THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

PROFILING PERSONALITY: A NON-COMPENSATORY, OPTIMALITY-BASED MEASUREMENT APPROACH

by
Christina Rachel Criswell

A dissertation submitted in partial fulfillment of

the requirements for the degree of Doctor of Philosophy

in the Discipline of Industrial/Organizational Psychology

The Graduate School

The University of Tulsa

2013

THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

PROFILING PERSONALITY: A NON-COMPENSATORY, OPTIMALITY-BASED MEASUREMENT APPROACH

by
Christina Rachel Criswell

A DISSERTATION

APPROVED FOR THE DISCIPLINE OF

INDUSTRIAL/ORGANIZATIONAL PSYCHOLOGY

By Dissertation Committee

______________________________, Chair
Robert Tett

____
Bradley Brummel

____
John McNulty

____
Mary Dana Laird

ABSTRACT

Christina Rachel Criswell (Doctor of Philosophy in Industrial/Organizational Psychology)

Profiling Personality: A Non-Compensatory, Optimality-Based Measurement Approach

Directed by Dr. Robert P. Tett

129 pp., Chapter 5: Discussion

(377 words)

Research over the last several decades has established that personality characteristics predict job performance (e.g., Barrick & Mount, 1991; Hogan & Holland, 2003; Tett, Jackson, & Rothstein, 1991). However, common methods for evaluating personality use approaches that primarily focus on individual scales; the field of Industrial/Organizational Psychology is missing a definitive approach to measuring personality holistically (i.e., at the profile level). Where personality profiles have been evaluated, most modern analytic approaches are either compensatory (combine scales additively such that one scale score can compensate for or mask another), linear (assume that, either in relative or absolute terms, it is important to have either a high or low score on a given raw or normative scale), or both.

In contrast, this study tested a uniquely non-compensatory, optimality-based approach to personality profile derivation, measurement, and evaluation. The approach here was non-compensatory in that the methods used were multiplicative. That is, scale scores were not combined additively: high scores on one scale were not allowed to mask or compensate for low scores on another. The approach in this study was also optimality-based versus linear; profile fit was evaluated with regard to how closely an individual's mean scale scores matched a target profile, rather than being evaluated linearly in terms of the sheer magnitude of the score itself. These two attributes of the current study—non-compensatoriness and optimality—represented an exploratory departure from mainstream personality profiling methods.

The present study attempted to (1) identify optimality-based, latent personality profiles, using a Latent Profile Analysis (LPA) methodology (2) assess the generalizability of those profiles within and across distinct industries and job levels, and (3) examine, via a unique non-compensatory fit score, the validity of those profiles for predicting performance, both alone, and incrementally (beyond existing predictive models).

Support for the hypotheses in this study was mixed. Using LPA, interpretable, optimality-based personality profiles were identified, with a few exceptions. However, these profiles were not generalizable across job level, industry, or a combination of the two. The non-compensatory fit score derived from this study did show modest relationships with job performance, but did not consistently demonstrate incremental validity over more established compensatory methods (D^2 and profile-fit correlation). Given the exploratory nature of this work, potential limitations and directions for future research are discussed at length.

ACKNOWLEDGEMENTS

First, I would like to thank my committee for their support. In particular, my undying gratitude goes out to Dr. Robert Tett for his patience and faith in my ability to cross the finish line. Over six different dissertation topics, he has helped me expand my knowledge base, refine my ideas, think critically about my core research questions, and evaluate what they can add to the field. Dr. Tett, you have been a constant source of reason, ideas, and creativity, and I cannot thank you enough.

To Drs. McNulty, Brummel, and Laird, I truly appreciate your patience and insight. I extend gratitude to Dr. Brent Holland for furthering my understanding of the applied side of our field, and to Drs. Robert and Joyce Hogan for wrapping meaning and values around it. Hogan Assessment Systems has always demonstrated an unwavering commitment to the field of I/O in general, and to the growth and development of I/O students and new graduates in particular, as demonstrated by the fact that they gave me my first *real* job, and more than a decade later, kindly furnished the data for this study.

Special thanks to Geoff Howland, Hugh Briggs, and B. Riley for their help, friendship, and specialized expertise. To the other friends and family too numerous to mention, thank you for your companionship, patience (or constructive lack thereof), and at times, lessons in overcoming adversity. Those lessons may not have always been welcome, but to the extent that I have learned something from you, I am grateful.

Thanks to my parents, Joyce and Harry Criswell. Dad, you taught me that “math always works.” Mom, you showed me how to face the world as a shy person and bend it to my will. You taught me to seek knowledge, question assumptions, and proceed logically before I could walk. You taught me never to be ashamed to “look smart,” which can be a tough lesson for a young girl… and when I failed to look smart, (sometimes miserably), you loved me anyway. You instilled in me the values that brought me here.

To my grandparents, Ray and LaVonne Louth, thank you for your endless supply of support and love. Grandpa, you instilled in me an appreciation for hard work and practicality, and have passed on to me much more of the wisdom you’ve accrued over your nine-plus decades of life than you realize. Grandma, you showed me how to approach life as a strong-willed, independent woman, and I miss you very much.

To Nathan and Maddie: you are not my blood, and I do not care in the slightest. You are each a small package of inspiration and joy, and my life is immensely better just for knowing you. The patience and understanding that you’ve extended to me as I’ve worked to finish this degree conveys a level of wisdom, maturity, and pure kindness achieved by few adults, much less kids of your age. I love you more than you can fathom.

Danny. My rock. I did not know you at the beginning of this journey, but I can say without question that I would not have reached the end without you. Since our beginning, you have demonstrated an uncanny ability for knowing when to support me, when to push me, and when to be patient with me. That ability has carried us through many joyous and difficult times, this effort included. Thank you for inspiring me to embrace life, chase down happiness, and tackle it with force when necessary. You have made my world a far better place, and me a far better person. I love you so, so much.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The discipline of personality assessment has grappled for decades with an identity crisis, struggling to define, defend, and attach consistent meaning to a century of research on the relationships between trait scores and workplace behavior (Barrick, Mount, & Judge, 2001; Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). After a particularly promising stretch of research that brought forth the development of the *California Psychological Inventory* in 1954 (Gough, 1987), the monograph that introduced the roots of the Five Factor Model (Tupes & Christal, 1961), and the first publication of the *Myers-Briggs Type Indicator* in 1962, (Myers, McCaully, Quenk, & Hammer, 1998), two devastating critiques of personality theory and measurement (Guion & Gottier, 1965; Mischel, 1968) effectively halted personality research for two decades (Hogan & Smither, 2001), made research funding for personality studies difficult to secure (Hogan & Roberts, 2001), and tarnished the reputation of those who still believed that personality measurement was a worthwhile endeavor (Digman, 1996; Hogan & Roberts, 2001).

Then, in the 1980s, personality research embarked on a renaissance, as several prominent researchers began to coalesce around a model of personality that was similar to that identified by Tupes and Christal decades earlier (Digman, 1996; Hogan & Roberts, 2001). In the years that followed, this model would come to be known as the Big

Five, or Five Factor Model (FFM). Researchers began to examine the assessments of the day for underlying evidence of the FFM (McCrae & Costa, 1985, 1989; Wiggins & Pincus, 1992), using the FFM structure as a Rosetta Stone of sorts to make new connections between existing assessments and associated publications and data sets. These findings were encouraging, bolstered arguments for a common five-factor structure, and helped lay the foundation for future studies to combine data across assessments. In relatively short order, several notable meta-analytic reviews in the early 1990s (Barrick & Mount, 1991; Hough, 1992; Schmitt, Gooding, Noe & Kirsch, 1984; Tett, Jackson, & Rothstein, 1991) shed new light on the viability of the FFM and its relationships with job performance.

In tandem with these developments, the applied usage of personality assessments in the workplace saw a marked increase. Practitioners were hungry for valid, reliable hiring methods that minimized group differences and enhanced fairness, particularly in the wake of the passage of the *Civil Rights Act of 1991* (CRA 1991), which strengthened certain provisions and protections that had been previously outlined in the *Civil Rights Act of 1964* (CRA 1964). As organizations worked to comply with the CRA 1991, and attempted to reconcile the potential contradictions between the 1964 and 1991 versions of the CRA, existing case law, and other relevant standards like *The Uniform Guidelines on Employee Selection Procedures* (hereafter *Uniform Guidelines*; Equal Employment Opportunity Commission [EEOC], 1978), they found that personality assessment afforded them a selection method that met the legal requirements for validity and reliability, while also minimizing differences between protected classes (Foldes, Duehr, & Ones, 2008; Van Landuyt, 2004).[1]

Specifically, personality assessment helped to address one particular mandate, found in the *Uniform Guidelines* and reiterated in the CRA 1991, which is critical to understanding and implementing psychological assessment for selection purposes in practice:

> Where two or more selection procedures are available which serve the user's legitimate interest in efficient and trustworthy workmanship, and which are substantially equally valid for a given purpose, the user should use the procedure which has been demonstrated to have the lesser adverse impact. Accordingly, whenever a validity study is called for by these guidelines, the user should include, as a part of the validity study, an investigation of suitable alternative selection procedures and suitable alternative methods of using the selection procedure which have as little adverse impact as possible, to determine the appropriateness of using or validating them in accord with these guidelines (EEOC, Section 3B).

Note that "adverse impact" is defined as "a substantially different rate of selection in hiring, promotion or other employment decision which works to the disadvantage of members of a race, sex or ethnic group" (EEOC, 1978, section 4D). The mandate that users seek out and use "alternative selection procedures" that minimize or eliminate adverse impact has had a far-reaching impact on the practice of I/O psychology, and the reiteration of such in the CRA 1991 did much to bolster and legitimize the use of personality assessments as selection instruments. Practitioners sought out creative ways to provide employers with valid, reliable selection tools that would minimize adverse impact, meet the "alternative selection procedure" requirement, or both. Because

[1] For a thorough review of the present legal environment surrounding the *Uniform Guidelines*, and its impact on research and practice in I/O psychology, see the focal article (McDaniel, Kepes, & Banks, 2001), commentaries, and author responses in the December 2011 volume of *Industrial and Organizational Psychology*.

personality assessments generally do not produce group differences among protected classes (Foldes et al., 2008; Gatewood & Feild, 1998; Hough, 1998; Sackett & Roth, 1996; Sackett & Wilk, 1994), pre-employment assessment programs with personality at their center flourished.

The dual impact of the legal atmosphere surrounding employee selection, combined with the new methodological frontiers afforded by meta-analysis, led to a marked increase in personality research. As researchers worked to respond to this new level of interest, improving observed validity coefficients became a goal in and of itself. Mischel (1968) had coined the derogatory term "personality coefficient," to describe "the correlation between .20 and .30 which is found persistently when virtually any personality dimension inferred from a questionnaire is related to almost any conceivable external criterion…" (p. 78). Personality researchers took as a challenge Mischel's assertion that correlations of this magnitude were "too low to have value for most individual assessment purposes." Armed now with a stable factor structure that was supported by meta-analytic evidence, the search began to identify higher validity estimates between personality scales and meaningful outcomes.

This renewed focus on improving validity sparked the development of new assessments like the *Hogan Personality Inventory* (Hogan & Hogan, 1995) and more focused approaches within personality research, such as exploring bandwidth-fidelity tradeoffs (see Holland, 2001; Ones & Viswesvaran, 1996; Tett, Steele, & Beauregard, 2000) and aligning predictors and criteria with one another (Hogan & Holland, 2003). Yet at times, these efforts seemed to focus on the discovery of stronger validity evidence on a scale-by-scale basis, putting forth a fragmented view of personality that was less than representative of how individuals may “experience” their own personalities and manifest their traits in behavioral form. Likewise, the increasingly common application of personality assessment in the workplace remained centered on the existing personality research at the time, and as such, maintained a similar scale-by-scale approach in practice.

The scale-level focus of both scientists and practitioners is understandable. Because the discovery and application of the FFM saw so much success, a five-scale focus usurped most alternative frameworks for personality research, and for the most part, still does. Concentrating on how best to measure, validate, and utilize each scale of the FFM has produced a bevy of informative research and data, allowed practitioners to make fair and valid hiring recommendations, and helped advance the field of I/O psychology in general (Hough, 2001). Yet the question of how traits work together within an individual—how each of these five traits act as a collection within, and on, a person—is a relatively unexplored frontier in the field.

As such, the present study attempts to assess the viability of a more holistic view of personality, based on personality *profiles* rather than personality *scales*. Such a configural approach to personality raises important concerns regarding interpretability, generalizability, and incremental prediction, each a key target of the current research.

As noted, applied personality research now generally conceptualizes normal personality in terms of the FFM. FFM-based assessments have been shown to predict performance across jobs and industries (Barrick & Mount, 1991; Bartram, 2005; Hogan & Holland, 2003; Hough, 1992; Hurtz & Donovan, 2000; Roberts, Chernyshenko, Stark, & Goldberg, 2005; Salgado, 1997; Tett, Jackson, Rothstein & Reddon, 1999), and as discussed previously, they tend to do so without producing the significant differences among protected groups typically seen with cognitive assessments. However, validity coefficients for personality assessments still rarely approach those of cognitive or General Mental Ability (GMA; see Spearman, 1904) measures, and most observed coefficients within the personality literature have only increased modestly since initial criticisms were leveled in the 1960s. For example, Schmitt, Gooding, Noe, and Kirsch (1984) examined validities for several different predictor-criterion pairs from published validation studies between 1964 and 1982. Averaged validity coefficients for personality ranged from .12 to .27, from which Schmitt et al. concluded, like Mischel years earlier, that the criterion validity of personality tests was uniformly poor. While the perspective of the field may have changed regarding whether a .27 correlation is considered poor, the bulk of published observed correlations between personality variables and various workplace criteria have not increased substantially, on average, since that time.

For these and other reasons, a more recent publication has renewed the decades-old call for the abandonment of traditional personality assessments in selection contexts. Morgeson et al. (2007) take issue with "the very low validity of personality tests for predicting job performance" (p. 683). In response to this call, other researchers have suggested that more attention needs to be focused on methodological issues ignored by Morgeson et al., such as the use of confirmatory strategies in validation (particularly those involving personality-based job analyses), attention to the potential for bidirectional relationships (e.g., scales or facets that relate to performance in a positive direction for some jobs and a negative direction for others), and the validation of cutoff scores or decision points in addition to assessment content (see Foster, Gaddis, & Hogan, 2012;

Holland & Van Landuyt, 2007; Ones, Dilchert, Viswesvaran, & Judge, 2007; Tett & Christiansen, 2007;).

Tett and Christiansen (2007) also point out that Morgeson et al. overlooked the potential of multivariate prediction in considering the criterion validity of personality tests. That is, much more so than GMA tests, personality assessments are multidimensional, and the five factors of personality (and their more specific facets) may each predict unique proportions of criterion variance. The multidimensionality of personality affords two distinct approaches to improving the prediction of performance over simplistic, single-trait models. First, multiple distinct personality scales can account for unique variance in a given criterion (Tett & Christiansen, 2007), as might be revealed by a simple hierarchical regression. Second, distinct personality traits may operate conjointly, whereby the impact of one trait (as a predictor of job performance) is magnified (or weakened) by the level of one or more other traits (Witt, 2002; Witt, Burke, Barrick, & Mount, 2002). Whereas the first approach is additive (in the form of a linear weighted sum), the second is interactive or configural.

Trait interactions are increasingly the focus of efforts to improve prediction and understanding of personality-job performance linkages. Witt et al. (2002), for example, found that the relationship between Conscientiousness and job performance is moderated by Agreeableness, such that high Conscientiousness contributes to performance only in those also high on Agreeableness. A logical next step in the study of trait interactions is to consider such interactions collectively at the level of the personality profile. That is, beyond a simple two-way trait interaction, can anything be gained by examining the joint contributions of multiple traits as a holistic set, i.e., as a *profile*?

Profile-level research in general, and profile fit evaluation methods in particular, are unique, in that analysis at the profile level carries an assumption of *optimality* rather than *linearity*. That is, a profile approach assumes it is important to be near (or ideally, at) a set of profile points (i.e., means, based on samples or subsamples); conversely, linear approaches, including both basic regression and interaction models, assume that, either in relative or absolute terms, it is important to have a high (or low) score on a given raw or normative scale, and an ideal score would be the highest (or lowest) score possible on said scale. One point regarding optimality bears further clarification. The word carries with it an implied value judgment via the root word "optimal." In this sense however, that value judgment is premature; optimality here does not refer to a better score—only a closer fit between an individual's score and a mean score on a target profile. The value judgment regarding whether this type of fit is good, or predictive of meaningful outcomes, is a separate issue evaluated in later stages of this study, but the distinction is relevant throughout this research. The evaluation of the optimality versus linearity contrast calls for different underlying methods of analysis, and is explored in more detail in later chapters.

Interactive or configural approaches to personality research are also non-compensatory. That is, a high score on one personality scale cannot compensate for, or mask, a low score on another scale, and vice versa. This is an important consideration in personality measurement; each scale of the FFM measures unique attributes and characteristics of a person. Non-compensatory strategies attempt to capture each of these attributes, and allow for the isolation of those traits or scales that are most important, whereas additive or compensatory strategies rely on methods that allow relative strength in one area to compensate for potential weakness in another. Consider the following example.

Assume that both Emotional Stability and Conscientiousness are found to predict performance positively in emergency medicine. A compensatory approach assumes that Applicant A with a normative score of 0 on Emotional Stability and 100 on Conscientiousness is as qualified as Applicant B with a normative score of 50 on each scale, even though scale-level research tells us that these individuals would behave differently on the job. Weighting methods can be used to offset this assumption, but the underlying methods are still additive, and consider scores in a compensatory fashion. On the other hand, non-compensatory selection strategies evaluate each scale independently, and require applicants to meet a particular threshold on each relevant scale to proceed to the next stage in a selection battery. The non-compensatory distinction is important to this study, and is discussed further in later chapters.

When considered in tandem, these two attributes of the current study—non-compensatoriness and optimality—represent a departure from mainstream personality profiling research methods. Consider some examples of where current configural analysis methods lie with respect to this duality. One of the most common metrics used to evaluate profile similarity, the *generalized distance function* (D^2; Osgood & Suci, 1952; Cronbach & Gleser, 1953), is optimality-based, but compensatory. Individuals are evaluated with respect to their distance from a mean (rather than the linear magnitude of their score), but differences on one scale may be masked by similarities on another. Traditional regression analyses are linear and compensatory; higher (or lower) scores increase R^2, and scale-by-scale discrepancies are combined. Analyses of trait-by-trait interactions are linear and non-compensatory; the magnitude of scores is the focus, rather than distances from a particular mean or profile point, while methods for evaluating interactions to not mask differences on one scale with similarities on another. When considering the duality in this way, it becomes clear that a non-compensatory, optimality-based methodology represents a new alternative for measuring personality at the profile level.

The present study attempts to (1) identify optimality-based personality profiles within and across several data samples, (2) assess the generalizability of those profiles within and across distinct industries and job levels, and (3) examine, via a non-compensatory fit score, the incremental validity of those profiles for predicting performance beyond existing predictive models. Several researchers have advocated methods that account for potential interactions between personality traits, or a configural or constellation approach (Kostman, 2004; Perry, Hunter, Witt, & Harris, 2010; Witt, 2002; Witt et al., 2002). To date, however, these approaches have focused primarily on identifying moderated relations among very few traits (usually, just two). In contrast, this study seeks to identify personality profiles in terms of recurring patterns of scores involving *all* traits within a specified taxonomy. The main research questions addressed in this study are elaborated next, with more specific hypotheses offered in a later section.

Main Research Questions

(1) Can personality profiles be identified and extracted from a typical work-oriented FFM measure? If so, how many distinct profiles emerge?

(2) How generalizable are personality profiles between samples drawn from within as well as across industries and job types? That is, do profiles demonstrate convergent and discriminant generalizability?

   a) How generalizable are profiles identified in a sample from one industry and job when applied to other samples from the same industry and job?

   b) How generalizable are profiles identified in a sample from one industry and job when applied to different industries and jobs?

(3) What is the validity of profile membership, and the incremental validity of profile membership (assessed along a non-compensatory continuum of fit) over

   a) the linear weighted sum of multiple scale predictors (i.e., in terms of multiple R)?

   b) existing compensatory indices of fit?

The third question addresses directly whether multivariate prediction using personality trait measures is most effectively undertaken by way of an additive (i.e., compensatory) or configural (non-compensatory) profile model.

Potential Benefits of the Current Research

The current study was designed to offer insights into how personality plays out in applied settings, which, in turn, could inform the field on ways to make better use of personality data. It was intended to extend research on configural interpretations from investigations of individual trait relationships to a more holistic profile approach. This could improve the correspondence between how I/O research often conceptualizes personality versus how people express or experience personality, shedding new light on how to create interpretations at the profile level while also helping to identify the limitations, if any, of profile generalizability.

The extent to which profiles generalize speaks to three key points. First, whether profiles demonstrate convergence with other theoretically similar profiles and divergence with theoretically dissimilar profiles helps answer the practical question of whether local profiling is necessary if practitioners attempt to apply this technique in the field. This is analogous to the concepts of local versus generalized norms, or situational specificity in validation, and can help researchers and practitioners design future studies or select appropriate methods in the use of profiling going forward. Much like research on the validity of personality scales progressed from assumptions of situational specificity to questions of generalizability, the emergence of coherent profiles of personality will raise questions of whether those profiles are robust descriptions across situations, or are specific representations within a job or industry. Second, generalizability, or the lack thereof, informs the conceptual question of how to think about and talk about personality at the profile level. Where profiles generalize, new frameworks may become available with which to understand and describe people and behaviors (e.g., workaholic or micromanager). Where situational specificity of profiles is supported, the use of profiles as a sense-making tool across situations becomes limited, but our understanding of how profiles are manifested in individuals and groups is enhanced. Generalizable profiles (even across only one variable, like job level) have implications for other related lines of research as well. For example, researchers interested in Person-Organization (P-O) fit, group cohesion, team performance, job satisfaction, withdrawal behaviors, or other organizational-level constructs may benefit from examining the extent to which personality profiles relate to other variables of interest within these particular literature domains.

