individual-level variable implies an evaluation of the self rather than a description. Further, the construct of individualism-collectivism meets the scope criteria in that it can apply to a broad range of feelings about others and group relations.

Triandis (1995) argues that, because individualism and collectivism are not mutually exclusive, they should be classified as two separate dimensions. In other words, this argument states that an individual can be both high in collectivism and high in individualism at the same point in time. Other researchers share this belief that collectivism and individualism are not a part of a continuum but rather should be viewed as orthogonal, domain-specific constructs (Ho & Chiu, 1994; Oyserman, Coon, & Kemmelmeier, 2002; Triandis, 1995). Therefore, the proposed research will focus on individualism-collectivism at the individual level and specifically examine the construct of collectivism. With regard to the construct of collectivism, there have been many studies linking collectivism to both team relations (Bell, 2007; Gundlach et al., 2006; Jackson et al., 2006; Tyran & Gibson, 2008) and job satisfaction (Foley, Ngo, & Loi, 2006; Hui & Yee, 1999; Hui, Yee, & Eastman, 1995).

1.4.4 Other Orientation

Other orientation was originally conceptualized to better understand altruistic behaviors (Korsgaard, Meglino, & Lester, 1997). Applied to the organizational setting, other orientation was conceived to better understand why individuals would engage in helpful behaviors that fell outside the range of their job responsibilities (Korsgaard, Meglino, & Lester, 1997). In fact, McNeely and Meglino (1994) compared concern for others (a facet of other orientation) to dispositional helping and found a strong correlation. In addition, concern for others was found to be more predictive of extra-role behaviors directed toward helping others at work. With such an emphasis on helping behaviors, one may wonder how this trait differs from altruism. It goes beyond altruism in the sense that it is not only concerned with helping behaviors but also measures an individual's concern for others in general (Korsgaard, Meglino, & Lester, 1997).

The theory of other orientation draws upon research on behavioral decision making and self-interest and describes the consideration of the self and self-enhancement when making decisions (Lester, Meglino, & Korsgaard, 2008). It is defined as "the dispositional tendency to be concerned with and helpful to other persons" (Meglino & Korsgaard, 2004, p. 948). It was developed as a part of the movement in organizational and behavioral sciences that was interested in alternatives to the notion that human motivation is based solely on self-interest. In other words, it purports that some individuals are not motivated to take action merely by the thought of personal gain, but rather may consider the effect on

13

others when making decisions. Researchers posit that those higher in other orientation will be less likely to rely on self-interested or rational thought processing, instead being more likely to consider the potential consequences to others when determining a course of action (Lester et al., 2008). More specifically, as stated by Lester et al. (2008):

> The theory of other orientation specifies that rational and self-interested choices involve the conscious pursuit of personal goals, which involves weighing the consequences of one's actions. That is, such choice is preceded and determined by a consideration of costs and benefits to the self. The theory of other orientation posits that individuals higher in other orientation are less apt to consider potential consequences to the self when making choices and acting (p. 831).

High other orientation is said to be related to such traits as high agreeableness (Graziano, Jensen-Campbell, & Hair, 1996), higher concern for others (Ravlin & Meglino, 1987), and higher association with social groups (Markus & Kitayama, 1991). Other orientation also has many implications for the organizational setting. Because helping and cooperative behaviors are said to be critical for the effective functioning of an organization (Korsgaard, Meglino, & Lester, 1997), other orientation serves as a beneficial variable to examine in regard to organizational outcomes. Researchers have found that productive organizational behaviors such as prosocial behavior and organizational citizenship behaviors are more likely to occur with highly other-oriented employees (Brief & Motowidlo, 1986; Katz & Kahn, 1966). In addition, studies have found a link between other-oriented behaviors and job satisfaction (e.g. Meglino & Korsgaard, 2007).

<u>1.5 Core Other-Evaluations</u>

The four variables described above comprise what will be collectively referred to as core other-evaluations (heretofore referred to as COE; see Figure 1 for a visual description). More specifically, these four variables are thought to have a common core which explains how people fundamentally feel about other people. It is recognized that there is some overlap between some of the variables. The same overlap was encountered with Judge, Locke, and Durham's (1997) construction of core self-evaluations. The rationale is that the construct of core other-evaluations is assumed to be a latent construct that underlies the included variables and cultivates their interrelationships. So, in other words, the commonality that is shared between these traits is said to be accounted for by core other-evaluations.

14

Someone who is high in core other-evaluations is said to trust the words and intentions of others, value being around other people, put others' needs above their own, and get along with others. In essence, this fundamental construct is claimed to be responsible for specific traits that manifest from it (i.e., agreeableness, trust, other orientation, and collectivism). It is not the proclamation of this paper to assert that these are the only manifest variables of core other-evaluations. In fact, I concede that there may be other appropriate trait manifestations of core other-evaluations out there. The claim is that, after scouring the literature and carefully considering the inclusion criteria, these traits emerged as being highly representative of core other-evaluations.
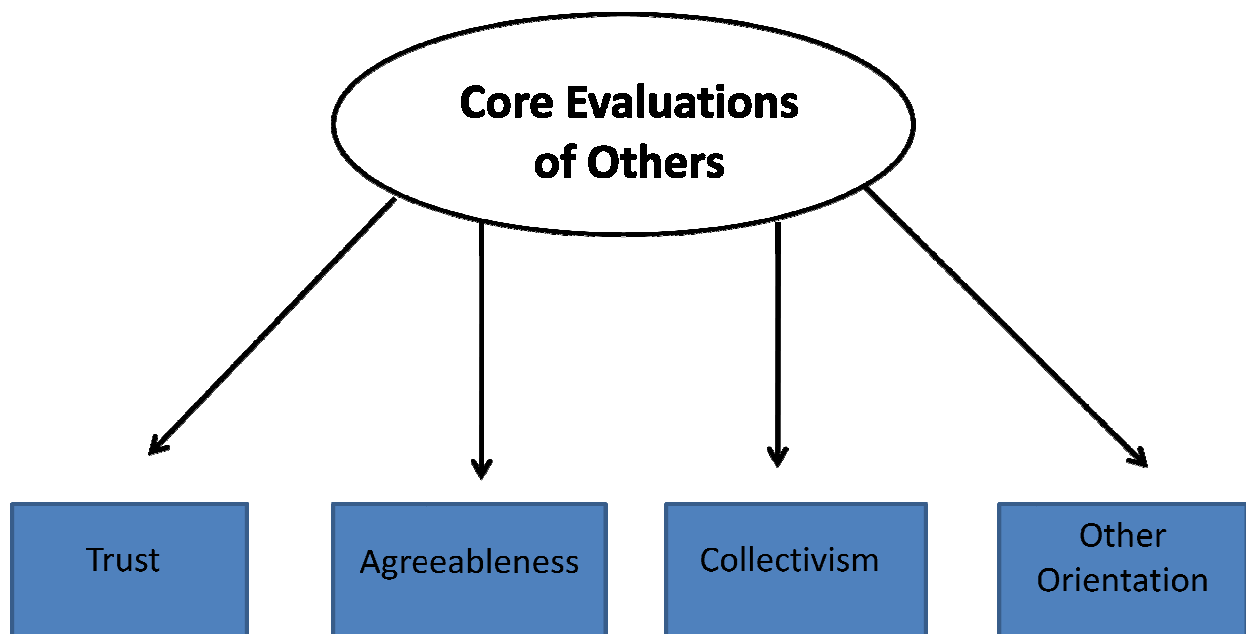
Figure 1. Visual Description of the Core Evaluations of Others Variables

Judge et al. (2003) noted that, though the associated traits of core self-evaluations are highly interconnected, they are not completely redundant. Each provides information that is unique and important. The same is true for COE. The assertion is that the redundancy of these traits (or the ways in which they are connected) can be explained by core other-evaluations. Also, COE is not limited to only the four traits described.

Judge and colleagues maintained that the construct of core self-evaluations is representative of a single, unidimensional construct (Bono & Judge, 2003, Erez & Judge, 2001; Heller, Judge, & Watson, 2002; Judge & Bono, 2001a; Judge, Bono, & Locke, 2000; Judge et al., 1998; Judge, Erez, Bono, & Thoresen, 2002, 2003) and should be conceptualized as such. Judge, Erez, Bono, and Thoresen (2002) examined the inter-relationships between self esteem, generalized self-efficacy, locus of control, and emotional stability in their meta-analysis to determine how closely related these traits were with each other. They found that the average (absolute) correlation between the traits was .60 (after being corrected for unreliability). Judge and Bono (1999) performed a similar analysis using 18 studies and found an average corrected correlation of .64. They note that this value is comparable to the average correlation between different measures of neuroticism (Digman, 1990). These strong relationships bolster the argument that CSE traits may be part of a higher-order factor. In fact, there have been numerous empirical studies that have found support for the notion that these variables comprise a super-ordinate factor. (e.g., Bono & Judge, 2003, Erez & Judge, 2001; Heller, Judge, & Watson, 2002; Judge & Bono, 2001a; Judge, Bono, & Locke, 2000; Judge et al., 1998; Judge, Erez, Bono, & Thoresen, 2002, 2003). Because the current study used the work of Judge and colleagues as a guide to developing the measure of core other-evaluations, a similar conceptualization of core other-evaluations was followed.

1.5.1 Construct Validity of COE

Construct validity determines whether the developed scale accurately measures the construct in which it is designed to measure (Cronbach & Meehl, 1955). In regard to core other-evaluations, the COE scale would show construct validity if it is indeed measuring the construct of core other-evaluations as it purported to do. Reliability is a necessary, but not sufficient component of this process (Judge et al., 2003). One method of establishing construct validity is to define the nomological network of the construct. Cronbach and Meehl (1955) explain that in order "to validate a claim that a test measures a construct, a nomological net surrounding the concept must exist." (p. 291). A common way to develop a nomological network is by examining convergent and discriminant validity. These components of the network work in tandem and cannot be examined in isolation. If a measure is shown to exhibit both convergent and discriminant validity, then it said to demonstrate evidence of construct validity. Each of these components of the nomological network is discussed in further detail.

1.5.1.1 Convergent Validity

As a means of evaluating construct validity and establishing a nomological network, convergent validity will be evaluated between COE and several key variables. Convergent validity is obtained when measures purporting to measure the same or a similar construct are statistically related (Campbell & Fiske, 1959). So, in other words, if the COE measure shares variance with a like measure, then COE is said to show convergent validity. The most apparent measure of convergent validity would be with the four indicator variables (i.e., trust, agreeableness, other orientation, and collectivism). A network between COE and its trait manifestations (i.e., trust, agreeableness, collectivism, and other orientation) has already been theoretically proposed. To add empirical evidence to this theoretical claim, statistical relationships between the COE measure and these trait manifestations will be explored. It was stated that COE exists in the same psychological space with these aforementioned variables, but resides at a far deeper, more meta-physical place. This deeper residence adds to the assertion that core other-evaluations are individuals' fundamental thoughts about other people. Because trust, agreeableness, other orientation, and collectivism are expected to be manifest traits of COE, they should be highly correlated.

Hypothesis 1: The COE measure will show a positive correlation with trust, agreeableness, collectivism, and other orientation.

Additionally, the Team Player Inventory (TPI; Kline, 1999) will be used to assess convergent validity. When scanning the literature for other like measures, the TPI was the closest measure found to COE. The TPI claims to measure an individual's predisposition toward working in a team-based environment (Kline, 1999). The two measures share a common goal in measuring a person's proclivity to fare well in a team environment, but also have some marked differences. One difference is that the TPI only focuses on predicting likelihood for success in team-based work environments, whereas COE is broad in nature and is, therefore, not limited by its practical application. COE can apply to the organizational setting, but can also go beyond work teams by describing an individual's likelihood to thrive in any group setting (e.g., school, sports teams, or other social groups). Additionally, COE applies not only to groups, but also to individual interactions with others. It describes whether someone can forge relationships easily with others or whether someone has difficulty with these types of interactions and

17

tends to be more distant and/or wary. Another distinction between the COE scale and the TPI is that, unlike the TPI, COE is grounded in solid theory and supportive research (i.e., CSE). The TPI lacks a theoretical foundation and bases its suppositions off mere conjecture or intuitive expectations. Despite its limitations, the TPI will still serve as a valuable measure of convergent validity. However, efforts will also be made to show the utility of the COE measure beyond that of the TPI.

Hypothesis 2a: The measure of COE will positively relate to the Team Player Inventory.

Hypothesis 2b: The measure of COE will show incremental validity beyond the Team Player Inventory in the prediction of work outcomes.

1.5.1.2 Discriminant Validity

Contrary to convergent validity, discriminant validity is concerned with whether or not a construct is statistically different from a construct to which it is claimed to be conceptually distinct (Campbell & Fiske, 1959). In other words, discriminant validity is realized if COE differs from scales that are not designed to measure the same construct. Judge et al. (2003) explored the discriminant validity of core self-evaluations with Big Five personality traits (e.g., conscientiousness, agreeableness, and openness) and found a theoretical distinction. Using their research as a guide, the current study seeks to explore the relationship between COE and these Big Five personality traits (with the obvious exclusion of agreeableness, seeing as it is proposed to be highly related to COE). In their research, Judge et al. (2003) found a moderate correlation with conscientiousness and little or no correlation with openness. The same effects are expected in the current study. Because of the altruistic aspects of COE and the inherent desire to go above and beyond in conscientiousness, it is likely that there is some relationship between conscientiousness and COE. Those who are more inclined to go out of their way for others are likely to also be highly conscientious (Organ & Konovsky, 1989). That is why these two constructs are expected to be moderately correlated. Even so, it is predicted that the two constructs will remain distinct and distinguishable. Moreover, the relationship between COE and its manifest variables (trust, agreeableness, collectivism, and other orientation) is predicted to be higher than the relationship between conscientiousness and COE. The same prediction was made with regard to openness.

Hypothesis 3: COE will have higher relations with trust, agreeableness, other orientation, and collectivism than it will with conscientiousness or openness.

18

Another measure of discriminant validity will be core self-evaluations. Like conscientiousness, COE and CSE are expected to be somewhat related to each other due to the fact that core self-evaluations are said to influence other core evaluations (Judge et al., 2003). Additionally, seeing as both CSE and COE are a part of the same tripartite model, they are likely to share some connection with each other. Nonetheless, the two constructs are thought to be conceptually diverse. Therefore, the following hypothesis will be tested:

Hypothesis 4: COE will have higher relations with trust, agreeableness, other orientation, and collectivism than it will with core self-evaluations.

1.5.1.3 Psychometric Approach to Construct Validity

With regard to the development of the CSES (Core Self-Evaluations Scale, Judge et al., 2003), the researchers used classical test theory (or CTT) methods to guide their development. To my knowledge, no efforts have been made to examine CSE using more modern analytic approaches (e.g., item response theory; IRT). IRT offers several advantages to CTT in regard to scale development and evaluation, and is growing increasingly more popular with psychologists and social scientists (Embretson & Reise, 2000). Due to its increasing application in the literature and well documented advantages over CTT in terms of scale construction and development, IRT may be suited to address some of the previously mentioned concerns surrounding the structural conceptualization of the COE construct.

1.5.1.3.1 Item Response Theory

Item response theory (IRT) focuses on latent trait functioning and has also been referred to as latent trait theory by some researchers. Ayala (2009) defines IRT as "a system of models that defines one way of establishing the correspondence between latent variables and their manifestations" (p. 4). Ayala (2009) further explains that the use of the word "theory" when describing this statistical technique is not the use of the word as most people know it to be. It does not explain why people behave the way that they do, rather it refers to a statistical estimation of how a person might respond to items designed to tap into a particular construct.

The theory came about in an effort to explain the differences between ability scores and true/observed scores. Ability scores are said to be independent of the test used to measure them, whereas true scores or observed scores are said to be test dependent (Hambleton & Jones, 1993). Lord

and Novick (1968) described this distinction in their book, *Statistical Theories of Mental Test Scores.* Rasch (1960) was also influential in defining the parameters of the linear logistic model (Embretson & Reise, 2000). As such, the one parameter model of IRT is more commonly referred to as the Rasch model. Numerous other psychometricians and researchers (e.g., Birnbaum, Bock, and Samejima, to name a few) were instrumental in the development of IRT. However, the purpose of this section is not to provide history of the development and use of IRT, but rather to describe some of the fundamentals of using IRT for scale development and some of the advantages it provides.

1.5.1.3.2 Fundamentals of IRT

IRT predicts individual ability levels based on item scores. Ability level is most commonly referred to by the Greek letter theta (θ) and is sometimes depicted on a z-score scale. In an effort to estimate theta and explore the psychometric properties of each item, IRT fits statistical models to individual response data. This is often done by estimating parameters. One parameter is the item difficulty parameter (or "b" parameter). This estimates the probability of an individual endorsing an item given their ability level. This parameter represents the location on the theta scale where the slope is the steepest. So, individuals with a high ability level on a particular construct will be more likely to endorse a more difficult item than someone with a lower ability (Embretson & Reise, 2000). If a COE item is determined to be "difficult", then only those individuals who possess high levels of COE will be likely to endorse that particular item. The other parameter is the discrimination parameter (or "a" parameter), which is similar to an item's item-total correlation in CTT. Items with high discrimination levels differentiate individuals differing in ability better than items with low discrimination levels (Reise, Ainsworth, & Haviland, 2005).

1.5.1.3.3 Advantages of IRT

Assuming that all of the IRT assumptions have been met, one of the advantages of IRT is its ability to provide item-level statistics that are independent of the sample from which they are derived. CTT scores, on the other hand, depend entirely upon the sample from which they are gathered (Hambleton & Jones, 1993). More specifically, CTT item discrimination statistics depend on whether the sample is heterogeneous or homogenous, with heterogeneous samples showing stronger discrimination and homogenous samples showing lower discrimination. Likewise, CTT item difficulty statistics depend on whether the sample has a high ability in the targeted construct versus a low ability. For instance, item

difficulty scores will be higher if individuals have a high ability, whereas, difficulty scores will be lower if individuals have a lower ability. Contrarily, IRT discrimination and difficulty indices are not affected by the sample from which they are drawn (Hambleton, 1989).

Another advantage of IRT is that it is test independent. As previously mentioned, IRT focuses on the latent scores, which are independent of the actual test or measure; whereas, CTT focuses on true scores based on observed scores, which are test/measure dependent. The fact that IRT examines ability at the item level enables researchers to create equivalent forms of tests that are matched on ability level. Equating items is possible using IRT due to the fact that IRT provides theoretical clearance for doing so. In essence, individuals can take completely different tests, but can still be matched on a common scale. This is a very important advantage in regard to test security. Because individuals can be given different tests (constructed using IRT) that can be equated on a common scale, replication and distribution of one test would not be detrimental. This also applies in regard to computer adaptive testing. IRT methods allow for such a practice where CTT methods would not. Computer adaptive testing (CAT) is most commonly used in the educational arena with ability tests such as the SAT or GRE. CAT allows flexibility in test-taking such that tests can actually be tailored or adjusted depending on the individual's particular ability level. In other words, if someone has a high ability, they can receive different questions that tap into their ability level than someone who has a lower ability level (Hambleton & Jones, 1993).

A third advantage of IRT is that it allows for examination of item functioning across groups. Differential item functioning (DIF) determines whether items on a particular measure are showing variance across groups. For example, DIF will denote whether an item is performing differently for one group (e.g., women) as opposed to another group (e.g., men). Because previous research has shown gender and race to be moderating variables in the relationship between CSE and job satisfaction (Lemelle & Scielzo, 2010), it is important to examine differential item functioning in this study in regard to COE. If measurement invariance is established, then the measure of COE would be considered safe to use across groups without the possibility of spurious results. Vandenberg and Lance (2000) called for increased attention to measurement invariance in their synthesis paper on the topic. They mention that it is essential to establish measurement invariance/equivalence before testing hypotheses due to the fact that premature examination of hypotheses regarding group differences may lead to faulty conclusions if

groups are in fact looking at the measure in psychometrically different ways. The presence of measurement invariance can be determined if item characteristics across different groups are the same. If item characteristics are not the same, then there is said to be differential item functioning, which could potentially show item bias. Item bias only becomes evident when differential item functioning that is said to be adverse as opposed to benign (Shealy & Stout, 1993). In other words, differential item functioning can exist without the item showing bias. DIF will be explored for gender, race, and age in the proposed study. Because the measure of COE will be suggested for use by organizations in selection, development, or team building capacities, it is imperative to ensure that included items are not showing bias across protected demographic groups.

1.5.1.3.4 Limitations and Assumptions of IRT

Even though IRT has some marked advantages over CTT, it is not without limitations (Hambleton & Jones, 1993; Reise, Ainsworth, & Haviland, 2005). One of the difficulties of using IRT is that it comes with much more stringent assumptions than CTT; thereby, making it more difficult for researchers to meet the required assumptions. In fact, some researchers have asserted that the assumptions of CTT are weaker and easier to meet than those of IRT (Hambleton & Jones, 1993).

The first assumption of IRT involves unidimensionality of the construct; hence, the construct under examination needs to be unidimensional in order to meet the requirements of IRT (unless a multi-dimensional IRT model is used). In other words, the shared variance between the items should be attributed to an underlying single dimension. Some psychometricians have challenged this assumption, claiming that it is too rigid for the purposes of most applications (Bryant & Wooten, 2006; Reckase, Ackerman, & Carlson, 1988). As a result, the assumption has been made more liberal by specifying that it be essentially unidimensional (Stout, Habing, Douglas, & Kim, 1996) or intentionally multidimensional (Reckase, 1985). The concept of essential unidimensionality suggests that multidimensional items may be used in cases where a dominant trait is present (Bryant & Wooten, 2006). Because COE is purported to be a unidimensional construct, this assumption and accommodation are highly appropriate for the purposes of the proposed scale development.

Research Question 1: Do the items of COE appear to form a unidimensional construct?

The second assumption of IRT pertains to local independence. If this assumption is met, items are said to show no other covariance with each other aside from their connection to the underlying latent dimension (Scherbaum, Finlinson, Barden, & Tamanini, 2006). In other words, the latent variable should be the only factor that links item together. Items that covary with each other aside from their relationship to the latent variable are said to violate the assumption of local independence. Researchers have stated that if the unidimensionality assumption is met, then the assumption of local independence is also met (Hambleton, Swaminathan, & Rogers, 1991).

A third, less explicit, assumption of IRT is that the model appropriately fit the data. If the previous assumptions are met, then further IRT analyses are permissible. The next step in the process would be to determine which IRT model best applies to the data. For instance, there are some IRT models that correspond to dichotomously scored items, whereas others are more appropriate for polytomous items. Because the COE items use a Likert-type response scale, with each response option being progressively higher than the previous, a model that takes into account ordered responses would be appropriate. For instance, with the COE scale items, the response option of "1" indicates that the individual strongly disagrees with that item. As the response options increase in number, so does the individual's endorsement of that item. This explains an ordered response. Also, because the data is polytomous in nature, dichotomous models that are often used with binary items would not apply. Based on this assessment, the model that covers both polytomous items and items that have ordered responses would be Samejima's (1969) Graded Response Model (GRM). GRM is often used for organizational research purposes where Likert rating scales exist (Scherbaum et al., 2006). Item characteristic curves and information functions for each item will be examined to determine item quality. In addition, test information functions will be examined to determine the value of the scale. Finally, model fit will be assessed.

Research Question 2: How well do the items of the COE scale fit the proposed model?

Another limitation of using IRT is that it requires relatively large sample sizes depending on the IRT model used. The sample sizes required often exceed what organizational researchers are typically used to obtaining (Scherbaum et al., 2006). Some models require upwards of at least 500, whereas others can be performed with a smaller sample of around 200 (Hambleton & Jones, 1993). Because it is not always feasible in organizational research to obtain this amount of data, IRT analyses may not be

possible in some cases. However, given a sufficient enough sample size, IRT methods can provide more robust analyses and a deeper understanding of scale functioning. An additional limitation of IRT is that it more complex in nature than CTT and requires different statistical software. IRT models have been said to be more time consuming and more difficult to perform (Scherbaum et al., 2006). Again, there is a trade-off here. With more time and more sophisticated techniques comes greater output.

In conclusion, IRT is a method of test construction and analysis that many psychologists are starting to adopt (Embretson & Reise, 2000). IRT offers numerous advantages to CTT and provides the ability to examine differential item functioning across groups. Given the advantages and more robust data output that IRT provides, the current study used IRT as the statistical method of choice for test construction.

1.5.2 Criterion-Related Validity of COE

Another important step in supporting the claim that a newly developed scale shows merit is establishing validity with a targeted criterion. Unlike construct validity, which is more focused on determining whether the scale measures the construct appropriately, criterion-related validity is more concerned with the predictive ability of the construct in question. To address this issue of predictive ability, several criteria will be explored. Firstly, Judge et al. (1997) predicted that COE would positively influence job satisfaction, so this study seeks to empirically test this proposition. Secondly, this study seeks to explore the relationship between COE and team relations. It is predicted that because COE implies an implicit liking for people, it would positively influence team satisfaction and team effectiveness. In addition, it is predicted that COE will show benefit beyond that of CSE; therefore, the incremental validity of COE will be explored. Each of these areas will be briefly discussed.

1.5.2.1 COE and Job Satisfaction

The theoretical mechanisms at play that link CSE to job satisfaction can also be referenced when relating COE to job satisfaction. Judge et al. (1997) reviewed the three theoretical techniques by which all three core evaluations could influence job satisfaction. One of their theories was emotional generalization. Through this process, individuals transfer their core feelings to other areas in their life, such as their jobs. Thus, in the case of COE, someone who is trusting, accepting, caring, and willing to put others' needs before their own will likely generalize their approving views about others to the workplace. Therefore, they

will be more likely to approve of their jobs as they do of other people. Contrarily, someone who is distrustful, cynical, disagreeable, and self-serving will likely transfer these negative views to their place of employment and, thus, be unsatisfied with their jobs. Additionally, because the work setting can be viewed as a highly social domain, COE can be easily transferred to this type of environment.

Another theory proposed by Judge et al. (1997) was that core evaluations influence the way in which the job is evaluated. They provide an example of an individual who is rewarded with a challenging new assignment with increased responsibility and a guaranteed pay raise at the end of six months. To justify the pay delay, the organization claims that the six month duration would give the employee time to prove that they could handle the job. Someone who is high in COE likely trusts the organization and may feel that the opportunity is fair, whereas a person low in COE may distrust the organization's intentions and may see this new opportunity as a means to exploit their talent. Hence, the satisfaction felt by the person with low COE would be lower than that of the person with high COE.

Another example could involve an organization that has just undergone a company-wide layoff and are trying to maintain similar levels of past productivity with a reduced work force (i.e., do more with less). Existing employees are encouraged to work harder and collaborate more in order to reach targeted goals. This example is especially fitting given the current state of economic affairs in the United States. An employee who is low in COE may dislike the prospect of having to work more intimately with other co-workers and may resent the organization for the decisions that it has made – potentially chalking it up to the fact that the organization cannot be trusted. In contrast, someone who is high in COE likely trusts the judgment of the organization, feels secure in the fact that they still have a job, and welcomes the idea of increased collaboration and cooperation with others.

Another theoretical explanation for the link between core evaluations and job satisfaction suggested by Judge et al., (1997) is defined by the actions that people take. Someone with a low regard for other people may be more withdrawn or disgruntled and, therefore, is less likely to form strong working relationships with co-workers than someone with a high regard for others. Their inability to connect with people in the workplace may cause them to feel out of their element at work, and they may hold negative feelings toward their place of employment as a result. On the opposite end, someone with a high regard

for other people may seek out opportunities to get to know others in the workplace and may look forward to interpersonal interactions at work, thereby increasing their level of satisfaction with their job.

As previously stated, empirical evidence shows that the traits that underlie core other-evaluations have strong relationships with job satisfaction. Trust (Dirks & Ferrin, 2002; Mulki, Jaramilla, & Locander, 2006; Requena, 2002; Aryee et al., 2002; Willenbrock & Kauffeld, 2010), agreeableness (Judge, Heller, & Mount, 2002; Ilies, Fulmer, Spitzmuller, & Johnson, 2009; Organ & Lindl, 1995), collectivism (Foley, Ngo, & Loi, 2006; Hui & Yee, 1999; Hui, Yee, & Eastman, 1995), and other orientation (Meglino & Korsgaard, 2007) all show a positive relationship with job satisfaction. Therefore, it is predicted that the construct of core evaluations of others will show a positive relationship with job satisfaction as well.

Hypothesis 5: Core other-evaluations will show a positive relationship with job satisfaction.

1.5.2.2 COE and Team Satisfaction

Because all the traits that underlie core evaluations of others involve interpersonal affiliation, it is plausible to conclude that core other-evaluations will show a strong relationship to team satisfaction. The rationale behind this thought process brings us back to the research of Packer (1985) and the theoretical roots in appraisal theory. According to Packer (1985), an individual's appraisal of a certain situation is guided in part by their core beliefs on the matter. In other words, situation-specific appraisals are affected by deeper, fundamental appraisals. Because COE is designed to be representative of fundamental, evaluation-focused, broad assessments; it is believed to lie at the core of many other situation-specific appraisals. For instance, when an individual determines their stance on a certain interpersonal encounter, they subconsciously refer to their core beliefs about others when determining their course of action. Given the interconnectedness that COE has with numerous subsequent social interactions, it is logical to think that COE can influence an individual's appraisal of their work team.

The empirical, existing literature is in agreement with this theoretical stance. The traits that underlie the construct of COE show strong relationships with teamwork and team relations. Agreeableness has been shown to be a valid predictor of teamwork and customer service (Mount et al., 1998). Additionally, the interpersonal relationships of disagreeable people are often defined by disagreement, conflict, and discord (Graziano, Jensen-Campbell, & Hair, 1996; Jensen-Campbell, Gleason, Adams, & Malcom, 2003; Jensen-Campbell & Graziano, 2001). In regard to other orientation,

26

those individuals said to have other-oriented values and beliefs are more likely to pursue affiliation with social groups (Markus & Kitayama, 1991). Also, many studies have linked collectivism (Bell, 2007; Gundlach et al., 2006; Jackson et al., 2006; Tyran & Gibson, 2008) and trust (Hansen, Morrow, & Batista, 2002; Mayer, Davis, & Schoorman, 1995; Ricketta & van Dick, 2005) to team relations. Given the quantity of research associating the scales that underlie COE with team relations, the following hypothesis is proposed:

Hypothesis 6: Core other-evaluations will show a positive relationship with team satisfaction.