Finally, the current study explores a relatively emergent method for identifying profiles in a given sample (Latent Profile Analysis, or LPA, described in later chapters). Profiles extracted using LPA are used to (a) test a uniquely non-compensatory approach to profile fit, (b) examine whether such a non-compensatory fix index relates to job performance, and (c) determine if the noted fit index provides incremental validity in predicting job performance relative to traditional compensatory metrics of D^2 and r_{pf},[2] which are fundamentally at odds with the interactive, multiplicative nature of profiles.

In the most general sense, this research seeks to determine whether coherent personality profiles naturally emerge from a set of samples drawn from a working adult population, and if so, the extent to which they are psychologically meaningful and/or useful. The current study is a first attempt to explore some key issues surrounding new ways to conceptualize and measure personality at a profile level. Little research has been

[2] Throughout this document, r_{pf} is used to distinguish correlational methods for evaluating profile fit (*pf*) from other correlations reported within.

conducted on the FFM via LPA, and few researchers have taken a purely non-compensatory approach to profile measurement. As such, this study is primarily exploratory in nature, with the aim of laying the groundwork for future confirmatory work on personality profiling in work settings.

CHAPTER 2

LITERATURE REVIEW

This chapter provides foundational background on the theories and prior research that inform the targeted research questions regarding personality profiles in the workplace. Literature is reviewed in sections regarding what constructs are being measured, ways those constructs can be used to make decisions about people, how they have been measured up to the present, and the ways that the current study differs from previous related works.

A Brief Overview of the FFM

Personality assessment research began in earnest in 1917 with the development of the Woodworth Personal Data Sheet (Gibby & Zickar, 2008). The predominant concern guiding much of personality research since that time has been what to measure. Historically, what was measured depended on practical needs (e.g., identifying psychopathology, as in the case of the *Minnesota Multiphasic Personality Inventory*; Hathaway & McKinley, 1943), or researchers' theoretical interests (e.g., Locus of Control; Rotter, 1966; *Myers-Briggs Type Indicator*; Myers, McCaulley, Quenk, & Hammer, 1998; *Thematic Apperception Test*; Morgan & Murray, 1935). Hogan and Smither (2001) note that many prominent personality theorists' research foundations reflected their own personal interests and histories. Multidimensional personality inventories emerged in the 1930s (e.g., the *Bernreuter Personality Inventory*; Bernreuter, 1931). Early efforts attempted to measure traits or dispositions. Allport (1961) defined a trait as "a generalized neuropsychic structure (peculiar to the individual), with the capacity to render many stimuli functionally equivalent, and to initiate and guide consistent (equivalent) forms of adaptive and stylistic behavior" (p. 363). With few exceptions (see notably Hogan, 1996, on the distinction between traits and characteristics), this conceptualization of a trait, and the focus on traits as the unit of measurement for personality, has persisted since that time.

Today, most current thinking in personality assessment converges on the idea that five personality dimensions describe the critical and consistent aspects of normal adult personality. Decades of research on the FFM (cf. Digman, 1990; Goldberg, 1992; John, 1990; McCrae & Costa, 1987; Norman, 1963; Thurstone, 1934; Tupes & Christal, 1961) suggest that personality can be described and conceptualized in terms of five major themes, as defined in Table 1.

Table 1: Scales and Definitions for the Five Factor Model

I.	Extraversion: the degree to which a person is outgoing, talkative, and seeks stimulation.
II.	Agreeableness: the degree to which a person is friendly, pleasant, and values getting along with others.
III.	Conscientiousness: the degree to which a person is self-disciplined and complies with rules, norms, and standards.
IV.	Neuroticism: the degree to which a person is vulnerable to stress and prone to experiencing negative emotions (opposite pole = Emotional Stability).
V.	Openness to Experience: the degree to which a person is creative, curious, and eager to learn.

The FFM is a taxonomy of normal personality traits (i.e., typical behavior and beliefs), offering a simple and understandable structure for describing people. The FFM mirrors the structure of observer ratings (Norman, 1963; Thurstone, 1934; Tupes & Christal, 1961) and, as the lexical approach holds, common terms that people use to describe themselves and others (Saucier & Goldberg, 1996a; for detailed reviews of the lexical approach, see John, Angleitner, & Ostendorf, 1988; John, 1990; Saucier & Goldberg, 1996b). When applied to someone's behavior, these words and phrases comprise an individual's personality as other people see it (cf. Argyris, 1957; Hogan, 1983;

Lundholm, 1940; Mead, 1934). Implicit in these descriptions is also an evaluative component that depicts how people perceive and judge others (Buss, 1991).

Many prominent personality inventories constructed in recent decades are based on the FFM (e.g., *NEO-PI*: Costa & McCrae, 1985; *Hogan Personality Inventory*: Hogan & Hogan, 1995; *Personal Characteristics Inventory*: Mount & Barrick, 2001). Evidence suggests that all existing multidimensional personality inventories can be mapped more or less completely onto these five dimensions (Wiggins & Pincus, 1992). One notable exception to the consensus around the FFM has been recent work on the HEXACO model (Ashton & Lee, 2007), though early results investigating its predictive power have been mixed (Ashton & Lee, 2010; De Raad et al., 2010). Consequently, the FFM remains the dominant paradigm for modern research in normal personality assessment today (Hogan & Hogan, 1995; Ones et al., 2007; Tett & Christiansen, 2007).

Challenges to Personality Assessment

The viability of personality assessment has ignited debate on at least three occasions. As discussed previously, Mischel (1968) and others responded to the proliferation of personality tests of the 1950s and 1960s by arguing that people's behavior reflected situational factors, not enduring personality characteristics (see also Dunnette, 1962; Guion & Gottier, 1965). Mischel asserted that the personality characteristics being identified and studied at the time were, in fact, not traits, but rather simple stereotypes or heuristics that people use to help understand others, which has little to do with actually explaining behavior. These criticisms reverberated through the personality and applied psychological fields for over two decades.

Meta-analyses in the 1980s and early 1990s helped personality research emerge from obscurity, and meta-analyses based on the FFM have improved understanding of the nature of personality prediction. The advent of meta-analysis helped usher personality assessment into the workplace, and demonstrated that personality tools can be viable decision-making aids in employment decisions. Barrick and Mount's (1991) FFM meta-analysis supported Conscientiousness as a valid predictor in work settings across occupational groups and performance criteria. Extraversion demonstrated useful validity for occupations that require social interaction (e.g., sales), and Extraversion and Openness were valid predictors of training proficiency. A few months later, Tett, Jackson, and Rothstein (1991) reported similar validities for the FFM. The estimates were somewhat stronger on the whole, owing to reliance on confirmatory strategies, including the use of job analysis to identify job-relevant traits. Salgado (1997) replicated Barrick and Mount's results, with the exception of Emotional Stability, using data from the

European community. Then, in 1999, Tett et al. used refined meta-analytic methods (cf. Ones, Mount, Barrick, & Hunter, 1994; Tett, et al., 1994) that account for bidirectionality in personality-performance relationships (i.e., that a given trait can predict performance positively under some conditions and negatively under others) to show that uncorrected mean personality test validity strength (for individual trait measures) is $r = .26$ (fully corrected mean $r = .38$), when based on confirmatory studies using job analysis.

During this resurgence in popularity, personality research came to face a second challenge. Unlike Mischel's questioning of the basic existence of personality traits, subsequent concerns focused on "protecting" the unknowing public from the duplicitous claims of personality test publishers, and on creating concerns in HR managers by using faulty information on laws, guidelines, and case law surrounding assessment usage (Arthur, Woehr, & Graziano, 2001; Ruiz, 2006). Paul (2004), in her "*Cult of Personality*," condemned personality assessment as an unregulated discipline that is often invalid, unreliable, and unfair. Her work targeted an audience of laypeople rather than academics or professional test users, and warned the public against relying on the results of personality tests to make decisions, or to attempt to understand someone, until additional research could unequivocally support the tools over time and across situations.

Such concerns about personality assessment may stem from differences that sometimes emerge between an assessment's intended purpose and its application. The MMPI (Hathaway & McKinley, 1942) provides a relevant example. Hathaway and McKinley designed the MMPI to help diagnose psychiatric patients based on an empirical approach comparing psychiatric patients' and the general public's responses on MMPI questions. Yet, by the 1960s, the MMPI appeared in the workplace as well as in clinical settings (Thumin, 1969, 1971), particularly for positions where employers believed that "maturity, emotional stability, and the capacity to act responsibly under stress [were] clearly requisite to satisfactory performance" (Thumin, 2002, p. 73). Recently, the courts have found that, without documented evidence showing that a psychiatric diagnosis is job related and consistent with business necessity, the use of the MMPI for job selection violates the Americans with Disabilities Act of 1990 (see *Karraker v. Rent-a-Center, Inc.*, 2005). Yet organizations continue to use the MMPI to screen applicants for public and private sector jobs, even in situations where no evidence exists that the constructs measured by the MMPI are job-relevant (Pearson Assessments, 2011).

Similarly, the Myers Briggs Type Indicator (MBTI: Myers et al., 1998) emerged from a desire to help people understand more about themselves, rather than as a tool for making employment-related decisions. Based on Jung's (1921) theory of personality, the MBTI classifies people into one of 16 personality types. Research has shown mediocre reliability and validity of the tool in employment settings (cf. McCrae & Costa, 1989, Pittenger, 1993; Hunsley, Lee, & Wood, 2004). Moreover, the MBTI manual itself, and the current test publisher, both expressly discourage the use of the tool as a predictor of job success (Myers et al., 1998; Thompson, 2013); yet it remains the most popular personality test in the world, with more than two million people completing it each year (CPP, Inc., 2011).

With usage volumes of this magnitude from one assessment alone, there is clearly a hunger for personality assessment in the general marketplace. Presently, personality assessment affects the lives of thousands of people every day. A survey commissioned by

Rocket Hire (Healy & Handler, 2009) estimated that over half of all U.S. companies use personality tools. With revenues exceeding $2 billion annually and growing at 15% a year, personality tools appear irrevocably intertwined into organizational life (Tahmincioglu, 2011).

Yet even more recently, a third line of criticism has been leveled at the field, which has echoed the decades-old argument (see Schmitt et al., 1984) that the validity of personality assessment is simply insufficient for predicting behavior or making personnel decisions. As described in Chapter 1, Morgeson et al. (2007) cited the low reported validity of personality tests in their call to abandon personality assessment in selection contexts. Researchers now seem to be in general agreement that the FFM offers a robust organizing framework for personality traits; the issue of contention is now one of validity—both statistical and practical—and more specifically, how to improve it (or some might say, how to uncover better estimates of true validity). The field has responded to this call aggressively, both by highlighting alternative methodologies, some of which Morgeson and colleagues overlooked (cf. Bartram, 2005; Hogan & Holland, 2003; Holland & Van Landuyt, 2007, Ones et al., 2007; Tett & Christiansen, 2007), and by considering alternative, and often innovative, methodological and data collection approaches (Connelly & Ones, 2010; Oh, Wang, & Mount, 2011). The current study extends this second line of research.

Factors Affecting the Criterion Validity of Personality Assessment

Recent studies have recognized that novel methodological approaches might improve the potential of personality measures for use in predicting performance criteria. For example, researchers have explored the tradeoff between broad versus narrow trait measures (Ashton, 1998; Bergner, Neubauer, & Kreuzthaler, 2010; Chapman, 2007; Christiansen & Robie, 2011; Holland, 2001; Ones & Viswesvaran, 1996; Paunonen, Rothstein, & Jackson, 1999; Salgado, 1997) and the value of using narrow versus broad measures on both the predictor and criterion sides of the equation (Tett et al., 2003). All told, the literature in this area supports the use of narrow over broad measures (Rothstein & Goffin, 2006; Tett & Christiansen, 2007).

Further efforts to apply theory to improve trait-performance linkages extend Tett et al.'s (1991, 1999) meta-analytic support for confirmatory over exploratory research strategies. Specifically, construct alignment is predicated on the assumption that particular personality constructs will best predict rationally and/or theoretically aligned criteria (Wernimont & Campbell, 1968; Campbell, 1990). Hogan and Holland (2003) attempted to overcome limitations of previous studies by evaluating the validity of FFM scales from a single measure of personality (i.e., Hogan Personality Inventory; Hogan & Hogan, 1995) and also by aligning predictors and criteria conceptually. They reported corrected criterion-aligned validities ranging from $\rho = .25$ (Learning Approach) to $\rho = .43$ (Adjustment). Across studies, their results supported the predictive value of Conscientiousness, Emotional Stability, and Agreeableness, and demonstrated the convergent and divergent nature of personality-based prediction. Hogan and Holland's results clearly support the value of confirmatory strategies in personality test validation.

Several studies have also explored situational moderators of the personality–performance relationship (Penney, David, & Witt, 2011). In fact, Mischel himself suggested decades ago that a link between a given trait and a behavior would only manifest in “weak” rather than “strong” situations (1977). Current research suggests, beyond situation strength, that situations vary in the degree they cue or “activate” the expression of a given trait. For example, Trait Activation Theory (TAT; Tett & Burnett, 2003; Tett & Guterman, 2000; see also Tett & Murphy, 2002) is founded on the premise that variance in trait expression (e.g., as high or low on Dominance) can be expected only in situations containing trait-relevant cues (e.g., opportunities to direct others). Moreover, expressing one’s traits is intrinsically rewarding and others’ evaluation of that expression (e.g., as job performance) can lead to extrinsic motivation. Similarly, other researchers have contended that possessing one individual difference characteristic can activate or trigger the expression of another. For example, Perry, Hunter, Witt, and Harris (2010) found that the Achievement facet of Conscientiousness can trigger people to express GMA in measures of task performance. Based on the theory that humans are naturally motivated to conserve their personal resources, the authors hypothesize that high levels of Achievement trigger individuals to expend their valued resources to express GMA.

Compensatory versus Non-compensatory Approaches

Many present-day approaches for both researching and utilizing personality assessments in the workplace focus on either single scale validity (e.g., using *r*) or additive methods (e.g., using *R*). The majority of current research (much of it meta-analytic) still reports a validity coefficient for each scale of the FFM (e.g., Morgeson et al., 2007). Many researchers then proceed to combine scales via regression techniques to arrive at a single coefficient that can then be used to determine the amount of variance accounted for by a given assessment. Likewise, practitioners frequently use regression to arrive at a single coefficient that can be used to make employment decisions.

A configural or profile approach differs from an additive model, such as that afforded by regression-based analyses, in that the derivation of a profile is a non-compensatory procedure—that is, it does not assume that a given score on one personality trait can compensate or be exchanged for a given score on a different trait. This consideration is relevant when considering the application of personality instruments in employee selection. A compensatory approach is employed when practitioners use regression-based analyses to assign weights to given assessments and scales, and then create selection systems that combine applicant scale scores (within and/or across instruments). In contrast, non-compensatory recommendations set cutoff scores on each job-relevant trait measure—cutoffs that an applicant must meet or exceed simultaneously to be selected outright or to progress to the next hurdle in the selection battery.

Compensatory cutoff models assume that an applicant's relative strength on one or more scales, or on one of many assessments, can help overcome weakness on another scale or assessment—and this is sometimes true. For example, the Watson-Glaser Critical

Thinking Appraisal (Watson & Glaser, 1980) measures, scores, and generates feedback reports across five areas of ability, but the manual recommends that only the total score be used to make employment decisions. This indicates that an applicant who scores poorly on the Inference subscale, for example, might compensate for this weakness by scoring well on another subscale or combination of subscales. This approach is common among cognitive tools, which are frequently designed to measure GMA rather than factor-level cognitive functioning.

The cutoff score literature often implies that a single recommendation across a compensatory battery is the best indicator of job performance (e.g., Cascio, Alexander, & Barrett, 1988; Martin & Raju, 1992). However, in contrast to cognitive tools like the Watson-Glaser, FFM personality assessments are not designed to measure a superfactor. Thus, non-compensatory cutoff scores, which set separate recommendations on each scale supported by job analysis and/or validation research, may be better suited to ensure that a given set of recommendations will account for (a) each trait that is associated with job performance (Hogan & Holland, 2003), and (b) the direction of the relationship between that trait and performance (Tett et al., 1994; 1999). The multi-dimensional nature of personality assessments suggests that practitioners should consider the possibility that a non-compensatory approach to modeling personality and making decisions with personality data is worth exploring.

The choice to use a compensatory or non-compensatory approach to developing and implementing recommendations for pass/fail decisions is a relatively unexplored issue in the literature. Although it has not yet been shown to directly influence an assessment's accuracy (Holland & Van Landuyt, 2007), Hogan and Holland's work on construct alignment (2003; discussed in more detail in the next section), supports the assertion that practitioners could maximize the accuracy of hiring decisions by adopting a non-compensatory approach. The current study aligns with this thinking as well; conducting analyses at the profile level implies that specific levels of *each* trait in an assessment will add value beyond an additive model where one trait is assumed to compensate for another. To be clear, this study does not venture into how a practitioner might set a cutoff score or make pass/fail recommendations at the profile level. However, the concept of profile generation is, by nature, non-compensatory, in that one must demonstrate similarity across a number of scales simultaneously to be classified into one profile over another. It is the interactive, multiplicative nature of profiles that distinguish them from traditional compensatory, additive correlation and regression approaches, and this study explores profile fit using appropriately multiplicative fit methods.

Trait Interactions and Personality Profiling

Most research on interactions between variables has examined the interaction of personality traits with non-personality constructs such as goal setting (Barrick, Mount, & Strauss, 1993) or GMA (Sackett, Gruys, & Ellingson, 1998; Ferris, Witt, & Hochwarter, 2001; Mõttus, 2006). However, recent research has taken these ideas a step further by exploring specific interactions among multiple distinct personality constructs, with researchers acknowledging that the way in which personality operates can depend on the pattern of other constructs *within* a personality profile (cf. Burke & Witt 2002; 2004; Foster & Macan, 2006, Hochwarter, Witt, Treadway, & Ferris, 2006; Hogan, Hogan, & Roberts, 1996; Hotard, McFatter, McWhirter, & Stegall, 1989; Perry, Dubin, & Witt, 2010; Perry, Hunter, Witt, & Harris, 2010). Witt et al. (2002) note that, "certain personality traits may interact with others to result in desirable, as well as undesirable, workplace behaviors" (p. 164). They found that Agreeableness moderates the relationship between Conscientiousness and job performance, such that managers high in Conscientiousness are rated highly if also high on Agreeableness, but are rated low on performance if low on Agreeableness. Thus, whether being high on Conscientiousness is valued positively or negatively depends on joint consideration of Agreeableness. This advances our understanding of the value of Conscientiousness in the workplace beyond that afforded by the far simpler notion that Conscientiousness is a universally positively valued trait (e.g., Barrick & Mount, 1991).

Further research suggests this is not an isolated phenomenon. Burke and Witt (2002) found that Extraversion and Emotional Stability each moderate the relationship between Openness to Experience and performance. Specifically, Openness contributes to performance (positively) only in Extroverts and/or those low on Emotional Stability; for Introverts and/or emotionally stable workers, Openness is unrelated to performance. Studies like this have broken new ground, given that, previously, Openness was considered a generally weak or inconsistent predictor of performance (e.g., Barrick & Mount, 1991).

Burke and Witt (2004) found that a combination of high Conscientiousness and low Agreeableness resulted in what the authors call, "High Maintenance Employee Behavior" or HMB. Like the other work described here, the authors were exploring interactions between two specific variables. However, Burke and Witt also examined a particular configuration of traits in relation to a heuristic, or narrative description, which goes beyond simple trait explanation to advance and promote sense-making when discussing personality and its correlates.

This body of research on trait interactions has helped confirm that the FFM is a valuable framework for conceptualizing and measuring normal personality, and moreover, that looking beyond relationships between single factors and overall job performance can advance our understanding of personality and its correlates. The current study takes this line of research a step further by exploring whether conceptualizing personality in terms of profiles, rather than in terms of just one or two traits, can identify interpretable personality characteristics in individuals, and demonstrate incremental validity in predicting job performance beyond single-scale prediction and multi-scale compensatory models (i.e., regression). Although researchers have explored interactions between selected personality *traits* in various ways, the potential merits of personality profiles, which consider the joint configurations of multiple traits *simultaneously*, remain largely unknown. Further, whereas interaction-based analyses are based on a *linear* continuum, analyses at the profile level carry an assumption of *optimality*. That is, profile similarity is assessed with regard to the distance between the individual's scores on each of several trait measures from their corresponding profile points (i.e., means), rather than the relative extent to which they are high or low. Thus, falling above a profile point lowers fit as much as falling below it.

Distinguishing the Profile Paradigm

The current study assigns individuals to given profiles via a fit score, effectively giving each individual a label or category that indicates the profile with which he or she is best aligned. However, this approach is distinct from the outcomes or analyses associated with measures of typology such as the MBTI (Myers et al., 1998). The underlying measure intended for use in this study is normative. That is, it provides scores for each personality dimension across a continuum, benchmarked against a large database of participants. By contrast, the MBTI and similar typology measures are ipsative tools. They compare individuals to themselves (e.g., an individual is more introverted than they are extroverted), but do not allow for comparisons between people on a normative or absolute scale. For this reason, ipsative tools are generally considered inappropriate for making decisions about people or identifying quantifiable similarities and differences among participants.

Likewise, the current research is distinct from most compound scoring efforts (see Frei & McDaniel, 1998; Holland, Hogan, & Van Landuyt, 2002; Ones, Viswesvaran, & Schmidt, 1993), in which researchers begin with a given construct (e.g., Customer Service; Integrity), and then, either rationally or empirically, create a customized scale or set of scales designed to align with the construct in question. In contrast, the present study begins with an exploratory approach designed to identify natural, psychologically identifiable profiles that are latent within a given sample, and then determine whether those profiles that emerge (a) demonstrate generalizability and (b) relate meaningfully to non-test behavior. Further details about the methods are discussed in later sections.

The current study also differs from profile-level research on P-O fit and variations thereof (e.g., person-job fit; person-workgroup fit). The personality profiles in this study are not derived through comparison to mean organizational or team profiles, etc., as is generally the procedure for studies focusing on P-O fit. Rather, the profiles here are derived from an exploratory analysis of the data, using a method that derives profiles from latent underlying variables.

The present research is also distinct from the current body of literature on profile matching in selection contexts, which, in general, has met with limited success (Edwards, 1993). Profile matching of this nature entails comparing applicants to an ideal target profile developed most often from job analysis or from the mean profiles of successful incumbents (Peters, Greer, & Youngblood, 1998). Peters et al. note that measuring fit in this context is difficult, given that, (a) both the magnitude of scores and overall shape of a profile can differ, and, (b) averaging differences across dimensions in a compensatory fashion can inflate measures of similarity, obscuring important deviations from a given profile.

Finally, the current research differs from efforts to identify code types in the clinical field, e.g., via scores on the MMPI (Hathaway & McKinley, 1943). Graham (2012) notes that Hathaway and McKinley conceptualized and promoted configural scoring of the MMPI from its conception. Configural MMPI scoring began by assigning a code to an individual based on his or her highest scale scores (hence the term "code type"). Efforts to create homogeneous groups of people based on configural scoring soon turned to more complex methods and algorithms, but these efforts were soon abandoned when researchers saw that (a) few individuals would be classifiable via these more intricate methods, and (b) the reliability of these classification systems over time suffered as a result of their specificity (Ben-Porath, 2006). Code typing today relies on identifying types via the highest one, two, or three scales on the MMPI. It is interesting to note that the absolute magnitude of the scales does not matter (which distinguishes the code type method from the present study). However, according to most researchers, the extent to which a code type's scales are "defined," i.e., sufficiently distant from other scale scores, does influence whether a particular code type should be interpreted or not, in that it affects the extent to which a code type is reliable over time, and has meaningful behavioral correlates (Graham, 2012; McNulty, Ben-Porath, & Graham, 1998).

A particularly serious concern with some of these past efforts is that fit with a given profile in such models has been routinely assessed using compensatory indices that are incompatible with a true profile approach. Describing this problem in more detail, Nunnally and Bernstein (1994) note that measures of profile similarity that simultaneously account for level (distance from the mean across scales), dispersion (how widely the points of each scale in a profile are distributed from each other), and shape (the rank order of scale scores) do not actually lend themselves well to mathematical analysis. However, the current study offers an alternative, emerging method of assigning people to profiles that retains the essentially interactive quality of profiles. The methods are based on Latent Profile Analysis (LPA) via Mplus statistical software (Muthén & Muthén, 2011). LPA, at least in its current form, is still in a relative infancy, having been facilitated by advanced technology and the development of the Mplus program in 1998. Further details on the profile classification method used in this study are offered in the next chapter.

Profile Measurement and Indices of Fit

Most previous studies that have attempted to analyze or match personality data at the profile level have used limited methods to establish profile similarity (Edwards, 1993; Cheung, 2009). As discussed in Chapter 1, many common measures of fit are either linear, compensatory, or both; an optimality-based non-compensatory strategy for evaluating fit has yet to emerge in mainstream research. Some of the most common methods for evaluating profile fit are discussed next.

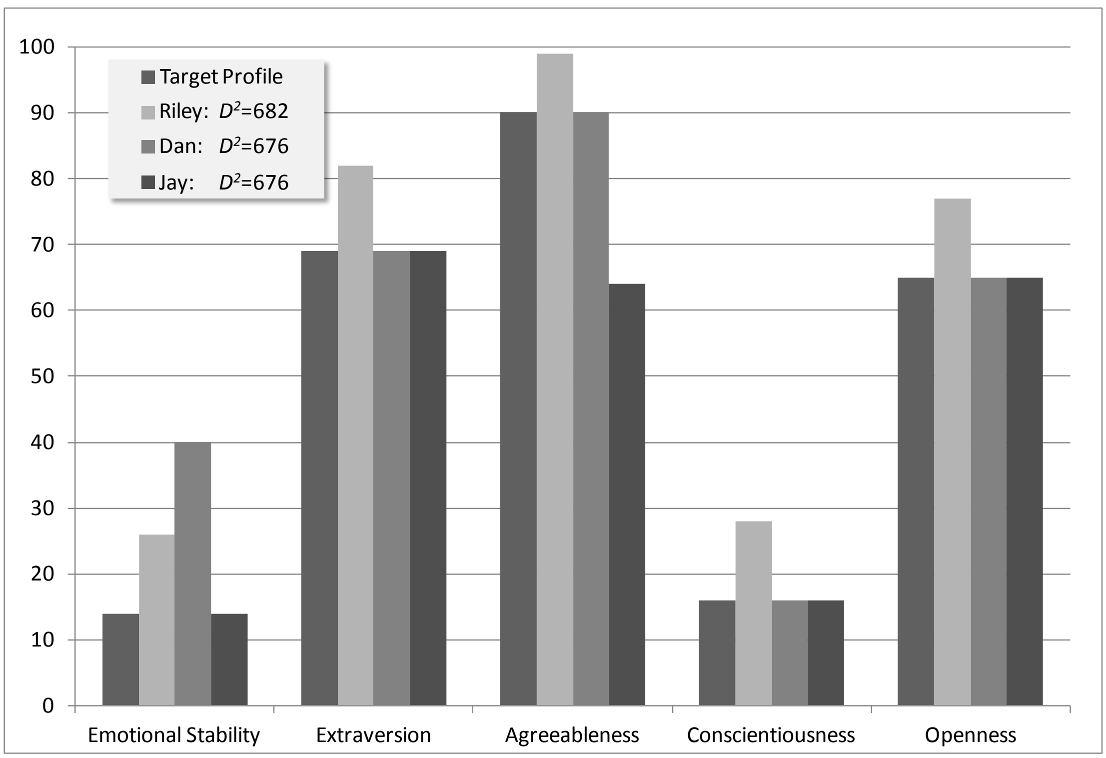
The Generalized Distance Function

Early work on congruence often relied on D^2 (Osgood & Suci, 1952; Cronbach & Gleser, 1953) to establish similarity between profiles (e.g., Drazin & Van de Ven, 1985). The formula for D^2 is

$$D^2 = (X_{a1} - X_{b1})^2 + (X_{a2} - X_{b2})^2 + \cdots + (X_{ak} - X_{bk})^2 = \Sigma\left(X_{aj1} - X_{bj}\right)^2$$

D^2 measures profile similarity with regard to level and dispersion, and allows for simple rank ordering of individuals, or between an individual and a given target profile. However, D^2 is nondirectional, and also combines differences across constructs in its calculation; like regression, D^2 is compensatory. Regression additively combines assessment scales, assuming that a participant's relative strength on one or more scales, or on one of many assessments, compensates for weakness on another scale or assessment. Similarly, D^2 assumes that differences across constructs in a given profile are additive and conceptually equal to one another (i.e., a better fit with one point on the target profile can mask, or compensate for, poorer fit with other profile points). Figure 1 shows an example of this phenomenon. The target profile is shown in blue. Riley exceeds the target by a small amount on each scale. Dan and Jay both match the target profile exactly, save for one scale each. Notably, the D^2 values for all cases are nearly identical. Yet you can see that Dan's Emotional Stability score is nearly double that of Jay, and Jay and Riley are a full 35 points apart on Agreeableness. D^2 results suggest these cases are highly similar, and that Dan and Jay are actually identical. Yet, what we know about the FFM and how each factor relates to behavior would suggest that these individuals are not likely to behave similarly.

Figure 1. Hypothetical Profiles with Similar D^2 Values



Correlational Methods

Correlations (r_{pf}) between profiles have often been used to establish congruence (e.g., Caldwell & O'Reilly, 1990). A correlational method of profile measurement is not easily categorized into the optimality/linear duality discussed earlier; rather it is best interpreted simply as the similarity between the rank order of scores on each scale in a profile. In fact, profile similarity with respect to r_{pf} is defined in terms of the similarity only between the *pattern* of scores common to the two profiles (Furr, 2008), rather than similarity between the patterns *and* magnitudes of scores. It is easy to imagine, however, a situation in which one person is relatively low on all scales and another relatively high, and yet the correlation between their patterns of scores (or the shape of their peaks and valleys) is strong. In terms of overall profile height (and the associated interpretation carried with their given scores), these two individuals would not be considered similar; but calculating the correlation between their profiles would lead a researcher to classify them as such. Moreover, correlational analyses are similar to D^2 in that it is a non-compensatory strategy; differences on one scale can be masked by similarities on other scales. The formula for computing a correlation makes this distinction clear:

$$r_{xy} = \frac{\sum z_x z_y}{(N-1)}$$

It is because the standardized scores in this formula are summed in the numerator that correlational analyses are compensatory. [3] Multiplicative methods, conversely, are indicative of non-compensatory metrics.

[3] This also explains how outliers can have a disproportionate impact on the magnitude of a correlation coefficient.

Manual and Judgment-based Methods

More recently, methods of analysis have emerged that allow researchers to measure and account for a pattern of scores from participants that includes the magnitude of each score, rather than just the size of an aggregated difference or rank order of scale scores. Using a simple median split technique, participants can be classified manually into groups based on whether they fall above or below the median on a variable or set of variables, and then differences among profiles in outcome variables are examined using ANOVA. This method is similar to the MMPI code-type technique discussed previously, in that the classification of individuals is a relatively manual task for the researcher, rather than relying on complex algorithms. However, there are limitations associated with the median split technique (Maxwell & Delaney, 1993; MacCallum, Zhang, Preacher, & Rucker, 2002). Many researchers' concerns with the method revolve around the dichotomization of continuous variables, defining a median within a given sample, and determining whether the resulting profile assignments result in sufficiently homogenous groups. The latter concerns can sometimes be alleviated by using large-sample norms from which medians are derived, or splitting each scale into more than two segments. Even so, no standardized method or significance test exists for using this technique to identify profiles, beyond a researcher’s own judgment and decision rules. As a result, profiles may be forced from a given dataset whether they exist or not.

Cluster analysis (Aldenderfer & Blashfield, 1984; Kaufman and Rousseeuw, 2005) is an exploratory technique used to divide participants into homogeneous subgroups based on multiple variables. It attempts to classify individuals into groups or clusters so as to minimize differences within clusters and maximize differences between clusters. Several methods for defining clusters and measuring distance are available, depending on data structure and the hypotheses in question. However, like the median split technique, a limitation of cluster analysis is that it tends to rely heavily on the researcher's judgment for final cluster creation and interpretation (Nunnally & Bernstein, 1994; Milligan & Cooper, 1995; Pastor, Barron, Miller, & Davis, 2006). That is, software packages generally output either a fixed number of clusters based on the user's input or a taxonomic visual output for a user to review and interpret, but individuals are assigned to those clusters without regard for whether any underlying structures or latent variables exist within the data.

Latent Profile Analysis

Latent Profile Analysis (LPA; Lazarsfeld & Henry, 1968) lends itself particularly well to meeting the aims of this study. LPA is a particular application of mixture modeling, used to find groups or subtypes of cases in multivariate data. LPA is used to determine the existence of homogeneous groups of individuals within a heterogeneous sample based on underlying patterns in a set of data (Magidson & Vermunt, 2004). LPA analyses are not constrained by assumptions of linearity, homogeneity of variances, or data being normally distributed (Magidson & Vermunt, 2002).

Like cluster analysis, LPA identifies similar groups of people (i.e., profiles) based on common characteristics (e.g., continuous FFM scores). However, in contrast to traditional cluster analytic techniques, LPA is model-based, whereas common applications of cluster analysis are not (Pastor et al., 2006). That is, where cluster analysis creates profiles simply by determining the distance between cases, LPA creates profiles based on the assumption that individuals respond similarly to one another due to an underlying latent trait. Researchers have noted that, in comparison to cluster analysis, mixture modeling techniques like LPA are the preferable device for exposing any latent grouping that may underlie a particular dataset (McLachlan & Chang, 2004; McLachlan & Peel, 2000).

In LPA, it is assumed that a mixture of underlying multivariate normal probability distributions generates the data, and the probability that an individual is classified into a given profile (referred to in this study as a "fit score") is estimated simultaneously with the overall model (e.g., the number of latent profiles identified within the data). The algorithms underlying LPA calculations use the maximum likelihood method for parameter estimation, which means the creation of profiles involves finding the means and variances of the multivariate normal distribution for each profile that maximize the probability of the observed cases being assigned to a given profile. LPA also differs practically from cluster analysis, in that it allows for a mixture of continuous and categorical variables without requiring that each variable be transformed or standardized prior to analysis (Muthén & Muthén, 2011).

Coherent profiles have emerged from LPA analyses in recent studies of business and marketing (Bassi, 2011; Magidson & Vermunt, 2002), clinical and consulting psychology (Herman, Ostrander, Walkup, Silva, & March, 2007; Ostrander, Herman, Sikorski, Mascendaro, & Lambert, 2008), and, more importantly for the current study, in mental health studies designed to derive clinical profiles from the FFM. In particular, Merz and Roesch (2011) indentified three distinct FFM profiles in a population of college students, which they labeled "well-adjusted," "reserved," and "excitable." In support of their uniqueness, profile scores related in different ways to positive and negative affect, self-esteem, depression, anxiety, and coping efficacy. This suggests that LPA can be used to effectively model commonalities among personality variables.

Research Questions and Hypotheses

The current study tests the following research questions and corresponding hypotheses. These questions are best interpreted in light of a 2 x 2 array of samples. Specifically, three independent samples offering data on the same personality instrument, the *Hogan Personality Inventory* (HPI: Hogan & Hogan, 1995), are nested within each of four cells defined by crossing job level (i.e., managerial versus non-managerial) with industry type (i.e., customer service versus sales). Thus, all told, the research questions are pursued using 12 independent samples (*N* range = 103 to 253).

(1) Within each of the 12 samples, do personality profiles emerge that are identifiable?

> *Hypothesis 1:* Patterns of scores will emerge from the data that are psychologically identifiable as particular latent profiles of personality when examined in aggregate.

(2) How well do profiles generalize within and across conditions? What proportion of a new sample is classifiable, via fit criteria derived from emergent profiles, into profiles developed from the original sample (i.e., in cross-validation)? Is profile generalizability moderated by sample similarity (e.g., job level or industry)? That is, do cross-validation results indicate that profiles from similar samples converge, and profiles from dissimilar samples diverge?

> *Hypothesis 2:* Profiles, where identified, will be more generalizable (i.e., the proportion classified in cross-validation will be higher) within job levels and industries (i.e., within the cells of the 2 x 2 matrix) than (a) across job levels, or (b) across industries.

(3) Within a given sample, do profile membership fit scores demonstrate validity in predicting job performance, and if so, do they offer incremental validity over (a) the linear weighted sum of multiple scale scores (in terms of multiple *R*), or (b) existing compensatory fit indices?

> *Hypothesis 3:* Profile membership (measured along a continuum of non-compensatory fit scores) will demonstrate incremental validity over (a) the linear sum of scale scores configured via regression in the prediction of rated job performance and (b) the compensatory fit indices of D^2 and r_{pf}.

The remainder of this document is organized as follows. The methods by which the hypotheses are assessed are described in Chapter 3. Chapter 4 summarizes the findings of this research, and Chapter 5 offers discussion of key results and suggestions for future research.

CHAPTER 3

METHODS

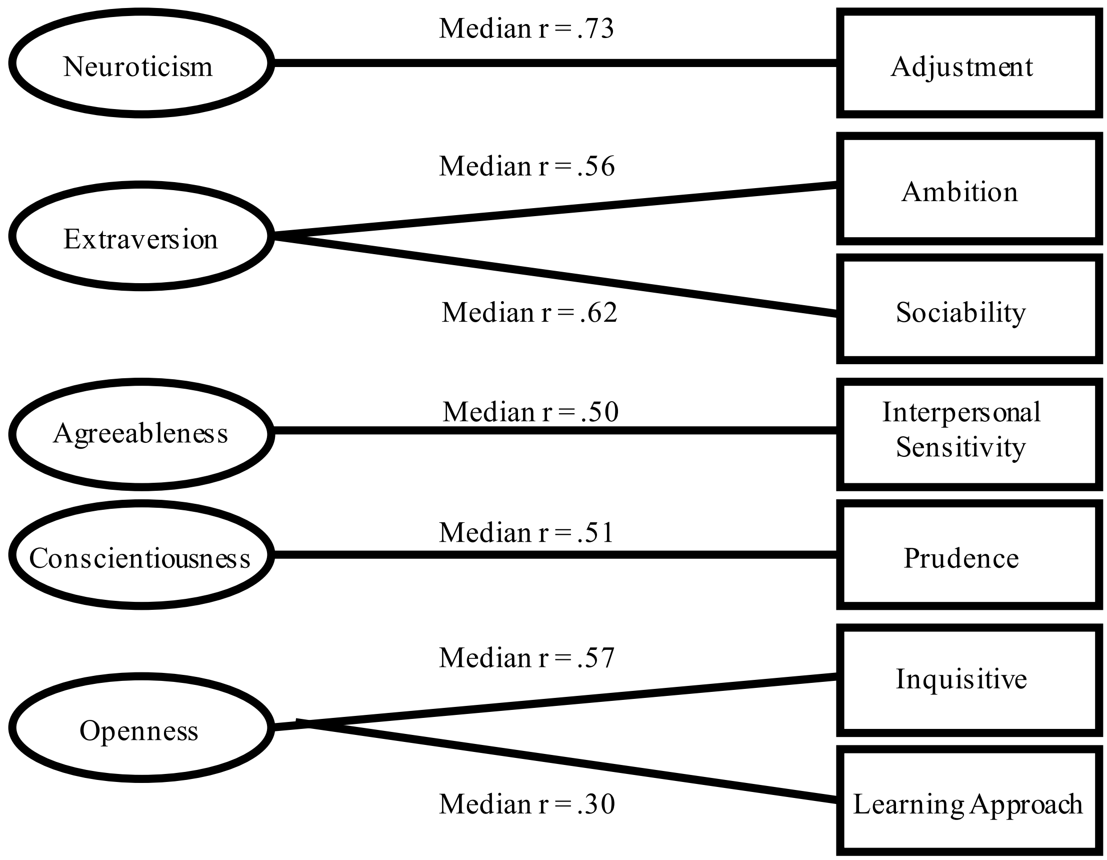
Data Sources and Approach

Hogan Assessment Systems, Inc. (HAS) provided data sets from 12 independent incumbent samples for the current study, all using the *Hogan Personality Inventory* (HPI; Hogan & Hogan, 1995). The HPI is a well-known FFM-based personality assessment that taps the five factors via seven personality scales. The traits targeted by the scales are defined in Table 2, and established relationships between the HPI scales and the FFM are detailed in Figure 2.

Table 2. HPI Scale Definitions

Scale	Definition
Adjustment	The degree to which a person appears calm and self-accepting.
Ambition	The degree to which a person seems socially self-confident, leader-like, competitive, and energetic.
Sociability	The degree to which a person seems to need and/or enjoy interacting with others.
Interpersonal Sensitivity	The degree to which a person is seen as perceptive, tactful, and socially sensitive.
Prudence	The degree to which a person seems conscientious, conforming, and dependable.
Inquisitive	The degree to which a person is perceived as bright, creative, and interested in intellectual matters.
Learning Approach	The degree to which a person seems to enjoy academic activities and to value educational achievement for its own sake.

Figure 2. Comparison between the Five Factor Model and Primary HPI Scales



Links between dimensions of the Big Five and the Hogan Personality Inventory (HPI). Median correlation coefficients summarize HPI relations with the NEO Personality Inventory—Revised (NEO–PI–R; Goldberg, 2000), Goldberg's (1992) Big Five Markers (R. Hogan & Hogan, 1995), Personal Characteristics Inventory (Mount & Barrick, 1995b), and the Inventario de Personalidad de Cinco Factores (Salgado & Moscoso, 1999). The ranges of correlates are as follows: Adjustment–Emotional Stability–Neuroticism (.66 to .81), Ambition–Extraversion–Surgency (.39 to .60), Sociability–Extraversion–Surgency (.44 to .64), Interpersonal Sensitivity Agreeableness (.22 to .61), Prudence–Conscientiousness (.36 to .59), Inquisitive–Openness– Intellect (.33 to .69), and Learning Approach–Openness–Intellect (.05 to .35). Figure reproduced with permission (Hogan & Holland, 2003).

Sample sizes for each of the 12 samples range from 103 to 253. As shown in Table 3, a 2x2 design with three samples (validation studies) per cell was used to contrast profiles within and across job level (entry level vs. management) and industry (sales vs.

customer service). The two-by-two design allows evaluation of profile emergence and generalization by sample, within each industry, within each job level, within each level-industry category (e.g., entry level sales), and between level, industry, and level-industry category.

Table 3. 2 x 2 Array of Available Samples

		Job Level	
		Entry level	Management
Industry	Service	A: Sample 1 B: Sample 2 C: Sample 3	A: Sample 7 B: Sample 8 C: Sample 9
	Sales	A: Sample 4 B: Sample 5 C: Sample 6	A: Sample 10 B: Sample 11 C: Sample 12

Criterion Treatment

Supervisor ratings of overall performance were available for each sample and additional item-level ratings were available for four samples. For those samples with item-level criterion ratings, internal consistency reliability coefficients ranged from .81 to .92. Because only ratings of overall performance were available for all samples involved, overall performance was the single criterion used in analyses. However, the rating scales differed among the samples. Accordingly, overall performance ratings were standardized within each sample and then combined across samples for use as a criterion (ZPERF) for the current analyses.

Question 1: Profile Emergence

Research question 1, which asks, “do profiles emerge that are identifiable?” was addressed by applying Latent Profile Analysis (LPA) via Mplus 6.1 software (Muthén & Muthén, 2006) to the 12 individual samples, two industries, two job levels, and four industry-by-level categories. A profile, in the current case, is defined as a set of seven means, one for each HPI scale.