1.5.2.3 COE and Team Effectiveness

Team effectiveness is another criterion that may be heavily influenced by core other-evaluations. So, beyond just the satisfaction that an individual derives from being in a team, it is also predicted that COE may affect how collaborative and cooperative an individual operates within a team. This is similar to the distinction between job satisfaction and job performance, which was explored with regard to CSE (e.g., Judge & Bono, 2001b). In this study, an effort was made to explore the relationship that COE has in reference to the performance of a team (i.e., team effectiveness). It is the belief that high levels of COE indicate better cooperation, collaboration, and overall team effectiveness due to the fact that people high in COE trust others, enjoy working with others, and put the needs of the team ahead of self-serving interests. On the other hand, those low in COE are often unconcerned with the well-being of others, cynical of peoples' ideas and intentions, and likely to put their own needs ahead of the group. With the objective of further determining the criterion-related validity of COE, the proposed study seeks to examine the functioning of teams and team effectiveness by surveying individuals in the field who work in some sort of team capacity. COE levels of individual team members will be collected and team outcomes (e.g., output effectiveness, interpersonal effectiveness, and overall effectiveness) will be explored. More specifically, it is believed that teams with higher levels of COE will also have stronger team outcomes. It is also believed that teams with less variability in COE scores will have better team outcomes.

Hypothesis 7a: Higher team COE averages will lead to higher team outcomes.

Hypothesis 7b: After controlling for team COE averages, teams with less variability in COE scores will have higher team outcomes.

1.5.2.4 Incremental Validity Beyond Core Self-Evaluations

Numerous studies have linked core self-evaluations to job satisfaction (Lemelle & Scielzo, 2010) and found a strong relationship. In fact, the average corrected correlation between CSE and job satisfaction was ρ_{corr} = .28. Furthermore, with the exclusion of one of the studies, the correlation increased to ρ_{corr} = .47, which represents a strong effect (Lemelle & Scielzo, 2010). It is undeniable that CSE contributes strongly to the prediction of job satisfaction. It is the contention of this paper that COE will have added value beyond that of CSE, particularly in the areas of job satisfaction that are more interpersonal in nature. For instance, satisfaction with supervisor and satisfaction with co-workers may be strongly influenced by a person's core beliefs about other people. It is predicted that core other-evaluations will show incremental validity beyond that of core self-evaluations in the facets of job satisfaction that involve relationships with other people. The rationale behind this notion is that a person's fundamental feelings about others are likely to affect the relationships that an individual forms with his/her supervisor and co-workers to a greater degree than their core beliefs about themselves. Therefore, given this basis, the following hypothesis will be tested:

Hypothesis 8: Core other-evaluations will show a relationship with the interpersonal facets of job satisfaction (i.e., satisfaction with supervisor and satisfaction with co-workers) over and beyond that of core self-evaluations.

Furthermore, it is believed that COE has unique value over and beyond that of CSE when it comes to the prediction of team satisfaction and team cohesiveness in groups. Because the manifest variables of COE are highly related to team satisfaction and cohesiveness, it is predicted that COE will show the same relationship. Even though CSE is likely to be related to team satisfaction as well, the hypothesis is that COE will show strength of prediction above and beyond that of CSE.

Hypothesis 9: Core other-evaluations will show a relationship with team satisfaction over and beyond that of core self-evaluations.

1.5.3 Content Validity

Content validity is another important consideration when trying to establish validity with a new measure. Also known as logical validity, content validity determines whether the scale covers the broad conceptualization of the construct in question (Schwab, 1980). One aspect of content validity involves a

more subjective evaluation of whether the items appear to be measuring the appropriate construct. This form of validity is known as face validity. Face validity, in and of itself, is not sufficient for establishing content validity. Because face validity is more subjective and less scientific, it is important to apply a more rigorous process when determining content validity. In this vein, the proposed study used conservative methods and enlisted the aid of subject matters experts in order to ensure that the requirement of content validity was met. The following sections describe the item generation and subject matter expert process.

1.5.3.1 Item Generation

When developing items for the COE scale, a content analysis of the traits said to comprise COE was reviewed for structure and details. In addition, the inclusion criteria detailed earlier (fundamentality, evaluation-focus, and scope) were applied to the individual items developed. Therefore, I sought to verify that the item met the criteria of being central and broad in nature and involving an evaluation rather than a description. It is important to note that all COE items were developed independent of the already established measures of trust, agreeableness, collectivism, and other orientation. In fact, COE items were written prior to even obtaining direct measures of the manifest traits. An effort was also made to include items that touched on more than one of the COE dimensions as well as items that transcended across all four areas. The goal of the process was to create more items than necessary so that poor performing items could be easily deleted without necessitating further item generation. Also, due to the fact that the Core Self-Evaluations Scale (CSES; Judge et al., 2003) contains 12 items, an effort was made in subsequent studies to reduce the final COE scale down to a comparable number. Based on this process, 40 items were written and reviewed by the item writer (see Appendix A for a list of all pilot items). As a further measure of refinement with the items, subject matter experts were called upon to examine each item's adherence to the inclusion criteria.

1.5.3.2 Subject Matter Experts

Item development and test construction experts from a small industrial-organizational consulting firm in the Dallas area were called upon to participate in the subject matter expert (SME) review process. During this process, SMEs were asked to rate each item on the three aforementioned inclusion criteria (fundamentality, evaluation-focus, and scope). SMEs were given detailed instructions on what each dimension measures, the definitions of all the inclusion criteria, and the manner in which their ratings

29

were to be performed (i.e., select the numerical rating that applies). An Excel spreadsheet was sent electronically to each SME to be completed and returned. They were also given my contact information if they had any questions or needed further clarification.

For fundamentality, SMEs were asked to rate the degree to which each item was central to one's core beliefs. They were asked to rate using the following scale: 1 = not fundamental at all, 3 = somewhat fundamental, 5 = highly fundamental. It is important to note that 2 and 4 were offered as answer choices but had no descriptive text. In order to determine the effectiveness of the item in terms of content validity, Lawshe's (1975) Content Validity Ratio (LCVR) was calculated. LCVR is a formula that calculates the proportion of SMEs who endorsed the item as being fundamental in nature. More specifically, the following formula was used in the LCVR calculation:

$$LCVR = (n_e - N/2)/(N/2)$$

In the formula, n_e represents the number of SMEs who endorsed the item as being fundamental in nature (i.e., the number of people who selected the response option of 3 or above) and N equals the total number of SMEs. The calculated value falls between -1 and 1. Items that generate a negative value are deemed unacceptable for the purposes of content validity and will be excluded from further analysis. In the case of fundamentality, none of the items were excluded. Table 1 shows the results of the LCVR analysis for all SME criteria.

Table 1 – *SME Survey Results for Lawshe Content Validity Ratios*

COE Item	Fundamentality	Eval Focus	Relevance	Readability	Coverage	Scope
1	0.67	0.17	1.00	1.00	0.50	0.00
2	1.00	0.33	1.00	1.00	0.67	1.00
3	1.00	0.17	0.83	-0.17	0.67	0.67
4	0.17	0.00	0.33	0.67	0.67	0.17
5	0.83	0.00	0.83	0.83	0.00	0.67
6	1.00	-0.17	0.67	1.00	0.17	0.83
7	1.00	0.17	0.83	0.50	0.17	0.67
8	0.50	-0.17	0.67	1.00	-0.50	0.50
9	0.50	-0.17	0.50	1.00	-0.67	0.67
10	0.50	-0.50	0.50	1.00	-0.17	0.00
11	0.50	-0.33	0.50	0.83	-0.17	0.17
12	1.00	0.17	0.83	1.00	0.50	1.00
13	0.67	0.00	1.00	0.83	0.33	0.83
14	1.00	0.17	1.00	1.00	0.33	0.67
15	0.67	-0.50	1.17	0.83	0.33	0.50
16	0.83	0.00	0.83	1.00	0.67	0.50
17	0.33	0.17	0.83	1.00	0.17	0.50

Table 1 – continued

18	0.17	0.00	1.00	1.00	-0.33	0.00
19	0.50	-0.33	1.00	1.00	1.00	0.67
20	0.67	-0.33	1.00	0.50	1.00	0.17
21	0.83	0.17	1.00	0.33	0.50	0.50
22	1.00	-0.17	1.00	0.83	0.50	0.67
23	1.00	0.17	0.83	1.00	0.50	1.00
24	1.00	0.17	0.83	1.00	0.50	1.00
25	0.83	-0.17	0.67	1.00	0.33	0.67
26	1.00	0.17	1.00	1.00	0.33	1.00
27	0.83	0.17	0.83	1.00	0.67	1.00
28	0.67	0.33	0.67	0.17	0.17	0.00
29	0.67	0.17	1.00	0.50	0.33	0.50
30	0.50	-0.17	1.00	1.00	1.00	0.50
31	1.00	0.33	0.83	0.50	0.50	1.00
32	0.83	0.33	0.50	0.50	0.33	1.00
33	0.67	0.00	1.00	0.67	0.67	0.83
34	0.50	0.00	0.67	0.67	0.33	0.17
35	0.33	0.00	0.33	0.83	0.17	0.50
36	0.83	0.00	1.00	1.00	0.67	1.00
37	0.00	-0.83	0.33	1.00	0.17	0.33
38	1.00	0.33	0.67	1.00	0.50	1.00
39	0.83	0.17	0.67	0.67	0.17	0.83
40	1.00	0.17	1.00	1.00	0.83	0.83

For evaluation-focus, SMEs were asked to rate whether the item involved an evaluation as opposed to merely being descriptive. They were told to select the rating of "1" if the item did not involve an evaluation and the rating of "2" if the item involved an evaluation. When determining whether items were acceptable as rated by SMEs, the same LCVR calculation was performed. In this case, n_e represented the number of SMEs who selected that the item was evaluation-focused. Again, items with a negative value did not meet the requirements of content validity and were, therefore, excluded. The following items were excluded based on this analysis: 6, 8 – 11, 15, 19, 20, 22, 25, 28, 30, and 37.

For scope, SMEs were asked to determine the degree of breadth with each item. They were instructed to rate whether each item could apply to a variety of situations or applications using the following response scale: 1 = situation-specific, 3 = somewhat broad, 5 = very broad. Again, 2 and 4 were viable answer choices, but had no descriptive text. The same LCVR calculation was performed to analyze the results. In this case, n_e pertained to the number of SMEs that endorsed the item as broad (i.e., assigned a rating of 3 or above). In analyzing the responses, LCVR calculations that yielded a negative

number were deemed unacceptable. All items showed acceptable agreement levels in reference to scope.

In addition to the inclusion criteria, SMEs were also asked to rate each item on three other dimensions: relevance, readability, and coverage. This effort was made to further ensure that the COE items could be practical not only for measuring the COE construct, but also for use in an organizational setting. Each dimension will be explained in turn. Relevance refers to how appropriate the item is for the purpose of the construct. In other words, this dimension seeks to ensure that all items are representative of the construct being measured (i.e., does the item measure COE as it is defined?). In essence, this dimension is examining face validity of the construct. SMEs were asked to rate how relevant each item was in regard to the COE construct using the following scale: 1 = not relevant, 3 = somewhat relevant, 5 = highly relevant (with 2 and 4 being response options). Once again, endorsement of relevance was classified as choosing the response option of 3 or above. In evaluating these responses, LCVR calculations that yielded a negative number were deemed unacceptable. All items showed acceptable agreement levels in reference to relevance for the construct.

Readability refers to the item's clarity and overall ability to be understood. This dimension asks SMEs to rate whether or not the item makes intuitive sense and can be understood by a broad audience. Readability is important to establish when creating new scales in order to determine whether the item will apply to a diverse audience. SMEs were asked to rate each item's readability using the following scale: 1 = very unclear, 3 = needs minor modification, 5 = easy to understand (with 2 and 4 being response options). The same method of calculation was applied to the readability criteria; however, endorsement of readability was classified as choosing the response option of 4 or above. All items met the criteria of readability except Item #3.

The last rated dimension was coverage. Coverage refers to the item's ability to span multiple traits said to comprise core evaluations of others rather than just focusing on one trait. In their development of the core self-evaluations scale, Judge and colleagues specified that because the construct is meant to be core across all traits involved, then it is important to focus on the general variance that explains the associations across all the traits (Judge, Erez, Bono, & Thoresen, 2003). Therefore, raters will be asked to determine how many trait(s) to which each item applies. Their answer

choices ranged from 1 = this item covers only one trait to 4 = this items covers all four traits. Those items that were rated as showing coverage across 2 or more traits by the majority of raters were deemed acceptable. Again, LCVR calculations were performed to determine acceptable fit. Based on this analysis, items 8, 9, 10, 11, and 21 were deemed unacceptable.

1.5.4 Generalizability

Another consideration when developing a new measure is to determine whether the scale will generalize beyond that of the population in which it was developed. IRT techniques aid in the generalization of a developed measure. In most cases, cross-validation efforts are made to ensure generalizability of the scale. In addition to cross-validating, the current study also seeks to examine the items in depth to determine if they are performing differently for one group as opposed to another. As previously stated, item response theory techniques allow for examining differential item functioning (DIF). Differential item functioning (DIF) determines whether items on a particular measure are showing variance across groups. For example, DIF will denote whether an item is performing differently for one group (e.g., women) as opposed to another group (e.g., men). The analysis is extremely important to conduct if efforts are made to use the tool in an applied setting. Due to pre-employment legal issues (such as adverse impact), it is important to verify that measures are functioning equally across all groups and not showing any adverse forms of DIF. Adverse impact is present when a seemingly neutral practice or measure causes a disproportionate number of protected class members to show unacceptable levels on that particular measure. If items in the COE scale were functioning differently across groups in an adverse way, then this occurrence could very well lead to adverse impact (if the COE scale was used in the hiring process). In order to examine this concern, the following research question is posed:

Research Question 3: Do any of the items of COE show differential item functioning?

Another concern with generalizability is whether there exist group differences in the prediction of targeted criteria (i.e., predictive bias). For instance, are there differences between racial groups on how effective COE is at predicting team satisfaction? Previous research has shown gender and race to be moderating variables in the relationship between CSE and job satisfaction (Lemelle & Scielzo, 2010). As a proactive measure, this study examined these potential moderators in the relationship between COE and targeted criteria. Age was not included in this analysis, because it was not previously found to be a

moderating variable, but will be explored nonetheless as a supplemental analysis. Given this objective, the following research questions are posed:

Research Question 4a: Does gender moderate the relationship between COE and job satisfaction?

Research Questions 4b: Does gender moderate the relationship between COE and team satisfaction?

Research Question 5a: Does race moderate the relationship between COE and job satisfaction?

Research Question 5b: Does race moderate the relationship between COE and team satisfaction?

Another potential moderating variable was sample population (i.e., field vs. students). As stated earlier, working individuals in the field will be included in the analysis to determine COE's ability to predict team effectiveness. Sample characteristics are often explored as potential moderators when exploring the relationship between two variables. In this case, it would be beneficial to explore whether the relationship between COE and work outcomes differ between students and working professionals. Therefore, the following research question will be posed:

Research Question 6a: Does sample type moderate the relationship between COE and job satisfaction?

Research Question 6b: Does sample type moderate the relationship between COE and team satisfaction?

<u>1.6 Pilot Studies</u>

The goal of this research was to highlight another facet of core evaluations besides core self-evaluations. Core self-evaluations add tremendous value to the field and have been shown to predict numerous work outcomes. However, core self-evaluations were originally described as being part of a tripartite model of core evaluations with core other-evaluations and core evaluations about the world. Core other-evaluations are defined as fundamental feelings that individuals hold about other people. This dispositional construct is expected to be effective in the prediction of not only interpersonal dynamics but also of job satisfaction. Like core self-evaluations, COE is believed to be a latent construct. The four variables that have been found to manifest core feelings about others are trust, agreeableness,

collectivism, and other orientation. These variables were chosen due to their adherence to the inclusion criteria and overall emphasis on core feelings about others. In an effort to determine if these variables actually showed some sort of common relationship with each other, a series of pilot studies were conducted.

More specifically, these pilot studies were conducted to examine the interrelationships between the variables of interest to determine whether there was any viability to the assertion that there is some interconnectedness between the variables. This process was broken into two phases. In the first phase, archival data from a human resource consulting firm was obtained to explore the interrelationships between variables similar to those of interest (i.e., those that are ostensibly comparable to trust, agreeableness, other orientation, and collectivism). In the second phase, data was administered to both students and non-students to understand the relationships between the factors said to underlie COE in their direct form (i.e., direct measures of trust, agreeableness, other orientation, and collectivism). This second phase was conducted in order to determine whether there was any value in linking these four traits together (i.e., is there a relationship between these variables?). The pilot studies also focused on item development, content analysis by use of subject matter experts (which was already discussed), and initial exploration into the functioning of the developed items. Phase One, which involved the examination of archival data, will be discussed first followed by the Phase Two results, where data administered to students and non-students was examined to determine the degree of interrelationships. Lastly, the item analysis will be discussed along with the retention of the final scale items for cross-validation.

1.6.1 Phase One

The overall goal of Phase One was to determine whether there was any credibility in linking the traits said to underlie core other-evaluations together. In order to address this issue, data was obtained from a human resource consulting firm to explore the interrelationships between variables similar to those described to underlie core other-evaluations. It was the hope that examination of these like variables would provide clearance to further pursue the interrelationships of trust, agreeableness, other orientation, and collectivism.

The Personal Characteristics Inventory (PCI; Barrick & Mount, 2007) was administered (via the aforementioned consulting firm) to individuals employed in a number of different industries, with greatest

representation by those in the restaurant industry. The PCI claims to measure the personality traits that are most important in the workplace and covers the Big Five personality dimensions (Extraversion, Agreeableness, Conscientiousness, Stability, and Openness to Experience, as well as twelve subscales (Sociability, Need for Recognition, Leadership Orientation, Cooperation, Consideration, Dependability, Efficiency, Even-Temperament, Achievement Striving, Self-Confidence, Abstract Thinking, and Creative Thinking), five occupational scores (Manager, Sales, Production, Clerical, and Driver), and four success scales (Teamwork, Commitment to Work, Learning Orientation, and Integrity).

Because the purpose of this phase of the Pilot Study was to analyze scales related to those said to underlie COE, only those that pertained to core other-evaluations were included in the analysis. The scales included were: Agreeableness, Cooperation, Consideration, and Teamwork. Due to the fact that agreeableness is one of the variables included in the COE model, this was an obvious choice. Consideration was chosen not only because it is a subscale of agreeableness, but also because it looks at the degree to which individuals assess the needs of others and show sensitivity to the feelings of others, which is very similar to other orientation. Those who score high in this area are described as "kind, empathetic, supportive, appreciative, tolerant, forgiving and straightforward in their dealings with others" (Barrick & Mount, 2007, p.7). Cooperation measures the extent to which individuals value working with other people. This trait has ties to agreeableness and other orientation. Individuals who score high in this area are said to show a genuine interest in helping people and likely to thrive in a team environment (Barrick & Mount, 2007). Finally, teamwork is concerned with an individual's ability to work effectively within a group and engage in behaviors that support the overall well-being of the group (Barrick & Mount, 2007). This definition of teamwork is very similar to that of collectivism.

In addition to the PCI, this archival dataset also included the Guilford-Zimmerman Temperament Survey (Guilford, Zimmerman, & Guilford, 1976), which includes scales useful in employee selection and placement. One of the scales, People Relations, was selected due to its similarity to the COE dimension of trust. Though this connection is probably the loosest, there is an element of trust at play in the People Relations scale. This area measures the optimism and tolerance that individuals have in their general outlook toward others as well as the degree of trust they exude toward other people (Guilford, Zimmerman, & Guilford, 1976). With the inclusion of this scale, all COE variables are covered (i.e.,

agreeableness, trust, other orientation, and collectivism). Correlations were run to determine the interrelationships amongst these variables.

1.6.1.1 Phase One Results

The sample number of participants for this phase was 297. Table 2 shows the correlations between these variables. As shown in Table 2, all correlations between these variables are substantially high (some being $r > .70$), with the majority of relationships demonstrating an $r > .40$. In fact, all but one (Cooperation and People Relations) of the interrelationships between these variables are statistically significant. It was previously stated that relating trust to the People Relations scale was somewhat of a stretch in comparison to the other connections. This stretch was apparent in the results of this analysis. All of the correlations involving People Relations were the lowest of the group, although still statistically significant in all cases but one. Due to the fact that such high relationships were found between these variables of interest, it is reasonable to say that there may be an underlying latent construct that is responsible for these interconnections. This possibility was further explored in Phase Two of the Pilot Study.

Table 2 – *Intercorrelations between Pilot Study/Phase One variables*

Scale	Mean	SD	1	2	3	4	5
1. Agreeableness	53.28	5.08	-	.91**	.76**	.78**	.25**
2. Consideration	25.90	3.60	-	-	.42	.76**	.28**
3. Cooperation	27.42	2.37	-	-	-	.52**	.11
4. Teamwork	83.69	7.74	-	-	-	-	.40**
5. People Relations	12.31	2.41	-	-	-	-	-

Note: $N = 297$. **$p < .01$

1.6.2 Phase Two

Because the results of Phase One warranted further exploration of the COE dimensions, Phase Two sought to provide a cursory glance into the interrelationships of the variables said to comprise COE (i.e., trust, agreeableness, other orientation, and collectivism) in an effort to give further credence to the exploration of this construct. In order to do this, data was collected from both students and non-students using an online survey platform (Survey Monkey).

1.6.2.1 Phase Two Measures

As part of the survey, participants were asked to respond to questions pertaining to their levels of agreeableness, trust, other orientation, and collectivism. The majority of the measures included have been published and widely used by researchers at large.

1.6.2.1.1 Trust

In order to measure trust at a fundamental level, Rotter's (1967, 1971) Interpersonal Trust Scale was utilized. This scale contains aspects of political and social elements in society. Due to the fact that this study is only interested in the social aspects of trust and that the political portion may prove to be a potential confound, only the items referencing the social elements involved in trust were utilized in this phase. Participants were asked to select their level of agreement with each question, and each item was rated on a six-point Likert-type scale with 1 = Strongly Disagree to 6 = Strongly Agree. A sample item reads, "Most people can be counted on to do what they say they will do."

1.6.2.1.2 Agreeableness

Items related to agreeableness from the Big Five Inventory (BFI; John, Donohue, & Kentle, 1991) were used for the purposes of this pilot study. The BFI contains 44 items in total, with 9 items pertaining to agreeableness. Participants were asked the level of agreement they have to each of the questions. Level of agreement is determined using a six-point Likert-type rating scale where 1 signifies that the participants disagrees strongly and 6 signifies that the participant agrees strongly. An example item reads, "I see myself as someone who is considerate and kind to almost everyone" (John et al., 1991). The reported reliability for the scale of agreeableness is $\alpha = .79$ (John & Srivastava, 1999). The authors note that agreeableness has one of the lowest reliabilities of the five personality dimensions. This finding is said to be consistent across all measures of the Big Five personality factors (John & Srivastava, 1999).

1.6.2.1.3 Collectivism

Collectivism was measured using the scale of Psychological Collectivism (Jackson, Colquitt, Wesson, & Zapata-Phelan, 2006). This scale contains a total of 15 items and measures five different facets of collectivism or feelings toward groups (preference, reliance, concern, norm acceptance, and goal priority). Participants were instructed to think about groups to which they currently belong or have belonged to in the past when answering the questions. A sample item reads, "Group goals are more

important to me than my personal goals" (Jackson et al., 2006). Participants were asked to choose their level of agreement on a six-point Likert-type scale with the following response scale: (1 = Strongly Disagree to 6 = Strongly Agree). The reported reliability of the overall scale is α = .84, which represents strong reliability (Jackson et al., 2006).

1.6.2.1.4 Other Orientation

In previous literature, other orientation has often been measured in terms of prosocial motivation (Grant & Wrzesniewski, 2010). Therefore, other orientation was measured using Grant's (2008) adapted measure of prosocial motivation. Grant (2008) adapted items from Ryan and Connell (1989). The total measure contains four items, and participants were asked to choose their level of agreement on a six-point Likert-type scale with the following response scale: (1 = Strongly Disagree to 6 = Strongly Agree).The items were adapted further to keep them in a broad, fundamental sense rather than being specific to the job. Participants were given an introductory question that reads, "Why are you motivated in life?" They then responded to the scale items. A sample item reads, "Because I want to help others" (Grant, 2008). Six distractor items were included to reduce the transparency of the scale. The reported reliabilities for the scale range from alpha = .90 (Grant, 2008) to alpha = .95 (Grant & Wrzesniewski, 2010).

1.6.2.2 Phase Two Results

The results of Phase Two of the Pilot Study can be found on Table 3. Data from 88 participants was collected and analyzed as a part of this phase. As Table 3 depicts, all but one of the correlations (Trust and Other Orientation) are statistically significant. Even though the correlations are not as high as those obtained in Phase One, these results are still promising. The lower correlations may be due, in part, to the fact that this sample is significantly smaller than the previous one ($N = 88$ as opposed to $N = 267$). Also, the scales used to measure the traits were not ideal. Unfortunately, it was difficult to obtain measures of trust and other orientation that were broadly used and highly validated. This may have contributed to the lower correlations obtained between the measures. Because of potential unreliability of these scales and issues with range restriction, correlations were corrected for attenuation using the following formula:

$$rx'y' = \frac{rxy}{\sqrt{rxx\,ryy}}$$

Table 4 shows the corrected correlations. With the applied correction, all correlations are statistically significant, thus signifying that the relationships between these variables are substantial and worthy of further attention. Correlations are high enough to show a common linkage, yet not too high to question whether they are completely redundant. Overall, based on the results of both phases, there seems to be something at play that can be contributing to the interconnections between these variables.

Table 3 – *Intercorrelations between Pilot Study/Phase Two variables*

Scale	Mean	SD	1	2	3	4
1. Agreeableness	4.71	.71	-	.37**	.41**	.44**
2. Trust	3.35	.87	-	-	.18	.36**
3. Other Orientation	5.10	.72	-	-	-	.42**
4. Collectivism	4.18	.93	-	-	-	-

Note: $N = 88$. **$p < .01$

Table 4 – *Corrected correlations between Pilot Study/Phase Two variables*

Scale	Mean	SD	1	2	3	4
1. Agreeableness	4.71	.71	1	.48**	.49**	.52**
2. Trust	3.35	.87	.37**	1	.22*	.43**
3. Other Orientation	5.10	.72	.41**	.18	1	.45**
4. Collectivism	4.18	.93	.44**	.36**	.42**	1

Note: $N = 88$. **$p < .01$, *$p < .05$; corrected values are above the diagonal and uncorrected values are below the diagonal

1.6.3 Item Analysis

The final contribution of the pilot studies was to take an initial glance at the item analysis of the COE scale items. The cross-validation study deeply explored each item using item response theory techniques, but the pilot study sought to initially investigate the item's functioning. Because IRT requires a large sample size, it was not practical to explore item analysis using IRT for the purposes of the pilot. Therefore, CTT methods were used to determine the item's mean, standard deviation, and discrimination index. Each of these values was calculated and is included in Table 5. The overall scale reliability with the

40 items was α = .92, which represents a high level of reliability. Items with a high mean (>4.50) were examined further and refined. For instance, Item 36 has a mean of 4.91 and reads, "It bothers me a great deal to see people treated unfairly." This item was modified to read, "Nothing bothers me more than to see people treated unfairly," thereby, potentially lowering the likelihood that people will endorse this item if they do not have a high level of COE. Items that were affected by the mean (or difficulty) examination included: items 15, 17, 20, 27, 30, 31, 36, 37, 38, 39, & 40. Also, items that had a low mean (<3.00) were examined for item content. These items included: items 1, 3, & 10. The item discrimination index measures the extent to which an item can distinguish between those with high levels of a trait versus those with low levels of a trait. Item discrimination is said to be acceptable at a value of .30 or above (Mehrens & Lehmann, 1973). In an effort to be more conservative and, thus, only retain items that show high discrimination levels; the bound of acceptable fit was increased to .40. Given this constraint, the items that were called into question include: items 1, 3, 8, 11, 12, 18, 25, 33, 35, 37, & 38.

Table 5 – *Item Analysis Results – Pilot Study*

Item	Item Wording	Mean	Std.Dev	Disc Index
1	I instantly trust people the first time I meet them.	2.68	1.37	.30
2	In general, most people can be trusted.	3.76	1.17	.54
3	You can't be too careful in dealing with other people.	2.92	1.19	.20
4	It is hard for me to work with people that I do not know.	4.10	1.38	.44
5	People tend to follow through with their word.	3.59	1.10	.45
6	In general, people keep the promises they make.	3.66	1.04	.42
7	Most people fail to follow through with their word.	3.78	1.09	.48
8	Teams can always accomplish more than individuals.	3.63	1.31	.36
9	The key to getting the best results is by working in a team.	3.83	1.27	.51
10	I can get more accomplished alone than in a group.	2.99	1.31	.45
11	I prefer working in a group rather than working independently.	3.34	1.37	.32
12	People tend to manipulate others in order to get what they want.	3.00	1.20	.12
13	I cannot rely on people, because most people let me down.	4.00	1.18	.53

Table 5 – continued

14	I believe that other people can be counted on to do what they say they will do.	3.72	1.10	.52
15	It hurts me to see other people hurting, even if I do not know them.	4.81	1.17	.51
16	I find it hard to care about people that I do not know.	4.25	1.50	.66
17	I find it rewarding to work with other people toward a common goal.	4.51	1.17	.64
18	I find it more rewarding to win an individual prize rather than a group prize.	3.20	1.50	.38
19	I've been told that I care too much about the well-being of others.	3.83	1.39	.47
20	I find it difficult to pass up someone in need without helping them.	4.64	1.14	.63
21	I'd rather take care of my personal goals than to devote my time to helping others.	3.97	1.22	.55
22	I consider others when making decisions before I consider myself.	4.08	1.39	.47
23	Other people generally have bad intentions.	4.50	1.05	.57
24	Other people generally have good intentions.	4.25	1.01	.57
25	I believe people when they say they will do something.	3.91	1.08	.35
26	I find it hard to trust other people.	3.63	1.38	.70
27	I consider myself to be a caring and nurturing person.	4.73	1.10	.48
28	Collaborating with others is fun to me.	4.25	1.18	.61
29	Dealing with other people can be quite a pain.	3.57	1.26	.54
30	I wouldn't help someone unless I knew them personally.	4.75	1.23	.44
31	Most people are up to no good.	4.53	1.07	.163
32	Everyone has kindness in their hearts, even if they don't use it.	4.36	1.32	.51
33	I am constantly considering the impact that my actions have on other people.	4.31	1.22	.36
34	People seldom pull their weight in group activities.	3.56	1.15	.42
35	I feel uncomfortable working with other people.	4.39	1.25	.21
36	It bothers me a great deal to see people treated unfairly.	4.91	1.08	.46
37	I have been told that I am rude to others at times.	4.68	1.35	.39
38	Most people are highly capable of doing great things.	4.64	1.06	.37
39	Most people are incompetent.	4.52	1.12	.43

Table 5 – continued

40	Doing things for others gives me more satisfaction than doing things for myself.	4.53	1.17	.48

When determining the final set of items in which to cross-validate, a sequence of criteria were followed. The first consideration was regarding content validity. Only the items that sufficiently met the criteria as rated by Subject Matter Experts (fundamentality, evaluation-focus, scope, relevance, readability, and coverage) were retained. Failure to meet these criteria eliminated 13 items from the scale (6, 8, 10, 11, 15, 19, 20, 21, 22, 25, 30, & 37). The next criterion was the discrimination index. Items with insufficient discrimination indices (i.e., <.40) were not retained. This resulted in the elimination of seven more items (1, 3, 12, 18, 33, 35, & 38). Finally, of the items that were left, those that had high mean levels were examined for word content and modified. The items affected by this change were 17, 27, 36, 39, & 40. Item 17 read, "I find it rewarding to work with other people toward a common goal," and was modified to read, "I find it more rewarding to work with other people toward a common goal than to work alone." Item 27 read, "I consider myself to be a caring and nurturing person," and was modified to read, "I consider myself to be an extremely caring and nurturing person." Item 36 read, "It bothers me a great deal to see people treated unfairly," and was modified to read, "Nothing bothers me more than seeing people treated unfairly." Item 39 read, "Most people are incompetent," was modified to read, "Other people are incompetent." Item 40 read, "Doing things for others gives me more satisfaction than doing things for myself," and was modified to read, "Doing things for others makes me happier than doing things for myself." The application of these criteria resulted in the elimination of 20 items, the retention of 15 items, and the modification of 5 items (see Appendix B for a list of the retained/modified items). The resulting 20 items showed a scale reliability of α = .90.

1.6.4 Conclusion

The Pilot Studies proved to be successful in warranting further exploration into the relationships between trust, agreeableness, other orientation, and collectivism. In both Phase One and Phase Two, significant relationships between the variables were found. This highlights the possibility of a higher-order, latent variable linking these traits together. Additionally, the Pilot Study enabled the generation of items designed to measure COE and analyzed each item in terms of subject matter expert ratings. Furthermore,

item analysis was performed on the pilot data to take an initial glance at the item functioning and determine a subset of 20 items to cross-validate. The reliability statistic proved to be very strong for the scale, and the items seem to be discriminating well between those low and high levels of COE.

<div align="center">1.7 Cross-Validation Study</div>

The overarching purpose of this dissertation study was to develop and validate a comprehensive measure of COE and determine its predictive validity in relation to organizational teams and job satisfaction. Efforts were already undertaken to define the construct, gain initial support for the interconnectedness of the trait-level manifestations, establish items designed to measure core other-evaluations, and examine the developed items to determine their psychometric performance. Based on the results of the Pilot Study, a select number of items were retained. The purpose of the cross-validation study was to further explore the reliability and validity of the COE scale by exploring the factor structure (thereby, addressing Research Question 1), examining item indices and model fit (hence, exploring Research Question 2), investigating convergent and discriminant validity (which addresses Hypotheses 1 – 4), establishing criterion-related validity (thus, exploring Hypotheses 5 & 6), assessing the incremental validity of COE beyond CSE and the TPI (which pertains to Hypotheses 2b, 7 & 8), and exploring group differences through differential item functioning (addressing Research Question 3) and moderation analysis (answering Research Questions 4 – 5). Each of these objectives will be discussed in turn.

1.7.1 Exploration of Factor Structure

The construct of COE is believed to be indicative of a higher-order latent construct that measures the variance shared between each of the traits that represent it. As Judge et al. (2003) describe, latent constructs are positioned at a deeper level than their indicators and, thus, control those manifestations to a certain degree. In other words, because COE is believed to be a latent construct that it is said to causally influence the traits that underlie it (Bollen & Lennox, 1991). If this assertion is supported, then the underlying traits should have some common connection between them (which was demonstrated in the Pilot Study). To determine whether the COE construct is unidimensional, an exploratory factor analysis was run on the data. If the COE items show unidimensionality, then the construct is said to be super-ordinate. Because unidimensionality is an assumption of IRT, this issue will be inherently addressed in the analysis. Local independence is typically covered if the assumption of unidimensionality is met.

1.7.2 IRT Model

If assumptions are met, then further IRT analyses are permissible. The next step in the process was to determine which IRT model best applies to the data. Because the COE items use a Likert-type response scale, with each response option being progressively higher than the previous, a model that takes into account ordered responses would be appropriate. Due to the nature of the scale, Samejima's (1969) graded response model (GRM) was used as the model of choice. GRM is the polytomous version of the two parameter logistic model (Thissen & Steinberg, 1986) and should be used when the item responses of the scale contain ordered categorical responses (Embretson & Reise, 2000) Information functions for each item were examined to determine item quality. In addition, test information functions were examined to determine the value of the scale. Finally, model fit was assessed.

1.7.3 Convergent and Discriminant Validity

Convergent and discriminant validity were also explored by examining the relationship between COE and other variables. Convergent validity was established if like constructs were found to be highly related. Because trust, agreeableness, collectivism, and other orientation are said to be manifest variables of the latent construct of COE, then they are expected to be highly related to one another, which was demonstrated in the Pilot Studies. The next step of this analysis involves the four manifest traits' relationship with COE. In addition, COE's relationship with a measure designed to examine an individual's likelihood to be a team player was explored.

Conversely, discriminant validity was established if distinct constructs show little or no relationship with each other. The constructs used in this analysis were core self-evaluations, conscientiousness, and openness to experience. Even though CSE and COE are theoretically related, they are expected to represent unique constructs and, therefore, should be moderately related at best. Likewise, though individuals high in COE may also be high in conscientiousness, it was expected that these constructs are conceptually distinct and, thus, not in the same nomological network. openness, on the other hand, is not proposed to show a relationship with COE and, therefore, should demonstrate a relatively small correlation with COE.

1.7.4 Criterion-Related Validity of COE

Another aspect of this study involved looking at the predictive validity of COE. Based on the theoretical propositions of Judge et al. (1997), COE was expected to be positively related to job satisfaction. In addition, due to the interpersonal nature of the construct, COE was expected to be more related to the interpersonal aspects of job satisfaction rather than the circumstantial aspects of job satisfaction. It was also proposed that COE would show strong relationships with team satisfaction and effectiveness at both the individual and team level.

1.7.5 Incremental Validity

Furthermore, it was predicted that COE will add unique predictive value over and beyond that of CSE. Though CSE shows strong relationships with important work outcomes and targeted criteria, it was expected that COE has something to offer beyond CSE, especially in regard to team dynamics and interpersonal relationships. In addition, COE was believed to better predict the interpersonal aspects of job satisfaction than CSE.

Also, in regard to incremental validity, it was predicted that COE will add unique predictive value over and beyond the Team Player Inventory (TPI). Because COE is a more fundamental construct, it is expected to affect other fundamental constructs (e.g., team satisfaction) to a greater extent than that of the TPI.

1.7.6 Group Differences

The last area of exploration involves the generalizability of the COE scale. Because this measure will be advocated for use in pre-employment settings, it is imperative to establish that items are not displaying adverse bias. As previously stated, items may show DIF without actually showing bias, so it is important to explore not only if items are showing DIF, but also if the DIF associated with items is benign or adverse. Demographic groups (i.e., age, race, and gender) were explored for differential item functioning.

In addition to exploring item functioning, moderation analyses were performed to determine if there are any group differences (e.g., age and race) in the prediction of COE and targeted criteria. Sample composition (students vs. working professionals) were also explored as a potential moderating variable.

CHAPTER 2

METHOD

2.1 Participants

For the purposes of the cross-validation study, two independent samples of participants were recruited. One sample contained participants who were undergraduate students in psychology classes at a southern university, and the second sample contained working professionals in a variety of industries (e.g., government, human resource consulting, healthcare, and technology). Participants were given an online survey that asked questions related to their personality type, attitudes about teams, and job satisfaction (if applicable).

2.1.1 Sample One

Student participants were enrolled in psychology courses and were required to participate in psychology studies as a part of their coursework. The survey for the current study was uploaded onto the SONA system for the university, and students were able to sign up for the study at their leisure. Participants were given informed consent about the details of the survey before deciding whether or not to participate. Students who participated in the current study (Personality and Work Attitudes) were given one hour of course credit as an incentive for their involvement. Those who chose not to participate were not penalized and were given other options to receive credit for their courses. Each participant was asked to complete an online survey that included questions related to their personality style, team attitudes, and attitudes about work. Only those students who indicated that they were currently working were given questions related to their attitudes about work. All student participants completed the personality portion and the team attitudes portion of the survey. Each participant was shown a debriefing statement at the conclusion of the survey.

There were a total of 506 student participants. Of the total 506 student participants, 46% indicated that they were currently working. The overwhelming majority of the student participants were females (80%), with a small percentage of males (19%) and missing data (1%). The race distribution was

pretty evenly dispersed with Whites being the majority (33%), followed by Hispanics (29%), then Asians (16%), then Blacks (15%), and finally those who specified that they are classified as belonging to a group other than those mentioned (5%). Two individuals failed to report their race on the survey (.4%). With regard to age, the overwhelming majority of participants were either 20 years of age or younger (71%), with only 1.4% of participants being 40 years of age or older.

2.1.2 Sample Two

The second sample of participants contained individuals in field, applied settings (i.e., working professionals). Companies with established work teams were recruited to participate in this study. Likely due to the Sarbanes Oxley Act and the full disclosure policy inherent in many large corporations, it was difficult to gain entry into and participation from larger, publicly traded companies. Therefore, the majority of participants in this sample were recruited from smaller, boutique-style firms in a variety of different industries. Some companies requested an incentive for their participation. For instance, one of the companies that participated in this process was a small healthcare organization. In return for their participation, they requested that an overall report and verbal presentation of the findings be provided to them at the conclusion of the research. No other companies requested any form of incentive for their participation. Participants in this sample were recruited on a voluntary basis. Initially, an effort was made to obtain data from the organization level, but when these efforts were futile, team volunteers were recruited on an individual basis.

Participants were asked questions related to their personality style, attitudes about teams, team effectiveness, and job satisfaction. All participants were required to input their team and company name for matching purposes. Participants were able to complete the online survey at their convenience and were not penalized if they chose to cease participating midway through the survey. Aside from the aforementioned report and verbal presentation, participants were not incentivized to participate in this study. Participants in this sample were able to review and accept the informed consent at the commencement of the study, and were informed that they have the option to discontinue at any point in the survey if they so choose. They were all also offered a debriefing statement at the close of the study.

The working professionals sample contained a total number of 105 participants. Within this sample, the majority of the participants were female (61%), with 36% being male (3% were missing). The

race breakdown was as follows: Whites (61.9%), Blacks (19%), Hispanics (8.6%), Asian (4.8%), Other (1.9), with 3.8% of participants failing to report their race. The age range was varied with 62.9% of participants being under 40 years of age, and 37.1% of participants being 40 years of age or older. With regard to team composition, a total of 32 teams were represented. On average, 60% of team members responded to the survey.

<div align="center">2.2 Measures</div>

The scales administered to the participants during the cross validation study included the measures previously discussed in the Pilot Study (i.e., trust, agreeableness, collectivism, other orientation) as well as measures of job satisfaction, team satisfaction, the Team Player Inventory, core self-evaluations, conscientiousness, openness to experience, and team effectiveness. Again, the majority of the measures included have been published and widely used by researchers.

2.2.1 Job Satisfaction

Job satisfaction was measured using the Job Descriptive Index (JDI; Smith, Kendall, & Hulin, 1969) measure of job satisfaction. The JDI measures job satisfaction using five facets: pay, promotions and promotion opportunities, coworkers, supervision, and the work itself. Participants were asked to choose their level of agreement with each item using a six-point Likert-type rating scale (with 1 = Strongly Disagree to 6 = Strongly Agree). A sample item reads, "All in all, I am satisfied with my job" (Smith et al., 1969). The reported reliability for this scale ranges from $\alpha = .84$ to $\alpha = .90$ (Saane, Sluiter, Verbeek, & Frings-Dresen, 2003). It is important to note that the original JDI was validated using a response option of "yes", "no", or "?". For the purposes of the current study, it was determined that a six-point Likert-type scale would show a greater range of variability in individual responses. The reliability of the scale used in the current study was $\alpha = .94$.

2.2.2 Team Satisfaction

Team satisfaction was measured using adapted items from the JDI (Smith et al., 1969) measure of overall job satisfaction. The five items pertaining to overall job satisfaction were adjusted to reflect team satisfaction. Participants were given the introductory statement that reads, "Think back to a time when you worked with a team to accomplish a goal. Please answer the following items based on your recollection of that team." An adapted sample item reads, "All in all, I was satisfied with my team." Again

and as previously mentioned, the response options given with this measure were transformed from the "yes", "no", "?" format to a six-point Likert-type scale. The reliability estimate calculated in the current study was α = .85.

2.2.3 Team Player Inventory

The Team Player Inventory (TPI) was created with the intent of measuring an individual's predisposition to succeed in a team-based environment (Kline, 1999). The measure contains 10 items and asks individuals to indicate their level of agreement with each item by selecting from a six point Likert-type rating scale. Individuals can specify that they disagree completely by selecting a 1 and can indicate that they agree completely by choosing a 6. The reported reliability of this scale is between α = .87 and .83 (Kline, 1999). The reliability coefficient obtained in this study was α = .80.

2.2.4 Core Self-Evaluations

Core self-evaluations was measured using the scale developed by Judge and colleagues (Judge, Erez, Bono, & Thoreson, 2003). The measure contains 12 items and is referred to as the Core Self-Evaluations Scale (CSES). The scale measures an individual's overall self-regard in terms of their self-esteem, emotional stability, locus of control, and generalized self-efficacy. An example item is, "Overall, I am satisfied with myself." Respondents were asked to choose their level of agreement with each item on a six-point Likert-type response scale (where 1 = Strongly Disagree and 6 = Strongly Agree). The reported average reliability estimate across four samples is α = .84 (Judge et al., 2003). In the current study, the reliability of the scale was α = .81.

2.2.5 Conscientiousness

Conscientiousness was measured by using the conscientiousness-related items from the Big Five Inventory (BFI; John, Donohue, & Kentle, 1991). As stated previously, the BFI contains 44 items in total, with 9 items pertaining to conscientiousness. Participants were asked their level of agreement with each of the questions. Level of agreement was determined using a six-point Likert-type rating scale where 1 signifies that the participant disagrees strongly and 6 signifies that the participant agrees strongly. An example item reads, "I see myself as someone who perseveres until the task is finished" (John et al., 1991). The reported reliability for the scale of Conscientiousness is α = .82 (John & Srivastava, 1999). The reliability coefficient obtained in the current study was α = .79.

2.2.6 Openness

Openness was measured using the openness to experience-related items from the Big Five Inventory (BFI; John, Donohue, & Kentle, 1991). As stated previously, the BFI contains 44 items in total, with 10 items pertaining to openness. Participants were asked their level of agreement with each of the questions. Level of agreement was determined using a six-point Likert-type rating scale where 1 signifies that the participant disagrees strongly and 6 signifies that the participant agrees strongly. An example item reads, "I see myself as someone who is curious about many different things" (John et al., 1991). The reported reliability for the scale of Openness is α = .81 (John & Srivastava, 1999). The reliability estimate calculated in this study was α = .77.

2.2.7 Team Effectiveness

Team effectiveness was measured using the items from both the Manz and Sims (1987) and Cox (1994) measures of multi-faceted aspects of team effectiveness (Pearce & Sims, 2002). This measure contains seven facets of effectiveness including: output effectiveness, quality effectiveness, change effectiveness, organizing and planning effectiveness, interpersonal effectiveness, value effectiveness, and overall effectiveness. The scale contained a total of 26 items. An example item from this questionnaire reads, "The team is highly effective." Participants were asked to select their level of agreement with each item on a six-point Likert-type rating scale ranging from 1 (Strongly Disagree) to 6 (Strongly Agree). The reported reliability for this scale is α = .85. The reliability obtained in the current study was α = .97.

<u>2.3 Data Analysis</u>

The factor structure of the COE scale was psychometrically investigated using an exploratory factor analysis with a Principal Components method of extraction to assess the unidimensionality assumption of IRT. The prediction was that the items of COE would load onto one higher-order factor. To determine whether unidimensionality existed, factor analysis results (eigenvalues, percentage of variance accounted, and scree plot results) were examined. At the very least, there was expected to be a dominant trait present in the factor structure which would fulfill the requirement of essential unidimensionality (Stout, 1987). In regard to the IRT analysis, the model that was found to best fit the data was Samejima's Graded Response Model (GRM; 1969).

Correlational analyses were performed to determine the presence of convergent and discriminant validity. Fisher's R to Z transformations were conducted to examine differences between correlations. The interrelationships between the four indicator traits (trust, agreeableness, collectivism, and other orientation) and COE were examined by reviewing inter-item correlations as well as scale correlations. High correlations were expected between each of the four indicator variables and COE. Conversely, low correlations were expected between COE and distinct constructs (e.g., conscientiousness, openness, and core self-evaluations). Hierarchical regression analyses were conducted to determine the predictive ability of the COE scale and its incremental validity beyond CSE and TPI.

Differential item functioning was explored using the Mantel-Haenszel statistic. As explained previously, measurement invariance is established when item characteristic curves from different groups are the same. If measurement invariance is found, then differential item functioning is not present. The Mantel-Haenszel statistic and logistic regression approaches are common DIF detection techniques (Schumaker, 2005). Moderated multiple regression analyses were performed to determine the presence of moderating variables. More specifically, demographic variables (gender and race) and sample composition (students vs. working professionals) were explored as potential moderators between COE and work outcomes. If significant interactions between the moderator variable and the independent variable were found in regard to the dependent variable, then moderators were said to be present. Age was not included in the research questions but was examined as a supplemental analysis.

CHAPTER 3

RESULTS

Before addressing the hypotheses and research questions, an effort was made to re-examine the

20-item COE scale to determine the effectiveness of the items given a larger, cross-validation study. With

that objective, an analysis was performed that was similar to the Item Analysis section of the pilot study.

More specifically, item means and discrimination indices were calculated and compared to the previously

mentioned inclusion criteria (excluding items with high means (>4.50) and low (<.40) discrimination index

values). For practical purposes, items with high means were excluded from the scale due to the issue with

the transparency of these items. A high mean (i.e., >4.50) is indicative of a negatively skewed item in

which most individuals are inclined to endorse the item in the positive direction. Because the COE scale

is recommended for use in organizational and pre-employment settings, less transparent items are ideal.

Also, in regard to the discrimination index, values that fell below .40 were considered unacceptable. Even

though some have claimed that discrimination values above .30 are adequate (Mehrens & Lehmann,

1973), an effort was made in this analysis to use a stricter criterion (i.e., .40) and pick highly

discriminating items; thereby, enabling the use of a shorter item bank (which is ideal for practical and

applied purposes) while still showing strong discrimination in COE ability. These criteria for inclusion were

determined a priori. In addition to these calculations, items were also measured on their ability to predict

some of the important work outcomes under study (e.g., job satisfaction and team satisfaction). Items with

low correlations to work criteria were eliminated. This practice is consistent with the procedure of Judge et

al. (2003) in their item analysis process. A total of six items were eliminated through this process, and

Table 6 displays the results of this analysis. Item 8 ("I find it more rewarding to work with other people

toward a common goal than to work alone") was eliminated due to its low discrimination value (.361). Item

12 ("I consider myself to be an extremely caring and nurturing person") was eliminated on the basis of

mean (4.86), discrimination (.339), and predictive ability (.075 for job satisfaction). Item 16 ("Everyone has

kindness in their hearts, even if they don't use it") was eliminated on the basis of mean (4.56),

discrimination (.395), and low predictive ability (.041 for job satisfaction). It should be noted that Item 16 had the lowest predictive ability for job satisfaction out of all 20 items. Item 17 ("People seldom pull their weight in group activities") was eliminated due to a low discrimination value (.332). Item 18 ("Nothing bothers me more than seeing people treated unfairly") was eliminated on the basis of mean (4.88), discrimination (.282), and low predictive ability (.093 for job satisfaction). It should be noted that Item 18 had the lowest predictive ability for team satisfaction (.101) out of all 20 items and had one of the lowest values for job satisfaction. Finally, Item 20 ("Doing things for others makes me happier than doing things for myself") was eliminated on the basis of mean (4.63) and discrimination value (.374). The remaining scale contains 14 items (which is comparable to Judge et al.'s (2003) parsimonious and practically useful 12 item measure of CSE; see Appendix C for a list of the final COE items).

Table 6 – *Item Analysis Results – Cross-Validation Study*

COE Item	Item Wording	R-coded	Mean	Std.Dev	Disc Index	Job Sat	Team Sat
1	In general, most people can be trusted.	No	3.58	1.23	.517	.260	.191
2	It is hard for me to work with people that I do not know.	Yes	4.04	1.43	.418	.155	.261
3	People tend to follow through with their word.	No	3.59	1.13	.555	.196	.222
4	Most people fail to follow through with their word.	Yes	3.80	1.27	.551	.266	.221
5	I cannot rely on people, because most people let me down.	Yes	4.01	1.32	.604	.185	.233
6	I believe that other people can be counted on to do what they say they will do.	No	3.65	1.19	.529	.239	.195

Table 6 – continued

7	I find it hard to care about people that I do not know.	Yes	4.35	1.35	.481	.135	.226
8	I find it more rewarding to work with other people toward a common goal than to work alone.	No	4.13	1.26	.361	.204	.222
9	Other people generally have bad intentions.	Yes	4.41	1.16	.543	.211	.290
10	Other people generally have good intentions.	No	4.08	1.01	.554	.226	.235
11	I find it hard to trust other people.	Yes	3.68	1.38	.544	.218	.221
12	I consider myself to be an extremely caring and nurturing person.	No	4.86	1.12	.339	.075	.152
13	Collaborating with others is fun to me.	No	4.42	1.16	.480	.272	.273
14	Dealing with other people can be quite a pain.	Yes	3.86	1.31	.511	.195	.273
15	Most people are up to no good.	Yes	4.46	1.18	.586	.247	.276
16	Everyone has kindness in their hearts, even if they don't use it.	No	4.56	1.27	.395	.041	.148
17	People seldom pull their weight in group activities.	Yes	3.49	1.20	.332	.176	.209
18	Nothing bothers me more than seeing people treated unfairly.	No	4.88	1.18	.282	.093	.101

Table 6 – continued

| 19 | Other people are incompetent. | Yes | 4.50 | 1.21 | .548 | .152 | .267 |
| 20 | Doing things for others makes me happier than doing things for myself. | No | 4.63 | 1.20 | .374 | .202 | .189 |

Note: R-coded = reverse coded

3.1 Exploration of Factor Structure

The next area of analysis involves the unidimensionality of the COE construct (which addresses Research Question 1). Because there is a lack of methods to test for unidimensionality using polytomous items, this assumption of unidimensionality was explored using an exploratory factor analysis in SPSS with a Principal Components method of extraction. As expected, the results indicated support for an essentially unidimensional solution. The first factor extracted showed an associated eigenvalue of 5.41, with the second showing a value of 1.67, and the third showing 1.17. In addition, the first factor accounted for 39% of the variance, whereas the second and third accounted for 12% and 8%, respectively. Kaiser (1970) recommends using the strict criteria that eigenvalues that are greater than 1 are indicative of separate factors; however, the second and third factors in this study did little in explaining additional variance. Moreover, Reckase (1979) specified that a value of total variance of 20% or more explained by the first principal component is necessary to satisfy the requirement of unidimensionality, which was accomplished in this analysis. In further support of a one-factor solution, an examination of the scree plot visually depicts a significant drop off after the first factor (see Figure 2 for the scree plot). The conclusion of this analysis indicates that there is a dominant factor that is accountable for the majority of the variance and, thus, the COE scale can be considered essentially unidimensional (Bryant & Wooten, 2006). Table 7 displays the factor loadings for the 14 items of the COE scale (none of which were below .400). Because the theory of essential unidimensionality states that unidimensionality can be claimed if there is a clear dominant trait in the presence of minor traits (Stout, 1987), then it is safe to conclude that COE sufficiently meets the requirement of essential unidimensionality, thereby permitting the subsequent IRT analyses.

Table 7 – *Factor Loadings for the 14 items of the COE Scale*

COE Item	Item Wording	Factor 1
1	In general, most people can be trusted.	.645
2	It is hard for me to work with people that I do not know.	.489
3	People tend to follow through with their word.	.651
4	Most people fail to follow through with their word.	.696
5	I cannot rely on people, because most people let me down.	.731
6	I believe that other people can be counted on to do what they say they will do.	.610
7	I find it hard to care about people that I do not know.	.509
8	Other people generally have bad intentions.	.670
9	Other people generally have good intentions.	.595
10	I find it hard to trust other people.	.688
11	Collaborating with others is fun to me.	.445
12	Dealing with other people can be quite a pain.	.580
13	Most people are up to no good.	.697
14	Other people are incompetent.	.622

Scree Plot

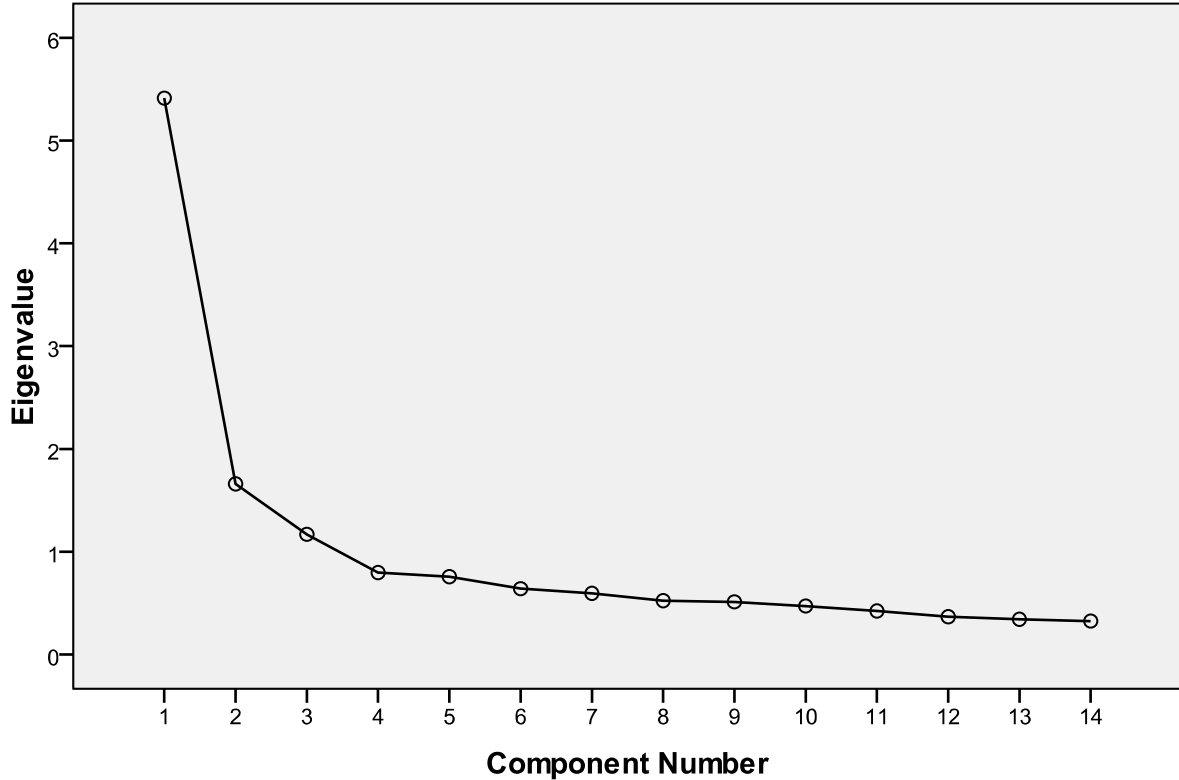*Figure 2*. Scree Plot of 14 item COE Factor Analysis

<u>3.2 IRT Model Fit</u>

As stated previously, the IRT model that best represented the data was Samejima's (1969) Graded Response Model (GRM). In order to test model fit and to address Research Question 2 (which asked, how well do the items of the COE scale fit the proposed model?), an IRT software program called R was used. More specifically, the latent trait model (ltm) package was used (Rizopoulous, 2006). This package is applicable for both dichotomous and polytomous data and calculates the Graded Response Model (GRM) under the Item Response Theory approach (Rizopoulous, 2006). Using this program, all 14 items were input, and the GRM model was selected with a marginal maximum likelihood method of estimation (with 21 quadrature points specified). The total number of participants used in this analysis

was 588. The cronbach's alpha found in this analysis was α = .873, which indicates solid reliability. The log likelihood value obtained was -11641.90 with AIC = 23451.70 and BIC = 23819.15. Table 8 displays the results of the IRT analysis along with discrimination values for each item. As reflected in Table 8, all discrimination values are all acceptable, with the only exception being Item 11 ("Collaborating with others is fun to me.") showing a discrimination value of .773.

Table 8 – *IRT Analysis Results*

Log Lik.	AIC	BIC
-16767.57	33775.13	34299.93

Item	Discrimination
1	1.589
2	1.015
3	1.656
4	1.964
5	2.226
6	1.461
7	1.060
8	1.806
9	1.360
10	1.882
11	0.773
12	1.193
13	1.960
14	1.454

IRT analyses were replicated in another software program in order to gain further detail about the model. EIRT software (Germain, Valois, & Abdous, 2007) was used to obtain fit statistics, parameter estimates, and ability estimates. Using this program, all 14 items were input and the GRM model was selected with a marginal maximum likelihood method of estimation. Ability estimates were calculated using an expected a posteriori (EAP; Bock & Mislevy, 1982) estimation of ability. Table 9 shows the chi-square test of fit for each item. According to the table, all items converged on the specified model (thus addressing Research Question 2), and the global test of fit shows convergence as well. Table 10 provides detailed parameter estimates for all 14 items and response options.