LPA analyses begin with the specification of a two-class model; class size is increased in subsequent iterations until fit declines. Mplus also provides options for post-hoc statistical tests to evaluate the optimal number of profiles in a model (Kupzyk, 2011; Merz & Roesch, 2011; Muthén & Muthén, 2011; Nylund, Asparouhov & Muthén, 2007). Model fit is first evaluated via the Vuong-Lo-Mendell-Rubin Likelihood Ratio Test (Lo, Mendell, & Rubin, 2001; Vuong, 1989). It compares the model to k classes to the model with (k–1) classes. When $p > .05$, the model with k classes is rejected and the model with k–1 classes is judged to fit. Akaike information criterion and Bayesian information criterion (AIC and BIC) may also provide a secondary evaluation of model fit, where lower AIC and BIC values are indicative of better fit (Vrieze, 2012). Entropy values are tertiary interpretive fit index, where higher entropy values indicate better fit (Kupzyk, 2011). For a pre-specified number of profiles, Mplus finds profile means and variances within a sample that maximize the likelihood of individuals being assigned to a given profile. The likelihood function that is maximized is a multiplicative product of terms, one for each individual, with each term being the probability of that individual’s scores. Given an arbitrary set of profile means as a starting point, Mplus uses an iterative two-step procedure to maximize the log likelihood. At the first step, individuals’ contributions to the log likelihood are maximized by assigning their data to the profile that makes their scores most likely given the current estimates of profile means and variances. At the second step, given the assignment of individuals to profiles, Mplus maximizes each profile's contribution to the log likelihood through the profile's mean and variance by assigning each profile the mean and variance for each scale score that are the mean and variance of the sample of individuals currently assigned to that profile.

Because (a) the assignment of subjects to profiles in the first step is contingent on the scale means and variances for each profile that are determined in the second step, and (b) the profile means and variances computed in the second step are contingent on the assignments of subjects to profiles determined in the first step, the maximization procedure iterates between these two steps until it converges on a set of assignments of subjects to profiles that does not change. Applied to the current data, this means that Mplus estimates profiles by finding commonalities among the seven HPI scale scores across individuals and simultaneously categorizes individuals into the best fitting profile. Note that when a single profile is identified, the profile points are simply the scale means drawn from the entire sample.

Question 2: Generalizability

Mplus allows a specified dataset to be tested for generalizability against a given profile generated from a different dataset (Muthén, 2004). Profiles generated within samples are evaluated for generalizability to other samples via Confirmatory Latent Profile Analysis (CLPA), again using Mplus. Consider Table 3 again. We would expect, for example, that, if interpretable profiles are generated from Sample 1, then Samples 2 and 3, which fall within the same quadrant as Sample 1, will be more likely to be classifiable into those Sample 1 profiles than would Samples 4, 5, and 6, owing to differences in industry. Likewise, Samples 4, 5, and 6 would be more likely to be classified into the Sample 1 profile than would be Samples 10, 11, and 12, owing to differences not only in industry but also in job level.

In contrast to the exploratory LPA discussed above, CLPA starts with a known set of profiles, again defined by score means and variances. The known profiles, called "reference profiles," are found by using LPA in one sample (e.g., Sample 1). Because Mplus uses the maximum likelihood method to generate profiles in LPA, this method can also be used to determine whether a second sample (e.g., Sample 2) has profiles that are statistically similar to the reference sample via a likelihood ratio test stating that, for any two sets of maximum likelihood estimates, if one set of estimates constrains the estimated parameters to values that are freely chosen to maximize the likelihood function in the other set of estimates, the statistic minus 2 times the log of the ratios of the likelihoods for the two sets of estimates follows a chi-square distribution with degrees of freedom equal to the number of constrained parameters (Rao, 1965).

The intuition behind the likelihood ratio test is that, if constraints imposed on the population parameters estimated from a sample are true statements about the population parameters, then any difference between the value of the likelihood function from an estimation that imposes those constraints and the value of the likelihood function from an estimation that does not impose the constraints should be relatively small. In this research, the null hypothesis (that the reference profiles are the same as the profiles in the test sample) imposes the constraint that the estimated latent profile means for the test sample are the same as the means for the latent profiles of the reference sample. If that null hypothesis is true, then any differences between the reference sample profile means and the unconstrained estimates of the test sample profile means are likely to be small, and as such, will lead to relatively small differences in the ratio of the likelihoods (because the likelihood function is a function of the means of the latent profiles).

Mplus provides the value of the log likelihood for any model it estimates, and the test statistic $2(\log L_U - \log L_C)$ for the likelihood ratio test is readily calculated from the output of two estimations performed on the test sample. The definition of the logarithm, $-2 \log(L_C / L_U) = 2(\log L_U - \log L_C)$, is used in the calculations because the value of the likelihood function itself is so small. The first estimation finds the unconstrained profile means and the value of the unconstrained log likelihood function, L_U, in the test sample by estimating a latent profile model with the same number of profiles that were found in the reference sample. The log of the likelihood function is reported because the value of a likelihood function is a probability that often has a value that is too small to represent accurately. The second estimation also uses the test sample data; Mplus iterates once from starting values equal to the latent profile means from the reference sample. This single iteration is used solely as a device to obtain the value of the constrained log likelihood, L_C, which results from imposing the constraint that the means of the latent profiles in the test sample are the same as means of the latent profile in the reference sample. Because the log of a ratio is equal to the difference of the logarithms of the numerator and denominator, these two estimations provide the statistics L_U and L_C that are used to construct the likelihood ratio test statistic as twice the *difference* between the *log* likelihoods, $2(L_U - L_C)$, rather than as twice the ratio of the likelihoods. The degrees of freedom for the test statistic (in the current case, with seven data points per profile) are equal to seven times the number of profiles that emerge from the estimated model. This reflects the number of restrictions placed on profile means in estimating the constrained model. Thus, if the null hypothesis that the test sample has the same latent profiles as the reference sample is true, then the test statistic $2(L_U - L_C)$ will follow a chi-square distribution with degrees of freedom equal to seven times the number of profiles in the reference sample.

Because the likelihood function is a function of the profile means, small differences in profile means will produce small difference between L_U and L_C, and therefore the value for $2(L_U - L_C)$ will be small. As described earlier, the likelihood ratio test will not reject the null hypothesis that reference and test samples have the same profiles (i.e., that they generalize), if the profile means that are estimated in the test sample are similar to the profile means in the reference sample (i.e., if the value for $2(L_U - L_C)$ is small).

Figure 3 shows the possible generalizability pairings among all 12 samples, organized by within- versus between-industry and within- versus between-job level.

Thirty-six cross-validation opportunities are available in each quadrant. Notably, in the within-within condition (upper left quadrant), 12 of the pairings are within-sample (e.g., Sample 1-Sample 1). These are omitted from the comparison set owing to complexities arising from the need to rely on half-samples in such cases, relative to whole samples in all other comparisons. Also noteworthy is that each pair of samples offers double cross-validation. Thus, for example, profiles from Sample 1 can be cross-validated against Sample 2 and profiles from Sample 2 can be cross-validated against Sample 1.

Figure 3. Generalizability Pairings

Within level / within industry

	Ref	Test
EL-Sales	1	1
	1	2
	1	3
	2	1
	2	2
	2	3
	3	1
	3	2
	3	3

	Ref	Test
M-Sales	4	4
	4	5
	4	6
	5	4
	5	5
	5	6
	6	4
	6	5
	6	6

	Ref	Test
EL=Service	7	7
	7	8
	7	9
	8	7
	8	8
	8	9
	9	7
	9	8
	9	9

	Ref	Test
M-Service	10	10
	10	11
	10	12
	11	10
	11	11
	11	12
	12	10
	12	11
	12	12

Between level / within industry

	Ref	Test
EL-Sales	1	4
M-Sales	1	5
	1	6
	2	4
	2	5
	2	6
	3	4
	3	5
	3	6

	Ref	Test
M-Sales	4	1
EL-Sales	4	2
	4	3
	5	1
	5	2
	5	3
	6	1
	6	2
	6	3

	Ref	Test
EL-Service	7	10
M-Service	7	11
	7	12
	8	10
	8	11
	8	12
	9	10
	9	11
	9	12

	Ref	Test
M-Service	10	7
EL-Service	10	8
	10	9
	11	7
	11	8
	11	9
	12	7
	12	8
	12	9

Within level / between industry

	Ref	Test
EL-Sales	1	7
EL-Service	1	8
	1	9
	2	7
	2	8
	2	9
	3	7
	3	8
	3	9

	Ref	Test
M-Sales	4	10
M-Service	4	11
	4	12
	5	10
	5	11
	5	12
	6	10
	6	11
	6	12

	Ref	Test
EL-Serivce	7	1
EL-Sales	7	2
	7	3
	8	1
	8	2
	8	3
	9	1
	9	2
	9	3

	Ref	Test
M-Service	10	4
M-Sales	10	5
	10	6
	11	4
	11	5
	11	6
	12	4
	12	5
	12	6

Between level / between industry

	Ref	Test
EL-Sales	1	10
M-Service	1	11
	1	12
	2	10
	2	11
	2	12
	3	10
	3	11
	3	12

	Ref	Test
M-Sales	4	7
EL-Service	4	8
	4	9
	5	7
	5	8
	5	9
	6	7
	6	8
	6	9

	Ref	Test
M-Service	10	1
EL-Sales	10	2
	10	3
	11	1
	11	2
	11	3
	12	1
	12	2
	12	3

	Ref	Test
EL-Service	7	4
M-Sales	7	5
	7	6
	8	4
	8	5
	8	6
	9	4
	9	5
	9	6

Ref = Reference Sample
Test = Test Sample
M = Managerial
EL = Entry Level

Question 3: Incremental Validity

Incremental validity of profile membership for predicting job performance was assessed via stepwise multiple regression analyses applied to the standardized supervisor ratings of overall job performance available per sample. Mplus generates the profiles but does not create an interpretable and appropriately non-compensatory fit score as is required for these analyses. As such, a new multiplicative fit score index (the Relative Root Product Index; RRPI) was created. The RRPI was designed to be non-compensatory, such that a high degree of fit on one scale cannot mask or compensate for a low degree of fit on another scale. The RRPI is scaled to range from 0 to 1, with 1 representing a perfect fit between an individual's data and a target profile. To calculate the RRPI per individual, each scale score is standardized (i.e., converted to a *z*-score) against the means for the target profile and then converted to *p* values using a basic normal curve table. The conversion thus yields the maximum *p* value (.50) for the highest similarity ($z = 0$) per profile point. The resulting *p* values are then multiplied (within persons) across the seven scales, and re-scaled to fit a 0-1 continuum for interpretability. Because a smaller *p* is indicative of poorer fit, and because the operations are multiplicative, the resulting value is very small (approaching zero) where any *single scale* deviates substantially from the target profile. In this way, the RRPI is non-compensatory, in that a poor fit on one scale cannot be masked or compensated for by a strong fit on other scales - i.e., the multiplicative operations penalize an individual's profile fit for a miss on any single scale.

The formula for the RRPI, as used in the current undertaking (involving seven scales), is as follows:

$$\text{RRPI} = \left((p_{ADJ} * p_{AMB} * p_{SOC} * p_{INT} * p_{PRU} * p_{INQ} * p_{LRN})^{\frac{1}{7}} \right) \div 0.5$$

where p_{ADJ} through p_{LRN} represent the values for each of the HPI scales following the conversion of each standardized scale score to the *p* distribution. The seven-way product of *p* values yields an extremely small number (even in the case of perfect fit: $.5^7 = .0078$). To simply expand the product, the *n*th root is taken, with $n = 7$ in the current case. For perfect fit, this would yield .50. The denominator of .50 thus sets an upper limit of 1.00; values approach 0 when fit is poor, owing to one or more *p* values being very small (in turn, due to one or more *z*s being very different from 0). Notably, the RRPI ignores whether poor fit is due to individual scores falling above or below corresponding profile points (means).

The RRPI was calculated for each individual, by profile, within each sample, industry, job level, and industry-by-level category. The RRPI was then used to facilitate the evaluation of the incremental validity of profiles over (a) the linear weighted sum of multiple scale predictors (as per multiple *R*) and (b) other common fit indices, where all personality scales are entered into block 1, the fit indices of D^2 and r_{pf} are entered into block 2, and the RRPI is entered into block 3. For each analysis, the change in R^2 indicates incremental validity.

Summary

The idea that personality traits may interact with one another is not new. Robert Bernreuter, the author of what has been argued (Gibby & Zickar, 2008) to be the first multidimensional personality assessment, noted almost 80 years ago that personality is "complex and dynamical" and that the study of single aspects in isolation had been "hazardous because the importance which becomes attached to a single aspect … threatens to obscure the still greater significance of the total integrated personality" (Bernreuter, 1933; p. 387). Yet, decades of research have attempted to evaluate the effectiveness of personality instruments primarily via methods originally designed for unidimensional assessments. The current study uses relatively new analytical methods to more precisely measure the dynamic and multidimensional nature of the personality constructs being assessed—methods that more closely approximate how personality manifests within individuals, and that attempt to simultaneously account for level, dispersion, and shape in a way that previously available methods have been unable to capture.

CHAPTER 4

RESULTS

Internal Characteristics of the HPI

Table 4 presents descriptive statistics for the HPI in each of the 12 current samples and compares these data to the descriptive statistics for the HPI found in the HPI manual (Hogan & Hogan, 1995). The data suggest that, although the samples differ from one another, none of the 12 samples is meaningfully restricted in range, and the sample as a whole is similar to the normative sample with respect to means and standard deviations. Item- and subscale-level data were not available; as such, the reliability of the HPI subscales per sample could not be assessed. Published internal consistency reliabilities for the HPI range from .71 (Interpersonal Sensitivity) to .89 (Adjustment), and test-retest reliabilities, from .74 (Prudence) to .86 (Adjustment and Learning Approach).

Table 4. Descriptive Statistics for the HPI

	Normative Sample		Entry Level Customer Service						Entry Level Sales					
Scale	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
			Sample 1 (N=253)		*Sample 2 (N=162)*		*Sample 3 (N=248)*		*Sample 4 (N=167)*		*Sample 5 (N=110)*		*Sample 6 (N=108)*	
ADJ	26.57	7.10	27.53	6.02	27.26	5.85	24.88	7.47	27.31	6.47	26.71	6.04	26.68	6.35
AMB	23.61	5.00	23.64	4.35	23.36	4.61	21.69	5.57	24.70	3.50	24.70	2.94	24.43	4.25
SOC	13.47	4.86	13.12	4.39	13.02	4.51	12.21	5.04	14.81	4.14	14.98	4.39	13.17	5.00
INT	19.62	2.36	19.72	2.26	19.52	2.46	19.43	2.44	19.47	2.53	19.94	2.83	19.09	2.71
PRU	20.28	4.62	20.32	4.16	20.23	4.21	20.38	4.05	20.41	4.38	20.28	4.02	20.47	4.19
INQ	14.68	4.90	14.69	5.16	14.28	4.83	12.86	5.28	14.91	4.22	14.69	4.58	13.96	4.67
LRN	8.66	3.14	7.53	3.61	7.57	3.63	8.81	3.26	9.40	2.90	8.97	3.12	8.02	3.36
	Normative Sample		**Management Customer Service**						**Management Sales**					
Scale	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
			Sample 7 (N=175)		*Sample 8 (N=180)*		*Sample 9 (N=167)*		*Sample 10 (N=140)*		*Sample 11 (N=103)*		*Sample 12 (N=210)*	
ADJ	26.57	7.10	23.93	6.62	28.86	5.10	22.33	7.58	27.79	6.39	30.20	4.35	26.35	6.29
AMB	23.61	5.00	24.07	4.95	24.52	4.03	21.12	5.62	24.70	4.23	27.61	2.13	24.70	4.13
SOC	13.47	4.86	12.87	5.25	12.24	4.94	14.23	4.56	11.79	5.46	15.50	4.37	12.10	5.31
INT	19.62	2.36	19.50	2.11	19.98	1.92	18.11	2.64	19.39	2.57	19.49	1.88	18.13	3.16
PRU	20.28	4.62	19.89	3.69	21.35	3.88	18.87	4.04	21.98	4.38	20.29	3.63	20.45	3.88
INQ	14.68	4.90	12.91	5.19	14.41	4.55	12.83	5.28	13.35	4.91	16.68	3.60	14.13	4.32
LRN	8.66	3.14	8.71	3.39	9.63	3.05	8.28	3.22	8.83	3.51	10.07	2.36	8.23	3.10

Note: ADJ = Adjustment; AMB = Ambition; SOC = Sociability; INT = Interpersonal Sensitivity; PRU = Prudence; INQ = Inquisitive; LRN = Learning Approach. Normative sample statistics taken from the *Hogan Personality Inventory Manual*, Hogan & Hogan, 1995.

Hypothesis Results

The first hypothesis concerned whether identifiable personality profiles would emerge using LPA.

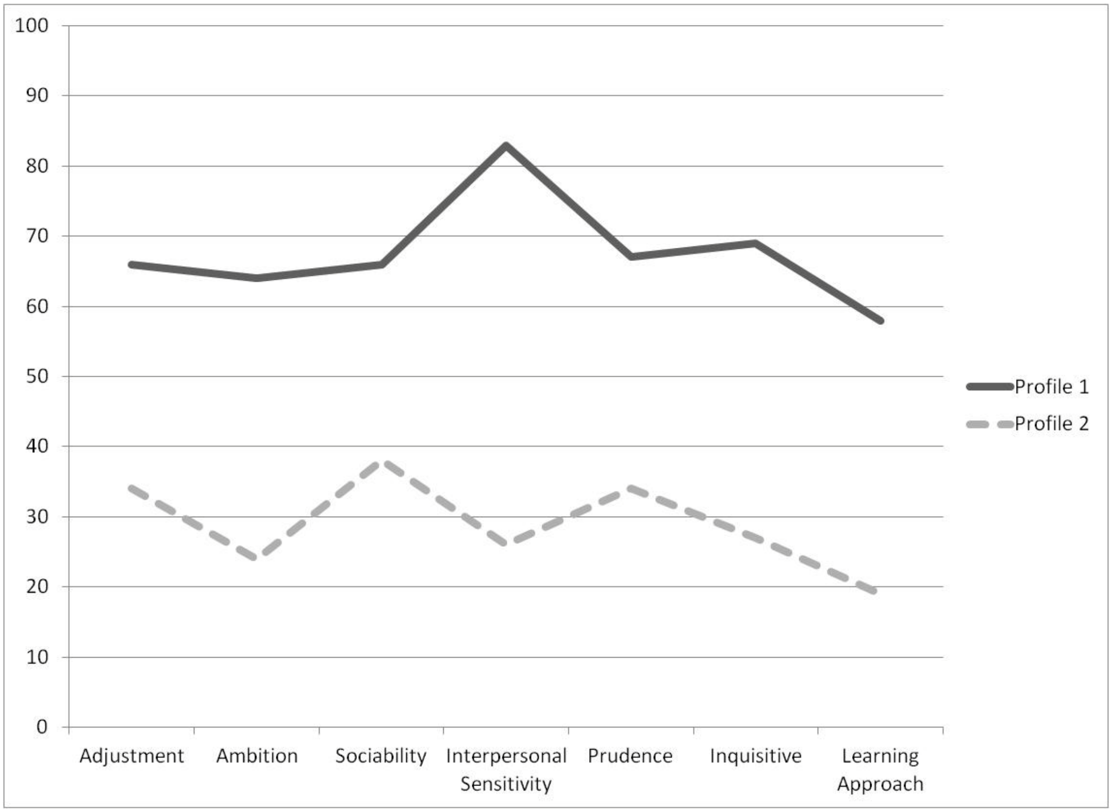
> *Hypothesis 1:* Patterns of scores will emerge from the data identifiable as particular profiles of personality when examined in aggregate.

Results for H1 are mixed. Of the 12 individual samples, multiple identifiable profiles (i.e., beyond a single array of scale means for the given sample) emerged from only four. However, when analyzed by industry (i.e. Customer Service or Sales), level (i.e., Entry Level or Management), and industry-by-level (e.g., Entry Level Sales, etc.), identifiable profiles emerged consistently, with only one exception (Management level when collapsing across industries).

Profiles that were identified are shown in Figures 4 through 14. To facilitate normative interpretations, profile scores have been converted to percentiles using the normative tables in the HPI manual (Hogan & Hogan, 1995). Likewise, the interpretations that accompany each figure are based on the definitions of each HPI scale (see Table 2) in conjunction with the scale-by-scale interpretive guidelines offered in the manual. Note that the manual defines low scores as ranging from 0% to 35%, average scores as falling between 36% and 64%, and high scores at or above 65%. Score differences within each range are also interpreted linearly (i.e., the ranges are guidelines, but elevations within each range are also interpretable; they are *not* bands where any score in a range is considered to be equivalent to any other score within the same range). Accordingly, these ranges are used in interpreting each profile. That is, even when examining data at a profile level, the scale-by-scale interpretations are absolute, such that a low score is interpreted as low, even if it is the highest score within a given profile, and the magnitude of a score is also interpreted, wherein a moderately high score and a very high score each have their own implications. Comparisons can still be made (e.g., individuals in Profile X are more conscientious than they are even tempered; or individuals in Profile Y are more conscientious than those in Profile X), but the absolute nature of the score interpretations is still grounded in the ranges provided by the test publisher.

Looking across all profiles, some similarities emerge. For almost all samples and sample groupings, the first profile that emerged (shown as the blue line on each graph) can be best described as a general "elevated score" profile. This profile shows average to high scores on most HPI scales, as well as elevations over the other profiles identified within their respective grouping. These profiles seem to represent employees who are stress-tolerant, goal-oriented, outgoing and friendly, conscientious, and relatively interested in learning. Some practitioners might refer to this as a generic "good employee" profile (relationships between emergent profiles and job performance are discussed in later sections). For those samples where more than one profile emerged, a secondary "low score" profile was also identified, which was characterized by average to low scores across most or all HPI scales. When interpreting these profiles, it is important to note that they are, in a sense, "purified." That is, individuals are only assigned to a single profile. As such, the low score profile has essentially been stripped of anyone who fits better with the elevated score profile, and vice versa. More than two profiles emerged from four samples. More in-depth interpretive descriptions follow each figure.