Table 9 – *IRT Analysis – Test of Model Fit*

COE Item	Item Wording	χ^2	DF	p-value
1	In general, most people can be trusted.	30.102	50	0.988
2	It is hard for me to work with people that I do not know.	31.272	50	0.982
3	People tend to follow through with their word.	43.630	50	0.725
4	Most people fail to follow through with their word.	26.033	50	0.998
5	I cannot rely on people, because most people let me down.	28.042	50	0.995
6	I believe that other people can be counted on to do what they say they will do.	33.482	50	0.965
7	I find it hard to care about people that I do not know.	35.780	50	0.935
8	Other people generally have bad intentions.	14.244	50	1.000
9	Other people generally have good intentions.	36.229	50	0.928
10	I find it hard to trust other people.	31.495	50	0.981
11	Collaborating with others is fun to me.	19.150	50	1.000
12	Dealing with other people can be quite a pain.	19.281	50	1.000
13	Most people are up to no good.	11.340	50	1.000
14	Other people are incompetent.	11.970	50	1.000

Global   372.047   700   1.000

Table 10 – *IRT Parameter Estimates*

Item	Option	Slope (a)	s.e.	Threshold (b)	s.e.
COE01	*1*	1.599	0.100	-2.383	0.237
COE01	*2*	1.599	0.100	-1.757	0.169
COE01	*3*	1.599	0.100	-0.632	0.077
COE01	*4*	1.599	0.100	0.437	0.066
COE01	*5*	1.599	0.100	1.848	0.183
COE01	*6*	1.599	0.100	2.691	0.276
COE02	*1*	1.023	0.086	-3.214	0.368
COE02	*2*	1.023	0.086	-2.492	0.278
COE02	*3*	1.023	0.086	-1.197	0.149
COE02	*4*	1.023	0.086	-0.152	0.081
COE02	*5*	1.023	0.086	1.087	0.141
COE02	*6*	1.023	0.086	1.855	0.222
COE03	*1*	1.665	0.102	-2.743	0.281
COE03	*2*	1.665	0.102	-1.981	0.192
COE03	*3*	1.665	0.102	-0.670	0.078
COE03	*4*	1.665	0.102	0.512	0.069
COE03	*5*	1.665	0.102	1.966	0.194
COE03	*6*	1.665	0.102	2.788	0.286
COE04	*1*	1.977	0.112	-2.166	0.208
COE04	*2*	1.977	0.112	-1.695	0.156
COE04	*3*	1.977	0.112	-0.764	0.079
COE04	*4*	1.977	0.112	0.165	0.046
COE04	*5*	1.977	0.112	1.268	0.122
COE04	*6*	1.977	0.112	1.902	0.186
COE05	*1*	2.243	0.123	-2.137	0.205
COE05	*2*	2.243	0.123	-1.669	0.153
COE05	*3*	2.243	0.123	-0.859	0.084
COE05	*4*	2.243	0.123	-0.092	0.040
COE05	*5*	2.243	0.123	0.903	0.089
COE05	*6*	2.243	0.123	1.472	0.143

Table 10 – continued

Item	Option	Slope (a)	s.e.	Threshold (b)	s.e.
COE06	*1*	1.461	0.096	-2.604	0.266
COE06	*2*	1.461	0.096	-1.982	0.195
COE06	*3*	1.461	0.096	-0.784	0.092
COE06	*4*	1.461	0.096	0.423	0.069
COE06	*5*	1.461	0.096	1.968	0.200
COE06	*6*	1.461	0.096	2.883	0.302
COE07	*1*	1.069	0.088	-3.740	0.435
COE07	*2*	1.069	0.088	-2.893	0.322
COE07	*3*	1.069	0.088	-1.601	0.185
COE07	*4*	1.069	0.088	-0.623	0.100
COE07	*5*	1.069	0.088	0.700	0.105
COE07	*6*	1.069	0.088	1.493	0.182
COE08	*1*	1.808	0.110	-2.902	0.307
COE08	*2*	1.808	0.110	-2.410	0.237
COE08	*3*	1.808	0.110	-1.545	0.150
COE08	*4*	1.808	0.110	-0.613	0.073
COE08	*5*	1.808	0.110	0.669	0.077
COE08	*6*	1.808	0.110	1.392	0.143
COE09	*1*	1.362	0.095	-3.736	0.429
COE09	*2*	1.362	0.095	-3.012	0.317
COE09	*3*	1.362	0.095	-1.662	0.173
COE09	*4*	1.362	0.095	-0.171	0.062
COE09	*5*	1.362	0.095	1.674	0.179
COE09	*6*	1.362	0.095	2.652	0.285
COE10	*1*	1.902	0.110	-2.059	0.198
COE10	*2*	1.902	0.110	-1.478	0.138
COE10	*3*	1.902	0.110	-0.554	0.067
COE10	*4*	1.902	0.110	0.178	0.048
COE10	*5*	1.902	0.110	1.281	0.125
COE10	*6*	1.902	0.110	1.995	0.197

Table 10 – continued

Item	Option	Slope (a)	s.e.	Threshold (b)	s.e.
COE11	*1*	0.774	0.082	-6.023	0.849
COE11	*2*	0.774	0.082	-4.912	0.648
COE11	*3*	0.774	0.082	-2.786	0.363
COE11	*4*	0.774	0.082	-0.925	0.152
COE11	*5*	0.774	0.082	1.081	0.171
COE11	*6*	0.774	0.082	2.243	0.309
COE12	*1*	1.204	0.089	-2.954	0.319
COE12	*2*	1.204	0.089	-2.253	0.235
COE12	*3*	1.204	0.089	-1.023	0.121
COE12	*4*	1.204	0.089	0.052	0.068
COE12	*5*	1.204	0.089	1.507	0.167
COE12	*6*	1.204	0.089	2.415	0.263
COE13	*1*	1.963	0.115	-2.701	0.281
COE13	*2*	1.963	0.115	-2.302	0.224
COE13	*3*	1.963	0.115	-1.497	0.143
COE13	*4*	1.963	0.115	-0.605	0.070
COE13	*5*	1.963	0.115	0.549	0.066
COE13	*6*	1.963	0.115	1.218	0.125
COE14	*1*	1.457	0.099	-3.313	0.365
COE14	*2*	1.457	0.099	-2.821	0.291
COE14	*3*	1.457	0.099	-1.755	0.179
COE14	*4*	1.457	0.099	-0.704	0.088
COE14	*5*	1.457	0.099	0.490	0.073
COE14	*6*	1.457	0.099	1.206	0.135

Finally, Figures 3 – 17 display the Item Characteristic Curves (ICC), Option Characteristic Curves (OCC), Item Information Functions (IIF) for each item, and Test Information Function (TIF) for the entire scale. Based on a review of the ICCs, OCCs, and IIFs; all items show visually acceptable difficulty and discrimination parameters. The only exception is item 11 ("Collaborating with others is fun to me."), which shows a lower slope value and higher item thresholds across response options. Still and all, each item shows convergence with the specified model.
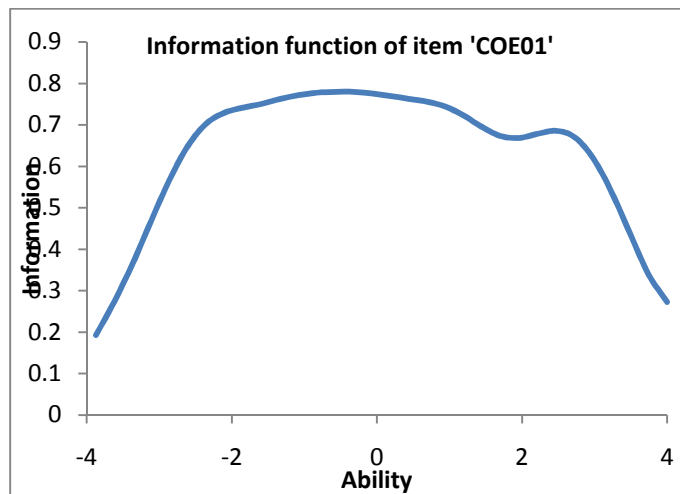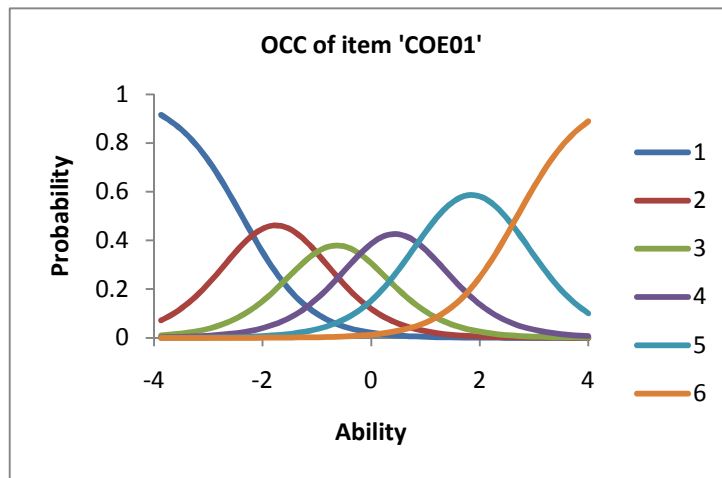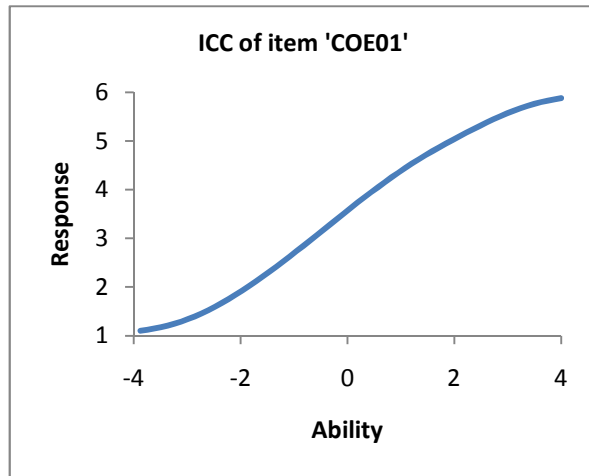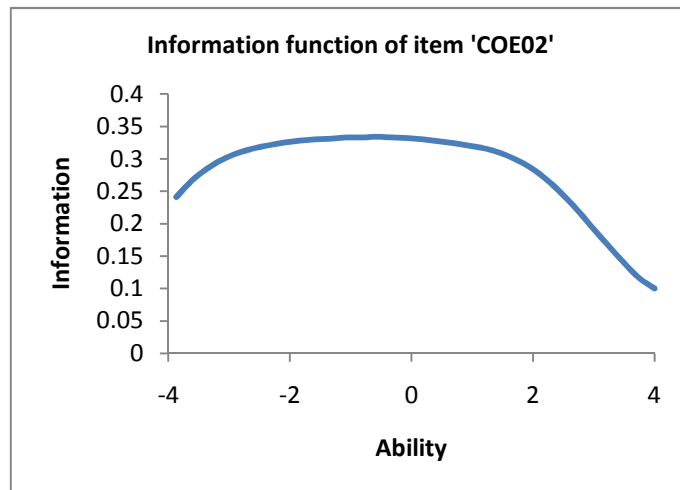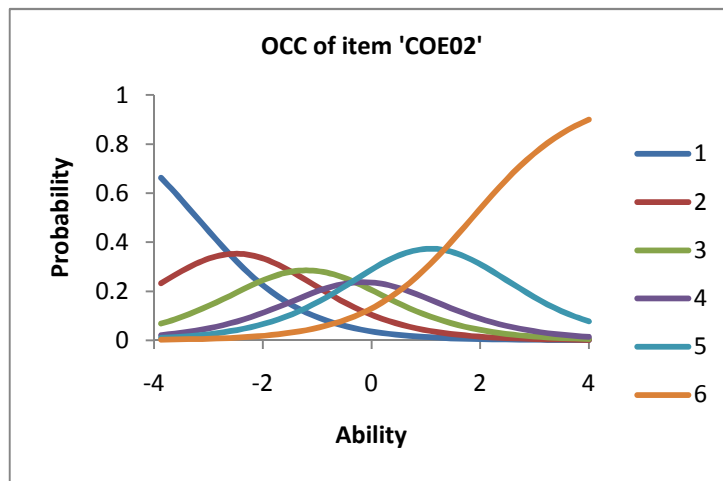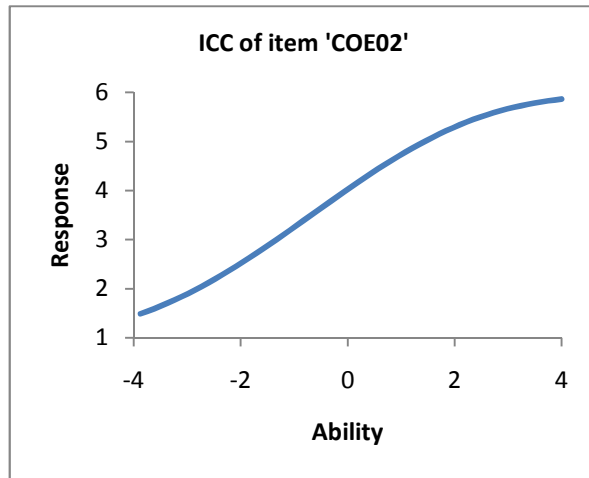
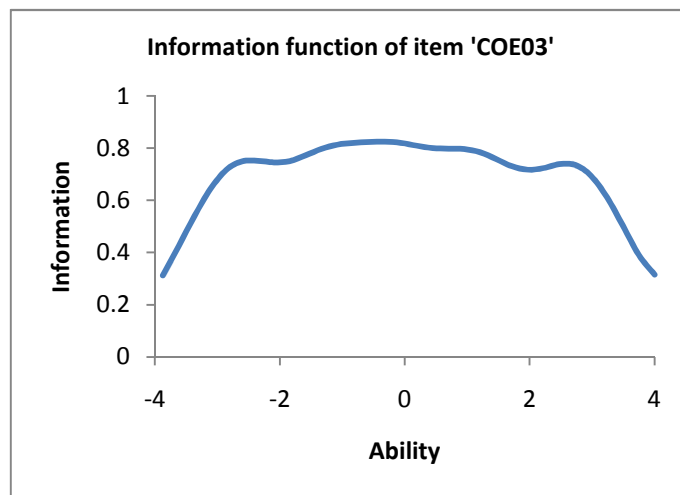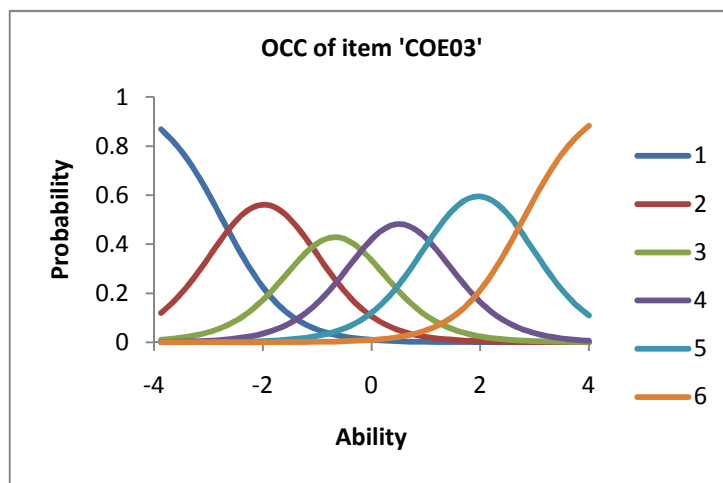*Figure 3.* ICC, OCC, and IIF Graphs for COE Item 1
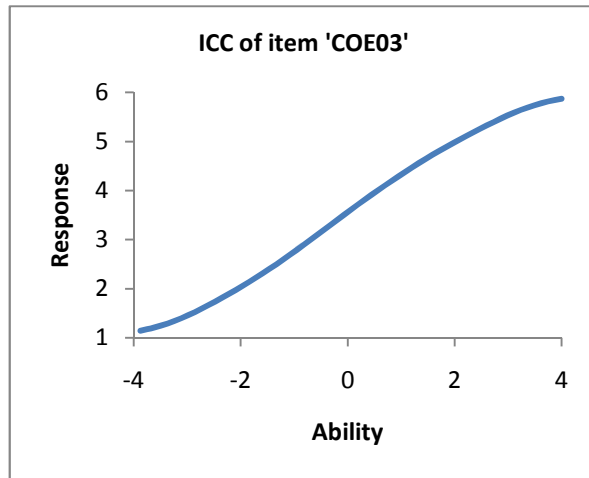
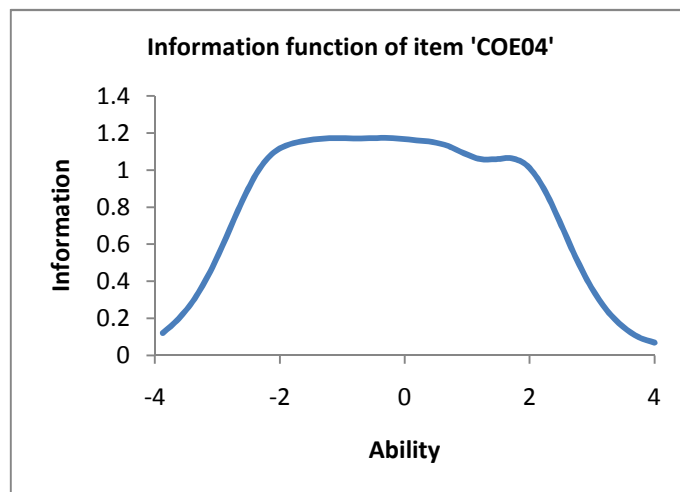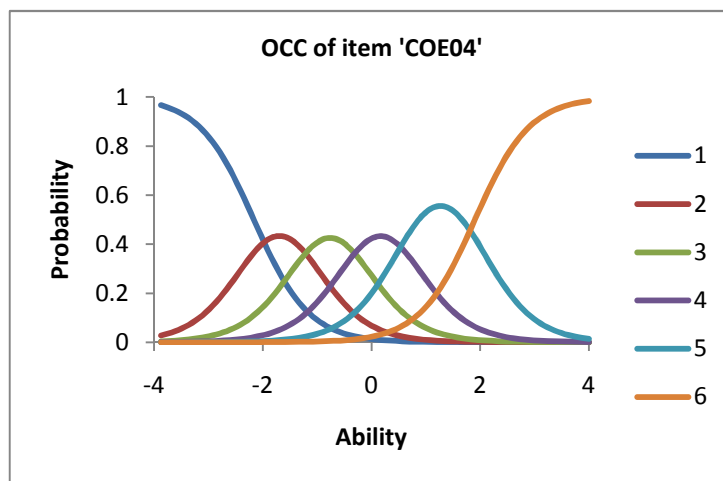*Figure 4.* ICC, OCC, and IIF Graphs for COE Item 2

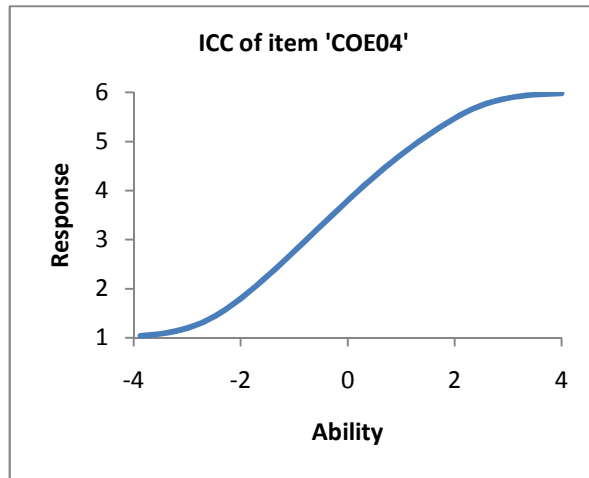*Figure 5.* ICC, OCC, and IIF Graphs for COE Item 3
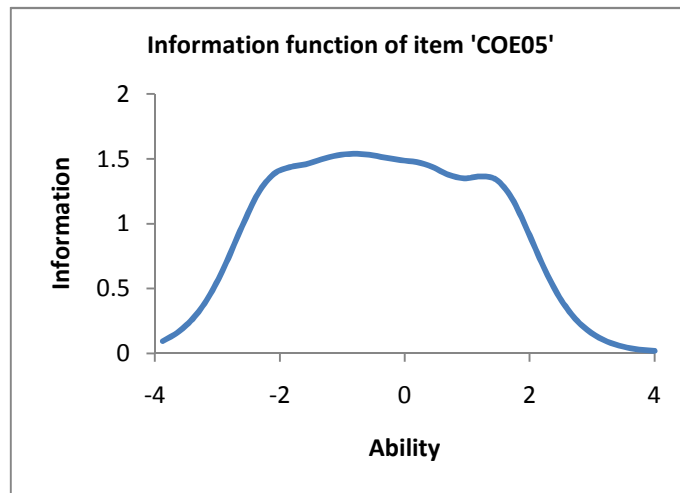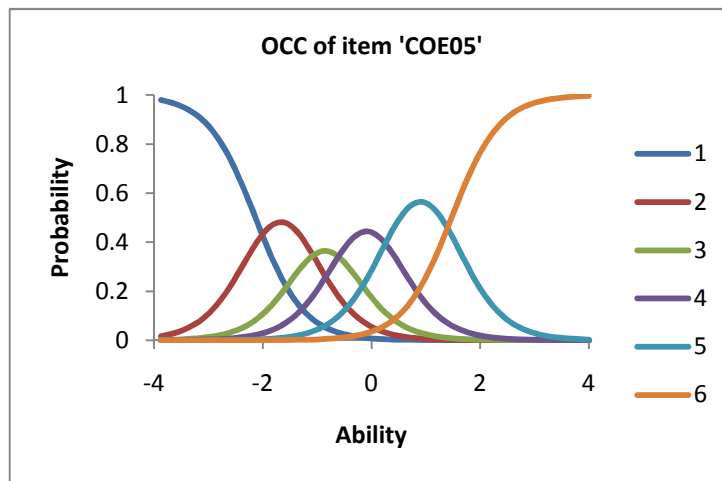
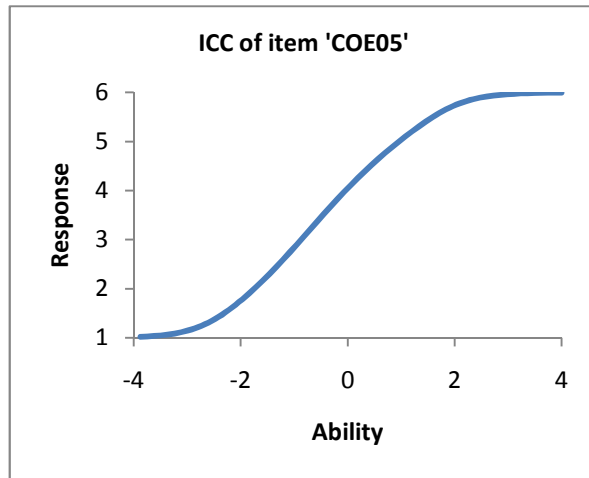*Figure 6.* ICC, OCC, and IIF Graphs for COE Item 4

Figure 7. ICC, OCC, and IIF Graphs for COE Item 5

Figure 8. ICC, OCC, and IIF Graphs for COE Item 6

Figure 9. ICC, OCC, and IIF Graphs for COE Item 7

Figure 10. ICC, OCC, and IIF Graphs for COE Item 8
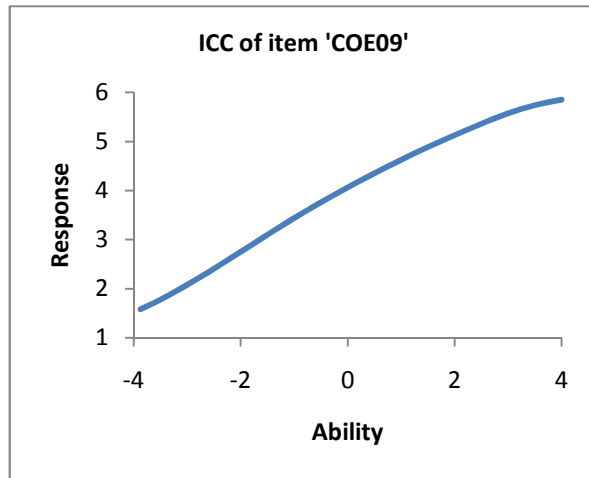
Figure 11. ICC, OCC, and IIF Graphs for COE Item 9
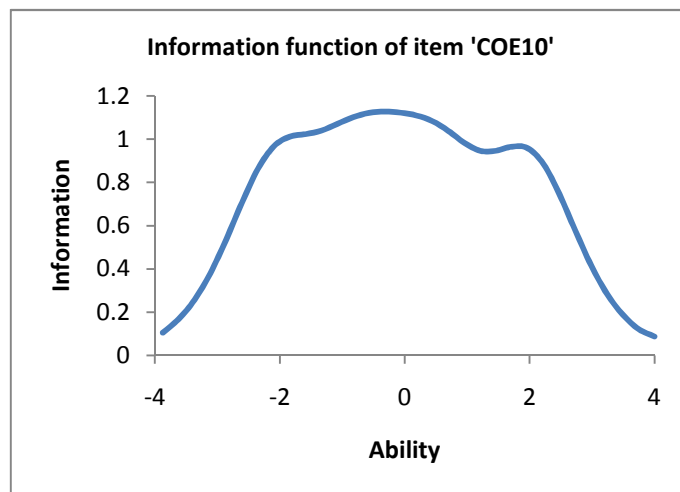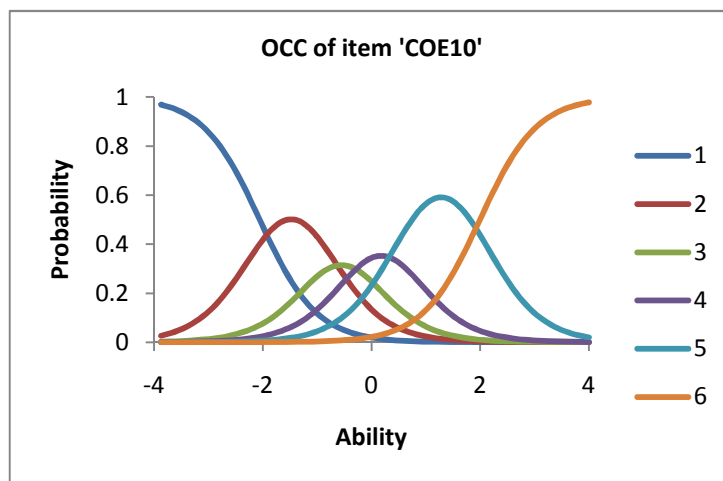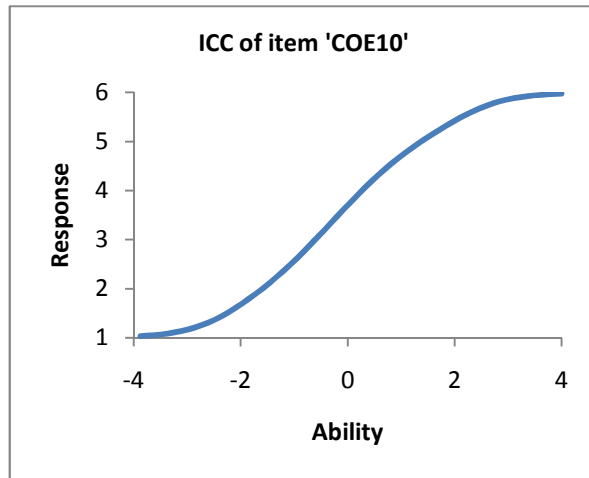
Figure 12. ICC, OCC, and IIF Graphs for COE Item 10
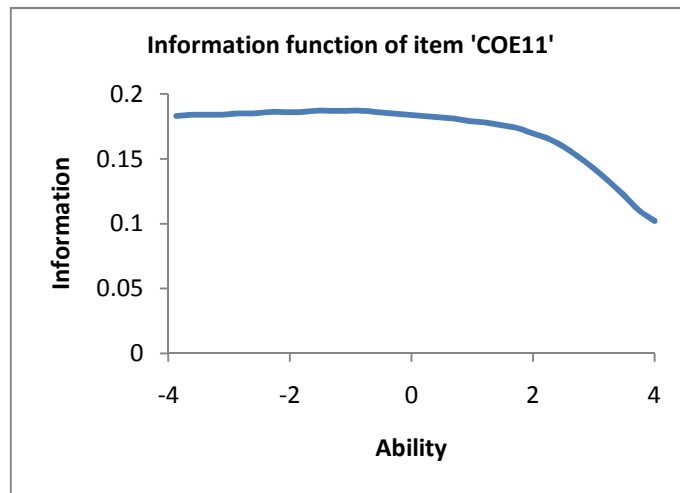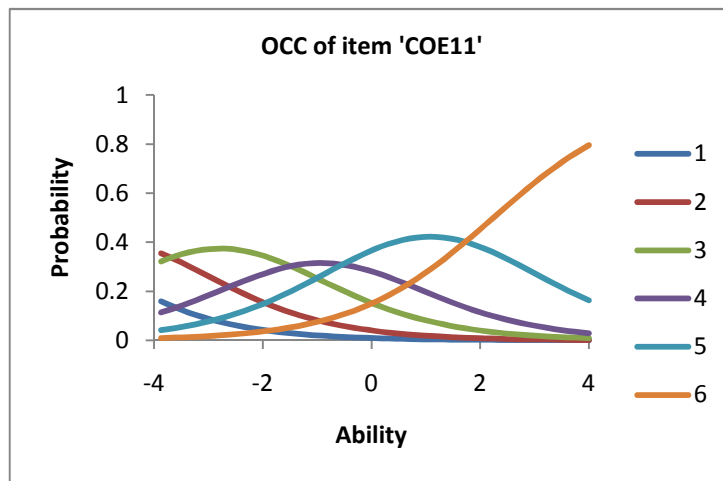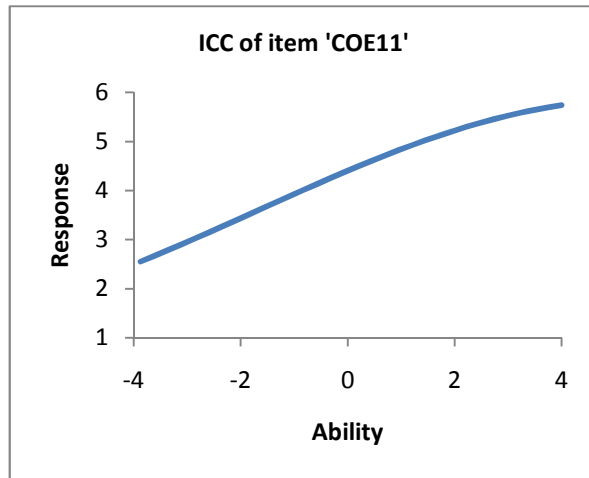
Figure 13. ICC, OCC, and IIF Graphs for COE Item 11
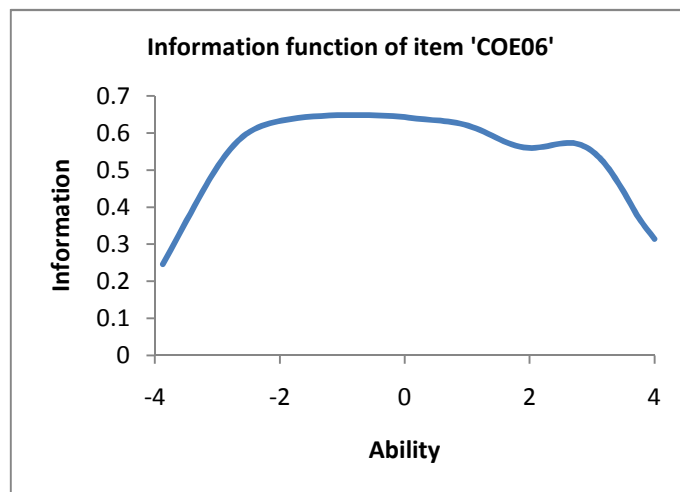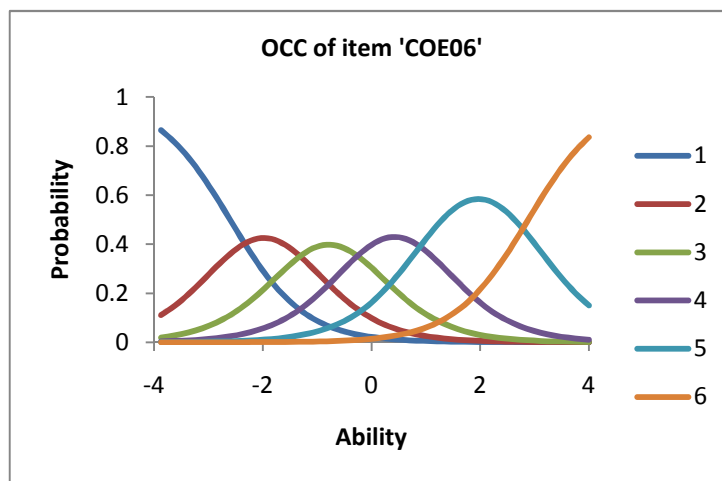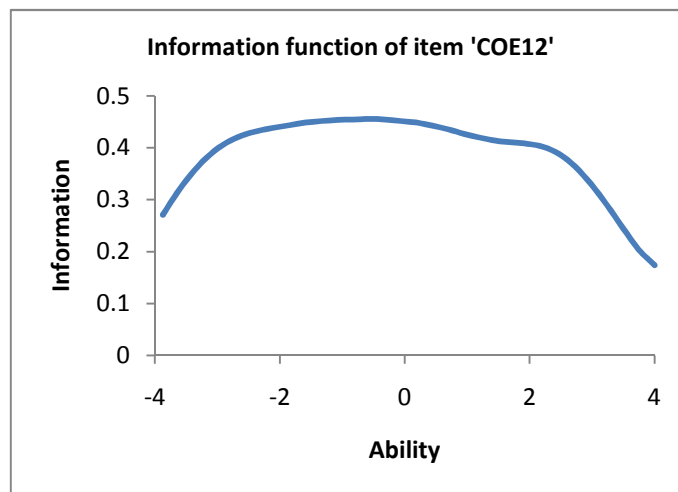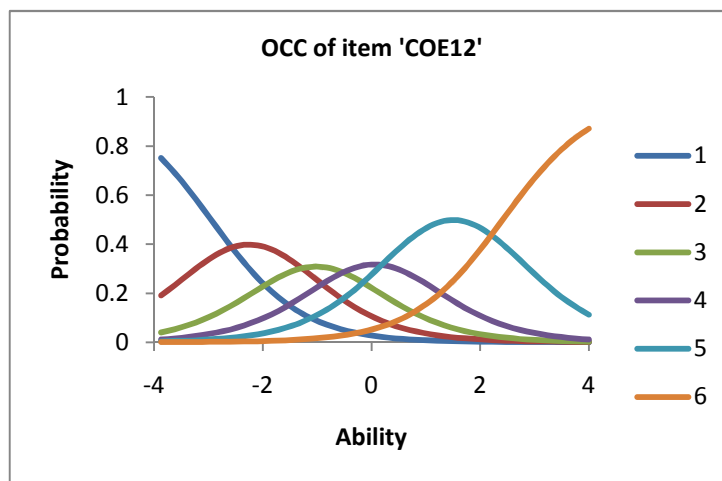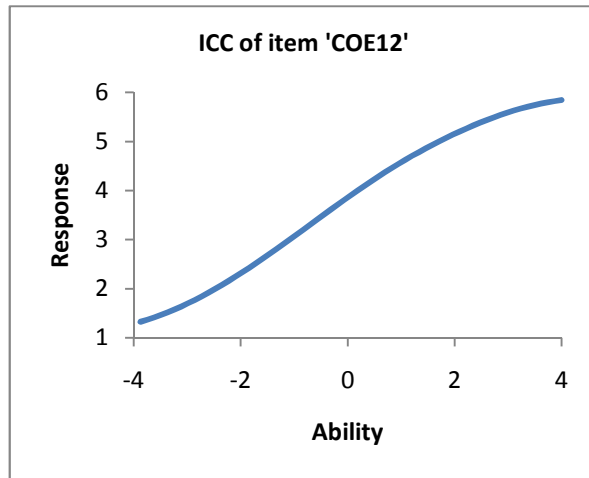
Figure 14. ICC, OCC, and IIF Graphs for COE Item 12
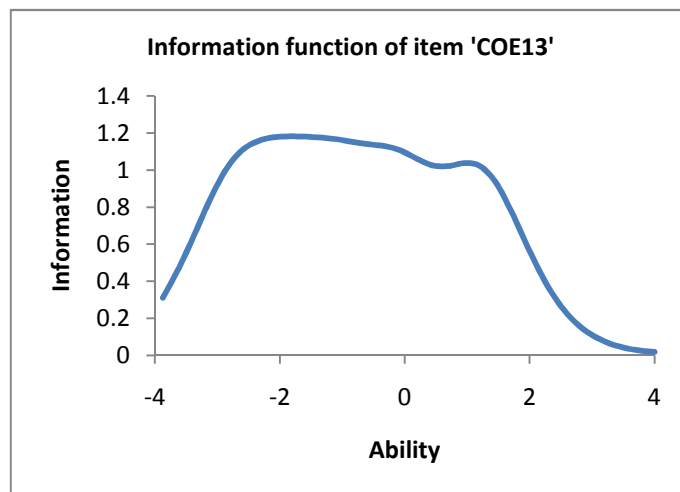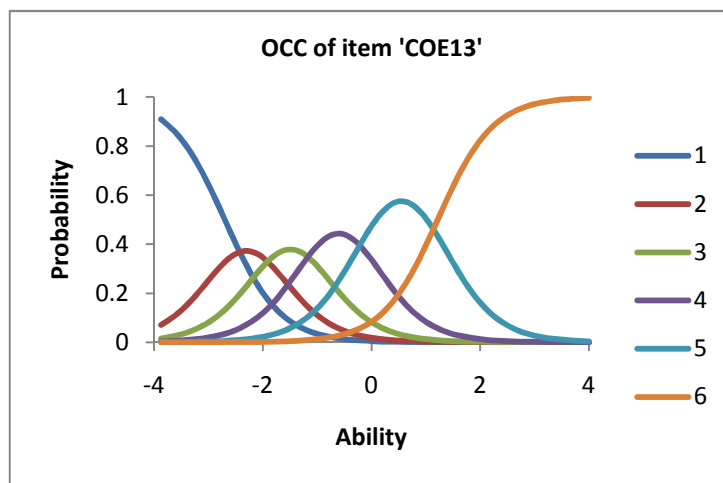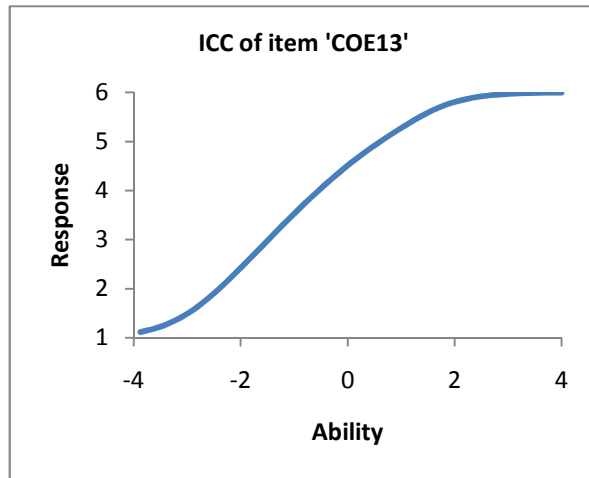
Figure 15. ICC, OCC, and IIF Graphs for COE Item 13

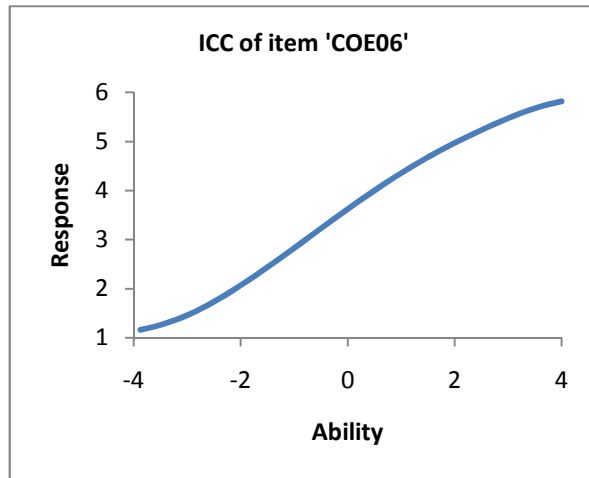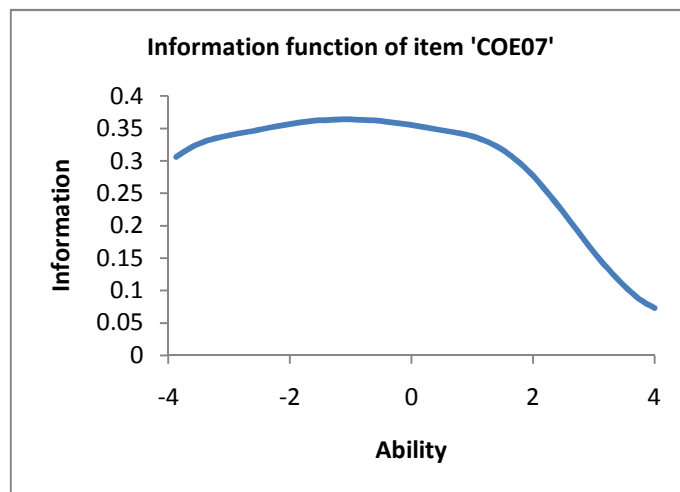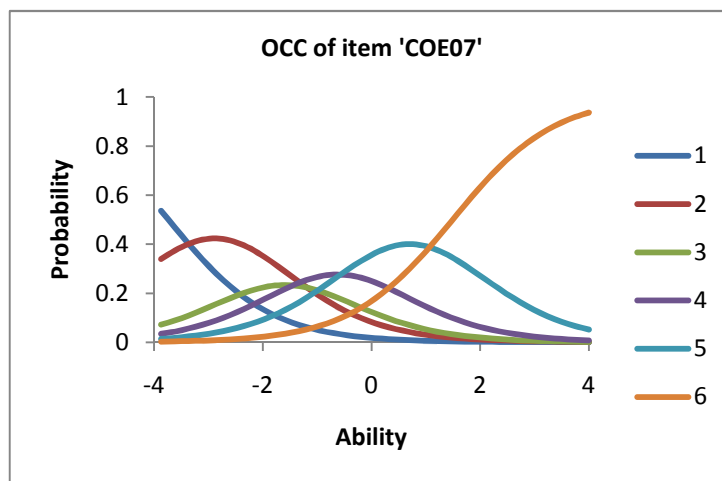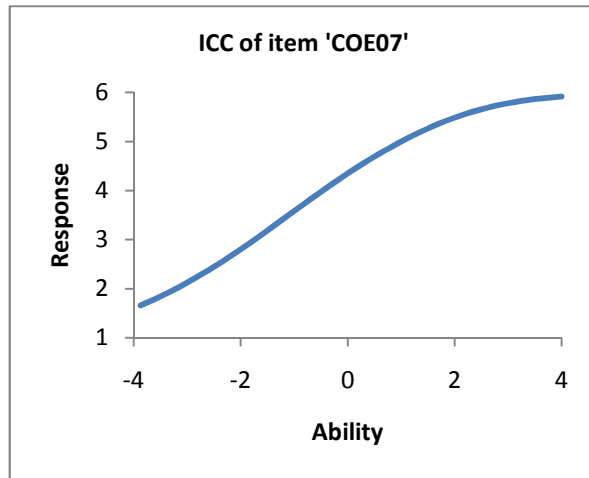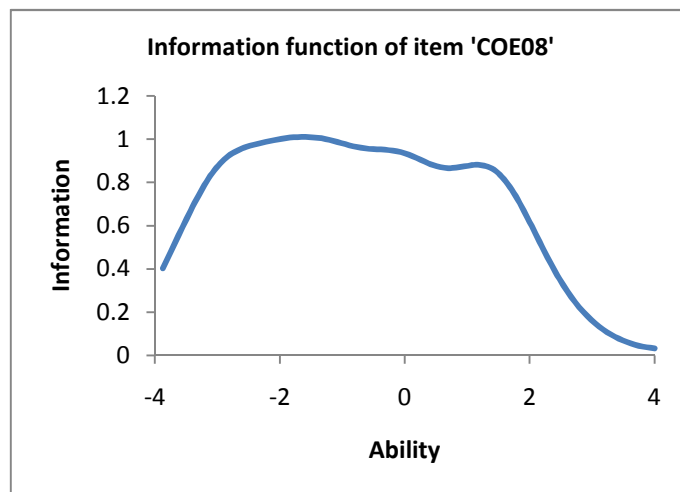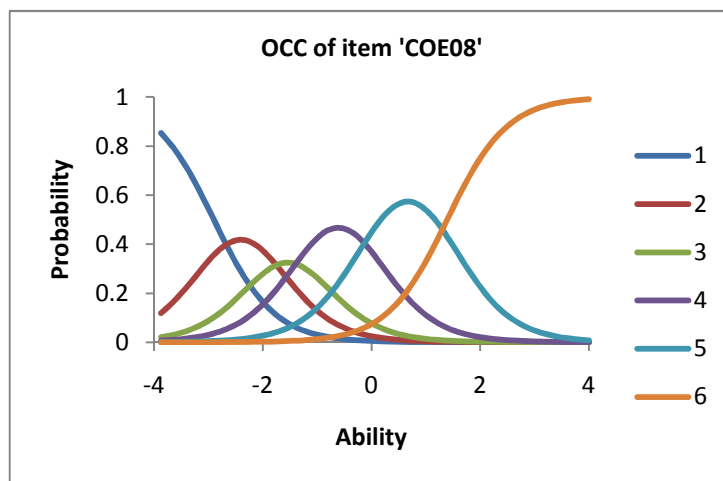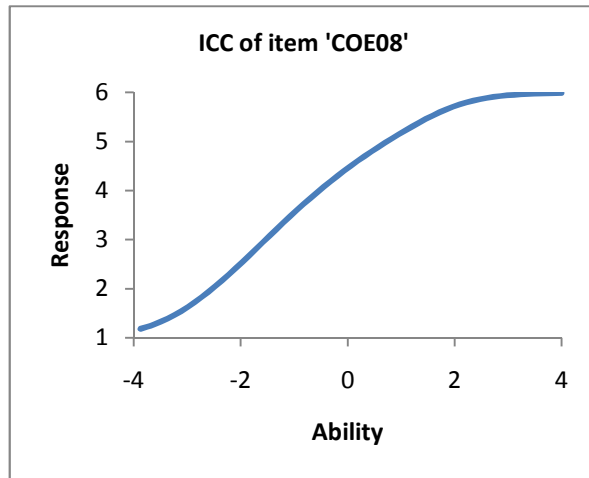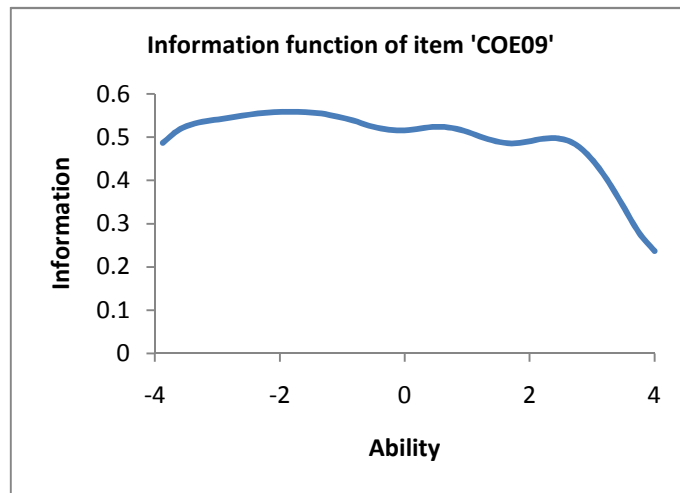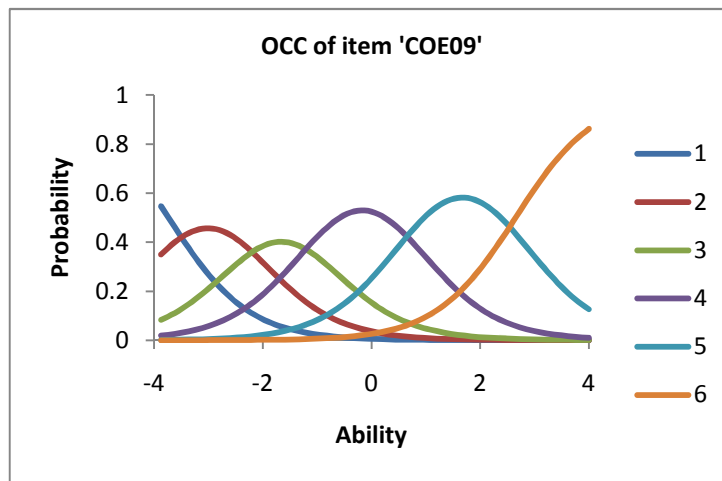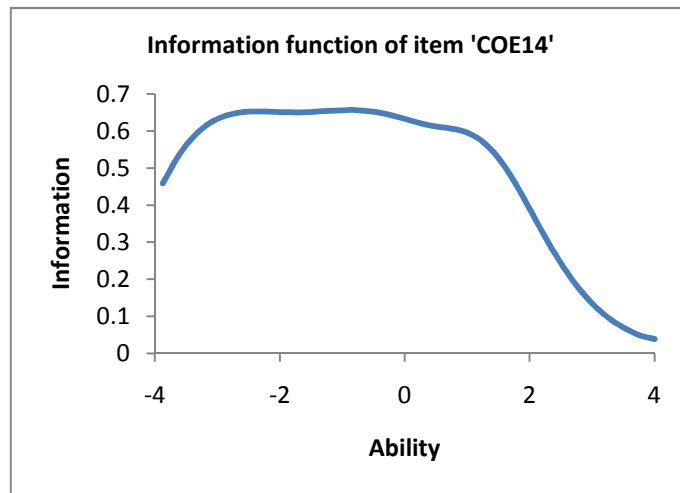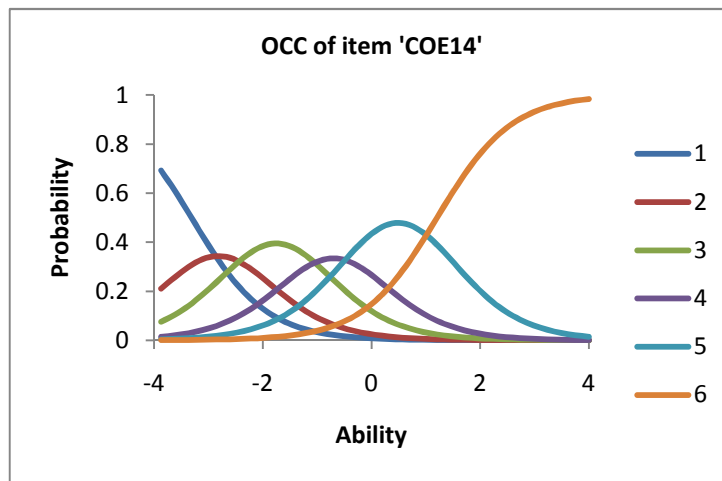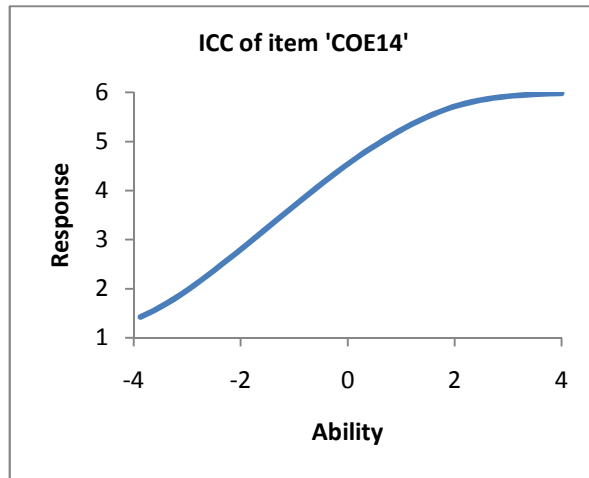*Figure 16.* ICC, OCC, and IIF Graphs for COE Item 14
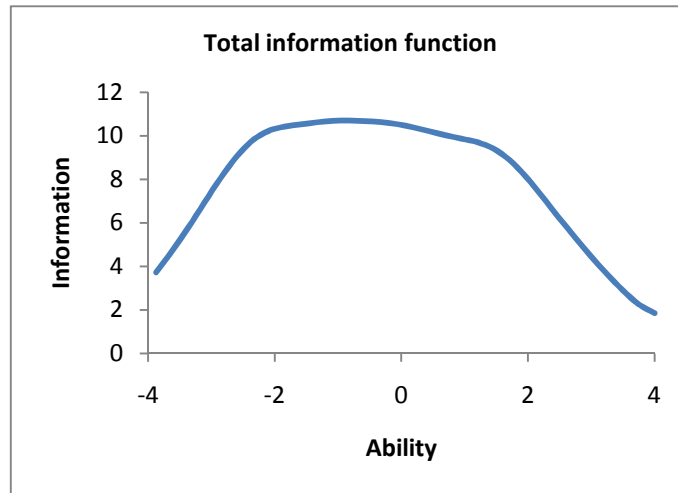
Figure 17. Test Information Function

3.3 Convergent and Discriminant Validity

In order to assist in developing a nomological network for COE, convergent and discriminant validity procedures were tested and analyzed. As previously stated, convergent validity is established if like constructs are found to be highly related. Hypothesis 1 claimed that the COE measure would be positively related to trust, agreeableness, collectivism, and other orientation. This hypothesis was tested in two different ways. First, an aggregate value of trust, agreeableness, collectivism, and other orientation was created in order to compare the aggregate of the four measures with that of the COE measure. After this variable was created, COE and the aggregate scale were correlated with each other. Results of this analysis support Hypothesis 1. The correlation obtained between COE and the aggregate of the four manifest traits was $r = .626$, $p < .001$, which represents a strong relationship according to Cohen's (1988) guidelines. In addition to examining the aggregate score between the four manifest variables and COE, each variable's relationship with COE was examined individually. The results indicated that trust ($r = .659$, $p < .001$) and agreeableness ($r = .509$, $p < .001$) showed strong significant correlations with COE, whereas collectivism ($r = .333$, $p < .001$) and other orientation ($r = .257$, $p < .001$) showed weaker correlations, though still statistically significant. A correction was applied for unreliability. When this correction for attenuation was calculated for these correlations, the resulting correlations for COE and trust ($r_c = .81$, $p < .001$) and COE and agreeableness ($r_c = .62$, $p < .001$) increased noticeably, and the correlations for COE and collectivism ($r_c = .36$, $p < .001$) and COE and other orientation ($r_c = .29$, $p < .001$) remained relatively

78

similar. The conclusion of this analysis is that COE is highly related to trust and agreeableness and related (but to a lesser degree) to collectivism and other orientation. Overall results of this analysis provide support for Hypothesis 1.

Hypothesis 2 was concerned with the relationship between the COE measure and the Team Player Inventory (TPI; Kline, 1999). The TPI is a dispositional measure of teamwork developed specifically for organizational team environments. Due to their common goal in measuring an individual's propensity to excel in a team environment, it was predicted that COE would be positively related to the TPI. The results showed support for Hypothesis 2a. The uncorrected correlation coefficient between COE and TPI was $r = .578$, $p <.001$, and the corrected correlation was $r_c = .617$, $p <.001$. Both of these values fall in accordance with Cohen's (1988) guidelines of a strong practical relationship, signifying that the measures of COE and TPI are highly related (thus supporting Hypothesis 2a), yet not completely redundant.

In addition to determining the relationship between COE and the TPI, Hypothesis 2b questioned whether COE would add value beyond the TPI in the prediction of work outcomes. The outcomes under analysis for this hypothesis were team satisfaction and job satisfaction. In order to test this hypothesis, a hierarchical regression was performed by entering the TPI into the first block and then the COE measure into the second block. Assumptions of linearity, homoscedasticity, and independence were examined prior to analysis.

Beginning with team satisfaction, Hypothesis 2b was supported by a significant R^2 change with the COE scale added to the regression model, $\Delta R^2 =.285$, $F_{\Delta r2}$ (1, 85) = 52.014, $p < .001$. Overall, 53.4% of the variance in team satisfaction was explained by COE in the second model, $F(2, 85) = 48.75$, $p < .001$, with an additional 28.5% of the variance being explained by COE over and beyond the TPI. In addition, the regression coefficient for COE was statistically significant in the second model at a $p < .001$ level. The results indicate that the COE scale adds considerably above and beyond the TPI scale in the prediction of team satisfaction (see Table 11 for a summary of the team satisfaction regression results).

Table 11 – *Hierarchical Regression Results – COE and TPI on Team Satisfaction*

Model		B	se_b	β	Sig
Step 1	Constant	3.177	3.352	--	.346
	TPI	0.478	0.089	.499	.000
Step 2	Constant	-2.247	2.760	--	.418
	TPI	0.063	0.091	.066	.492
	COE	0.396	0.055	.688	.000

With regard to job satisfaction, the same pattern of results emerged. A significant R^2 change was evidenced when COE was added to the regression model, $\Delta R^2 = .179$, $F_{\Delta r2}$ (1, 85) = 21.33, $p < .001$. Overall, 28.5% of the variance in job satisfaction was explained by COE in the second model, $F(2, 85) = 16.93$, $p < .001$, with an additional 18% of the variance being explained by COE over and beyond the TPI. In addition, the regression coefficient for COE was statistically significant in the second model at a $p < .001$ level. Once again, these results imply that the COE scale accounts for unique variance beyond the TPI scale in the prediction of job satisfaction (see Table 12 for a summary of the job satisfaction regression results). The overall results indicate support for Hypothesis 2b.

Table 12 – *Hierarchical Regression Results – COE and TPI on Job Satisfaction*

Model		B	se_b	β	Sig
Step 1	Constant	7.091	3.672	--	.057
	TPI	0.312	0.098	.325	.002
Step 2	Constant	2.772	3.433	--	.422
	TPI	-.019	0.113	-.019	.871
	COE	0.315	0.068	.546	.000

Shifting to discriminant validity, it was deemed necessary to show the empirical distinction between COE and other theoretically unrelated constructs. In an effort to show discriminant validity, two dimensions of the Big Five personality dimensions (conscientiousness and openness) were compared to COE. Hypothesis 3a stated that COE would have a stronger relationship with trust, agreeableness, other orientation, and collectivism than it would with conscientiousness or openness. In order to test this hypothesis; correlations between COE, conscientiousness, openness, and the aggregate COE score (of trust, agreeableness, other orientation, and collectivism) were calculated and compared. Starting with conscientiousness, the correlation coefficient obtained between conscientiousness and COE was $r = .349$, $p < .001$, which signifies a moderate correlation, and the correlation coefficient obtained between COE and the aggregate score was $r = .626$, $p < .001$, which signifies a large correlation. A Fisher's R to Z transformation was performed on the correlation coefficients of interest to determine if there was a significant different between the two. The results indicated partial support for Hypothesis 3 ($z = 5.88$, $p < .05$). Shifting to openness, the correlation between COE and openness was low ($r = .128$, $p < .004$). The Fisher's R to Z transformation showed a significant difference between COE and openness ($z = 9.74$, $p < .05$). It is safe to conclude that there is clear distinction between these two constructs. Hypothesis 3 was fully supported.

Another measure of discriminant validity was CSE. Even though CSE and COE are both a part of the core evaluations tripartite model, they are still predicted to be disparate constructs. Because of their common development and theoretical background, it is understood that COE and CSE are likely related. However, Hypothesis 4 stated that COE would have higher relations with trust, agreeableness, other orientation, and collectivism than it would with CSE. This hypothesis was not supported. The relationship between COE and CSE was relatively high ($r = .565$, $p < 001$) and was not found to be distinct from the relationship between COE and the aggregate COE score (which was $r = .626$, $p < .001$). The Fisher's R to Z transformation revealed a non-significant value ($z = 0.84$, $p = .200$).

<u>3.4 Criterion-Related Validity</u>

In order to test the practical significance of COE, the scale's ability to predict important work outcomes (such as job satisfaction, team satisfaction, and team effectiveness) was examined. Hypothesis 5 stated that COE would show a positive relationship with job satisfaction. In order to test this hypothesis,

a regression analysis was performed with COE serving as the independent variable and job satisfaction serving as the dependent variable. Based on this analysis, Hypothesis 5 was supported (R^2=.171, F (1, 337) = 69.44, p < .001). Overall, 17% of the variance in job satisfaction was explained by COE. In addition, the regression coefficient for COE was statistically significant at a p < .001 level (B = .249, se_b = .030, β = .413). This indicates that COE successfully predicts job satisfaction. In addition, an examination of the correlations with all facets of job satisfaction was explored. The facets of job satisfaction included in this study were as follows: overall job satisfaction, task satisfaction, satisfaction with pay, satisfaction with promotions, satisfaction with supervisor, and satisfaction with co-workers (Smith, Kendall, & Hulin, 1969). As depicted in Table 13, all correlations with COE show significance at the p <.001 level. It should be noted that, though not statistically different, the interpersonal facets of job satisfaction (satisfaction with supervisor and satisfaction with coworkers) showed stronger relationships with COE in terms of magnitude than did the non-interpersonal facets of job satisfaction (task satisfaction, satisfaction with pay, and satisfaction with promotions).