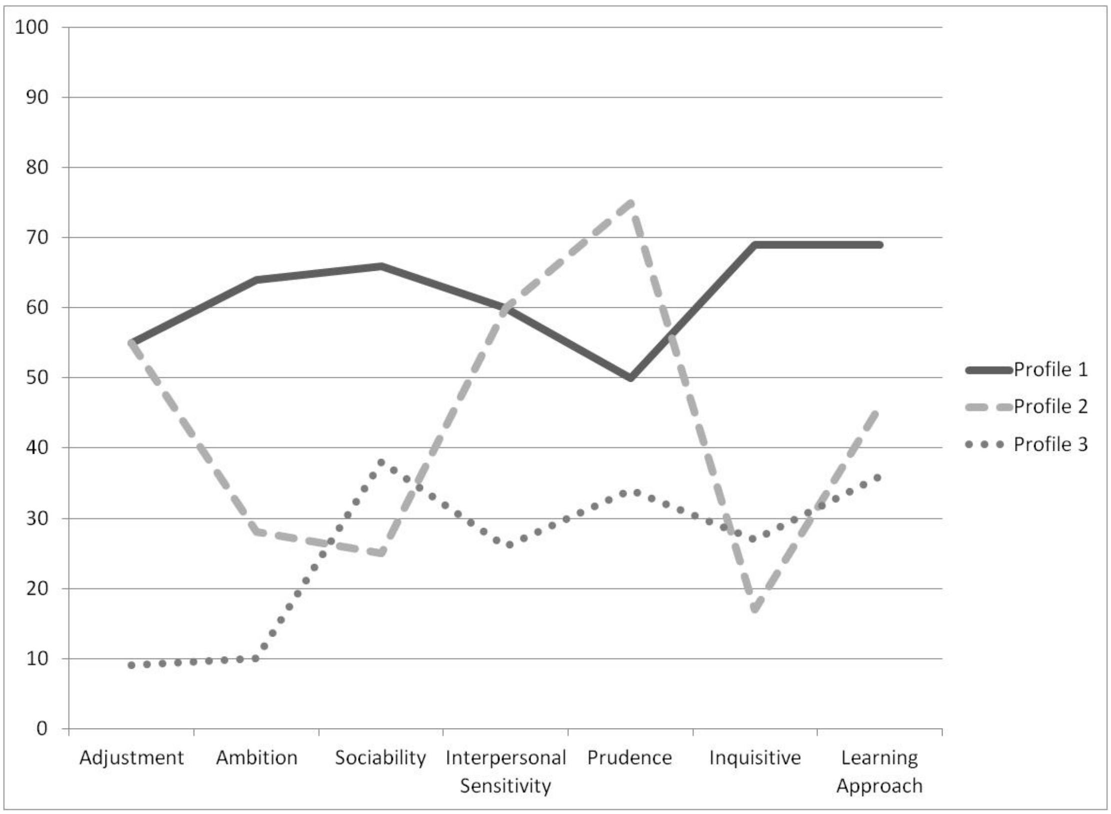
Figure 4. Profiles Identified in Sample 1



Sample 1 falls in the Entry Level Customer Service quadrant. The primary profile here fits the elevated score description mentioned previously, demonstrating above-average scores across the HPI scales, and elevated Interpersonal Sensitivity in particular. This is the profile of an employee who is resilient and embraces goals, is outgoing and makes positive connections with customers, and is relatively curious and interested in learning new skills. The secondary profile suggests an employee who is fairly susceptible to the stressors involved in a customer service environment, lacks ambition, is moderately introverted, relatively unfriendly, unconscientious, and shows little curiosity or aptitude for training. This profile encapsulates the low score profile mentioned earlier. As an example of the interpretive guidelines and decision rules noted previously, the secondary profile has Sociability as its highest score and Learning Approach as its lowest score.

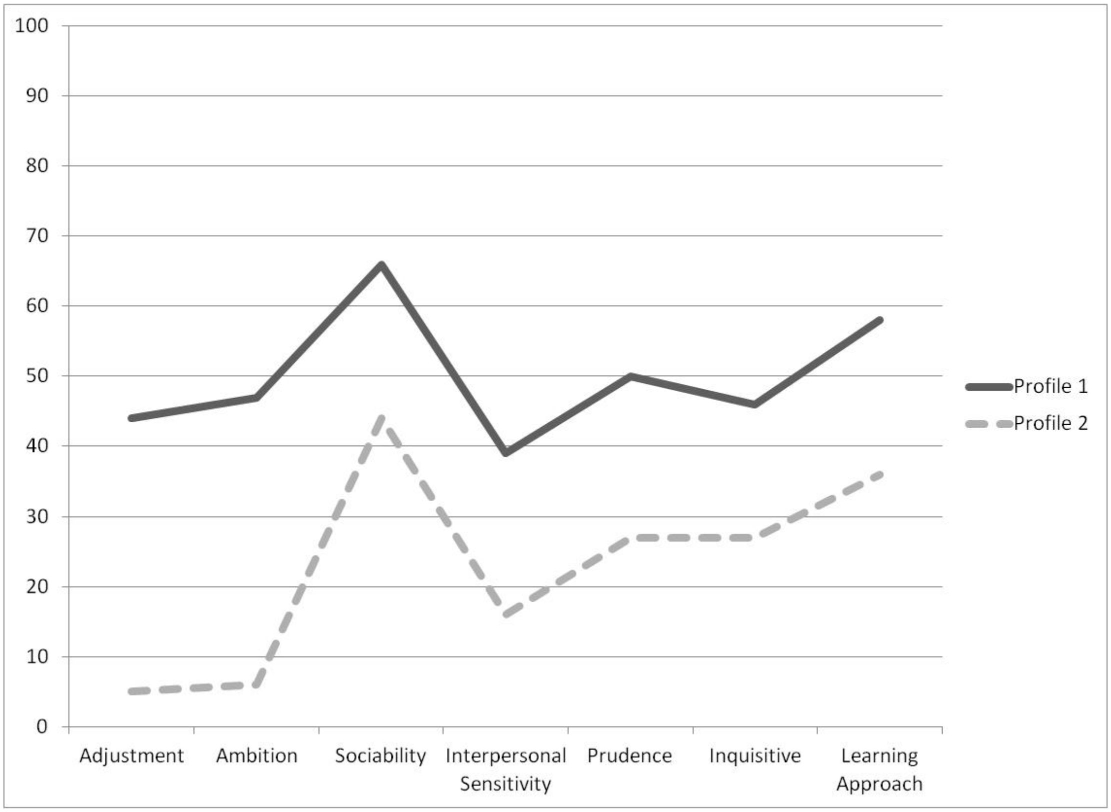
Though this may mean that individuals fitting this profile are more outgoing than they are interested in training and development, both scores still fall into the low score range provided by the test publisher. As such, these individuals would not be described (and more importantly, would not be likely to behave) as extroverted or outgoing simply because the highest elevation here is on the Sociability scale. They are simply less introverted than they are un-ambitious.

Figure 5. Profiles Identified in Sample 3



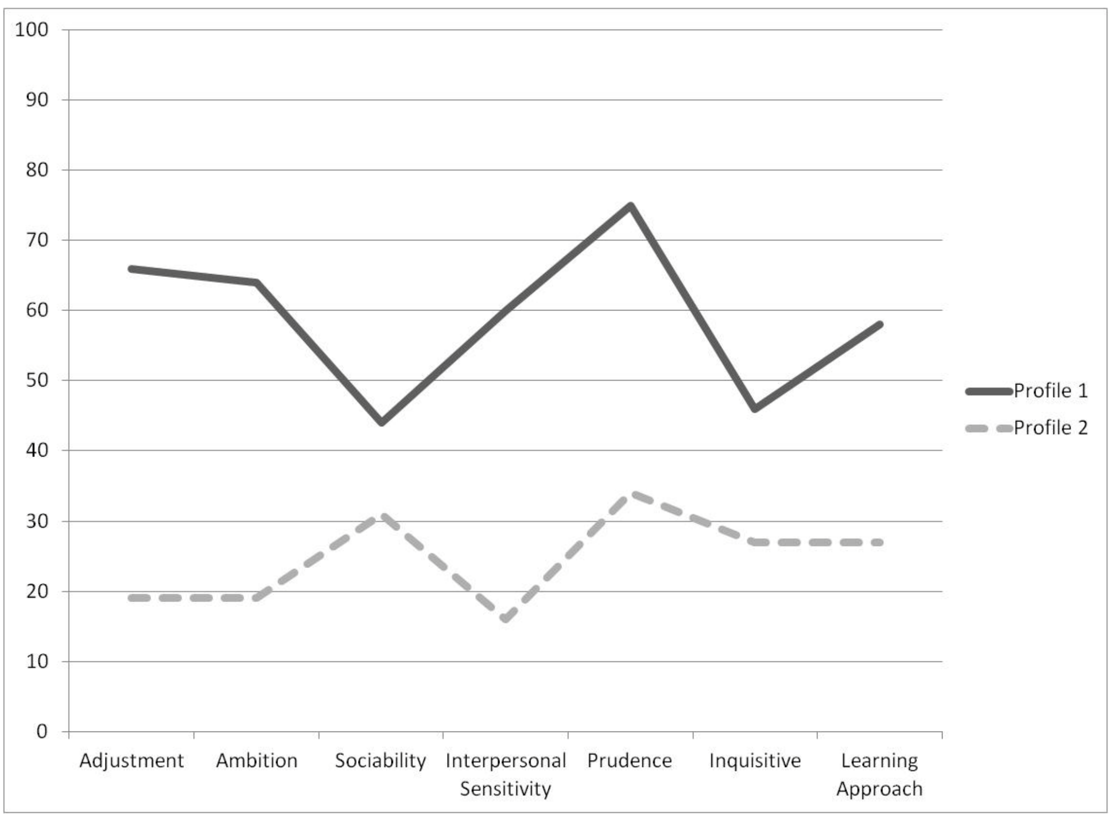
Sample 3 also came from the Entry Level Customer Service quadrant. The primary profile here still fits the elevated score description mentioned previously, demonstrating above average scores across the HPI scales. In contrast to the first profile in Sample 1, however, elevations can be seen here on the Inquisitive and Learning Approach scales, with a lower score on Prudence. This pattern suggests employees who are curious, generate ideas, have a positive attitude toward training and learning new skills, exercise balance between conscientiousness and flexibility, are stress tolerant, ambitious, and friendly. Interestingly, Profile 2 shows a peaked elevation on the Prudence scale, with above average scores on Adjustment and Interpersonal Sensitivity as well. Taken together, the first and second profiles suggest that, while this sample was drawn from a service-type job, perhaps this sample is drawn from a more heterogeneous incumbent population that is primarily characterized as either (a) very detail-oriented and pragmatic (high Prudence and low Sociability and Ambition – head down and focused), or (b) tasked with generating unique solutions to customers' problems, and/or building relationships (high Ambition, Sociability, Inquisitive, and Learning Approach). It is also possible that the organization combined two similar but distinct jobs into a single job family for the purposes of a validation study. Profile 3 represents the low score profile mentioned previously, and is indicative of employees who are susceptible to stress and show less interest in goal achievement, helping others, or learning new skills than their counterparts in Profiles 1 or 2.

Figure 6. Profiles Identified in Sample 9



Sample 9 comes from the Management Customer Service quadrant. The two emergent profiles, shown in Figure 6, are of interest because they share a similar shape, yet with differing magnitudes. The primary profile is similar to the elevated score profiles discussed previously, but with lower overall elevations in comparison, particularly on Adjustment and Ambition. This may indicate an overall flaw or deficiency in the organization's hiring strategy, or could point toward a particular job or industry characteristic. It is also possible that, even though this is a managerial customer service job, there may be an additional job requirement, where a bias toward urgency and action versus stress tolerance or resilience (lower Adjustment), and a preference for flexibility over conscientiousness (lower Prudence), are adaptive qualities, yet being outgoing, developing one's direct reports, and maintaining a knowledge base about one's products (high Sociability and Learning Approach) is still important. This would help make sense of Profile 2 as well, which still shows relatively elevated Sociability and Learning Approach scores, but with significant deficits on stress tolerance, goal achievement, and relationship building (low Adjustment, Ambition, and Interpersonal Sensitivity). This profile warrants attention from a methodological standpoint as well. As discussed in the introductory chapters, methods that account for the shape of a profile (such as correlational methods) without consideration for the magnitude of scores would likely have categorized these profiles as being similar, given their highly similar shapes.

Figure 7. Profiles Identified in Sample 10



Sample 10 comes from the Management Sales quadrant. These profiles are of interest because they demonstrate almost a mirror-image shape. The primary profile is again similar to the elevated score profiles discussed previously, with particular elevations on Adjustment, Ambition, Interpersonal Sensitivity, and Prudence. Employees fitting this profile may appear resilient to stress and pressure, goal-oriented, friendly, and conscientious. The Sociability and Inquisitive scores show moderate decreases in elevation, but are near the 50th percentile. The secondary low score profile for Sample 10 is characterized by lower scores across all seven HPI scales, but particularly on Adjustment, Ambition, and Interpersonal Sensitivity. This profile is indicative of an employee who is easily stressed, lacks motivation, detail orientation, and curiosity, is withdrawn and hostile when interacting with others, and has reduced interest in training or development.

Figure 8. Profiles Identified within Entry Level Customer Service

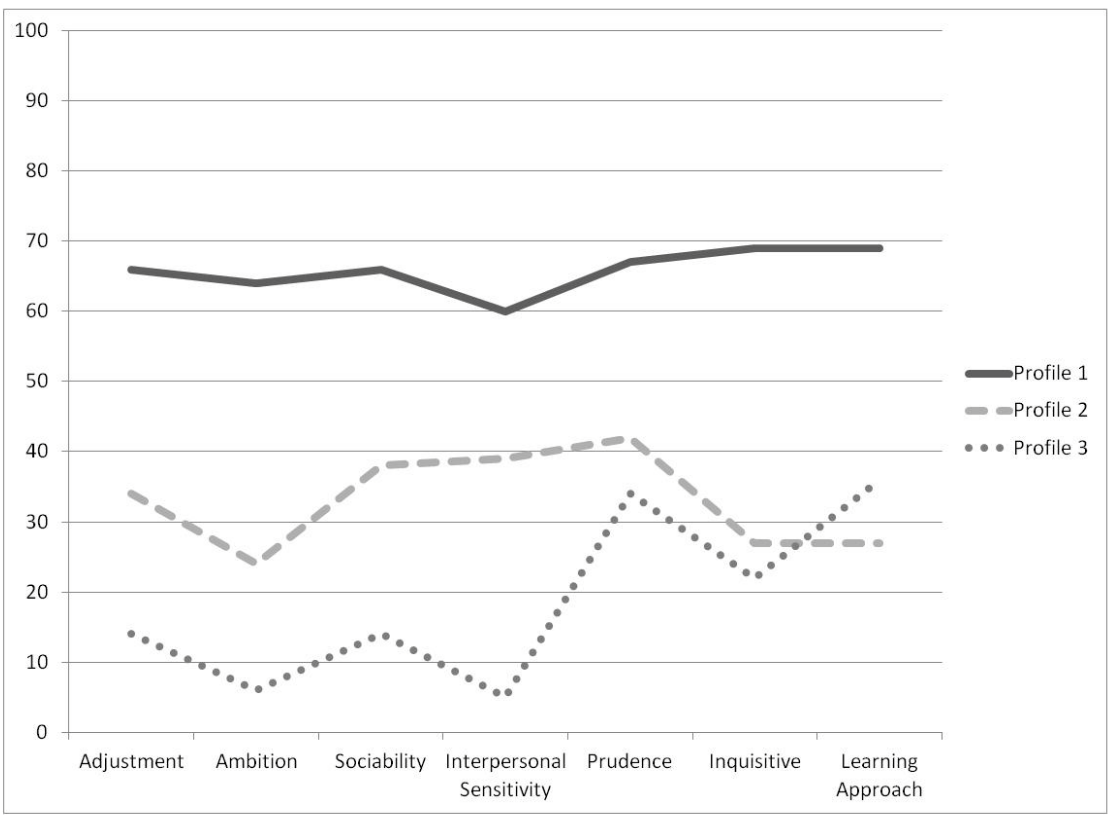


Figure 8 depicts three profiles that emerged when all three samples from the Entry Level Customer Service quadrant (Samples 1, 2, and 3) were combined. As previously mentioned, even after combining across samples, the primary profile is again similar to the elevated score profile, showing high scores across all HPI scales. Profile 2 is defined by moderate to low scores across scales, and Profile 3 shares some similarities with Profile 2, but with further decrements on the Adjustment, Ambition, Sociability, and Interpersonal Sensitivity scales. It is possible that the second profile represents an average employee, the third represents the low score profile seen previously, and that selection strategies and/or organizational attrition have acted to exclude the bulk of individuals who would have contributed very low scores on Prudence, Inquisitive, or Learning Approach to this combined sample. Both Profile 2 and Profile 3 are indicative of employees who would experience a good deal of stress, lack strong drive for achievement, and exhibit low to moderate levels of conscientiousness, curiosity, and interest in learning. However, those fitting Profile 3 would seem noticeably more stress-prone, lacking in drive, withdrawn, and unfriendly than those fitting Profile 2.

Figure 9. Profiles Identified within Entry Level Sales

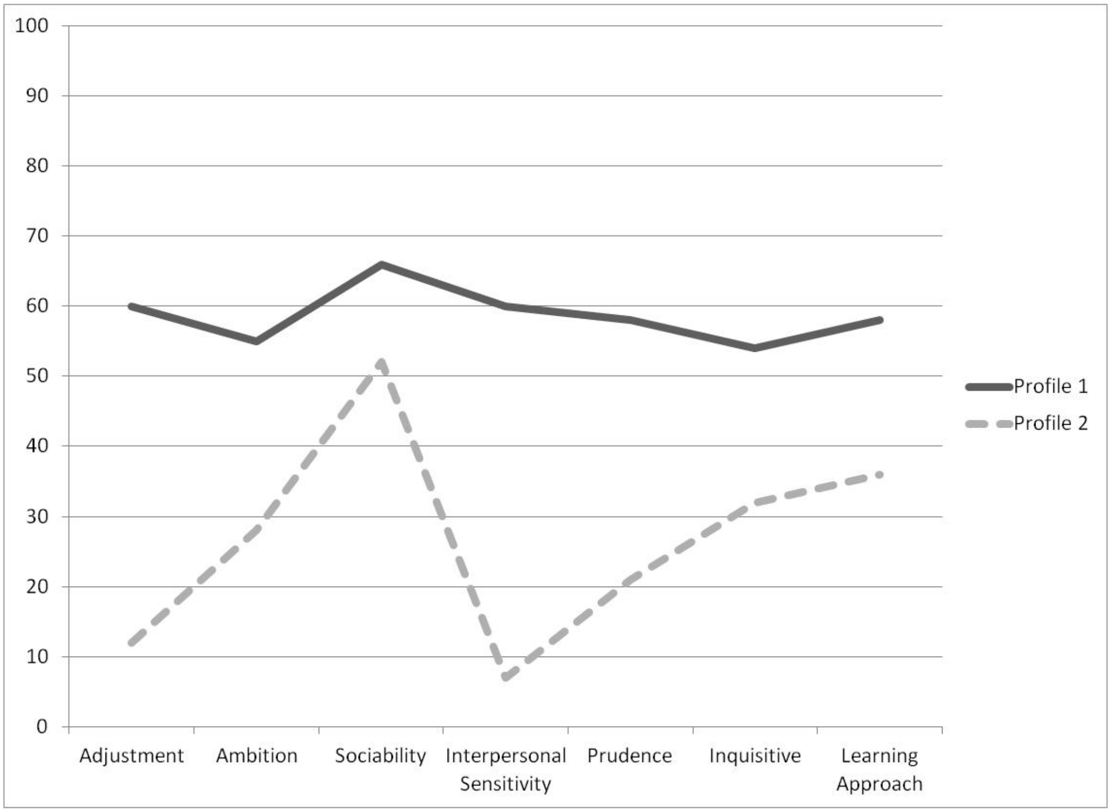


Figure 9 identifies the profiles that emerged when the data from the three Entry Level Sales samples (4, 5, and 6) were combined. The primary profile fits the noted elevated score profile, characterized by high scores across the HPI scales. The secondary profile shows marked decrements on the Adjustment, Ambition, Interpersonal Sensitivity, and Prudence scales, low to moderate scores on the Inquisitive and Learning Approach scales, but with a noticeable (relative) elevation on the Sociability scale. The low levels of Adjustment, Ambition, Interpersonal Sensitivity, and Prudence indicate that employees fitting the secondary profile may seem susceptible to stress, lacking in drive and motivation, unfriendly or hostile, and generally unconcerned with details or rules. The relative lack of separation between the two profiles' Sociability scores could be a function of a particular selection strategy or cultural element of the organization, wherein individuals with Sociability scores far outside the average range are either not selected, or self-select out of the applicant pool.