Table 13 – *Correlations between COE and Job Satisfaction Facets*

Scale	Mean	SD	1	2	3	4	5	6	7
1. COE	56.27	11.09	1	.41**	.32**	.21**	.26**	.39**	.41**
2. Job Sat	20.10	6.69	-	1	.71**	.46**	.67**	.68**	.62**
3. Task Sat	18.58	6.32	-	-	1	.27**	.55**	.53**	.55**
4. Pay Sat	18.85	5.46	-	-	-	1	.39**	.40**	.38**
5. Prom Sat	18.21	5.78	-	-	-	-	1	.58**	.51**
6. Supervisor	21.50	6.22	-	-	-	-	-	1	.68**
7. Coworkers	22.09	5.39	-	-	-	-	-	-	1

Note: N = 339. **p < .01

Hypothesis 6 was concerned with COE's ability to predict team satisfaction, stating that there would be a positive relationship between COE and team satisfaction. Because of the fundamental premise behind COE and the inherent feelings toward others that are associated with the construct, COE is expected to show a strong prediction for satisfaction with teams. More specifically, individuals high in core other-evaluations are expected to value teamwork and, thus, gain more satisfaction from working within this type of team framework. Conversely, individuals lower in core other-evaluations are expected to show less satisfaction with teams due to their fundamental, negative feelings about others. This

hypothesis was also supported (R^2=.269, F (1, 337) = 124.11, $p < .001$). Overall, 27% of the variance in team satisfaction was explained by COE. Like job satisfaction, the regression coefficient for COE and team satisfaction was statistically significant at a $p < .001$ level ($B = .262$, $se_b = .024$, $\beta = .519$). Again, this result signifies that COE successfully predicts team satisfaction.

The final work outcome under examination was team effectiveness. It was predicted that COE would add value in the prediction of work team success and, therefore, should predict team effectiveness in addition to team satisfaction. In this vein, it was hypothesized that higher team COE averages would lead to higher team outcomes (Hypothesis 7a). The measure of team effectiveness used in the current study contained 7 facets of team effectiveness (output effectiveness, quality effectiveness, change effectiveness, planning and organizing effectiveness, interpersonal effectiveness, value effectiveness, and overall effectiveness; Pearce & Sims, 2002). First, an effort was made to examine the relationship between COE and team effectiveness. Results showed a significant relationship between COE and all 7 facets of team effectiveness: output ($r = .424$, $p < .001$), quality ($r = .347$, $p < .001$), change ($r = .339$, $p < .001$), planning and organizing ($r = .444$, $p < .001$), interpersonal ($r = .310$, $p < .001$), value ($r = .505$, $p < .001$), and overall ($r = .361$, $p < .001$). All of these correlations show moderate to large practical significance. In addition, COE is a strong predictor of overall team effectiveness, $R^2 = .129$, $F(1, 103) = 15.22$, $p < .001$, accounting for 13% of the variance in overall team effectiveness ($B = .100$, $se_b = .026$, $\beta = .359$). Table 14 displays the correlations obtained between COE and all facets of team effectiveness.

Table 14 – *Correlations between COE and Team Effectiveness Facets*

Scale	Mean	SD	1	2	3	4	5	6	7	8
1. COE	56.27	11.09	1	.42**	.35**	.35**	.45**	.31**	.51**	.36**
2. Output	24.19	4.00	-	1	.80**	.75**	.82**	.63**	.72**	.84**
3. Quality	13.96	2.67	-	-	1	.76**	.82**	.61**	.64**	.82**
4. Change	14.10	3.20	-	-	-	1	.78**	.64**	.62**	.73**
5. Planning	18.87	3.70	-	-	-	-	1	.71**	.68**	.83**
6. Interpersonal	17.15	4.77	-	-	-	-	-	1	.40**	.58**
7. Value	15.89	2.91	-	-	-	-	-	-	1	.72**
8. Overall	19.56	3.58	-	-	-	-	-	-	-	1

Note: $N = 105$. **$p < .01$

Shifting gears to address the hypothesized prediction (Hypothesis 7a); Table 15 shows the relationship between the COE team average and the seven team effectiveness variables. Based on this data, Hypothesis 7a is partially supported. All of the correlations are positive and moderate in magnitude; however, only 3 of them are statistically significant. However, a smaller sample size may have affected this finding ($N = 32$). Despite the lower sample size, the trend clearly demonstrates COE's value in regard to team effectiveness. The value in this analysis was its ability to provide an initial, cursory glance into the effect that COE has on team effectiveness.

Table 15 – *Correlations between Average Team COE and Team Effectiveness Facets*

Scale	Mean	SD	1	2	3	4	5	6	7	8
1. COE	3.65	.865	1	.25	.27	.39*	.36**	.25	.43**	.22
2. Output	4.81	.569	-	1	.86**	.83**	.94**	.66**	.72**	.90**
3. Quality	4.62	.747	-	-	1	.82**	.89**	.67**	.61**	.89**
4. Change	4.63	.736	-	-	-	1	.82**	.62**	.66**	.80**
5. Planning	4.67	.632	-	-	-	-	1	.74**	.75**	.91**
6. Interpersonal	4.33	.717	-	-	-	-	-	1	.47**	.59**
7. Value	5.22	.811	-	-	-	-	-	-	1	.74**
8. Overall	4.79	.713	-	-	-	-	-	-	-	1

Note: $N = 32$. * $p < .05$; **$p < .01$

Hypothesis 7b stated that after controlling for team COE averages, teams with less variability in COE scores would have higher team outcomes. This hypothesis was tested using a hierarchical multiple regression. Team averages were input into the first block (as a control), and the standard deviation scores for each team were input into the second block (as a representation of the variability within the team). Overall team effectiveness served as the dependent variable. Prior to analysis, teams with less than two members represented were excluded from the analysis. The test of the main effect would signify whether teams with less variability in COE scores would show more positive team outcomes (as measured by overall team effectiveness) after controlling for team average. The results did not show support for Hypothesis 7b. Even though there was a significant change in R^2 when team variability was added to the model after controlling for team average, $\Delta R^2 = .165$, $F_{\Delta r2}$ (1, 24) = 4.84, $p < .05$, the main effect of team variability was not significant, $F(2, 26) = 2.70$, $p = .088$. Also, the beta value of variability went in the opposite direction of the predicted relationship. As variability in team COE levels increased, so

did positive work outcomes. Again, it should be noted that sample size may have been an issue in this analysis due to the low representation of teams.

3.5 Incremental Validity

Beyond the predictive validity of the COE scale, it was important to establish incremental validity beyond other established measures. The incremental validity of COE over and beyond the TPI in regard to job satisfaction was already established in Hypothesis 2b. It is also important to consider the unique contribution that COE makes above and beyond that of CSE. Because CSE has been shown to produce meaningful effects in the prediction of job satisfaction (Lemelle & Scielzo, 2010) and other important work constructs (Judge, 2009), it is important to determine whether COE can provide added value. The initial examination centered around overall job satisfaction. To determine whether COE added value in the prediction of job satisfaction over and beyond CSE, a hierchical regression was performed by entering CSE into the first block and then COE into the second block. Assumptions of linearity, homoscedasticity, and independence were examined prior to analysis. The results show support for the notion that COE adds value in the prediction of job satisfaction beyond CSE. This was evidenced by a significant change in R^2 when COE was added to the regression model, $\Delta R^2 = .091$, $F_{\Delta r2}$ (1, 336) = 36.84, $p < .001$. Overall, 17.4% of the variance in job satisfaction was explained by COE in the second model, $F(2, 336) = 35.41$, $p < .001$; with an additional 9% of the variance being explained above and beyond CSE. In addition, the regression coefficient for COE was statistically significant in the second model at a $p < .001$ (see Table 16 for the results of this regression analysis). Due to the fact that CSE has strong established predictions of job satisfaction, the result that COE can add value in the prediction above and beyond CSE is a powerful finding.

Table 16 – *Hierarchical Regression Results – COE and CSE on Job Satisfaction*

Model		B	se_b	β	Sig
Step 1	Constant	8.230	2.170	--	.000
	CSE	0.225	0.041	.289	.000
Step 2	Constant	4.590	2.148	--	.033

Table 16 – continued

CSE	0.055	0.048	.070	.252
COE	0.224	0.037	.372	.000

_____

Beyond overall job satisfaction, it was predicted that COE would show strong effects with the interpersonal facets of job satisfaction. More specifically, Hypothesis 8 stated that COE would show a relationship with the interpersonal facets of job satisfaction (e.g., satisfaction with supervisor and satisfaction with co-workers) over and beyond that of core self-evaluations. Beginning with satisfaction with supervisor, Hypothesis 2b was supported by a significant R^2 change with the COE to the regression model, $\Delta R^2 = .060$, $F_{\Delta r2}$ (1, 336) = 24.25, $p < .001$. Overall, 17% of the variance in satisfaction with supervisor was explained by COE in the second model, F(2, 336) = 34.14, $p < .001$; with an additional 6% of the variance being explained above and beyond CSE. In addition, the regression coefficient for COE was statistically significant in the second model at a $p < .001$ (see Table 17 for the results of this regression analysis).

Table 17 – *Hierarchical Regression Results – COE and CSE on Satisfaction with Supervisor*

Model		B	se$_b$	β	Sig
Step 1	Constant	8.895	1.990	--	.000
	CSE	0.239	0.037	.330	.000
Step 2	Constant	6.135	2.004	--	.002
	CSE	0.110	0.044	.152	.014
	COE	0.170	0.034	.303	.000

_____

Examining satisfaction with co-workers, the same pattern emerged with COE showing incremental validity over and beyond CSE, thus fully supporting Hypothesis 8. The R^2 change was significant in the regression model, $\Delta R^2 = .075$, $F_{\Delta r2}$ (1, 336) = 30.45, $p < .001$. Overall, 17.5% of the

variance in satisfaction with coworkers was explained by COE in the second model, $F(2, 336) = 35.67$, p < .001; with an additional 8% of the variance being explained above and beyond CSE. In addition, the regression coefficient for COE was statistically significant in the second model at a p < .001 (see Table 18 for the results of this regression analysis.

Table 18 – *Hierarchical Regression Results – COE and CSE on Satisfaction with Coworker*

Model		B	se_b	β	Sig
Step 1	Constant	11.595	1.734	--	.000
	CSE	0.199	0.032	.317	.000
Step 2	Constant	8.922	1.732	--	.000
	CSE	0.074	0.038	.118	.055
	COE	0.164	0.030	.338	.000

Finally, Hypothesis 9 stated that COE would predict team satisfaction over and beyond core self-evaluations. Even though how you feel about yourself may influence how satisfied you are with your team, it was believed that how you feel about other people would show added predictive value above and beyond core self-evaluations. Once again, this hypothesis was supported. A hierarchical regression showed a significant change in R^2 when COE was added to the regression model, $\Delta R^2 = .101$, $F_{\Delta r2}$ (1, 336) = 48.25, p < .001. Overall, 30% of the variance in team satisfaction was explained by COE in the second model, $F(2, 336) = 71.85$, p < .001; with an additional 10% of the variance being explained above and beyond CSE. In addition, the regression coefficient for COE was statistically significant in the second model at a p < .001 level (see Table 19 for the results of this regression analysis).

Table 19 – *Hierarchical Regression Results – COE and CSE on Team Satisfaction*

Model		B	se_b	β	Sig
Step 1	Constant	8.364	1.699	--	.000
	CSE	0.291	0.032	.446	.000

Table 19 – continued

Step 2	Constant	5.145	1.658	--	.002
	CSE	0.140	0.037	.216	.000
	COE	0.198	0.029	.392	.000

_____

Overall, the results of these analyses show support for the claim that COE demonstrates incremental validity beyond CSE. Through an examination of overall job satisfaction, satisfaction with supervisor, satisfaction with coworkers, and overall satisfaction; results have shown the added value of COE above and beyond CSE. Because CSE has been proven to be effective in determining important work criteria, the results found in the current study suggest that the combination of both CSE and COE may be highly beneficial to utilize in employment settings.

<u>3.6 Group Differences</u>

3.6.1 Differential Item Functioning

Besides establishing validity of the newly developed measure of COE, it is critical to determine whether the measure is susceptible to significant differences between groups. To explore this issue at an item level, differential item functioning was explored for each item of the scale (thereby addressing Research Question 3). Differential item functioning (DIF) is said to be present when items have a different relationship to the latent variable across different groups (e.g., men and women; Embretson & Reise, 2000). In other words, DIF occurs when items function differentially across groups that have the same ability on the related construct. From an IRT perspective, DIF is present when item characteristic curves are different across groups. It is important, however, to distinguish between DIF and test bias (Shealy & Stout, 1993). Just because an item shows DIF does not imply that the associated test is bias. An item could show DIF that is based on distributional differences between groups and, therefore, would be considered benign. Roussos and Stout (1996) distinguish between auxiliary dimensions and nuisance dimensions. It was stated that, in cases where a construct is either essentially unidimensional or intentionally multidimensional, there is a possibly to have secondary (or auxiliary) dimensions that are related to the construct in question and intentionally assessed as part of the construct. On the other hand, nuisance dimensions are those that are unintentionally assessed as a part of the construct. In other

88

words, auxiliary dimensions show relevance to the construct, whereas nuisance dimensions do not. Items that show DIF as a result of auxiliary dimensions are considered to be benign and, contrarily, items that show DIF as a result of nuisance dimensions are considered to be adverse. Test bias only comes into play with the detection of adverse DIF.

When examining DIF in a particular measure, Camilli and Shepard (1994) recommend a three-step approach where you first statistically analyze whether DIF exists in any items of the specified scale. Then you further examine items that show DIF using logical judgment to determine whether there is a valid reason why groups would understand the item differently. Finally, you classify whether the associated DIF can be attributed to bias or whether the DIF is benign. In order to detect DIF in the current study, the Mantel-Haenszel statistic (Mantel & Haenszel, 1959) was calculated and chi-square tests were examined for non-significance. Significant chi-square tests indicate that an item contains DIF, therefore, non-significance in these tests is ideal. DIF is often graded on a scale from A – C, with "A" signifying that the item contains little or no DIF, "B" signifying moderate DIF, and "C" specifying large DIF. Items that showed "C" levels of DIF were flagged as potentially bias items. The groups under examination in the current study were gender, age, and race. Because COE is predicted to be of benefit in applied work settings, it is important from a legal standpoint to ensure that test items do not show potential adverse impact across groups. In regard to age, participants were dichotomized as either over 40 years of age or under 40 years of age. This cutoff was created to be in alignment with the ADEA (Age Discrimination in Employment Act, 1967) for practical and applied purposes. With regard to race, participants were dichotomized as either White or Non-White (in accordance with the protected classes of the Civil Rights Act 1964, 1991). Finally, gender was maintained as either male or female. The focal group in each analysis was represented by the marginalized group (e.g., females, non-Whites, those over 40), and the reference group was the majority group. All items were stratified and matched on observed scores with the binary item effect size being the odds ratio. The results of these analyses are depicted in Tables 20 – 22.

Table 20 – *DIF Analysis Results for Race*

COE Item	Item Wording	χ^2	p-value	DIF Class
1	In general, most people can be trusted.	3.08	0.08	AA
2	It is hard for me to work with people that I do not know.	1.01	0.32	AA
3	People tend to follow through with their word.	1.06	0.30	AA
4	Most people fail to follow through with their word.	1.06	0.30	AA
5	I cannot rely on people, because most people let me down.	0.81	0.37	AA
6	I believe that other people can be counted on to do what they say they will do.	0.04	0.84	AA
7	I find it hard to care about people that I do not know.	0.00	0.96	AA
8	Other people generally have bad intentions.	4.35	0.04	AA
9	Other people generally have good intentions.	4.36	0.04	AA
10	I find it hard to trust other people.	1.10	0.29	AA
11	Collaborating with others is fun to me.	10.73	0.00	BB-
12	Dealing with other people can be quite a pain.	19.38	0.00	CC-
13	Most people are up to no good.	2.86	0.09	AA
14	Other people are incompetent.	2.15	0.14	AA

Table 21 – *DIF Analysis Results for Gender*

COE Item	Item Wording	χ^2	p-value	DIF Class
1	In general, most people can be trusted.	9.88	0.00	CC+
2	It is hard for me to work with people that I do not know.	0.76	0.38	AA
3	People tend to follow through with their word.	3.58	0.06	AA
4	Most people fail to follow through with their word.	0.34	0.56	AA
5	I cannot rely on people, because most people let me down.	0.44	0.51	AA
6	I believe that other people can be counted on to do what they say they will do.	2.25	0.13	AA
7	I find it hard to care about people that I do not know.	13.18	0.00	CC-
8	Other people generally have bad intentions.	0.94	0.33	AA
9	Other people generally have good intentions.	0.09	0.77	AA
10	I find it hard to trust other people.	8.77	0.00	BB+
11	Collaborating with others is fun to me.	2.12	0.14	AA
12	Dealing with other people can be quite a pain.	3.53	0.06	AA
13	Most people are up to no good.	0.10	0.75	AA
14	Other people are incompetent.	7.45	0.01	BB-

Table 22 – *DIF Analysis Results for Age*

COE Item	Item Wording	χ^2	p-value	DIF Class
1	In general, most people can be trusted.	2.33	0.13	AA
2	It is hard for me to work with people that I do not know.	0.25	0.62	AA
3	People tend to follow through with their word.	1.42	0.23	AA
4	Most people fail to follow through with their word.	1.64	0.20	AA
5	I cannot rely on people, because most people let me down.	0.01	0.94	AA
6	I believe that other people can be counted on to do what they say they will do.	0.10	0.76	AA
7	I find it hard to care about people that I do not know.	0.01	0.93	AA
8	Other people generally have bad intentions.	0.17	0.68	AA
9	Other people generally have good intentions.	0.78	0.38	AA
10	I find it hard to trust other people.	0.55	0.46	AA
11	Collaborating with others is fun to me.	2.38	0.12	AA
12	Dealing with other people can be quite a pain.	4.11	0.04	BB+
13	Most people are up to no good.	0.80	0.37	AA
14	Other people are incompetent.	0.20	0.66	AA

In reference to age, none of the items show significant differentiation between the two groups (i.e., no significant DIF was detected). With regard to race, one item showed significant DIF (item 12 – "Dealing with others can be quite a pain"). No other items showed significant differentiation between the two groups. For gender, two items were flagged as potentially bias items (item 01 – "In general, most people can be trusted"; item 07 – "I find it hard to care about people that I don't know"). Overall, only 3 of the items showed potential for bias across these three groups.

Alas, though methods for DIF detection (from a statistical standpoint) are prevalent, there are not many established substantive methods for identifying why certain items show high levels of DIF (Gierl, Bisanz, & Li, 2004). In fact, it was stated that, despite serious deliberation and cogitation, there is often a lack of clarity as to why some seemingly reasonable items show high levels of DIF (Angoff, 1993). An effort was made to use logical judgment to determine why the items flagged in the current study would show such high levels of DIF. The only conclusion that was ascertained on the basis of this analysis was that the item that showed DIF across race could be considered as somewhat colloquial in nature, thereby, causing confusion between races. Other than that, no substantive conclusions could be drawn that would explain why items would show high levels of DIF. This is in line with many other researchers who struggle to find substantive meaning behind the high DIF levels obtained in their study (Angoff, 1993; Camilli & Shepard, 1994; Roussos & Stout, 1996).

In addition to the logical judgment analysis, an effort was made to explore the items from a factor analytic approach to see if there were any distributional differences across groups. This analysis was explored only for gender and race (which were the only demographic categories to show items with significant DIF). A factor analysis with a principal components method of extraction was performed for each group separately (i.e., separately for men and women, and separately for Whites and non-Whites) with a Varimax rotation. The results showed that the difference between these items is likely due to distributional differences along relevant constructs, which is considered benign in terms of DIF.

3.6.2 Moderator Analyses

In addition to exploring whether items showed DIF, it was important to determine whether there was predictive bias across groups along the work outcomes under study. More specifically, Research Questions 4 – 5 asked whether gender and race moderated the relationship between COE and job/team

satisfaction. If demographic moderators were found to be present, then this would suggest that COE predicts important work criteria differently across demographic groups and, thus, could be potentially problematic for use in employment settings. In order to test whether gender or race were moderating variables in the prediction of job/team satisfaction, a moderated multiple regression analysis was performed in SPSS. Moderators were deemed present if the interaction term in the model was significant. First, the continuous independent variable (COE) was centered to reduce multicollinearity between predictors (which is a typical practice when you have a continuous predictor; Holmbeck, 2002). Then the interaction terms between the centered COE value and the potential moderators were calculated. Finally, variables were input into the regression model in a hierarchical fashion to determine the significance of the interaction. Assumptions of homoscedasticity, linearity, and independence were examined prior to analysis.

Starting with the research question surrounding gender (Research Question 4), there were no moderating effects found for either job satisfaction or team satisfaction. With regard to job satisfaction, the change in R^2 was non-significant when the interaction term was entered into the model, the ΔR^2 = .002, $F(1, 332)$ = .762, p = .383. Even though the omnibus F-test showed significance, $F(3, 332)$ = 22.77, $p <$ .001, the interaction between COE and gender was non-significant (B = -.062, se_b = .070, β = -.184, $t(3, 332)$ = -.873, p = .383). For team satisfaction, the same pattern emerged. The change in R^2 was found to be non-significant, ΔR^2 = .003, $F(1, 603)$ = 1.99, p = .159. Even though the omnibus F-test showed significance, $F(3, 603)$ = 55.94, $p <$ .001, the interaction between COE and gender was non-significant (B = -.058, se_b = .041, β = -.217, $t(3, 603)$ = -1.410, p = .159). So, even though two items showed DIF across gender, there was found to be no predictive bias in terms of gender on the relationship between COE and job/team satisfaction.

The same analysis was performed with regard to race (thereby, addressing Research Question 5). Like gender, race was not found to moderate either job or team satisfaction. With regard to job satisfaction, the change in R^2 was not significant when the interaction term was added to the model, ΔR^2 = .002, $F(1, 330)$ = .976, p = .324. Even though the omnibus F-test revealed significance $F(3, 330)$ = 22.38, $p <$ .001, the interaction term was not significant in the model, B = -.061, se_b = .062, β = -.155, $t(3, 330)$ = -.988, p = .324. Shifting to the examination of team satisfaction, again, the change in R^2 was not

significant when the interaction term was added to the model, ΔR^2 = .000, F(1, 601) = .032, p = .858. The omnibus F-test revealed significance, F(3, 330) = 38.71, p < .001. However, the interaction term was not significant in the model, B = .006, se_b = .036, β = .021, t(3, 601) = .179, p = .858. These results signify that, even though one item was found to show DIF in COE, there were no differences found in the predictive validity of COE with regard to race.

Although DIF was not present in regard to age nor were there any moderating effects found for age in previous studies, age was still examined as a potential moderator in the current study as a supplemental analysis. The results showed that there were no moderating effects of age for job satisfaction. The change in R^2 was not significant when the interaction term was added to the model, ΔR^2 = .001, F(1, 319) = .405, p = .525. The omnibus F-test revealed significance, F(3, 319) = 25.48, p < .001. However, the interaction term was not significant in the model, B = .053, se_b = .083, β = .110, t(3, 319) = .636, p = .525. Surprisingly, however, age was found to be a significant moderator in the prediction of COE and team satisfaction, ΔR^2 = .007, F(1, 577) = 5.05, p < .05 (B = .132, se_b = .059, β = .318, t(3, 577) = 2.247, p < .05). These findings suggest that the prediction of COE and team satisfaction is stronger for those 40 years of age and older than it is for those under 40 years old.

Finally, the last moderation analysis performed was in regard to the sample population. The research question asked whether or not the sample population (students vs. working professionals) had any influence on the relationship between COE and job/team satisfaction (Research Question 6). Beginning with team satisfaction, a significant interactive effect was found between COE and participant status in relation to team satisfaction, ΔR^2 = .043, F(1, 607) = 35.32, p < .001 (B = .237, se_b = .040, β = .639, t(3, 577) = 5.94, p < .001). These results suggest that the relationship between COE and team satisfaction was stronger for working professionals than it was for students. The same pattern emerged with regard to job satisfaction. There was a significant moderating effect of participant status on the relationship between COE and job satisfaction, ΔR^2 = .028, F(1, 335) = 11.97, p < .01 (B = .205, se_b = .059, β = .513, t(3, 335) = 3.460, p < .01). Again, this finding indicates that the relationship between COE and job satisfaction was stronger for working professionals than it was for students. For a review of all hypotheses and indication of whether or not they were supported, please see Table 23.

95

Table 23 – *Results of Hypotheses*

Hypothesis Number	Hypothesis Wording	Supported Yes or No
1	The COE measure will positively relate with trust, agreeableness, collectivism, and other orientation.	Yes
2a	The measure of COE will positively relate to the Team Player Inventory.	Yes
2b	The measure of COE will show incremental validity beyond the Team Player Inventory in the prediction of work outcomes	Yes
3	COE will have higher relations with trust, agreeableness, other orientation, and collectivism than it will with conscientiousness or openness.	Yes
4	COE will have higher relations with trust, agreeableness, other orientation, and collectivism than it will with core self-evaluations.	No
5	Core other-evaluations will show a positive relationship with job satisfaction.	Yes
6	Core other-evaluations will show a positive relationship with team satisfaction.	Yes
7a	Higher team COE averages will lead to higher team outcomes.	Partial
7b	After controlling for team COE averages, teams with less variability in COE scores will have higher team outcomes.	No
8	Core other-evaluations will show a relationship with the interpersonal facets of job satisfaction (i.e., satisfaction with supervisor and satisfaction with co-workers) over and beyond that of core self-evaluations.	Yes
9	Core other-evaluations will show a relationship with team satisfaction over and beyond that of core self-evaluations.	Yes

CHAPTER 4

DISCUSSION

The construct of core self-evaluations, while relatively new to the I/O field, has made some

significant contributions since its development. It continues to thrive in the literature and provides valuable

insights in relation to some important job outcomes (Judge, 2009). When originally proposed, the

construct of core self-evaluations was a part of a trichotomy that also included core other-evaluations and

core evaluations about the world. Until now, no research has been done to empirically explore either of

the other two constructs. The focus of this paper was to address that gap in the literature and to

conceptualize, create, and analyze a measure of core other-evaluations.

It was predicted that this measure could be useful in applied settings by determining an

individual's disposition as it relates to their interpersonal/team effectiveness. The results found in this

study have provided compelling evidence toward this claim. The COE scale is not only effective at

predicting job satisfaction, team satisfaction, and team effectiveness; but it has also been proven to

provide predictive ability above and beyond core self-evaluations. In addition, the COE scale adds

incremental value in the prediction of team satisfaction and team effectiveness over and beyond another

like measure (which also claims to be a dispositional measure of teamwork). Further, empirical evidence

suggests that teams comprised of individuals high in COE are more likely to produce positive team results

than teams comprised of members low in COE. Moreover, the COE scale shows minimal, benign

differential item functioning across demographic groups and shows no predictive bias in terms of gender

or race groups. Based on these significant findings, it is encouraged that COE be further explored in

future studies and that it be considered for use in employment settings seeking to better understand an

individual's likelihood to succeed in a team environment.

Each of the findings related to the development and validation of the COE scale will be further

discussed in more detail. I will begin with a review of the item development and generation process and

findings (including the pilot study results), discuss the examination of the factor structure, summarize the

97

IRT model fit and findings, review the construct validity results, describe the findings for predictive and incremental validity, and close with a review of findings on group differences. In addition to each of these sections, limitations and future directions will be discussed. I begin with a review of the item development process.

<h2 style="text-align:center">4.1 Item Development</h2>

As proclaimed throughout the entire paper, this research was grounded on the work of Judge et al. (1997). Many of their theoretical assumptions and inclusion criteria were used not only to create the COE items, but also to support the conceptualization and theory behind the construct. It was important that COE be aligned to CSE and, by doing, be an accurate representation of a core evaluation. With this objective, assiduous steps were taken in order ensure that the conceptualization and development of COE mirrored the work of Judge and colleagues (e.g., Judge et al., 1997; Judge et al., 2003). The same inclusion criteria employed by Judge et al. (i.e., fundamentality, evaluation-focus, and scope; 1997) were used in the current study to determine inclusion as a manifest trait of COE. After perusing the literature for possible trait inclusions, four traits emerged as manifest indicators of COE (namely trust, agreeableness, collectivism, and other orientation). Because COE was thought to be a fundamental appraisal about how individuals viewed and related to other people, these traits seemed indicative of how COE would manifest itself, and they each met the aforementioned inclusion criteria. More specifically, those high in COE were expected to trust the intentions and motivations of others, value being around other people, put others needs ahead of their own, and show concern for others' well-being. On the other hand, someone low in COE values independence, dislikes collaboration, seems cynical and distrustful of others, and often considers personal gain before group gain. It has been said previously that the assertion of this paper is not to imply that the four variables said to underlie COE are the only trait level manifestations of COE. There could potentially be other traits that fall within the nomological network of COE. During the thorough review process, the four traits described earlier were found to be the best trait-level depictions of how COE would manifest.

After conceptualizing the construct, the next step of the process involved generating items for which to measure this broad construct. When developing items, a strong effort was made to create items that transcended across multiple manifest traits rather than just focusing on one trait in isolation. This was

done in an effort to capture the interconnectedness of these variables, thereby, capturing the essence of COE. Moreover, efforts were made to ensure coverage across the domain by enlisting the aid of subject matter experts and analyzing the results of their ratings. Results of the SME analysis warranted the need to eliminate items that did not meet the inclusion criteria or other specified criteria. Once items were generated, they were piloted to check to see if the construct of COE was worthy of further review. In two separate pilot studies, the results revealed that there was considerable overlap between the variables of interest (i.e., trust, agreeableness, collectivism, and other orientation), thus indicating the potential for a higher-order, latent construct. The proclamation of the paper and subsequent conclusion drawn was that COE was responsible for the interrelationships between these variables and describes how individuals fundamentally view other people.

4.2 Exploration of Factor Structure

With regard to CSE, the question of factor structure has been a contentious topic. The primary researchers argue for a superordinate factor (e.g., Bono & Judge, 2003, Erez & Judge, 2001; Heller, Judge, & Watson, 2002; Judge & Bono, 2001; Judge, Bono, & Locke, 2000; Judge et al., 1998; Judge, Erez, Bono, & Thoresen, 2002, 2003), whereas others have claimed that there are better classifications for CSE (e.g., collective set or aggregate score; Dormann, Fay, Zapf, & Frese, 2006; Johnson, Rosen, & Levy, 2007). The stance in the current paper was to remain in alignment with the theory and conceptualization of CSE, thus, classifying the construct of COE as broad and superordinate. Because COE is said to be broad and fundamental and reside at a deeper space, then it is expected to be an unidimensional latent construct. As stated previously; trust, agreeableness, collectivism, and other orientation are not subfactors of COE, but rather different ways in which COE manifests itself.

Because IRT techniques were employed when analyzing the data, and given the fact that one of the major assumptions of IRT is unidimensionality, the factor structure of COE was examined prior to analyzing any of the data. To my knowledge, few methods exist to test the unidimensionality of polytomous items (whereas, tests for dichotomously scored items abound). In fact, research using IRT in its multidimensional form is relatively new and constantly evolving. Until relatively recently, most of the focus of IRT has been with regard to binary test items. In addition, the strict assumption of unidimensionality with IRT models has been called into question by many psychometricians and

researchers (e.g., Bryant & Wooten, 2006; Stout, 1987). Alternatives to this assumption (e.g., essential unidimensionality) have been proposed and used in analyses (Stout, 1987). Essential unidimensionality is said to exist where there is a clear presence of a dominant factor, even though there may also be minor factors present (Stout, 1987).

In order to determine whether COE showed evidence of being essentially unidimensional, factor analytic techniques were used to explore the dimensionality of COE. The results of this analysis showed support for the claim that COE is essentially unidimensional. Not only was the majority of variance in the model explained by the first factor (38%), but the scree plot also indicated a significant drop-off after the first factor. In conclusion, the exploration of the factor structure yielded results that support the fact that COE is representative of an (essentially) unidimensional latent construct.

4.3 IRT Model Fit

Once clearance was given on the basis of meeting the IRT assumptions, the next step of the process was to examine how well each of the items fit the model. To test the functionality of the items, a Graded Response Model (GRM; Samejima, 1969) analysis was conducted. Based on the chi-square test of fit, all items were said to converge onto the IRT model. However, some items showed better fit than others. An examination of discrimination (a) parameter indicated that item 5 ("I cannot rely on people, because most people let me down.") had the highest slope value (2.243), whereas item 11 ("Collaborating with others is fun to me.") had the lowest slope value (0.774). In fact, item 11 showed the least connection to all other items in the scale. However, all items showed acceptable fit according the IRT model (GRM). Also, when dealing with slope parameters, it is typical to find values between .5 and 2.5 (Reeves & Fayers, 2005). The item threshold (b parameter) signifies the difficulty or severity level of a particular item response. In the current study's model, slope parameters were held constant across item responses, whereas item thresholds differed across response options. Examining the item thresholds for all items, the only item that showed an unusual pattern was item 11 ("Collaborating with others is fun to me."). The difficulty parameters for option 1 and option 2 (-6.02 and -4.9 respectively) signify a low likelihood that individuals would select that response. As with the discrimination parameter, this item had the highest negative threshold values of all the items in the scale. Information functions, which display the range over ability levels where the item has the optimal ability to distinguish between individuals (Reeves & Fayers,

100

2005), were examined to determine the reliability of that given item. Items with low information functions may be measuring something different than the other items. All items showed acceptable information functions; however, item 11 showed the weakest functioning of the fourteen. Given these findings, item 11 may be a candidate for elimination in future research due to its barely adequate psychometric performance. However, it was retained in the current study because, even though it was the poorest performing item in relation to the others, it still fit the model and the item parameters were within acceptable ranges. In addition, COE is claimed to cover a broad spectrum of core beliefs about others, so efforts were made to retain items that transcended across multiple manifest traits (e.g., trust, agreeableness, collectivism, and other orientation) rather than focusing on just one. Because many of the items focus on trust and agreeableness, efforts were made to retain items that measured other facets of COE (i.e., Item 11).

4.4 Construct Validity

When exploring the construct validity of COE, efforts were made to begin establishing a nomological network for COE by showing connections with COE and like constructs and distinguishing it from dissimilar constructs. With regard to convergent validity, the focus of analysis here was to determine whether COE was related in a significant way to the four traits said to underlie it (i.e., trust, agreeableness, collectivism, and other orientation). Beginning with the aggregate value between these four scales, COE showed a strong, significant relationship with the summation of these variables (r = .63). In fact, as Judge and Bono (1999) note, this value is comparable to the relationship obtained between different measures of neuroticism (Digman, 1990). Given this finding, there is no doubt that COE is highly related to these traits.

The next effort involved looking at COE's relationship to each of the variables individually. Results show that COE is very highly correlated with both Trust and Agreeableness, but less so with Collectivism and Other Orientation. All relationships were significantly significant, but the magnitude of the trust (r_c = .81) and agreeableness (r_c = .62) relationships were significantly higher than both collectivism (r_c = .36) and other orientation (r_c = .29). These findings indicate that, as currently measured, COE relates to trust and agreeableness to a greater extent than it does to collectivism and other orientation. As stated previously, an examination of the item content revealed that most items focus more on trust and

101

agreeableness related factors more than anything else. Many of the items that corresponded more closely to collectivism and other orientation were eliminated in previous phases of the process due to poor psychometric performance. However, concerted efforts were made during the item analysis process to retain items that did not solely focus on trust and agreeableness. It is understood that trust has the greatest degree of influence on COE due to the fact that it is the most fundamental of the traits (Judge et al., 1997), however, it is claimed that COE has something to offer beyond trust alone.

In addition to these aforementioned scales, the Team Player Inventory (TPI; Kline, 1999) was also used a measure of convergent validity. The results show support for the notion that the COE scale and the TPI scale are strongly related. An examination of the incremental validity that COE provides above and beyond the TPI indicated that COE provides a considerable amount of unique variance above and beyond that of the TPI. In addition, though the TPI is very useful in workplace settings, it was not designed to generalize beyond that. The COE scale, however, is not limited only to work populations but can be used to determine an individual's likelihood to succeed in any type of team/interpersonal environment (e.g., school, sport's team, etc.).

In regard to discriminant validity, several constructs were proposed to differ from COE. Beginning with the Big Five personality dimensions, both conscientiousness and openness were predicted to show an empirical distinction from COE. For conscientiousness, the relationship with COE was moderate ($r =$ .35). Due to the altruistic aspects inherent in both conscientiousness and COE, they were expected to be somewhat related. The findings were in line with this theory. However, conscientiousness was related to COE to a lesser degree than were the manifest variables (i.e., trust, agreeableness, collectivism, and other orientation), thus, adding support for the hypothesis. For openness, the same pattern emerged in the data. Openness was significantly less related to COE than were trust, agreeableness, collectivism, and other orientation.

In addition, CSE was examined for discriminant validity purposes. Like with conscientiousness, there was expected to be a moderate relationship between CSE and COE due to their shared development and theoretical framework. The results indicated that CSE and COE are related to each other at a higher magnitude than expected ($r =$ .57), and this value was not significantly different from the correlation obtained between COE and the manifest traits. This signifies that the two constructs' shared

102

development methods and theoretical backing influenced the relationship between COE and CSE in some way. Judge et al. (1997) mentioned that CSE is the most fundamental of all the core evaluations and, thus, influences the others to a significant degree. From a logical standpoint, it makes sense that how you feel about yourself could affect how you feel about other people. If someone lacks confidence in themselves, devalues their self-worth and effectiveness, and has difficulty maintaining a sense of self-composure; it stands to reason that they will struggle connecting with others, appreciating the value of being around other people, and trusting the words and intentions of others.

In conclusion, COE has demonstrated significant connections to like constructs and dissimilarity from distinct constructs. These results are critical to the development of the nomological network for COE. The only finding contrary to predictions was regarding the relationship between COE and CSE. Despite being highly related, there was still enough differentiation in the relationship to conclude that the two constructs are not redundant.

<u>4.5 Criterion-Related Validity</u>

Beyond the development and construct validity of COE, it was imperative to establish criterion-related validity with COE in order to support the utility of the scale. Several important work criteria were explored in this analysis (e.g., job satisfaction, team satisfaction, and team effectiveness). COE was found to predict all three of the work outcomes to a significant degree. With job satisfaction, those higher in COE were more likely to be satisfied with their jobs than those low in COE. This goes in line with the theory and predictions backing this hypothesis. COE was also found to add unique variance in the prediction of job satisfaction above and beyond CSE. This finding is worthy of attention due to the fact that CSE has been promoted as the best predictor of job satisfaction by some researchers (c.f., Judge & Bono, 2001a). Given the results of this study, it seems that COE may be a strong contender in the prediction of job satisfaction. Furthermore, as hypothesized, COE showed a stronger prediction (although not statistically different) of the interpersonal aspects of job satisfaction (e.g., satisfaction with coworkers, satisfaction with supervisor) as opposed to the non-interpersonal aspects of job satisfaction (e.g., task satisfaction, satisfaction with pay, and satisfaction with promotion opportunities). Due to the inherent interpersonal nature of the COE construct, this finding is in line with theory and intuitive thought.

COE's ability to predict team satisfaction and team effectiveness were also explored. Because COE is being advocated as a beneficial selection tool for an individual's propensity to succeed in a work team, then these outcomes are significant dependent variables of interest. The thought was that because the COE scale measures an individual's proclivity to get along with others and trust other people, then those higher in COE would also appreciate working in a team more. Thus, they would show higher satisfaction with teams. Furthermore, those who feel more comfortable working in team environments and relating to other people would be more likely to contribute to the efficacy of the team, thereby, enhancing team effectiveness. Both of these lines of thought held up to the statistical test. COE was a strong predictor of both team satisfaction and team effectiveness. In addition, examining COE at the team level, it was found that teams comprised of individuals with higher levels of COE (with higher team COE averages) were more effective than teams with low team COE averages.

However, the only criterion-related result that was contrary to prediction was the hypothesis dealing with variability of COE levels within teams. It was predicted that teams with lower levels of COE variability would show higher levels of team outcomes. Even though the result was not statistically significant, the trend went in the opposite direction of the one predicted. The results implied that teams comprised of individuals with varied levels of COE showed higher team outcomes than those with less variability. Again, the sample size was not large enough to draw meaningful conclusions. It would be interesting to explore this finding further in future research. Position in the team (i.e., team member or leader) was not taken into account with this analysis. A potential explanation could be that team effectiveness may be affected by the team leader's COE level.

In summary, the predictive effects of COE are pretty substantial. Not only was COE able to predict job satisfaction, team satisfaction, and team effectiveness to a significant degree, but it was also able to show incremental validity beyond both CSE and the TPI. These findings make COE a worthy candidate for inclusion in a selection system when one wants to predict job satisfaction, team satisfaction, and/or team effectiveness. In addition, even though it was not reported in this study, COE was found to be a strong, significant predictor of both employee engagement and affective commitment. This indicates that individuals higher in COE are likely to be more productive in their work and remain more committed to the organization than those low in COE. Due to the fact that employee turnover is a difficult problem

that many organizations face (during smooth economic climates); this finding is highly useful for practical purposes.

<div align="center">4.6 Group Differences</div>

4.6.1 Differential Item Functioning

In order to determine whether measurement invariance existed with the COE scale, differential item functioning was performed on all fourteen items. It was important to examine measurement invariance with the COE scale before testing subsequent analyses regarding group differences in order to avoid spurious results (Vandenberg & Lance, 2000). With this objective, DIF was performed to determine whether individuals belonging to separate demographic groups were interpreting the scale in the same manner. DIF was run on the demographic variables of interest for employment settings (i.e., age, race, and gender). With regard to age, none of the items showed DIF between the two age groups (over 40 and under 40). This result signified that individuals of differing age groups interpreted the items in the same fashion. With regard to race, one item was found to show a significant amount of DIF (item 12 – "Dealing with other people can be quite a pain"). Because race was dichotomized as White or non-White, this finding suggests that there is a difference in how Whites interpret this item as compared to non-Whites. When trying to determine why this difference would be present, the logical judgment method was used. It was suggested that DIF could potentially exist due to the colloquial nature of the item. With regard to gender, two items were found to show significant DIF (Item 1 – "In general, most people can be trusted" and Item 7 – "I find it hard to care about people that I do not know"). When applying logical judgment, it was difficult to determine why these items may show DIF between gender groups. DIF analysis can be limited in a sense due to the fact that, even though there are well established methods for detecting the presence and magnitude of DIF, there is a dearth of substantive methods to further understand why an item is showing DIF (Gierl, Bisanz, & Li, 2004). In order to determine whether the DIF found in the current study was adverse in nature and, thus, problematic for use in the employment arena, a separate factor analysis was performed for each of the demographic groups (race and gender). The results of the factor analysis suggested that the factor loadings of some of the items in question may have more to do with why they are showing DIF than the item itself. In other words, the factor analysis shows that Item 7 loads

differently for women than it does for men. Because this type of DIF is more so due to distributional differences rather than actual item bias, the DIF is said to be benign (Shealy & Stout, 1993).

4.6.2 Predictive Bias

The second exploration of group differences pertained to whether predictive bias occurred. Analyses were performed to determine whether gender or race significantly moderated the relationship between COE and job/team satisfaction. In all cases, neither demographic characteristic was found to moderate the relationship between COE and job or team satisfaction. These findings suggest that there is a similar level of predictive validity for COE with job and team satisfaction across demographic groups. These results further support the utilization of the COE scale in employment settings.

A supplemental analysis was performed to explore whether age was a significant moderator for the relationship between COE and job/team satisfaction. Because there has been no past concern with regard to age being a moderating variable nor was there found to be in differential item functioning for age, this analysis was not included as a question of research. Nonetheless, the analysis was performed to capture all the demographic variables of employment interest to determine if predictive bias existed. Surprisingly, age was the only demographic variable that resulted to be a significant moderator. It was in the prediction of team satisfaction that age was found to be a significant moderating variable. The results suggested that the relationship between COE and team satisfaction was stronger for those 40 or over than it was for those under 40. Because this finding is in support of the protected class group (i.e., those over 40), then this finding would not be adverse for the work environment. However, an effort was still made to better understand this result. Because a student population was used as a part of this analysis, and because the overwhelming majority of students in the sample were below 20 years of age, this result may have been likely due to differences in the sample rather than actually age differences. Team satisfaction was measured regardless of whether the participant was working or not. If they were working, then participants were encouraged to refer to their work team when responding to the team satisfaction items; whereas, if they were not working, individuals were asked to describe a recent team to which they have been a part of and respond according to their recollection of that team. This difference in team type or sample type may have influenced the results more so than age specifically. In fact, when students were removed from the analysis, and only working professionals were considered, age was not found to

be a significant moderator in the relationship between COE and team satisfaction (ΔR^2 = .000, F(1, 93) = .105, p = .746). In summary, age was not found to be a significant moderator for job satisfaction nor was it found to be a significant moderator for team satisfaction in the work context.

The age discussion makes a perfect segue into the analysis regarding sample population. Type of sample (students vs. working professionals) was examined to determine whether this characteristic influenced the relationship between COE and job/team satisfaction. In both cases (job and team satisfaction), sample population was found to be a significant moderator. The results suggested that the relationship between COE and job/team satisfaction was stronger for working professionals than it was for students. For team satisfaction, this falls in line with the previous discussion around age. The results show that the distinction between predictive power for COE and team satisfaction has less to do with age and more to do with whether the participant was a student or a working professional. Because working students were encouraged (and not mandated) to refer to their work team when responding to team satisfaction items, they may not have complied with these instructions. With regard to job satisfaction, the difference between students and working professionals may have been due to the fact that students are often not fully engaged in their jobs at this level. They typically tend to view their jobs as a means to an end rather than as a career outlet. Additionally, many students who work and attend school do so on a part-time basis, so their satisfaction with work may be limited as a result of their part-time status. For those who work full-time and attend school, their job satisfaction could be attenuated due to the fact that they are trying to balance two time-consuming aspects of their life and potentially struggling to do so. Because it has been advocated throughout the entire paper for COE to be used within an employment context, these results actually add support for doing this. The results suggest that the predictive power of COE is significant in both contexts but statistically stronger for work contexts.

<div align="center">4.7 Limitations and Future Directions</div>

As with any research study, there are always associated limitations. This research is no exception. The main areas of limitations for discussion are sample size, sample composition, and statistical methods. I will discuss each in turn beginning with sample size.

4.7.1 Sample Size

The sample size limitation was most pronounced in the analyses for work teams and their associated COE levels. Because it was difficult to gain entry into organizations for data collection purposes, a smaller number of work teams participated in the survey. Even though I had substantial participation from working professionals, the number of teams included in the analysis was only *N* = 32. Also, after eliminating teams that only had representation from one member, that number decreased to *N* = 26. Furthermore, it was difficult to obtain full participation from each team member, thus, resulting in representation from incomplete teams. As mentioned previously, many organizations prohibit surveys from circulating through their departments and denied me entry into their organizations. With such a small sample of teams, I was limited in my ability to analyze trends and occurrences from a team level. As a result, the majority of the analyses performed in this study were centered on individual level outcomes as opposed to team level outcomes. It is recommended that future studies examine the predictive power and effectiveness of COE from the team level. Because COE is so predictive and powerful at the individual level, it is likely to also be effective and predictive at the team level. Also, even though not statistically significant, many of the team-level analyses ran in the current study show trends in the hypothesized directions. With the appropriate amount of data, it seems that much can be explained from COE at the team level.

In addition, sample size is a potential limitation when it comes to the IRT analyses performed. Because there are no clear or definitive answers regarding sample size requirements (Embretson & Reise, 2000; Orlando, 2005), it was difficult to determine whether the sample size used is sufficient for analyses. In fact, some researchers have claimed that a sample size of 500 is appropriate for accurate parameter estimates (Tsutakawa & Johnson, 1990), others have suggested that as few as 250 can be sufficient (Reise & Yu, 1990), and others have contended that over 500 participants are necessary for some IRT models (Hambleton & Jones, 1993). The sample size of the current dataset was 588. Embretson and Reise (2000) claim that Graded Response Model parameters can be estimated with around 500 participants. Despite the confusion regarding the appropriate sample size requirements, the sample size used in this study is likely sufficient for accurate estimation of parameters due to Embretson and Reise's guidelines (2000).

4.7.2 Sample Composition

The limitation regarding sample composition stems from the previous discussion around differences between the student sample and the working professionals sample. Because students were not mandatorily required to specify a work team when reporting their team satisfaction, this may have influenced some of the results found in the study (particularly regarding the significant moderation effect found for age). However, having both samples of participants bolstered the argument that COE would be more beneficial for use in an employment setting. Still and all, it would be interesting to continue to study COE in the school context to further understand the impact of COE with non-work teams. Because one of the advantages of COE over the Team Player Inventory was that it can be useful in non-work settings, exploring areas outside of the work arena is permissible. One area in particular that may benefit from the use of COE is graduate programs that require heavy team-based projects and assignments (e.g., MBA programs). COE was found to be predictive in student samples, showing that students with higher levels of COE also had higher levels of team satisfaction. One could also study the impact that COE has on school outcomes (e.g., success in grad school, completion of an MBA program, etc.) to determine its effectiveness in this domain.

4.7.3 Statistical Methods

The final major limitation of the study dealt with some of the statistical methods used for analysis. Some of the analyses run in the current study are relatively new in the world of psychoanalytics and, therefore, have not been fully vetted with regard to specific procedures and interpretations (e.g., DIF interpretation). The sample size issue was already discussed. However, the ambiguity here stems more from the fact that a conclusive number is hard to provide (due to the fact that it depends on the model used, the number of response options, the study purpose, etc.), rather than from the fact that there is no clear guidance on how many participants one would need. In addition, there was an issue with finding an appropriate method to test for unidimensionality with polytomously-scored items. Clear methods exist for dichotomously-scored items, but there is not a great deal of guidance when it comes to polytomous items. Nonetheless, exploratory factor analytic techniques were found to be a sufficient test of the unidimensionality assumption.

The other area of ambiguity with regard to statistical analyses pertains to DIF. The Mantel-Haenszel (M-H) statistic was used in the current study as the method of choice for DIF detection. Even though it has been described as one of the most popular DIF-detection methods (Embretson & Reise, 2000), it is not without its limitations. The M-H procedure has been criticized for its inability to detect DIF that is not uniform across the range of ability scores (Swaminathan & Rogers, 1990). Pertaining to the detection of uniform DIF, the M-H procedure does so with great accuracy. Uniform DIF occurs when two item characteristic curves are different, but function in a parallel fashion (i.e., the item characteristic curves may have similar discrimination parameters but different difficulty parameters; Swaminathan & Rogers, 1990). On the other hand, non-uniform DIF is present when an interaction is found between ability levels and group membership. Non-uniform DIF is typically present when item characteristic curves have similar difficulty parameters and different discrimination parameters (Swaminathan & Rogers, 1990). Even though the lack of power to detect non-uniform DIF is a limitation of the M-H procedure, it has still been one of the most widely using DIF detection methods (Embretson & Reise, 2000). Also, the M-H technique does not rely solely on the chi-square statistic, which has been known to be overly sensitive when dealing with larger sample sizes (Wheaton, Muthen, Alwin, & Summers, 1977), thus, making it more ideal as compared to other methods. Furthermore, the M-H statistic allows you to not only determine whether DIF exists, but it also allows inferences to the degree of DIF present.

The other limitation as it pertains to statistical methods was regarding the lack of substantive procedures for truly understanding why an item shows significant DIF. Logical judgment and factor analytic techniques were used in the current study to determine the source and severity of the items that showed significant DIF, however, there was not a clear method used for understanding why DIF was present. One proposed method for bridging the gap between DIF detection and clarity of why DIF exists is the multidimensionality-based DIF analysis paradigm (Roussos & Stout, 1996). However, this method was not used in the current study. Future studies may benefit from exploring the benefit of using the multidimensionality-based DIF analysis paradigm with regard to COE.

4.8 Implications

4.8.1 Theoretical Implications

Until now, research on core evaluations has been isolated only to core self-evaluations. To my knowledge, core other-evaluations or core evaluations about the world have not yet been explored in the extant literature. The current research has addressed this gap in the literature. Judge et al. (1997) made reference to the amalgamation of two disparate lines of research in two different psychological disciplines around core beliefs. A clinical psychologist (Packer, 1985) and a social psychologist (Rockeach, 1972) both explored the notion of a trichotomy of core beliefs. Each stated that this model of core beliefs included core evaluations of the self, core evaluations of others, and core evaluations of the world. The core beliefs were said to be fundamental and, thus, would influence other subsequent beliefs. Since Judge et al.'s (1997) original paper on core evaluations, a significant amount of research has been devoted to studying core self-evaluations. Compelling results have been obtained through research on core self-evaluations. The construct of core other-evaluations has a great deal to add to the field as well in terms of theoretical contributions.

Based on the work of the current study, COE has been proven to be a valid construct with strong predictive ability. The conceptualization of the construct is similar to the theoretical propositions for core self-evaluations. Assertions made in regard to CSE were also contended and replicated in the current study. For instance, CSE is claimed to be a higher-order, latent construct. The same is true for COE. Even though the measurement of COE (i.e., the COE scale) is not perfectly unidimensional, the latent construct is claimed to be unidimensional in nature. In addition, COE has demonstrated evidence of a nomological network containing constructs that are similar in nature (e.g., trust, agreeableness, and the Team Player Inventory) and disparate constructs (e.g.,conscientiousness and openness). COE also shows a strong connection with CSE, likely due to the similarities in theoretical development.

4.8.2 Practical Implications

In addition to continuing the core evaluations line of research, COE has also provided some compelling results for applied psychology. Organizations can stand to benefit by using the COE scale in a hiring, developing, or team building capacity. Not only has this study proven that COE is a strong predictor of job satisfaction, team satisfaction, and team effectiveness; it has also shown practical utility

above and beyond other predictive measures. Furthermore, it has demonstrated initial support for scale fairness with regard to demographic groups (namely age, race, and gender).

Because teams are quite customary in the organizational arena, many organizations may find benefit in utilizing the COE scale as a potential screening measure. Organizations could know before hiring someone whether or not they have what it takes to succeed in a team environment. Because some individuals are naturally inclined to thrive in a team setting, whereas others are not, the COE scale provides a measurable indicator of a person's likelihood to excel in team or group functions. In addition to the team application of COE, it is also believed to be useful in predicting an individual's likelihood to succeed in a customer service-oriented position. Because those high in COE are better able to get along with others, show concern for others, and be helpful toward others than those low in COE; it is likely that COE can add value in determining a prospective employee's propensity to display a high level of customer service. In fact, any job that requires significant people skills is likely to benefit from using the COE scale in the selection process. Future research should further explore the empirical usefulness of the COE scale within customer service positions.

<u>4.9 Overall Conclusion</u>

To briefly summarize the findings, COE is undoubtedly a construct worthy of further attention and research. It has been proven to be a valid construct, a strong predictor, and a substantial contributor to the field. It has filled several gaps in the literature; specifically, empirically exploring a never-before-examined core evaluation trait, using IRT methodology in the development of the core evaluation trait, and providing a dispositional measure of an individual's likelihood to succeed in a team environment. To my knowledge, little or no research has been done in any of these areas. With regard to workplace teams, the plethora of research surrounds defining types of teams and examining team effectiveness. However, little attention has been paid to the characteristics that individuals bring to the team that predisposes them to be an effective team member. COE serves as a valuable outlet for that purpose. The study of COE also leads credence to the benefit of the core evaluation constructs. The construct of core self-evaluations has well-established predictive findings and, due to the results of this study, the construct of core other-evaluations has significant benefits as well. Furthermore, COE shows potential to contribute

112

to other disciplines of psychology. Given the degree of significant findings included in this research study, the field of I/O psychology would be remiss to ignore the construct of core other-evaluations.

APPENDIX A

CORE OTHER-EVALUATION PILOT ITEMS

1. I instantly trust people the first time I meet them.
2. In general, most people can be trusted.
3. You can't be too careful in dealing with other people (R).
4. It is hard for me to work with people that I do not know (R).
5. People tend to follow through with their word.
6. In general, people keep the promises they make.
7. Most people fail to follow through with their word (R).
8. Teams can always accomplish more than individuals.
9. The key to getting the best results is by working in a team.
10. I can get more accomplished alone than in a group (R).
11. I prefer working in a group rather than working independently.
12. People tend to manipulate others in order to get what they want (R).
13. I cannot rely on people, because most people let me down (R).
14. I believe that other people can be counted on to do what they say they will do.
15. It hurts me to see other people hurting, even if I do not know them.
16. I find it hard to care about people that I do not know (R).
17. I find it rewarding to work with other people toward a common goal.
18. I find it more rewarding to win an individual prize rather than a group prize (R).
19. I've been told that I care too much about the well-being of others.
20. I find it difficult to pass up someone in need without helping them.
21. I'd rather take care of my personal goals than to devote my time to helping others (R).
22. I consider others when making decisions before I consider myself.
23. Other people generally have bad intentions (R).
24. Other people generally have good intentions.
25. I believe people when they say they will do something.
26. I find it hard to trust other people. (R)
27. I consider myself to be a caring and nurturing person.
28. Collaborating with others is fun to me.
29. Dealing with other people can be quite a pain (R).
30. I wouldn't help someone unless I knew them personally (R).
31. Most people are up to no good (R).
32. Everyone has kindness in their hearts, even if they don't use it.
33. I am constantly considering the impact that my actions have on other people.
34. People seldom pull their weight in group activities (R).
35. I feel uncomfortable working with other people. (R)
36. It bothers me a great deal to see other people treated unfairly.
37. I have been told that I am rude to others at times (R).
38. Most people are highly capable of doing great things.
39. Most people are incompetent (R).
40. Doing things for others gives me more satisfaction than doing things for myself.

Note: *R* = reverse-coded

APPENDIX B

RETAINED CORE OTHER-EVALUATION ITEMS

COE Retained Items

1. (Old item #2) In general, most people can be trusted.
2. (Old item #4) It is hard for me to work with people that I do not know (R).
3. (Old item #5) People tend to follow through with their word.
4. (Old item #7) Most people fail to follow through with their word (R).
5. (Old item #13) I cannot rely on people, because most people let me down (R).
6. (Old item #14) I believe that other people can be counted on to do what they say they will do.
7. (Old item #16) I find it hard to care about people that I do not know (R).
8. *(Old item #17) I find it more rewarding to work with other people toward a common goal than to work alone.*
9. (Old item #23) Other people generally have bad intentions (R).
10. (Old item #24) Other people generally have good intentions.
11. (Old item #26) I find it hard to trust other people. (R)
12. *(Old item #27) I consider myself to be an extremely caring and nurturing person.*
13. (Old item #28) Collaborating with others is fun to me.
14. (Old item #29) Dealing with other people can be quite a pain (R).
15. (Old item #31) Most people are up to no good (R).
16. (Old item #32) Everyone has kindness in their hearts, even if they don't use it.
17. (Old item #34) People seldom pull their weight in group activities (R).
18. *(Old item #36) Nothing bothers me more than seeing people treated unfairly.*
19. *(Old item #39) Other people are incompetent (R).*
20. *(Old item #40) Doing things for others makes me happier than doing things for myself.*

Note: *R* = reverse-coded ; Items in italics were modified

APPENDIX C

FINAL CORE OTHER-EVALUATION ITEMS

	Final COE Items

1. In general, most people can be trusted.
2. It is hard for me to work with people that I do not know (R).
3. People tend to follow through with their word.
4. Most people fail to follow through with their word (R).
5. I cannot rely on people, because most people let me down (R).
6. I believe that other people can be counted on to do what they say they will do.
7. I find it hard to care about people that I do not know (R).
8. Other people generally have bad intentions (R).
9. Other people generally have good intentions.
10. I find it hard to trust other people. (R)
11. Collaborating with others is fun to me.
12. Dealing with other people can be quite a pain (R).
13. Most people are up to no good (R).
14. Other people are incompetent (R).

Note: *R* = reverse-coded

REFERENCES

Albrecht, S.L., & Traviglione, A. (2003). Trust in public senior management during times of turbulent change. *International Journal of Human Resource Management*, *14*, 1 – 17.

Angoff, W. (1993). Perspective on differential item functioning methodology. In P.W. Holland & H. Wainer (Eds.). *Differential item functioning* (p 3 – 24). Hillsdale, NJ: Lawrence Erlbaum.

Aryee, S., Budhwar, P.S., & Chen, Z.X. (2002). Trust as a mediator between the relationship of organizational justice and work outcomes: Test of a social exchange model. *Journal of Organizational Behavior*, *23*, 267 – 285.

Ayala, R.J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.

Barrick, M.R., & Mount, M.K. (2007). Personal characteristics inventory (PCI). *Wonderlic*, 1 – 10.

Barrick, M.R., Mount, M.K., & Judge, T.A. (2001). Personality and job performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9 – 30.

Barrick, M.R., Stewart, G.L., Neubert, M.J., & Mount, M.K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*, 377 – 391.

Bell, S.T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology*, *92*, 595 – 615.

Bennis, W., & Nanus, B. (1985). *Leaders*. New York: Harper and Row.

Best, R.G., Stapleton, L.M., & Downey, R.G. (2005). Core self-evaluations and job burnout: The test of alternative models. *Journal of Occupational Health Psychology*, 10, 441-451.