Figure 10. Profiles Identified within Management Service

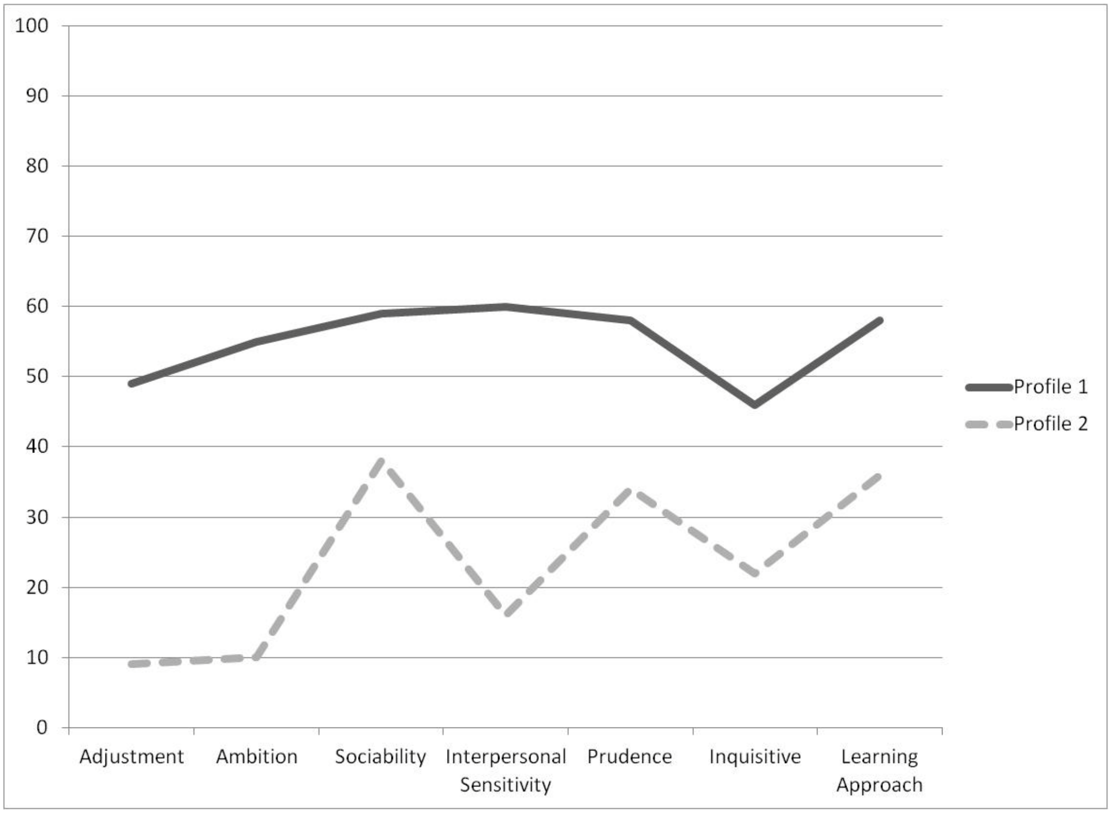


Figure 10 identifies the profiles that emerged when the data from the Management Customer Service quadrant (Samples 7, 8, and 9) were combined. The primary profile shows the familiar elevated score pattern, save for Inquisitive falling just below the 50th percentile. The secondary low score profile indicates lower scores overall with markedly low scores on the Adjustment, Ambition, and Interpersonal Sensitivity scales. As with many of the secondary profiles discussed, this profile is indicative of employees who lack stress tolerance, drive, and friendliness.

Figure 11. Profiles Identified within Management Sales

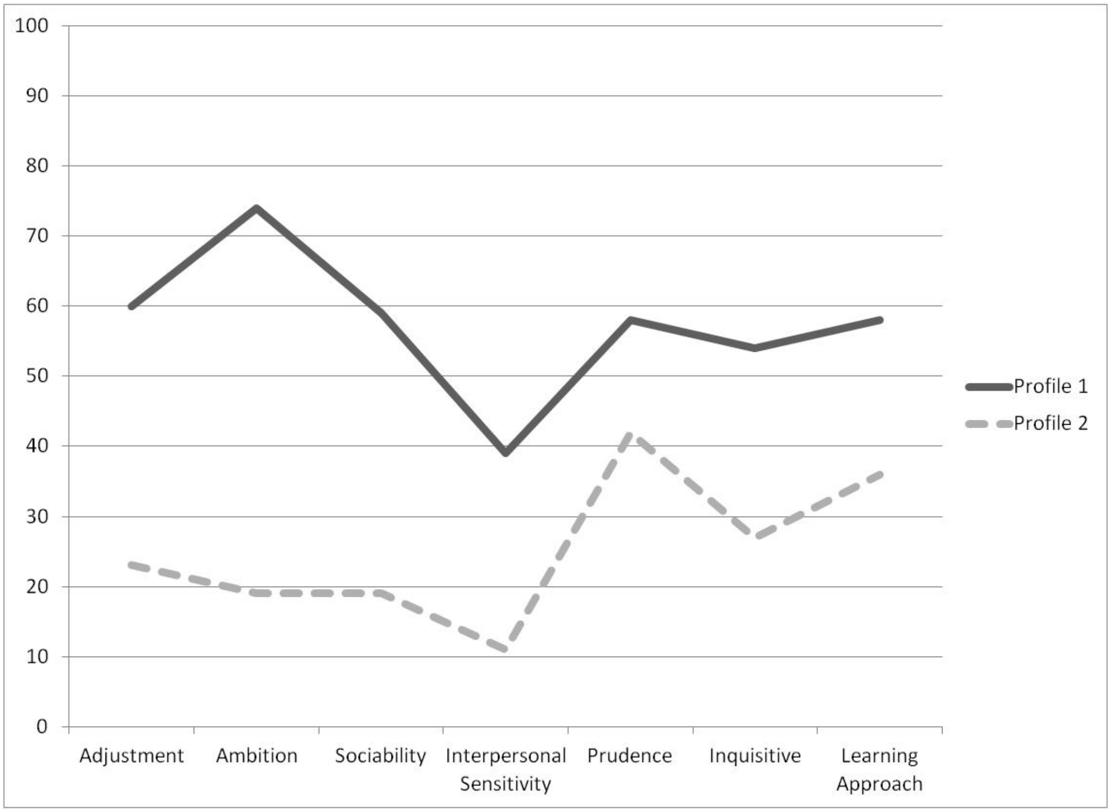


Figure 11 identifies the profiles that emerged when the data from the Management Sales quadrant (Samples 10, 11, and 12) were combined. The primary profile shows some similarities to the elevated score profile mentioned previously, but with a spike on the Ambition scale and a decreased elevation on the Interpersonal Sensitivity scale. This is still a profile that would be representative of employees who are stress tolerant, outgoing, conscientious, open to experiences, and interested in development opportunities. In addition, they are likely to be driven and motivated to get ahead, but less warm and friendly compared to the other elevated score profiles identified previously. Employees who fit the secondary low score profile are again likely to lack even-temperedness and motivation, to be withdrawn and unfriendly, and to have low levels of curiosity or openness to new experiences and ideas. However, they should show moderate levels of conscientiousness and occasional interest in learning opportunities.

Figure 12. Profiles Identified by Industry: Sales

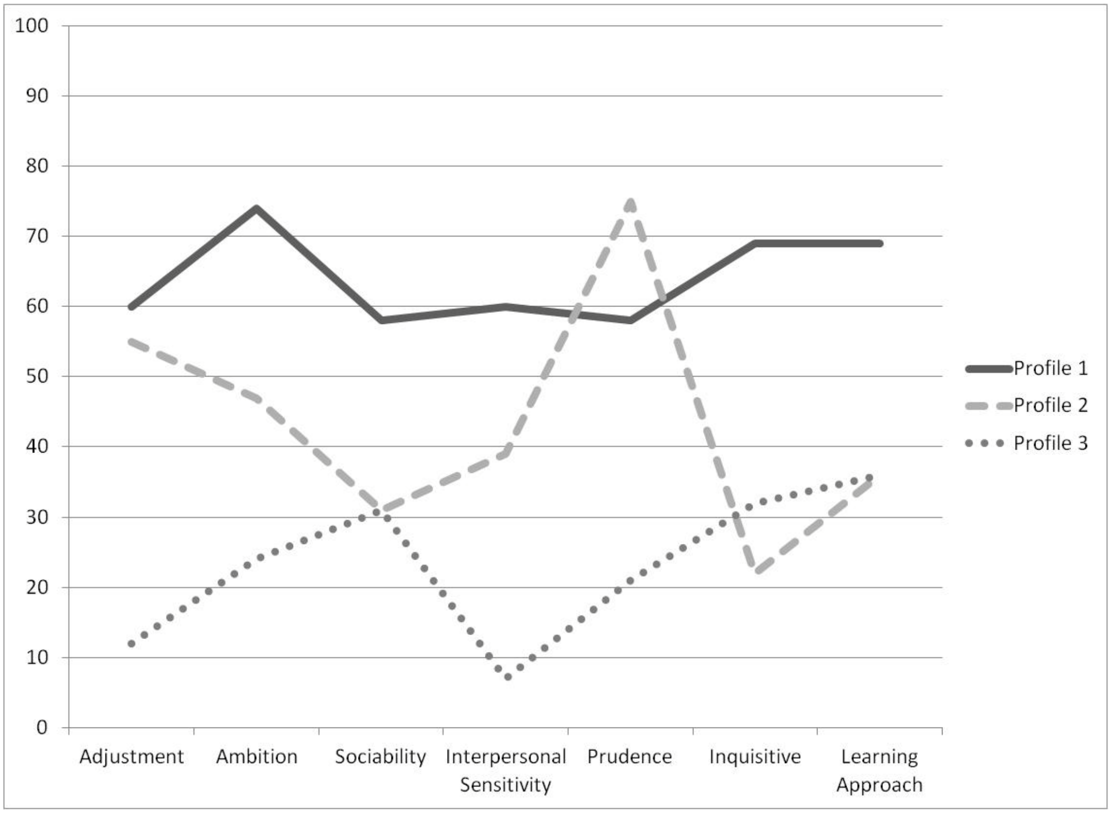


Figure 12 identifies the profiles that emerged when the data from the Entry Level and Management Sales quadrants (Samples 4, 5, 6, 10, 11, and 12) were combined. The primary profile is, again, similar to the elevated score profile seen previously, with high scores across the HPI scales. Scores on the Ambition, Inquisitive, and Learning Approach scales are notably high, suggesting that employees fitting this profile may be ambitious, curious, and interested in professional development. The second profile is of particular interest. As seen previously, most HPI scores in Profile 2 are lower than the primary profile. However, the Adjustment score is still above average, the Ambition score reaches nearly the 50th percentile, and the Prudence score is higher than its counterpart in the primary profile. Employees fitting Profile 2 may be likely to be resilient, relatively driven, and very conscientious; yet they will also be more withdrawn, less friendly, less interested in learning, and much less open and curious when compared to employees who fit the primary profile. The third low score profile is representative of employees who are prone to stress and worry, lack drive, are withdrawn, exhibit open hostility toward others, and lack conscientiousness. However, those fitting Profile 3 are likely to exhibit levels of openness and curiosity that are similar to employees who fit Profile 2.

Figure 13. Profiles Identified by Industry: Service

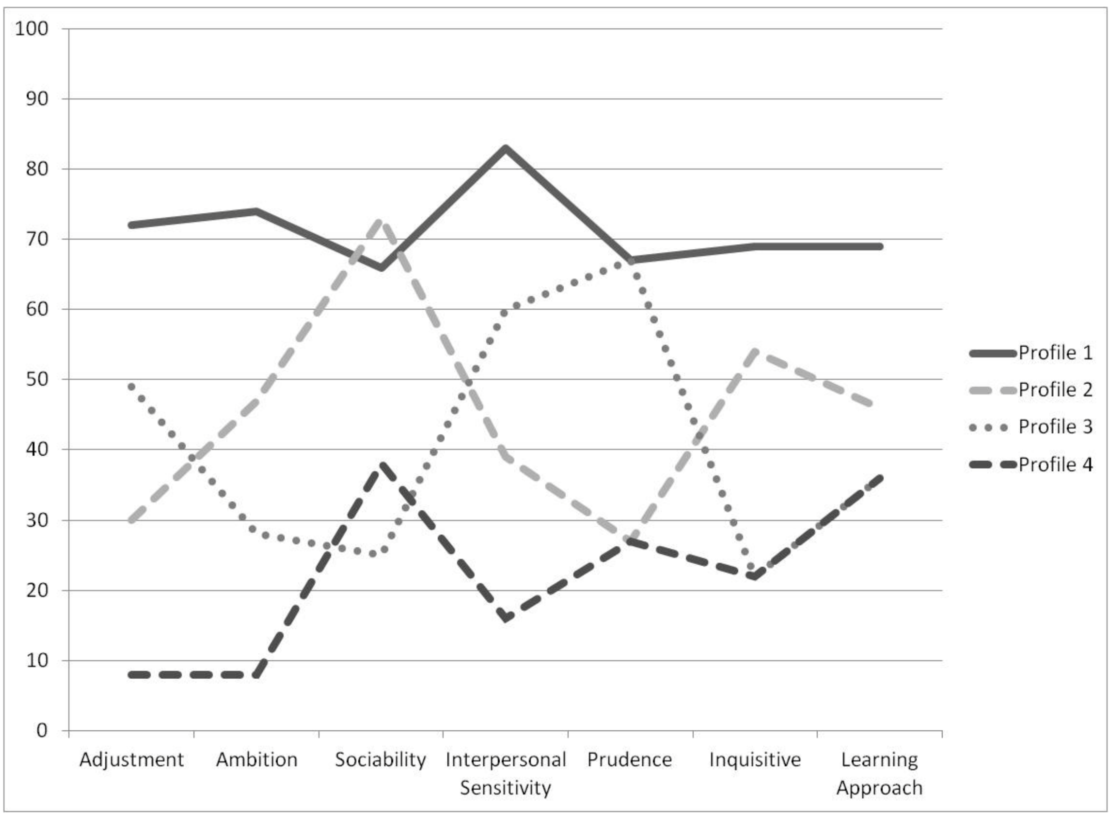


Figure 13 identifies the profiles that emerged when the data from the Entry Level and Management Service quadrants (Samples 1, 2, 3, 7, 8, and 9) were combined. The primary profile is much like the elevated score profile already discussed, with high scores across the HPI scales and a noticeable elevation on Interpersonal Sensitivity. Profile 4 is similar to most of the low score profiles discussed previously, showing low scores across the HPI scales. However, Profiles 2 and 3 represent an interesting middle ground. Profile 2 is characterized by a spike on Sociability and decrements on Adjustment, Interpersonal Sensitivity, and Prudence, suggesting employees who are engaging but stress-prone, moderately unfriendly, and inattentive to detail. Conversely, Profile 3 demonstrates a decrement on Sociability, an Adjustment score that approaches the 50th percentile, and spikes on Interpersonal Sensitivity and Prudence—a near-reversal of Profile 2, and an indication of employees who are quiet, fairly resilient, friendly, and conscientious. As suggested previously, it is possible that these two types of profiles may be driven by data from two different types of Customer Service environments or roles—one that is based on up-selling, relationship building, and/or blending other sales functions into a service job, and another that is strictly focused on providing service in a more structured or administrative fashion.

Figure 14. Profiles Identified by Level: Entry Level

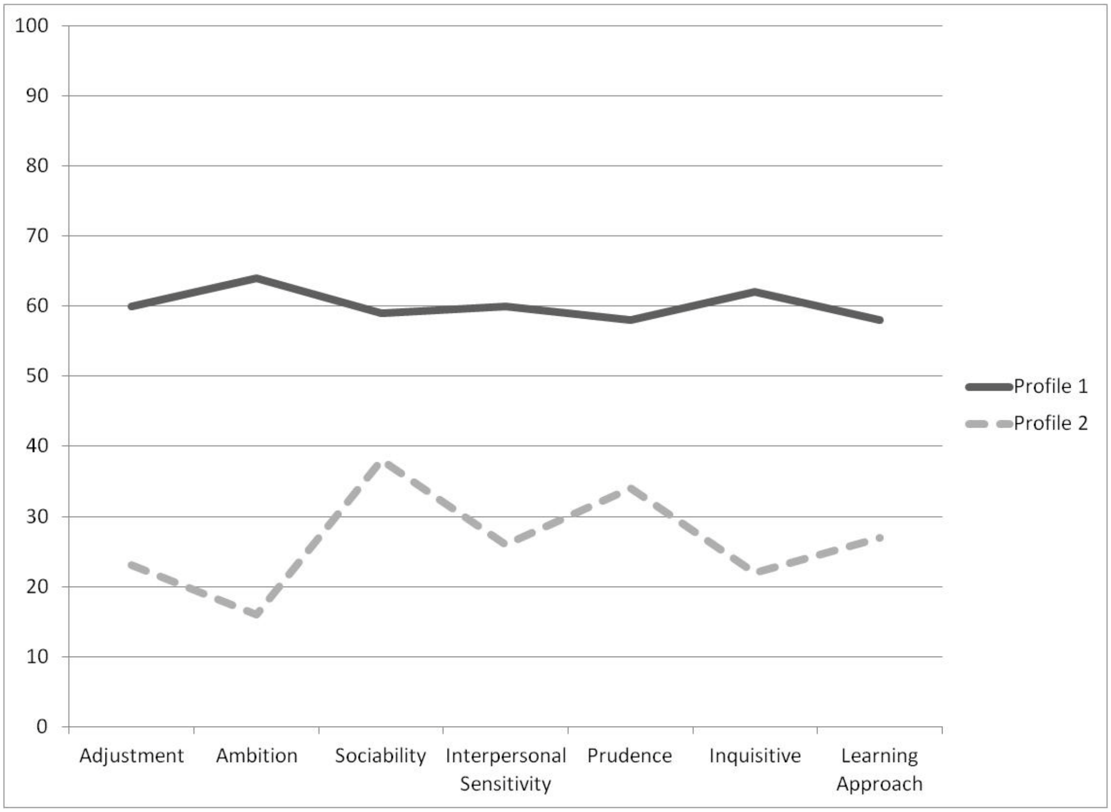


Figure 14 identifies the profiles that emerged when the data from the Entry Level Customer Service and Sales quadrants (Samples 1, 2, 3, 4, 5, and 6) were combined. Again, the primary profile echoes the elevated score profile found previously, with high scores across the HPI scales. The secondary profile is similar to the bulk of the low score secondary profiles discussed previously, and is characterized by low to moderate scores across the HPI scales.

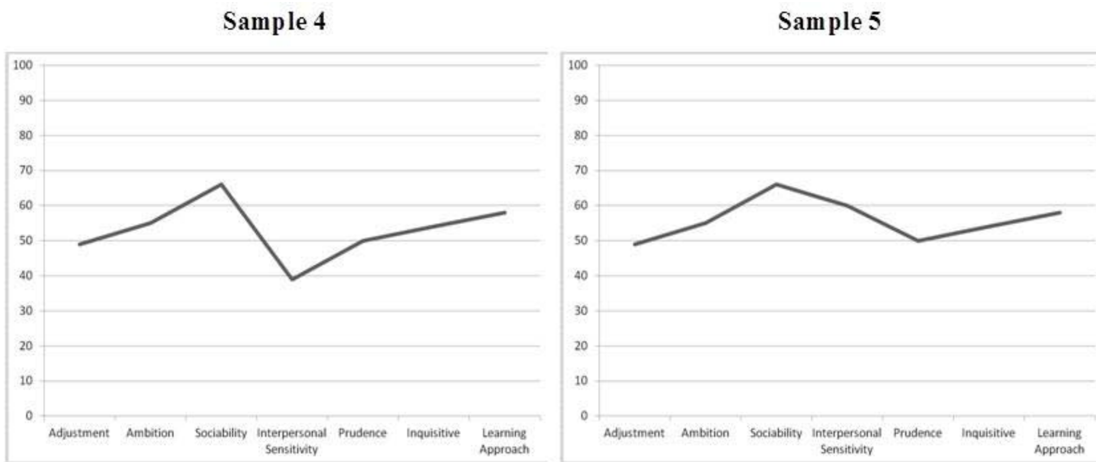
Overall, these results demonstrate that, across samples, industries, and job types, when profiles are identified within the data, two general profiles emerge, one characterized by moderate to high elevations across most or all HPI scales (dubbed here a "elevated score" profile), and the other by low scores across HPI scales (referred to here as a “low score” profile). Implications for how these profiles relate to performance are discussed in later sections.

The second hypothesis concerned whether profiles would generalize within and across conditions.

> *Hypothesis 2:* Profiles will be more generalizable (i.e., proportion classified in cross-validation will be higher) within job levels and industries (i.e., within cells of the 2 x 2 matrix) than (a) across job levels, and (b) across industries.

The results of these analyses do not support H2. A CLPA was conducted on each of the 144 pairings described in Figure 3, as well as within and across industry and level (16 additional pairings), for a total of 160 separate analyses. Each pairing was subject to CLPA, even if profiles were not identified from a particular sample, as the means for the entire sample offer a general profile. This allowed a more complete test of H2 beyond the findings in H1. However, of the 160 pairings tested, only two samples generalized into one another (Samples 4 and 5 from the Entry Level Sales quadrant), and these were two samples where more than one identifiable profile did not emerge (see Figure 15). While not a confirmation of H2, the high degree to which the mean scores from Samples 4 and 5 are similar does suggest that CLPA works as expected, and is a viable method to identify samples where data from one latent profile can generalize to a profile generated from a separate sample. The lack of support for H2 via CLPA is discussed in more detail in the next chapter.

Figure 15. Profiles with Confirmed Generalizability via CLPA



The third hypothesis concerned whether profile membership would offer incremental validity in terms of predicting job performance.

> *Hypothesis 3*: Profile membership (measured along a continuum of fit scores) will demonstrate incremental validity over (a) the linear weighted sum of predictor scales, and (b) other indices of fit or profile similarity, as configured via regression in the prediction of rated job performance.

The results of these analyses do not support H3. As mentioned previously, profile fit was measured via a new fit index, the RRPI. The RRPI does show a moderate relationship with standardized performance for many samples in the expected direction (positive relationships between fit to the primary elevated score profile and performance; see Table 5 for full correlations between the RRPI, D^2, r_{pf},); yet the RRPI added incrementally to the prediction of job performance for only three profiles (Profiles 2 and 4 within the Customer Service industry, and Profile 2 within the Entry Level Customer Service category; see Table 6). These results demonstrate that, on the whole, the RRPI does not add incremental validity beyond other predictors or measures of fit. However, across samples and sample groupings, these results do demonstrate that the RRPI shows moderate relationships with performance on its own, and also relates to other fit indices as expected. The RRPI correlates significantly with D^2 and r_{pf} (r = -.32 and .24, respectively across samples), but the relationships are not so strong as to preclude the RRPI from contributing uniquely to prediction of performance. It might be noted that the correlations between D^2 and the RRPI for the primary profiles are moderately strong, suggesting limited opportunity to operate independently (r = -.80 to -.85). However, even correlations of this magnitude allow correlations in the range of .50 to .60 with other variables (e.g., $.80^2 + .60^2 = 1.00$). Considerations for the usage of the RRPI in future research are discussed in Chapter 5.

Table 5. Correlations between the RRPI, r_{pf}, D^2, and Standardized Performance

Sample 1	ZPerf	r_{pf} Profile 1	r_{pf} Profile 2	D^2 Profile 1	D^2 Profile 2				
RRPI Profile 1	.062**	.621**	.556**	-.840**	.088**				
RRPI Profile 2	-.033	.182**	.275**	-.007	-.867**				
ZPerf	--	.057*	.056*	-.076**	.013				
Sample 3	**ZPerf**	**r_{pf} Profile 1**	**r_{pf} Profile 2**	**r_{pf} Profile 3**	**D^2 Profile 1**	**D^2 Profile 2**	**D^2 Profile 3**		
RRPI Profile 1	.065**	.567**	.220**	.047*	-.829**	-.224**	.297**		
RRPI Profile 2	.033	.334**	.674**	.670**	-.247**	-.854**	-.238**		
RRPI Profile 3	-.068**	-.268**	-.051*	.323**	.276**	-.222**	-.866**		
ZPerf	--	.063**	.043	.035	-.072**	-.036	.060**		
Sample 9	**ZPerf**	**r_{pf} Profile 1**	**r_{pf} Profile 2**	**D^2 Profile 1**	**D^2 Profile 2**				
RRPI Profile 1	.045*	.499**	.094**	-.808**	-.007				
RRPI Profile 2	-.083**	-.335**	.389**	.137**	-.859**				
ZPerf	--	.072**	.018	-.065**	.068**				
Sample 10	**ZPerf**	**r_{pf} Profile 1**	**r_{pf} Profile 2**	**D^2 Profile 1**	**D^2 Profile 2**				
RRPI Profile 1	.058*	.662**	.627**	-.848**	.141**				
RRPI Profile 2	-.050*	-.008	.143**	.125**	-.878**				
ZPerf	--	.059**	.060**	-.077**	.042				
Industry: Sales	**ZPerf**	**r_{pf} Profile 1**	**r_{pf} Profile 2**	**r_{pf} Profile 3**	**D^2 Profile 1**	**D^2 Profile 2**	**D^2 Profile 3**		
RRPI Profile 1	.061**	.595**	.321**	.486**	-.850**	-.307**	.374**		
RRPI Profile 2	.039	.463**	.654**	.561**	-.370**	-.854**	-.184**		
RRPI Profile 3	-.054*	-.200**	-.133**	-.001	.312**	-.04	-.850**		
ZPerf	--	.068**	.059**	.083**	-.079**	-.059**	.050*		
Industry: Service	**ZPerf**	**r_{pf} Profile 1**	**r_{pf} Profile 2**	**r_{pf} Profile 3**	**r_{pf} Profile 4**	**D^2 Profile 1**	**D^2 Profile 2**	**D^2 Profile 3**	**D^2 Profile 4**
RRPI Profile 1	.073**	.597**	.583**	.391**	.100**	-.827**	-.371**	-.166**	.583**
RRPI Profile 2	.02	.207**	.545**	-.044*	.018	-.421**	-.869**	-.246**	-.256**
RRPI Profile 3	.02	.401**	.192**	.630**	.679**	-.248**	-.242**	-.861**	-.300**
RRPI Profile 4	-.080**	-.357**	-.283**	-.150**	.238**	.545**	-.124**	-.170**	-.861**
ZPerf	--	.059**	.075**	.044	.038	-.079**	-.038	-.028	.065**

Note: * = p < .05; ** = p < .01

Table 5, continued. Correlations Between The RRPI, r_{pf}, D^2, and Standardized Performance

Level: Entry Level	ZPerf	r_{pf} Profile 1	r_{pf} Profile 2	D^2 Profile 1	D^2 Profile 2
RRPI Profile 1	.057*	.620**	.449**	-.823**	.132**
RRPI Profile 2	-.052*	-.029	.226**	.131**	-.880**
ZPerf	--	.062**	.053*	-.075**	.041

Entry Level Customer Service	ZPerf	r_{pf} Profile 1	r_{pf} Profile 2	r_{pf} Profile 3	D^2 Profile 1	D^2 Profile 2	D^2 Profile 3
RRPI Profile 1	.068**	.594**	.461**	.276**	-.841**	.013	.523**
RRPI Profile 2	-.023	.209**	.385**	.329**	-.089**	-.876**	-.668**
RRPI Profile 3	-.076**	-.269**	-.131**	.062**	.546**	-.439**	-.823**
ZPerf	--	.059**	.052*	.026	-.076**	.005	.071**

Entry Level Sales	ZPerf	r_{pf} Profile 1	r_{pf} Profile 2	D^2 Profile 1	D^2 Profile 2
RRPI Profile 1	.050*	.592**	.587**	-.840**	.099**
RRPI Profile 2	-.042	-.175**	.122**	.137**	-.842**
ZPerf	--	.062**	.099**	-.075**	.037

Management Customer Service	ZPerf	r_{pf} Profile 1	r_{pf} Profile 2	D^2 Profile 1	D^2 Profile 2
RRPI Profile 1	.060**	.580**	.351**	-.825**	.100**
RRPI Profile 2	-.069**	-.198**	.236**	.194**	-.859**
ZPerf	--	.069**	.046*	-.073**	.060**

Management Sales	ZPerf	r_{pf} Profile 1	r_{pf} Profile 2	D^2 Profile 1	D^2 Profile 2
RRPI Profile 1	.055*	.630**	.424**	-.838**	.121**
RRPI Profile 2	-.041	-.022	.179**	.126**	-.876**
ZPerf	--	.070**	.050*	-.080**	.041

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

Table 6. Summary of Multiple Regression Analysis for Samples and Categories with Two or More Emergent Profiles

Model	*R*	*R Square*	*R Square Change*	*Model*	*R*	*R Square*	*R Square Change*
Sample 1, Profile 1				***Industry 2, Profile 3***			
1	.110	.012	.012	1	.162	.026	.026*
2	.124	.015	.003	2	.166	.028	.001
3	.124	.015	.000	3	.171	.029	.002
Sample 1, Profile 2				***Industry 2, Profile 4***			
1	.127	.016	.016	1	.159	.025	.025*
2	.137	.019	.003	2	.161	.026	.001
3	.200	.040	.021	3	.178	.032	.005*
Sample 3, Profile 1				***Level 1, Profile 1***			
1	.299	.089	.089*	1	.100	.010	.010
2	.314	.099	.009	2	.106	.011	.001
3	.335	.112	.014	3	.106	.011	.000
Sample 3, Profile 2				***Level 1, Profile 2***			
1	.299	.090	.090*	1	.107	.011	.011
2	.314	.098	.009	2	.113	.013	.001
3	.333	.111	.013	3	.131	.017	.004
Sample 3, Profile 3				***Entry Level Customer Service, Profile 1***			
1	.299	.089	.089*	1	.144	.021	.021
2	.313	.098	.009	2	.161	.026	.005
3	.314	.098	.000	3	.172	.029	.003
Sample 9, Profile 1				***Entry Level Customer Service, Profile 2***			
1	.416	.173	.173*	1	.144	.021	.021
2	.427	.182	.009	2	.156	.024	.004
3	.433	.187	.005	3	.194	.037	.013*
Sample 9, Profile 2				***Entry Level Customer Service, Profile 3***			
1	.402	.162	.162*	1	.149	.022	.022
2	.406	.164	.003	2	.160	.026	.004
3	.406	.165	.000	3	.163	.026	.001
Sample 10, Profile 1				***Entry Level Sales, Profile 1***			
1	.175	.031	.031	1	.184	.034	.034
2	.228	.052	.022	2	.192	.037	.003
3	.236	.056	.004	3	.192	.037	.000
Sample 10, Profile 2				***Entry Level Sales, Profile 2***			
1	.176	.031	.031	1	.183	.033	.033
2	.243	.059	.028	2	.187	.035	.002
3	.259	.067	.008	3	.187	.035	.000

* = $p < .05$

Dependent variable: Standardized Overall Performance.

For all samples, Model 1 enters the seven HPI scales in Block 1; Model 2 adds the fit statistics of D^2 and r_{pf} in Block 2; Model 3 adds the RRPI in Block 3.

Table 6, continued. Summary of Multiple Regression Analysis for Samples and Categories with Two or More Emergent Profiles

Model	*R*	*R Square*	*R Square Change*	*Model*	*R*	*R Square*	*R Square Change*
Industry 1, Profile 1				***Management Customer Service, Profile 1***			
1	.204	.041	.041*	1	.259	.067	.067
2	.206	.043	.001	2	.268	.072	.005
3	.207	.043	.000	3	.281	.079	.007
Industry 1, Profile 2				***Management Customer Service, Profile 2***			
1	.186	.035	.035*	1	.252	.064	.064
2	.187	.035	.000	2	.255	.065	.002
3	.187	.035	.000	3	.256	.066	.000
Industry 1, Profile 3				***Management Sales, Profile 1***			
1	.199	.040	.040*	1	.251	.063	.063
2	.200	.040	.000	2	.252	.063	.000
3	.202	.041	.001	3	.252	.063	.000
Industry 2, Profile 1				***Management Sales, Profile 2***			
1	.154	.024	.024*	1	.255	.065	.065
2	.156	.024	.001	2	.256	.066	.001
3	.158	.025	.001	3	.264	.069	.004
Industry 2, Profile 2							
1	.159	.025	.025*				
2	.161	.026	.001				
3	.182	.033	.007*				

* = $p < .05$
Dependent variable: Standardized Overall Performance.
For all samples, Model 1 enters the seven HPI scales in Block 1; Model 2 adds the fit statistics of D^2 and r_{pf} in Block 2; Model 3 adds the RRPI in Block 3.

CHAPTER 5

DISCUSSION

Decades of research have shown the value of personality assessment in the workplace; this study was designed to extend that research, and to expand on the measurement strategies available to personality researchers and practitioners. The current study sought to explore whether personality measurement at the profile level would yield psychologically interpretable profiles, and if found, whether those profiles were generalizable vs. situationally specific, whether they are related to performance on the job, and, if so, whether they add validity incrementally to other indices.

This study is unique in its focus on the duality of non-compensatoriness and optimality. As mentioned previously, most methods for approaching profile analysis are either compensatory, linear, or both. This study targeted non-compensatory, optimality-based methods because they capture a plausible, yet understudied, configural understanding of personality traits based on the FFM. Unlike many constructs (e.g., cognitive ability), the five factors of personality each measure unique attributes and characteristics of an individual. It is unlikely that a high score on one personality scale would compensate for or offset a low score on a different personality scale. Likewise, in terms of profile measurement, and more specifically, the measurement of profile fit, the concept of optimality, defined here as having a score that is near (or at), a target profile point or mean, should be more meaningful than simply being high or low on a given scale, because it is the similarities of those means that indicate a better fit between an individual and a target profile. As such, the non-compensatory, optimality-based methods used in this study represent a departure from mainstream research methods, and an effort to align research methods with how personality profiles are manifested in reality.

Overall, with some qualifications, current hypotheses were not strongly supported. Hypothesis 1, regarding whether profiles would emerge, found modest support, and stronger support at an aggregate level than at the level of individual samples. Hypothesis 2, regarding whether profiles would be generalizable, was not confirmed; likewise, only weak support was found for Hypothesis 3, which tested whether profile fit is a valid predictor of job performance and/or demonstrates incremental validity beyond more established fit indices of D^2 and r_{pf}.

The inconsistent findings with regard to H1 permit several interpretations. Datasets with a mixture of samples were almost consistently confirmed as having latent profiles, while the majority of the single-sample datasets were not. It is possible that this particular method of extracting profiles relies on a certain level or threshold of heterogeneity within a sample for profiles to emerge as distinct enough to be significant. It is also possible that the sample sizes within each sample (N range = 103 to 253) were simply not sufficient, either for the analysis or to create the necessary heterogeneity required by the program.[4] Finally, it is possible that, given a typical sample of incumbents drawn from one job within one organization, only one profile is likely to emerge, given that such a sample is comprised of people who were presumably attracted

[4] Although the Mplus documentation does not provide guidelines for minimum LPA sample sizes, methods for calculating lower-bound sample size requirements in SEM suggest that the sample sizes in this study were adequate (Westland, 2010).

to the same job and company, and have been successful enough and satisfied enough to remain employed in that job. This is in keeping with the Attraction-Selection-Attrition (ASA) theory, which hypothesizes that those attracted to particular organizations are more homogeneous than the applicant pool in general (Schneider, 1987). As such, it is important to consider that, to the extent that an applicant population is more heterogeneous than an incumbent population, conducting similar analyses with applicants rather than incumbents could lead to derivation of more profiles per sample.

The near-complete lack of generalizability found with regard to H2 is compelling, and suggests that, where profiles emerge, be they by company, industry, or job level, they carry a good deal of situational specificity with them. However, it is also possible that the CLPA function found within Mplus is so sensitive that it capitalizes on statistical significance to the exclusion of practical significance. This is discussed further under future research directions.

With regard to H3, multiple regression maximized the predictive value of the HPI within the current dataset. This result suggests that there may be benefits associated with linear, compensatory methods that were not captured in the optimality-based, non-compensatory profile approach used in this study. However, recall that Mplus generated the profiles found in H1 without regard to performance data. This was an important and deliberate first step in this stream of research, which sought to determine whether latent profiles emerged within FFM personality data at all, and if so, whether those profiles are inherently related to performance. That being said, the fact that these profiles were not derived from performance data is important to understanding and interpreting the H3 findings. That is, Mplus did not rely on, account for, or even access, the criterion data available when generating the profile scores that emerged. As such, the emergent profiles discussed in this study can be considered to be representative of the entire sample, category, or grouping to which they pertain, rather than being anchored to any given level of performance.

Alternatively, the additive approach may have out-predicted the other fit indices in this study, and the RRPI in particular, because the methods employed here included all seven HPI scales and treated them all as being equally important. This approach was also deliberate, as this study was a first step in profile exploration with regard to the FFM and job performance, and considering all scales simultaneously seemed a reasonable, if exploratory, starting point. Adding in the complexity of targeted scale selection or weighting could potentially have obscured any overall profile level findings that might have emerged. However, in practice, it is unlikely that every scale on a personality inventory is equally important for profile interpretation or the prediction of performance. As such, calculating fit indices in the manner employed with the RRPI, which (a) is non-compensatory and (b) treats each scale equally, has the potential to actually obscure the true predictive value of a non-compensatory approach. This is because, as noted previously, a miss on any one scale will drastically (and intentionally) reduce an individual's RRPI fit score—even if that one scale is unrelated to job performance. Where every scale is important, this method could still enhance the precision with which high-performing individuals are identified; but where any single scale is irrelevant, this method introduces error that is likely to mask relationships involving profiles based on only job-relevant traits. As such, the RRPI warrants further targeted research as an emergent non-compensatory fit index.

Likewise, as previously discussed, a profile fit approach is predicated on an assumption of *optimality* rather than *linearity*. That is, falling above a profile point mean lowers fit as much as falling below it. Where a given personality construct shares a linear relationship with a performance variable (e.g., where a higher score is better), non-linear approaches such as the RRPI, D^2, and r_{pf}, will fail to account for the linear nature of the predictor-criterion relationship.

Finally, it is possible that this is simply an issue of distinguishing between false positives and false negatives. That is, regression is most effective at identifying "hits" (potential good performers) but may target for selection the occasional candidate who has a favorable score compensating for a poor score; conversely, a profile approach could more effectively weed out "misses" (potential poor performers), but be less effective at identifying acceptable candidates overall. For example, recall the hypothetical data of two candidates, provided in the introduction: one with average scores on Conscientiousness and Emotional Stability, and one with an extremely high Conscientiousness score and an extremely low Emotional Stability score. Assume that both Emotional Stability and Conscientiousness are job-relevant. Only a non-compensatory strategy would be certain to identify the candidate with extreme scores and distinguish him from the candidate with average scores. However, this type of extreme profile would, by normative standards, be very uncommon. As such, research studies may not find value in a non-compensatory approach when analyzing data in aggregate; yet in practice, practitioners who are hiring for a high-stakes or executive-level position may benefit from using a non-compensatory assessment strategy to ensure that any applicant with an uncommon but undesirable profile such as the one described here is screened out.

Limitations

As mentioned, the data used in this research were exclusively drawn from concurrent validation studies involving incumbent samples and the HPI. An applicant population would likely have had more variability, which could have increased the likelihood of profile emergence, and could have also increased the magnitude of observed relationships in all analyses. Additionally, only HPI scale scores were made available. Subscale- or item-level data could open up new avenues for analysis. Larger samples would enable more granular analyses and contribute to more robust conclusions.

The only common performance measure among samples was a supervisor rating of Overall Performance. Facet-level ratings and/or objective performance measures would have added depth to the analyses and could have increased the accuracy and specificity of the current findings. Likewise, job analysis data were not available. These data, particularly data from a personality-oriented job analysis, would facilitate methods requiring predictor-criterion alignment and assist in targeting specific scales in the future.

Finally, LPA and the associated software packages that allow for this type of analysis are in relative infancy with regard to I/O psychology research. As such, it was not possible to determine or adjust the sensitivity of the CLPA procedure. Likewise, the fit score produced by the LPA analyses was not non-compensatory.

Directions for Future Research

Sampling

Future studies should attempt to identify latent profiles within large-sample applicant populations in order to overcome the inherent limitations of working with small incumbent samples (e.g., lack of significant findings; homogeneity amongst participants). Gathering data from applicants rather than incumbents would further facilitate a true predictive design when evaluating any relationships with criterion data.

Sense-making

If profiles emerge in subsequent studies, it would be valuable to assess as a next step, whether conceptualizing personality in terms of profiles contributes to a greater understanding of people and of how personality operates to predict performance in practice. SME ratings could be employed to evaluate whether profiles carry any additional meaning beyond a conglomeration of scale scores, or might contribute to sense-making when laypersons are asked to interpret complex personality data. A draft of an SME survey that was originally planned to collect additional data for the present study, and which might be utilized for this follow-up research, appears in Appendix A. Participants would ideally come from a mix of Industrial/Organizational Psychology, Human Resources, and Business backgrounds, and could be drawn from academic, industry, and government settings. Additionally, assessing whether performance levels are significantly different among participants in given profiles (e.g., by comparing group means via *t*-test or ANOVA) would help provide a simple descriptive context for how profiles differ (or fail to differ) from one another. Future studies could also directly evaluate the optimality versus linearity distinction noted previously. By comparing directly the optimality-based non-compensatory approach employed in this study with another method that is non-compensatory but linear (e.g., interactions testable using moderated regression), future research could help inform whether closeness of fit to a particular profile is important, or whether simply exceeding a given score threshold simultaneously on multiple scales is sufficient.

Researchers may also evaluate whether profiles are differentially useful for particular types of selection strategies. For example, if data from predictive (applicant-based) validation studies are available in the future, including a measure of false positives and negatives (or "hits" and "misses") would allow for the coding and quantification of good / poor / no hires. In selection work, different types of jobs and/or selection strategies may rely on minimizing one type of error over another. For instance, in selecting entry level workers, most companies want to maximize the number of good hires they make, and will accept some number of poor hires in order not to miss a number of good hires. Conversely, when selecting high-level executives, there is little room for error and companies will risk passing over a potential top performer to ensure they do not make a mistake. These two different strategies may call for different types of analysis. Good and poor hires are easily identified through criterion measures, and false positive selection errors refer to hires who perform poorly; false negative errors in selection typically refer to applicants who would have performed adequately but were not hired, and this is a more difficult type of error to capture accurately. However, indirect measures are available that at a minimum, allow researchers to determine the extent to which an organization favors a select-in (target and hire only high-potential candidates) versus select-out (use selection tools to weed out those with very low potential) strategy for a particular opening. For example, measures that account for the time it takes to fill an open position, the stringency or pass rate of a selection battery, and the number of applicants who progress to the final stage of a selection process are all indirect indicators of what type of hiring strategy is in place.

Profiling Methodology – Criteria

As noted, this study did not target or weight individual scales when evaluating fit to a given profile. Understanding the extent to which particular personality traits are relevant to a job is important in assessing the value of personality as a predictor of job performance (Hogan & Holland, 2003; Tett & Christiansen, 2007), and personality-oriented job analyses instruments can help drive the alignment of predictors and criteria (Raymark, Schmit, & Guion, 1997; Tett & Christiansen, 2007). Future research might seek to align specific personality scales with performance constructs, either theoretically or empirically, and then evaluate profile fit with regard to those scales that are expected to predict performance. This type of scale-specific profile analysis could be accomplished via a weighting algorithm, or simply by excluding from the RRPI calculations those predictor scales that are not hypothesized to relate specifically to the criteria and/or population of workers in question.[5]

[5] Although the RRPI could be easily adapted to this sort of research, job analysis data (personality-oriented or otherwise) were not available for the data in the current study, limiting opportunity for follow-up analyses here.

Likewise, researchers may choose to utilize LPA to generate performance-based profiles by restricting the data available to Mplus to high- or low-performing individuals and then evaluating profile emergence based on those constrained samples. This line of performance-based profile research could also begin, much like the current study, by treating all scales of an instrument equally, and then proceed to target specific scales within performance-based profiles (either theoretically or empirically, as described in the previous paragraph).

When considering future research on optimality-based fit methodology, other criteria may warrant attention in conjunction with personality data. For example, job satisfaction, commitment, engagement, and organizational citizenship are all constructs that, conceptually, may relate to a participant's distance, or lack thereof, from an organizational or work group mean. Future studies might include the *Motives, Values, Preferences Inventory* (MVPI; Hogan & Hogan, 1996), a measure of motivators that can help capture organizational culture when evaluated at the group level. The MVPI is also designed to work with the HPI, which may facilitate data gathering or analysis. Cultural profiles may emerge via LPA on the MVPI, and/or MVPI scores may help interpret profiles that emerge—or do not emerge—at the HPI level. Conversely, future research may benefit from employing a different measure of normal personality, or potentially multiple personality tools within an analysis.