Bettenhausen, K.L. (1991). Five years of group research: What we have learned and what needs to be addressed. *Journal of Management*, *17*, 345 – 381.

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J.P. Robinson, P.R. Shaver, & L.S.

Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. San Diego:

Academic Press.

Bock, R.D., & Mislevy, R.J. (1982). Adaptive EAP estimation of ability in a microcomputer environment.

Applied Psychological Measurement, *6*, 431 – 444.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation

perspective. *Psychological Bulletin*, *110*, 305 – 314.

Bono, J.E., & Colbert, A.E. (2005). Understanding responses to multi-source feedback: The role of core

self-evaluations. *Personnel Psychology*, *58*, 171 – 203.

Bono, J.E., Judge, T.A. (2003). Core self-evaluations: A review of the trait and its role in job satisfaction

and job performance. *European Journal of Personality*, 17, 5-18.

Boyar, S.L. & Mosley, D.C. (2007). The relationship between core self-evaluations and work and family

satisfaction: The mediating role of work-family conflict and facilitation. Journal of Vocational

Behavior, 265-281.

Brief, A.P., & Motowildo, S.J. (1986). Prosocial organizational behavior. *Academy of Management

Review, 11*, 710 – 725.

Brunborg, G.S. (2008). Core self-evaluations: A predictor variable for job stress. *European Psychologist*,

13, 96-102.

Bryant, D., & Wooten, W. (2006). Developing an essentially unidimensional test with cognitively designed

items. *International Journal of Testing*, *6*, 205 – 228.

Camilli, G., & Shepard, L. (1994). Methods for identifying biased test items. Newbury Park: Sage.

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-

multimethod matrix. *Psychological Bulletin*, *56*, 81 – 105.

Cattell, R.B. (1965). *The scientific analysis of personality*. Baltimore, MD: Penguin.

Chen, C.C., Chen, X., & Meindl, J.R. (1998). How can cooperation be fostered? The cultural effects of

individualism – collectivism. *Academy of Management Review*, *23*, 285 – 304.

Chiaburu, D.S., & Byrne, Z.S. (2009). Predicting OCB role definitions: Exchanges with the organization

and psychological attachment. *Journal of Business Psychology, 24*, 201 – 214.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Lawrence Erlbaum
　　　Associates. Hillsdale: New Jersey.

Cohen, S.G. & Bailey, D.E. (1997). What makes teams work: Group effectiveness research from the shop
　　　floor to the executive suite. *Journal of Management*, *23*, 239 – 290.

Colquitt, J.A. (2004). Does the justice of the one interact with the justice of the many? Reactions to
　　　procedural justice in teams. *Journal of Applied Psychology, 89*, 633 – 646.

Colquitt, J.A., Noe, R.A., & Jackson, C.L. (2002). Justice in teams: Antecedents and consequences of
　　　procedural justice climate. *Personnel Psychology, 55*, 83 – 110.

Costa, P.T. Jr., & McCrae, R.R. (1994). Set like plaster? Evidence for the stability of adult personality. In
　　　T.F. Hetherton & J.L. Weinberger (Eds), *Can personality change*? Washington, DC: American
　　　Psychological Association.

Costa, P.T., Jr., & McCrae, R.R. (1988). Personality in adulthood: A six-year longitudinal study of self-
　　　reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social
　　　Psychology, 54*, 853-863.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual
　　　Differences*, *13*, 653-665.

Cox, J.F. (1994). *The effects of superleadership training on leader behavior, subordinate self-leadership
　　　behavior and subordinate citizenship*. Unpublished doctoral dissertation, University of Maryland,
　　　College Park.

Cox, T.H., Lobel, S.A., & McLeod, P.L. (1991). Effects of ethnic group cultural differences on cooperative
　　　and competitive behavior on a group task. *Academy of Management Journal, 34*, 827 – 847.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*,
　　　281 – 302.

de Pater, I.E., Schinkel, S., & Nijstad, B.A. (2007). Validation of the Dutch core self-evaluations scale.
　　　Gedrag en Organisatie, *20*, 82-100.

Devine, D.J. (2002). A review and integration of classification systems relevant to teams in organizations.
　　　Groups Dynamics: Theory, Research, and Practice, *6*, 291 – 310.

Devine, D.J., Clayton, L.D., Phillips, J.L., Dunford, B.B., & Melner, S.B. (1999). Teams in organizations: Prevalence, characteristics, and effectiveness. *Small Group Research*, *30*, 678 – 711.

Dickson, P.H., & Weaver, K.M. (1997). Environmental determinants and individual-level moderators of alliance use. *Academy of Management Journal, 40*, 404 – 425.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Dilchert, S., Ones, D. S., Van Rooy, D. L., & Viswesvaran, C. (2006). Big Five factors of personality. In J. H. Greenhaus & G. A. Callanan (Eds.), *Encyclopedia of career development*. Thousand Oaks, CA: Sage.

Dirks, K.T., & Ferrin, D.L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology*, *87*,   611 – 628.

Dormann, C., Fay, D., Zapf, D., & Frese, M. (2006). A state-trait analysis of job satisfaction: On the effect of core self-evaluations. *Applied Psychology: An International Review*, 55, 27-51.

Earley, P.C., & Gibson, C.B. (1998). Taking stock in our progress on individualism-collectivism: 100 years of solidarity and community. *Journal of Management*, *24*, 265 – 304.

Eby, L.T., & Dobbins, G.H. (1997). Collectivistic orientation in teams and individual and group-level analysis. *Journal of Organizational Behavior*, *18*, 275 – 295.

Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Erez, A., & Judge, T.A. (2001). Relationship of core self-evaluations to goal-setting, motivation, and performance. *Journal of Applied Psychology*, 86, 1270-1279.

Erikson, E.H. (1950). *Childhood and society*. New York: Norton.

Erikson, E.H. (1963). *Childhood and society* (2nd ed.). New York: Norton.

Foley, S., Ngo, H., Loi, R. (2006). How do cultural types affect work-related attitudes? The mediating role of perceived organizational support. *International Journal of Employment Studies*, *14*, 37 – 62.

Germain, Valois, & Abdous (2007, September 7). Eirt – Item response theory assistant for Excel. Retrieved from http://libirt.sourceforge.net/eirt-en/index.html.

Gierl, M.J., Bisanz, J., & Li, Y.Y. (2004). Using the multidimensionality-based DIF analysis paradigm to study cognitive skills that elicit group differences: A critique. Paper presented at the *Annual Meeting of the National Council on Measurement in Education (NCME)*. San Diego, CA.

Goldberg, L. R. (1990). An alternative "description of personality": The Big Five factor structure. *Journal of Personality and Social Psychology, 59*, 1216-1229.

Grant, A.M. (2008). Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology*, *93*, 48-58.

Grant, A.M., & Wrzesniewski, A. (2010). I won't let you down…or will I? Core self-evaluations, other-orientation, anticipated guilt and gratitude, and job performance. *Journal of Applied Psychology, 95*, 108 – 121.

Graziano, W.G., Hair, E.C., & Finch, J.F. (1997). Competitiveness mediated the link between personality and group performance. *Journal of Personality and Social Psychology*, *73*, 1394 – 1408.

Graziano, W.G., Jensen-Campbell, L.A., Hair, E.C. (1996). Perceiving interpersonal conflict and reacting to it: The case for agreeableness. *Journal of Personality and Social Psychology*, *70*, 820 – 835.

Guilford, J.P., Guilford, J.S., & Zimmerman, W.S. (1978). *The Guilford-Zimmerman Temperament Survey: Manual of Instructions and Interpretations*. Sheridan Psychological Services, Inc.

Guilford, J.S., Zimmerman, W.S., Guilford, J.P. (1976). *The Guilford-Zimmerman Temperament survey handbook: Twenty-five years of research and application.* San Diego, CA: EdITS Publishers.

Gully, S.M. (2000). Work team research: Recent findings and future trends. In M.M. Beyerlein (Ed.), *Work teams: Past, present, and future* (p. 25 – 44). The Netherlands: Kluwer Academic Publishers.

Gundlach, M., Zivnuska, S, & Stoner, J. (2006). Understanding the relationship between individualism-collectivism and team performance through an integration of social identity theory and the social relations model. *Human Relations*, *59*, 1603 – 1632.

Guzzo, R.A., & Dickson, M.W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*, 307 – 338.

Guzzo, R.A., & Shea, G.P. (1992). Group performance and intergroup relations in organizations. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, p. 269 – 313). Palo Alto, CA: Consulting Psychologist Press.

Hackman, J.R. (1992). Group influences on individuals in organizations. In M.D. Dunnette & L.M. Hough

   (Eds.), *Handbook of industrial and organizational psychology* (Vol.3, p. 199 – 267). Palo Alto:

   Consulting Psychologist Press.

Halfhill, T., Sundstrom, E., Lahner, J., Calderone, W., & Nielsen, T.M. (2005). Group personality

   composition and group effectiveness: An integrative review of empirical research. *Small Group

   Research*, *36*, 83 – 105.

Hambleton, R.K. (1989). Principles and selected applications of item response theory. In R.L. Linn (Ed.),

   Educational Measurement (3rd ed., pp. 147 – 200). New York: Macmillan.

Hambleton, R.K., & Jones, R.W. (1993). Comparison of classical test theory and item response theory

   and their applications to test development. *Educational Measurement: Issues and Practice*, *12*,

   253 – 262.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory*.

   Beverly Hills, CA: Sage.

Hansen, M.H., Morrow, J.L., & Batista, J.C. (2002). The impact of trust on cooperative membership

   retention, performance, and satisfaction: An exploratory study. *International Food and

   Agribusiness Management Review*, *5*, 41 – 59.

Harter, S. (1990). Causes, correlates, and the functional role of global self-worth: A life-span perspective.

   In R.J. Sternberg & J. Kolligan, Jr. (Eds), *Competence Considered*. New Haven: Yale University

   Press.

Heller, D., Judge, T.A., & Watson, D. (2002). The confounding role of personality and trait affectivity in the

   relationship between job and life satisfaction. *Journal of Organizational Behavior*, *23*, 815.

Ho, D.Y., & Chiu, C. (1994). Component ideas of individualism, collectivism, and social organization. In

   U.Kim, H.C. Triandis, C. Kagitcibasi, S.C. Choi, & G. Yoon (Eds.), *Individualism and collectivism:

   Theory, method, and applications* (p. 200 – 210). Thousand Oaks, CA: Sage.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly

   Hills, CA: Sage.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and

   organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.

Holmbeck, G.N. (2002). Post-hoc probing of significant moderational and meditational effects in studies of

pediatric populations. *Journal of Pediatric Psychology*, *27*, 87 – 96.

Holt, D.T., Jung, H. (2008). Development of a Korean version of a core self-evaluations scale.

Psychological Reports, *103*, 415-425.

Hough, L.M., & Oswald, F.L. (2000). Personnel selection: Looking toward the future-remembering the

past. *Annual Review of Psychology*, *51*, 631 – 664.

Hui, C.H., & Yee, C. (1999). The impact of psychological collectivism and workgroup atmosphere on

Chinese employees' job satisfaction. *Applied Psychology: An International Review*, *48*, 175 –

185.

Hui, C.H., Yee, C., & Eastman, K.L. (1995). The relationship between individualism-collectivism and job

satisfaction. *Applied Psychology: An International Review*, *44*, 276 – 282.

Hurtz, G.M., & Donovan, J.J. (2000). Personality and job performance: The Big Five revisited. *Journal of

Applied Psychology*, *85*, 869 – 879.

Ilgen, D.R., Hollenbeck, J.R., Johnson, M., Jundt, D. (2005). Teams in organizations: From input-process-

output models to IMOI models. *Annual Review of Psychology*, *56*, 517-543.

Ilies, R., Fulmer, I.S., Spitzmuller, M., & Johnson, M.D. (2009). Personality and citizenship behavior: The

mediating role of job satisfaction. *Journal of Applied Psychology*, *94*, 945 – 959.

Jackson, C.L., Colquitt, J.A., Wesson, M.J., & Zapata-Phelan, C.P. (2006). Psychological collectivism: A

measurement validation and linkage to group member performance. *Journal of Applied

Psychology, 91*, 884 – 899.

Jensen-Campbell, L.A., & Gleason, K.A., & Adams, R., & Malcolm, K.T. (2003). Interpersonal conflict,

agreeableness, and personality development. *Journal of Personality*, *71*, 1059 – 1085.

Jensen-Campbell, L.A., & Graziano, W.G. (2001). Agreeableness as a moderator of interpersonal conflict.

Journal of Personality, *69*, 323 – 362.

John, O., Donahue, E., & Kentle, R. (1991). *The "Big Five" Inventory—Versions 4a and 54*. Technical

Report, Institute of Personality Assessment and Research, Berkeley, CA: University of California,

Berkeley.

John, O.P., & Srivastava, S. (1999). The big-five trait taxonomy: History, measurement, and theoretical

    perspectives. In L. Pervin & O.P. John (Eds.), *Handbook of personality: Theory and research* (2nd

    ed.). New York: Guilford.

Johnson, R.E., Rosen, C.C., & Levy, P.E. (2007). Getting to the core of core self-evaluation: A review and

    recommendations. *Journal of Organizational Behavior*, 29, 391-413.

Judge, T.A. (2009). Core self-evaluations and work success. *Current Directions in Psychological Science*,

    18, 58-62.

Judge, T.A., & Bono J.E. (2001a). A rose by any other name : Are self-esteem, generalized self-efficacy,

    neuroticism, and locus of control indicators of a common construct? In B.W. Roberts & R. Hogan

    (Eds.), *Personality psychology in the workplace* (pp. 93–118). Washington, DC: American

    Psychological Association.

Judge, T.A., & Bono, J.E. (2001b). Relationship of core self-evaluations traits—self-esteem, generalized

    self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance:

    A meta-analysis. *Journal of Applied Psychology*, 86, 80–92.

Judge, T.A., & Bono, J.E. (1999, April). *Core self-evaluations and construct breadth: Effects on job*

    satisfaction and job performance. Paper presented at the 14th annual conference of the Society

    for Industrial and Organizational Psychology, Atlanta, GA.

Judge, T.A., Bono, J.E., Erez, A., & Locke, E.A. (2005). Core self-evaluations and job and life satisfaction:

    The role of self-concordance and goal attainment. *Journal of Applied Psychology*, 90, 257-268.

Judge, T.A., Bono, J.E., & Locke, E.A. (2000). Personality and job satisfaction: The mediating role of job

    characteristics. *Journal of Applied Psychology*, 85, 237–249.

Judge, T.A., Erez, A., & Bono, J.E. (1998). The power of being positive: The relationship between positive

    self-concept and job performance. *Human Performance*, *11*, 167 – 187.

Judge, T.A., Erez, A., Bono, J.E., & Thoresen, C.J. (2003). The core self-evaluations scale: Development

    of a measure. *Personnel Psychology*, *56*, 303–331.

Judge, T.A., Erez, A., Bono, J.E., & Thoresen, C.J. (2002). Are measures of self esteem, neuroticism,

    locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of*

    Personality and Social Psychology, *83*, 693–710.

Judge, T.A., Heller, D., & Mount, M.K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, *87*, 530 – 541.

Judge, T.A. & Hurst, C. (2007). Capitalizing on One's Advantages: Role of core self-evaluations. *Journal of Applied Psychology*, *92*, 1212-1227.

Judge, T.A. & Hurst, C. (2008). How the rich (and happy) get richer (and happier): Relationship of core self-evaluations to trajectories in attaining work success. *Journal of Applied Psychology, 93*, 849-863.

Judge, T.A., Locke, E.A., & Durham, C.C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior*, *19*, 151-188.

Judge, T.A., Locke, E.A., Durham, C.C., & Kluger, A.N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, *83*, 17–34.

Judge, T.A., Thoresen, C.J., Pucik, V., & Welbourne, T.M. (1999). Managerial coping with Organizational Change : A dispositional perspective. *Journal of Applied Psychology*, *84*, 107-122.

Judge, T.A., Van Vianen, A.E.M., & De Pater, I.E. (2004). Emotional stability, core self-evaluations, and job outcomes: A review of the evidence and an agenda for future research. *Human Performance*, *17*, 325-346.

Kaiser, H.F. (1970). A second generation little jiffy. *Psychometrika*, 35, 401 – 415.

Kamdar, D., & Van Dyne, L. (2007). The joint effects of personality and workplace social exchange relationships in predicting task performance and citizenship performance. *Journal of Applied Psychology, 92*, 1286 – 1298.

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

Katz, H.A., & Rotter, J.B. (1969). Interpersonal trust scores of college students and their parents. *Child Development*, 40, 657 – 661.

Kirkman, B.L., & Shapiro, D.L. (2001). The impact of cultural values on job satisfaction and organizational commitment in self-managing work teams: The mediating role of employee resistance. *Academy of Management Journal, 44*, 557 – 569.

Kline, T.J.B. (1999). The team player inventory: reliability and validity of a measure of predisposition toward organizational team-working environments. *Journal of Specialists in Group Work*, *24*, 102 – 112.

Kline, T.J.B., & O'Grady, J.K. (2009). Team member personality, team processes and outcomes: Relationships within a graduate student project team sample. *North American Journal of Psychology, 11*, 369 – 382.

Korsgaard, M.A., Meglino, B.M., & Lester, S.W. (1997). Beyond help: Do other-oriented values have broader implications in organizations? *Journal of Applied Psychology*, *82*, 160 – 177.

Kozlowski, S.W.J., & Bell, B.S. (2003). Work groups and teams in organizations. In W.C. Borman, D.R. Ilgen, & R.J. Klimoski (Eds.), *Handbook of psychology: Vol. 12: Industrial and organizational psychology* (p. 333 – 375). London: Wiley.

Kozlowski, S.W.J., Gully, S.M., Nason, E.T., & Smith, E.M. (1999). Developing adaptive teams: A theory of compilation and performance across time and levels. In D.R. Ilgen & E.D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, personnel actions, and development* (p. 240 – 292). San Francisco: Jossey – Bass.

Kozlowski, S.W.J., & Ilgen, D.R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*, 77 – 124.

Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*, 563 – 575.

Lemelle, C.J., & Scielzo, S.A. (under review). How you feel about yourself can affect how you feel about your job: A meta-analysis examining the relationship of core self-evaluations and job satisfaction. Submitted for publication at the *Journal of Organizational Behavior* (2010).

Lester, S.W., Meglino, B.M., & Korsgaard, M.A. (2008). The role of other orientation in organizational citizenship behavior. *Journal of Organizational Behavior*, *29,* 829 – 841.

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In H.M. Lefcourt (Ed.), *Research with the locus of control construct*. New York: Academic Press.

Locke, E.A., McClear, K., & Knight, D. (1996). Self-esteem and work. *International Review of Industrial/Organizational Psychology*, *11*, 1-32.

Lord, F.N., & Novick, M.R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*, 719 – 748.

Manz, C.C., & Sims, H.P. Jr. (1987). Leading workers to lead themselves: The external leadership of self-managing work teams. *Administrative Science Quarterly*, *32*, 106 – 129.

Markus, H.R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224 – 253.

Mayer, R.C., Davis, J.H., & Schoorman, F.D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*, 709 – 734.

McNeely, B.L., & Meglino, B.M. (1994). The role of dispositional and situational antecedents in prosocial behavior: An examination of the intended beneficiaries of prosocial behavior. *Journal of Applied Psychology, 79*, 836 – 844.

Meglino, B.M., & Korsgaard, M.A. (2004). Considering rational self-interest as a disposition: Organizational implications of other orientation. *Journal of Applied Psychology*, *89*, 946 – 959.

Meglino, B.M., & Korsgaard, M.A. (2007). The role of other orientation in reactions to job characteristics. *Journal of Management, 33*, 57 – 83.

Mehrens, W.A., & Lehmann, I.J. (1973). *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart, and Winston.

Moorman, R.H., & Blakely, G.L. (1995). Individualism – collectivism as an individual difference predictor of organizational citizenship behavior. *Journal of Organizational Behavior, 16*, 127 – 142.

Mount, M.K., Barrick, M.R., & Stewart, G.L. (1998). Five factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance*, *11*, 145 – 165.

Mulki, J.P., Jaramillo, F., & Locander, W.B. (2006). Effects of ethical climate and supervisory trust on salesperson's job attitudes and intention to quit. *Journal of Personal Selling and Sales Management*, *1*, 19 – 26.

Murphy, K.R. (1999). The challenge of staffing a postindustrial workplace. In D.R. Ilgen & E.D. Pulakos

    (Eds.), *The changing nature of performance: Implications for staffing, motivation, and*

    development (p. 295 – 324). San Francisco: Jossey – Bass.

Neuman, G.A., & Wright, J. (1999). Team effectiveness: Beyond skills and cognitive ability. *Journal of*

    Applied Psychology, *84*, 376 – 389.

O'Neill, T.A., & Kline, T.J.B. (2008). Personality as a predictor of teamwork: A business simulator study.

    North American Journal of Psychology, *10*, 65 – 78.

Organ, D.W, & Konovsky, M.S. (1989). Cognitive versus affective determinants of organizational

    citizenship behavior. *Journal of Applied Psychology*, 74, 157 – 164.

Organ, D.W., & Lindl, A. (1995). Personality, satisfaction, and organizational citizenship behavior. *The*

    Journal of Social Psychology, *135*, 339 – 350.

Orlando, M., (2005). Critical issues to address when applying item response theory models.

Oyserman, D., Coon, H.M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism:

    Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3 – 72.

Packer, E. (1985). Understanding the subconscious. *The Objectivist Forum*, *6* (1), 1-10.

Pearce, C.L., & Sims, H.P. (2002). Vertical versus shared leadership as predictors of the effectiveness of

    change management teams: An examination of aversive, directive, transactional,

    transformational, and empowering leader behaviors. *Group Dynamics: Theory, Research, and*

    Practice, *6*, 172 – 197.

Perrewe, P.L., & Spector, P.E. (2002). Personality research in the organizational sciences. In G.R. Ferris

    (Ed.), *Research in personnel and human resource management* (Vol. 21, p. 1 – 63). New York:

    Elsevier Science.

Piccolo, R.F., Judge, T.A, Takahashi, K., Watanabe, N., & Locke, E.A. (2005). Core self-evaluations in

    Japan: Relative effects on job satisfaction, life satisfaction, and happiness. *Journal of*

    Organizational Behavior, 26, 965-984.

Pillai, R., & Meindl, J.R. (1998). Context and charisma: A "meso" level examination of the relationship of

    organic structure, collectivism, and crisis to charismatic leadership. *Journal of Management, 24*,

    643 – 671.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Ravlin, E.C., & Meglino, B.M. (1987). Effect of values on perception and decision making: A study of alternative work values measures. *Journal of Applied Psychology*, *72*, 666 – 673.

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

Reckase, M.D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9, 401 – 412.

Reckase, M.D., Ackerman, T.A., Carlson, J.E. (1988). Building a unidimensional test using multidimensional items. *Journal of Educational Measurment*, 25, 193 – 203.

Reeves, B.B., & Fayers, P. (2005). Applying item response theory modeling for evaluating questionnaire item and scale properties. *In Assessing Quality of Life in Clinical Trial: Methods and Practice* (Eds. Fayers, B. & Hays). Oxford: Oxford University Press, 55 – 74.

Reise, S.P., Ainsworth, A.T., & Haviland, M.G. (2005). Item response theory: Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science*, *14*, 95 – 101.

Reise, S.P., & Yu, J. (1990). Parameter recovery in the graded response model using MULTILOG. *Journal of Educational Measurement, 27*, 133 – 144.

Requena, F. (2002). Social capital, satisfaction, and quality of life in the workplace. *Social Indicators Research*, *61*, 331 – 360.

Ricketta, M., & Van Dick, R. (2005). Foci of attachments in organizations: A meta-analytic comparison of the strength and correlates of work group versus organizational commitment and identification. *Journal of Vocational Behavior*, *67*, 490 – 510.

Rizopoulos, D. (2006). LTM: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, *17*, 1 – 25.

Robinson, S.L., & Morrison, E.W. (1995). Psychological contracts and OCB: The effect of unfilled obligations on civic virtue behavior. *Journal of Organizational Behavior*, *16*, 289 – 298.

Rokeach, M. (1972). *Beliefs, attitudes, and values: A theory of organization and change*. San Francisco: Jossey-Bass.

Rosenberg, M.J., Schooler, C., & Schoenbach, C. (1989). Self-esteem and adolescent problems: Modeling reciprocal effects. *American Sociological Review, 54*, 1004-1018.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*.

Rotter, J.B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35*, 651 – 665.

Rotter, J.B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist, 26*, 443 – 452.

Roussos, L., & Stout, W. (1996). A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement, 20*, 355 – 371.

Ryan, R.M., & Connell, J.P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*, 749 – 761.

Saane, N.V., Sluiter, J.K., Verbeek, J.H.A.M., Frings-Dresen, M.H.W. (2003). Reliability and validity of instruments measuring job satisfaction – a systematic review. *Occupational Medicine, 53*, 191-200.

Sager, J.K., Strutton, H.D., & Johnson, D.A. (2006). Core self-evaluations and salespeople. *Psychology & Marketing, 23* (2), 95-113.

Salvaggio, A.N., Schneider, B., Nishii, L.H., Mayer, D.M., Ramesh, A., & Lyon, J.S. (2007). Manager personality, manager service quality orientation, and service climate: Test of a model. *Journal of Applied Psychology, 92*, 1741 – 1750.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, No. 17.

Scherbaum, C.A., Finlinson, S., Barden, K., & Tamanini, K. (2006). Applications of item response theory to measurement issues in leadership research. *The Leadership Quarterly, 17*, 366 – 386.

Schoorman, F.D., Mayer, R.C., Davis, J.H. (2007). An integrative model of organizational trust: Past, present, and future. *Academy of Management Review, 32*, 344 – 354.

Schumaker, R.E. (2005). Test bias and differential item functioning. *Applied Measurement Associates*, 1 – 2.

Schwab, D.P. (1980). Construct validity in organization behavior. In B.M. Staw & L.L. Cummings (Eds.), *Research in Organizational Behavior*. Vol 2, Greenwich, CT: JAI Press.

Seibert, S.E., & Kraimer, M.L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behavior, 58,* 1 – 21.

Shealy, R. & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DIF as well as item bias/DIF. *Psychometrika, 58*, 159 – 194.

Smith, P.C., Kendall, L.M., & Hulin, C.L. (1969). *The measurement of satisfaction in work and retirement: A strategy for the study of attitudes.* Oxford, England: Rand McNally.

Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups from the Hawthorne studies to work teams of the 1990s and beyond. *Group Dynamics: Theory, Research, and Practice*, *4*, 44 – 67.

Stout, W.F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, *52*, 589 – 617.

Stout, W., Habing, B., Douglas, J., & Kim, H.R. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 20 331 – 354.

Swaminathan, H., & Rogers, H.J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361 – 370.

Taggar, S. (2000). Personality, cognitive ability, and behavior: The antecedents of effective autonomous work teams. *Dissertation Abstracts International*, *60*(9 – A).

Tan, H.H., & Tan, C.S. (2000). Toward the differentiation of trust in supervisor and trust in organization. *Genetic, Social, and General Psychology Monographs*, *126*, 241 – 260.

Tett, R.P., Jackson, D.N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*, 703 – 742.

Tharenou, P. (1979). Employee self-esteem: A review of the literature. *Journal of Vocational Behavior*, *15*, 316-346.

Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, *51*, 567 – 577.

Triandis, H.C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

Triandis, H.C., Leung, K., Villareal, M., & Clack, F.L. (1985). Allocentric vs idiocentric tendencies:
Convergent and discriminant validation. *Journal of Research in Personality*, *19*, 395 – 415.

Tsutakawa, R.K., & Johnson, J.C. (1990). The effect of uncertainty of item parameter estimation on ability
estimates. *Psychometrika, 55*, 371 – 390.

Tyler, T.R. (2003). Trust within organizations. *Personnel Review*, *32*, 556 – 568.

Tyran, K.L., & Gibson, C.B. (2008). Is what you see, what you get?: The relationship among surface- and
deep-level heterogeneity characteristics, group efficacy, and team reputation. *Group &
Organzation Management, 33*, 46 – 76.

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*, 4-69.

Van Dyne, L., Vandewalle, D., Kostova, T., Latham, M.E., & Cummings, L.L. (2000). Collectivism,
propensity to trust, and self-esteem as predictors of organizational citizenship in a non-work
setting. *Journal of Organizational Behavior, 21*, 3 – 23.

Wagner, J.A. (1995). Studies of individualism – collectivism: Effects on cooperation in groups. *Academy
of Management Journal, 38*, 152 – 172.

Wagner, J.A., & Moch, M.K. (1986). Individualism-collectivism; Concept and measurement. *Group and
Organizational Studies, 11*, 280 – 304.

Wanberg, C.R., Glomb, T.M., Song, Z., & Sorenson, S. (2005). Job-search persistence during
unemployment: A 10-wave longitudinal study. *Journal of Applied Psychology*, *90*, 411 – 430.

Wheaton, B., Muthen, B., Alwin, D.F., & Summers, G.F. (1977). Assessing reliability and stability in panel
models. *Sociological Methodology*, *8*, 84-136.

Willenbrock, N.L., & Kauffeld, S. (2010). Development and construct validation of the German Workplace
Trust Survey (G – WTS). *European Journal of Psychological Assessment*, *26*, 3 – 10.

Yagil, D., Luria, G., & Gal, I. (2008). Stressors and resources in customer service roles: Exploring the

    relationship between core self-evaluations and burnout. *International Journal of Service Industry*

    Management, *19*, 575-595.

Zhao, H., & Seibert, S.E. (2006). The Big Five personality dimensions and entrepreneurial status: A meta-

    analytic review. *Journal of Applied Psychology, 91*, 259 – 271.

BIOGRAPHICAL INFORMATION

Chloe Lemelle received her doctoral degree in Experimental Psychology with an emphasis on Industrial/Organizational Psychology on December 16, 2010. Previously, Dr. Lemelle received her master's degree in Industrial/Organizational Psychology in May 2008 from the University of Texas at Arlington. Dr. Lemelle maintained a 4.0 GPA in nine of the ten semesters of her graduate program. From a research standpoint, Dr. Lemelle's interests include organizational teams, core self-evaluations, diversity issues, and cognitive ability testing in the workplace. Dr. Lemelle has presented research at many conferences including the Society for Industrial and Organizational Psychologists' Annual Conference 2007 - 2010, Psi Chi Annual Conference 2007, the Dallas Area Industrial and Organizational Psychologists Annual Conference 2008, and the Annual Celebration of Excellence in Students Conference 2008 and 2009. Dr. Lemelle is currently employed as a consultant at a human resource consulting firm, Batrus Hollweg International. She has been with Batrus Hollweg International since May 2007.