Profiling Methodology – Alternatives

Researchers might also reconsider LPA altogether, and instead, evaluate alternatives such as the code type method discussed previously, which has seen success in providing interpretable MMPI profiles (Graham, 2012; McNulty, Ben-Porath, & Graham, 1998). The results found in this study echo the findings of initial code type researchers. That is, when the method for identifying a code type profile became too complex, fewer individuals were meaningfully classified and the profiles themselves related less strongly to behaviors of interest. It is possible that the profile concept is not flawed, but rather that LPA is simply more complex than necessary for assigning people to a given profile. Classifying individuals by their top two or three scores would be a relatively simple next step. Conversely, the mechanisms by which Mplus works and the nuances of the program might become more clear by using Mplus to conduct latent profile analyses on MMPI data for which code types have already been identified. This type of confirmatory strategy could help researchers better understand the potential strengths and limitations of the software for generating interpretable profiles.

Finally, the CLPA function of Mplus may simply be more sensitive than is necessary for these analyses. In the early stages of this study, a manual process for evaluating generalizability was devised, in the event that Mplus did not provide a formal test of generalizability itself. This original method was eschewed in the present study in favor of the parsimony provided by using the built-in Mplus function to test generalizability. However, future research might investigate how the results found via a more manual—and configurable—process might compare to the present findings. The previous method considered for evaluating generalizability rests on using Chi square in conjunction with the Incremental Fit Index (IFI; Bollen, 1989; 1990) to define generalizability via expected classification rates and user-specified decision rules. More information on this method can be found in Appendix B.

Conclusion

The aims of this study were to identify optimality-based personality profiles, assess the generalizability of those profiles within and across levels and industries, and calculate a non-compensatory fit score with which to evaluate the validity and incremental validity of those profiles for predicting job performance. Corresponding hypotheses received mixed support. Where profiles emerged, they were non-generalizable, and were rarely predictive of performance beyond more established methods. Further research should attempt to inform whether non-compensatory, optimality-based approaches such as this one add value in other ways, relate to performance when different methods are employed, or contribute value by assisting in sense-making activities, or conversely, whether this approach is simply more complex than necessary without an associated benefit.

REFERENCES

Aldenderfer, M. S., & Blashfield, R. K. (1984). Cluster analysis. In *Sage University Paper Series on Quantitative Applications in the Social Sciences*, series no. 07-044. Newbury Park, CA: Sage Publications.

Allport, G. W. (1961). *Pattern and growth in personality.* New York: Holt, Rinehart & Winston.

Americans with Disabilities Act of 1990, Pub. L. No. 101-336, §2, 104 Stat. 328 (1991).

Argyris, C. (1957). *Personality and organization: The conflict between system and the individual*. New York: Harper.

Arthur, W., Woehr, D. J., & Graziano, W. G. (2001). Personality testing in employment settings: Problems and issues in the application of typical selection practices. *Personnel Review, 30,* 657-676.

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior, 19*, 289–303.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11,* 150-166.

Ashton, M. C., & Lee, K. (2010). On the cross-language replicability of personality factors. *Journal of Research in Personality, 44,* 436–441.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and 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., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78,* 715–722.

Bartram, D. (2005). The Great Eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90,* 1185–1203.

Bassi, F. (2011). Latent class analysis for marketing scale development. *International Journal of Market Research, 53,* 209-230.

Ben-Porath, Y. S. (2006). Differentiating normal from abnormal personality with the MMPI-2. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (2nd ed., pp. 311-335). New York: Springer.

Bergner, S., Neubauer, A. C., & Kreuzthaler, A. (2010). Broad and narrow personality traits for predicting managerial success. *European Journal of Work & Organizational Psychology, 19*, 177-199.

Bernreuter, R. G. (1931). *The Personality Inventory.* Stanford, CA: Stanford University Press.

Bernreuter, R. G. (1933). The theory and construction of The Personality Inventory. *Journal of Social Psychology, 4,* 387-405.

Burke, L. A., & Witt, L. A. (2002). Moderators of the openness to experience-performance relationship. *Journal of Managerial Psychology, 17,* 712-721.

Burke, L. A., & Witt, L. A. (2004). Personality and high-maintenance employee behavior. *Journal of Business and Psychology, 18,* 349-363.

Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology, 45*, 459-491.

Caldwell, D. F., & O'Reilly, C. A. (1990). Measuring person-job fit with a profile comparison process. *Journal of Applied Psychology, 75,* 648-657.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.

Cascio, W. F., Alexander, R. A., & Barrett, G. (1988). Setting cutoff scores: Legal, psychometric, and professional issues and guidelines. *Personnel Psychology, 41,* 1-24.

Chapman, B. P. (2007). Bandwidth and fidelity on the NEO-Five Factor Inventory: Replicability and reliability of Saucier's (1998) item cluster subcomponents. *Journal of Personality Assessment, 88,* 220-234.

Cheung, G. W. (2009). Introducing the Latent Congruence Model for improving the assessment of similarity, agreement, and fit in organizational research. *Organizational Research Methods, 12,* 6-33.

Christiansen, N., & Robie, C. (2011). Further consideration of the use of narrow trait scales. *Canadian Journal of Behavioural Science, 43,* 183-194.

Civil Rights Act of 1964 § 7, 42 U.S.C. § 2000e *et seq* (1964).

Civil Rights Act of 1991 § 109, 42 U.S.C. § 2000e *et seq* (1991).

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analysis integration of observers' accuracy and predictive validity. *Psychological Bulletin, 6,* 1092-1122.

Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory Manual*. Odessa, FL: Psychological Assessment Resources.

CPP, Inc. (2011). *Myers-Briggs Type Indicator data sheet.* Minneapolis: Author.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin, 50,* 456-473.

De Raad, B., Barelds, D. P. H., Mlačić, B., Church, A. T., Katigbak, M. S., Ostendorf, F., Hřebíčková, M., Di Blas, L., & Szirmák, Z. (2010). Only three personality factors are fully replicable across languages: Reply to Ashton and Lee. *Journal of Research in Personality, 44,* 442-445.

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

Digman, J. M. (1996). The curious history of the five-factor model. The five-factor model of personality: Theoretical perspectives. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 1-20). New York, Guilford.

Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative Science Quarterly, 30,* 514-539.

Dunnette, M. D. (1962). Personnel management. *Annual Review of Psychology, 13,* 285-314.

Edwards, J. R. (1993). Problems with the use of profile similarity indices in the study of congruence in organizational research. *Personnel Psychology, 46,* 641-665.

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice (1978). *Uniform guidelines on employee selection procedures. Federal Register, 43,* 38,290–38,315.

Ferris, G. L., Witt, L. A., & Hochwarter, W. A. (2001). Interaction of social skill and general mental ability on job performance and salary. *Journal of Applied Psychology, 86,* 1075-1082.

Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group differences in personality: Meta-analyses comparing five U.S. racial groups. *Personnel Psychology, 61,* 579-616.

Foster, J., & Macan, T. (2006). *The use of interactions between personality variables to predict performance.* Paper presented at the Twenty-first Annual Meeting of the Society for Industrial and Organizational Psychology, Dallas, Texas.

Foster, J., Gaddis, B., & Hogan, J. (2012). Personality-based job analysis. In A. Wilson, W. Bennett Jr., S. G. Gibson, & G. M. Alliger (Eds.), *The handbook of work analysis: Methods, systems, applications and science of work measurement in organizations* (pp. 247-264). New York: Routledge.

Frei, R., & McDaniel, M. A. (1998). The validity of customer service orientation measures in employee selection: A comprehensive review and meta-analysis. *Human Performance, 11,* 1-27.

Gatewood, R. D., & Feild, H. S. (1998). *Human resource selection (4th ed.)*. Orlando, FL: The Dryden Press.

Gibby, R. E., & Zickar, M. J. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. *History of Psychology, 11,* 164-184.

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

Goldberg, L. R. (1992). The development of markers for the Big Five factor structure. *Psychological Assessment*, *4*, 26–42.

Gough, H. G. (1987). *California Psychological Inventory administrator's guide.* Palo Alto, CA: Consulting Psychologists Press.

Graham, J. R. (2012). *MMPI-2: Assessing personality and psychopathology (5th ed.)*. New York: Oxford University Press.

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 8*, 135-164.

Hathaway, S. R., &McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory.* Minneapolis: University of Minnesota Press.

Healy, M. C., & Handler, C. A. (2009). Rocket Hire. *2009 Rocket-Hire pre-employment assessment usage survey: Complete findings*. Retrieved March 20, 2011 from http://rocket-hire.com/_pdf/whitepapers/Rocket-Hire-Assessment-Usage-Survey-2009.pdf

Herman, K., Ostrander, R., Walkup, J., Silva, S., & March, J. (2007). Empirically derived subtypes of adolescent depression: latent profile analysis of co-occurring symptoms in the Treatment for Adolescents with Depression Study (TADS). *Journal of Consulting and Clinical Psychology, 75,* 716-728.

Hochwarter, W. A., Witt, L. A., Treadway, D. C., & Ferris, G. R. (2006). The interaction of social skill and organizational support on job performance. *Journal of Applied Psychology, 91*, 482-489.

Hogan, J., & Hogan, R. (1996). *Motives, Values, Preferences Inventory*. Tulsa, OK: Hogan Assessment Systems.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88,* 100-112.

Hogan, R. & Smither, R. (2001). *Personality: Theories and applications.* Boulder, CO: Westview Press.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *1982 Nebraska Symposium on Motivation* (pp. 55-89). Lincoln: University of Nebraska Press.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 163-179). New York: Guilford Press.

Hogan, R., & Hogan, J. (1995). *Hogan Personality Inventory Manual* (2nd ed.). Tulsa, OK: Hogan Assessment Systems.

Hogan, R., & Roberts, B. (2001). Introduction: Personality and industrial and organizational psychology. In B. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 3-16). Washington, DC: American Psychological Association.

Hogan, R., Hogan, J., & Roberts, B. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, *51*(5), 469–477.

Holland, B. (2001). *Specificity of prediction: Beyond the Big-Five job performance relations.* Unpublished doctoral dissertation: The University of Tulsa.

Holland, B., & Van Landuyt, C. (2007). Evaluating Non-Compensatory Cutoffs in Personality-Based Selection. In P. M. Mangos (chair), *Cut score development as an extension of the validation process.* Symposium presented at the Twenty-second Annual Meeting of the Society for Industrial and Organizational Psychology, New York, New York.

Holland, B., Hogan, J., & Van Landuyt, C. (2002). *How to measure sociopolitical IQ.* Paper presented at the Seventeenth Annual Meeting of the Society for Industrial and Organizational Psychology, Toronto, Canada.

Hotard, S. R., McFatter, R. M., McWhirter, R. M., Stegall, M. E. (1989). Interactive effects of extraversion, neuroticism, and social relationships on subjective well-being. *Journal of Personality and Social Psychology, 57*, 321-331.

Hough, L. M. (1992). The "Big Five" personality variables—Construct confusion: Description versus prediction. *Human Performance, 5,* 139-155.

Hough, L. M. (1998). Personality at work: Issues and evidence. In M. Hake1 (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131-159). Hillsdale, NJ: Erlbaum.

Hough, L. M. (2001). I/Owes its advances to personality. In B. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 19-44). Washington, DC: American Psychological Association.

Hunsley J., Lee, C. M., & Wood, J. M. (2004). Controversial and questionable assessment techniques. In S. O. Lilenfeld, J. M. Lohr, & S. J. Lynn (Eds.), *Science and Pseudoscience in Clinical Psychology* (pp. 39-76). New York: Guilford.

Hurtz, G. M., & Donovan, J. J., (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85,* 869-879.

John, O. P. (1990). The “Big-Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality theory and research* (pp. 66–100). New York: Guilford.

John, O. P., Angleitner, A., & Ostendorf, F. (2006). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality, 4,* 89-118.

Jung, C. G. (1921). *Psychological Types*. (H. G. Baynes, 1971 Trans.). Princeton, NJ: Princeton University Press.

Karraker v. Rent-A-Center, Inc., 411 F.3d 831 (7th Cir. 2005).

Kaufman, L., & Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis*. New York: Wiley Interscience.

Kostman, J. T. (2004). *Multi-dimensional performance requires multi-dimensional predictors: Predicting complex job performance using cognitive ability, personality and emotional intelligence assessment instruments.* Unpublished doctoral dissertation, The City University of New York.

Kupzyk, K. A. (2011). *Introduction to mixture modeling.* Paper presented at the National Center for Research on Rural Education (R2Ed) and Nebraska Center for Research on Children, Youth, Families and Schools (CYFS), University of Nebraska-Lincoln.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent Structure Analysis.* Boston: Houghton Mifflin.

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88,* 767-778.

Lundholm, H. (1940). Logic and beauty. *Journal of Personality, 8*(4), 281-291.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19-40.

Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: A comparison with K-Means. *Canadian Journal of Marketing Research, 20,* 37-44.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences* (Chapter 10, pp. 175-198). Thousand Oakes: Sage Publications.

Martin, S. L., & Raju, N. S. (1992). Determining cutoff scores that optimize utility: A recognition of recruiting costs. *Journal of Applied Psychology, 77*(1), 15-23.

Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin, 113,* 181–190.

McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49,* 710-721.

McCrae, R. R., & Costa, P. T. (1987). Validity of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90.

McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the Five-Factor Model of personality. *Journal of Personality, 57,* 17-40.

McDaniel, M. A., Kepes, S., & Banks, G. C. (2011). The Uniform Guidelines are a detriment to the field of personnel selection. *Industrial and Organizational Psychology, 4,* 494-514.

McLachlan, G. & Peel, D. (2000). *Finite mixture models.* New York: Wiley.

McLachlan, G., & Chang, S. (2004). Mixture modeling for cluster analysis. *Statistical Methods in Medical Research, 13,* 347-361.

McNulty, J. L., Ben-Porath, Y. S., & Graham, J. R. (1998). An empirical examination of the correlates of well-defined and not defined MMPI-2 code types. *Journal of Personality Assessment, 71,* 393-410.

Mead, G. H. (Ed. C. W. Morris). (1934). *Mind, self, and society*. Chicago: University of Chicago.

Merz, E. L., & Roesch, S. C. (2011). A latent profile analysis of the Five Factor Model of personality: Modeling trait interactions. *Personality and Individual Differences, 51,* 915-919.

Milligan, G. E., & Cooper, M. C. (1985). Examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50,* 159–179.

Mischel, W. (1968). *Personality and assessment.* Wiley, New York.

Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N. Endler (Eds.), *Personality at the Crossroads: Current Issues in Interactional Psychology* (pp. 333-352). Hillsdale, NJ: Erlbaum.

Morgan, C. D., & Murray, H. A. (1935). A method of investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry, 34,* 289–306.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60,* 683–729.

Mõttus, R. (2006). Investigating personality in different levels of cognitive ability. *Baltic Journal of Psychology, 7,* 60-67.

Mount, M. K., & Barrick, M. R., (2001). *Manual for the Personal Characteristics Inventory*. Libertyville, IL: Wonderlic Personnel Test, Inc.

Muthén, B. O. (2004). *Mplus technical appendices.* Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed.). Palo Alto: Consulting Psychologists Press.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, *66*, 574–583.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Nylund, K. L., Asparouhov. T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535-569.

Oh, I. S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the Five-Factor Model of personality traits: A meta-analysis. *Journal of Applied Psychology, 96,* 762-773.

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17,* 609-626.

Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology, 60,* 995–1027.

Ones, D. S., Mount, M. K., Barrick, M. R., & Hunter, J. E. (1994). Personality and job performance: A critique of the Tett, Jackson, and Rothstein (1991) meta-analysis. *Personnel Psychology, 47,* 147-156.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78,* 679-703.

Osgood, C. E., & Suci, G. J. (1952). A measure of relation determined by both mean differences and profile information. *Psychological Bulletin, 49,* 251-262.

Ostrander, R., Herman, K., Sikorski, J., Mascendaro, P., & Lambert, S. (2008). Patterns of psychopathology in children with ADHD: A latent profile analysis. *Journal of Clinical Child & Adolescent Psychology, 37,* 833–847.

Pastor, D., Barron, K., Miller, B., & Davis, S. (2006). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology, 32*(1), 8-47.

Paul, A. M. (2004). *The cult of personality: How personality tests are leading us to miseducate our children, mismanage our companies, and misunderstand ourselves.* New York, Simon & Schuster, Inc.

Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior, 20,* 389-405.

Pearson Assessments (2011). *7th Circuit Appeals Court Ruling in the Karraker v. Rent-A-Center case confirms the appropriate use of the MMPI-2™ test in employment settings.* Retrieved April 17, 2011 from http://www.reedpetersen.com/portfolio/pe/pa/resources/appealscase.htm.

Penney, L. M., David, E., & Witt, L. A. (2011). A review of personality and performance: Identifying boundaries, contingencies, and future research directions. *Human Resource Management Review, 21,* 297-310.

Perry, S. J., Dubin, D. F., & Witt, L. A. (2010). The interactive effect of extraversion and extraversion dissimilarity on exhaustion in customer-service employees: A test of the asymmetry hypothesis. *Personality and Individual Differences, 48*, 634-639.

Perry, S. J., Hunter, E. M., Witt, L. A., & Harris, K. J. (2010). P = f (conscientiousness x ability): Examining the facets of conscientiousness. *Human Performance, 23*, 343-360.

Peters, L. H., Greer, C. R., & Youngblood, S. A. (1998). *Blackwell encyclopedic dictionary of human resource management.* Ipswich, MA: Blackwell Publishing Ltd.

Pittenger, D. J. (1993). Measuring the MBTI…and coming up short. *Journal of Career Planning and Placement, 54,* 48-52.

Rao, C. R. (1965). *Linear statistical inference and its applications.* New York: John Wiley & Sons.

Raymark, P. H., Schmit, M. J., & Guion, R. M. (1997). Identifying potentially useful personality constructs for employee selection. *Personnel Psychology, 50,* 723–736.

Roberts, B. W., Chernyshenko, O. S., Stark, St., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58,* 103-139.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review, 16,* 155-180.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*(1).

Ruiz, G. (2006). Use care when conducting pre-employment tests. *Workforce Management*. Retrieved from http://www.workforce.com/section/news/article/use-care-conducting-pre-employment-tests.html

Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies: A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology, 49,* 549-572.

Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist, 11,* 929-954.

Sackett, P. R., Gruys, M. L., & Ellingson, J. E. (1998). Ability–personality interactions when predicting job performance. *Journal of Applied Psychology*, *83*, 545–556.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82,* 30-43.

Saucier, G. & Goldberg, L. R. (1996a). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality, 10*, 61-77.

Saucier, G. & Goldberg, L. R. (1996b). The language of personality; Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 21-50). New York: Guilford Press.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37,* 407-422.

Schneider, B. (1987). The people make the place. *Personnel Psychology, 40,* 437-453.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15,* 201-293.

Tahmincioglu, E. (2011). Employers turn to tests to weed out job seekers. *msnbc.com.* Retrieved from http://www.msnbc.msn.com/id/44120975/ ns/business-personal_finance/t/employers-turn-tests-weed-out-job-seekers/#.TmF9zl28-uI

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88,* 500-517.

Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt. *Personnel Psychology, 60,* 967–993.

Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality, 34,* 397-423.

Tett, R. P., & Murphy, P. J. (2002). Personality and situations in coworker preference: Similarity and complementary in worker compatibility. *Journal of Business and Psychology, 17,* 223-241.

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.

Tett, R. P., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1994). Meta-analysis of personality-job performance relations: A reply to Ones, Mount, Barrick, and Hunter (1994). *Personnel Psychology, 47,* 157-172.

Tett, R. P., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1999). Meta-analysis of bidirectional relations in personality-job performance research. *Human Performance, 12,* 1-29.

Tett, R. P., Steele, J. R., & Beauregard, R. S. (2000). *Broad and narrow measures on both sides of the personality-job performance relationship*. Paper presented in J. Hogan (Chair), Specificity versus generality in personality-job performance linkages: Data speak louder than words. Symposium conducted at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Thompson, R. (2013). No, we haven't been "duped" by the world's most popular personality assessment [Blog post]. Retrieved from www.cppblogcentral.com/

cpp-connect/no-we-havent-been-duped-by-the-worlds-most-popular-personality-assessment/

Thumin, F. J. (1969). MMPI scores as related to age, education, and intelligence among male job applicants. *Journal of Applied Psychology, 53,* 404-407.

Thumin, F. J. (1971). A comparative study of the MMPI profiles of salesmen and technical managers. *Personnel Psychology, 24,* 481-487.

Thumin, F. J. (2002). Comparison of the MMPI and MMPI-2 among job applicants. *Journal of Business & Psychology, 17,* 73-86.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, *41*, 1–32.

Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (ASD-TR-61-97). Lackland Air Force Base, TX: Aeronautical Systems Division, Personnel Laboratory.

Van Landuyt, C. (2004). *Albemarle Paper Co. v. Moody: Implications for the field of Industrial/Organizational Psychology in practice.* Unpublished manuscript, The University of Tulsa, Tulsa, OK.

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods, 17,* 228-243.

Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57,* 307-333.

Watson, G. B., & Glaser, E. M. (1980). *WGCTA Watson-Glaser Critical Thinking Appraisal Manual: Forms A and B.* San Antonio: The Psychological Corporation.

Wernimont, P. & Campbell, J. (1968). Sign, sample, and criteria. *Journal of Applied Psychology, 52,* 372-376.

Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications, 9,* 476-487.

Whittaker, T. A., & Stapleton, L. M. (2006). The performance of cross-validation indices used to select among competing covariance structure models under multivariate nonnormality conditions. *Multivariate Behavioral Research, 41,* 295-335.

Wiggins, J. S., & Pincus, A. L. (1992). Personality structure and assessment. *Annual Review of Psychology*, *43*, 473–504.

Witt, L. A. (2002). The interactive effects of extraversion and conscientiousness on performance. *Journal of Management, 28*, 835-851.

Witt, L. A., & Ferris, G. R. (2003). Social skill as moderator of the conscientiousness-performance relationship: Convergent results across four studies. *Journal of Applied Psychology, 88,* 809-820.

Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology, 87,* 164-169.

APPENDIX A

SURVEY MATERIALS FOR FUTURE RESEARCH

Draft of SME Survey 1

About You

Age ____________

Gender ____________

Ethnicity ____________

*Profession

- o Industry, Govt., Consulting
- o Academic
- o Student
- o Other

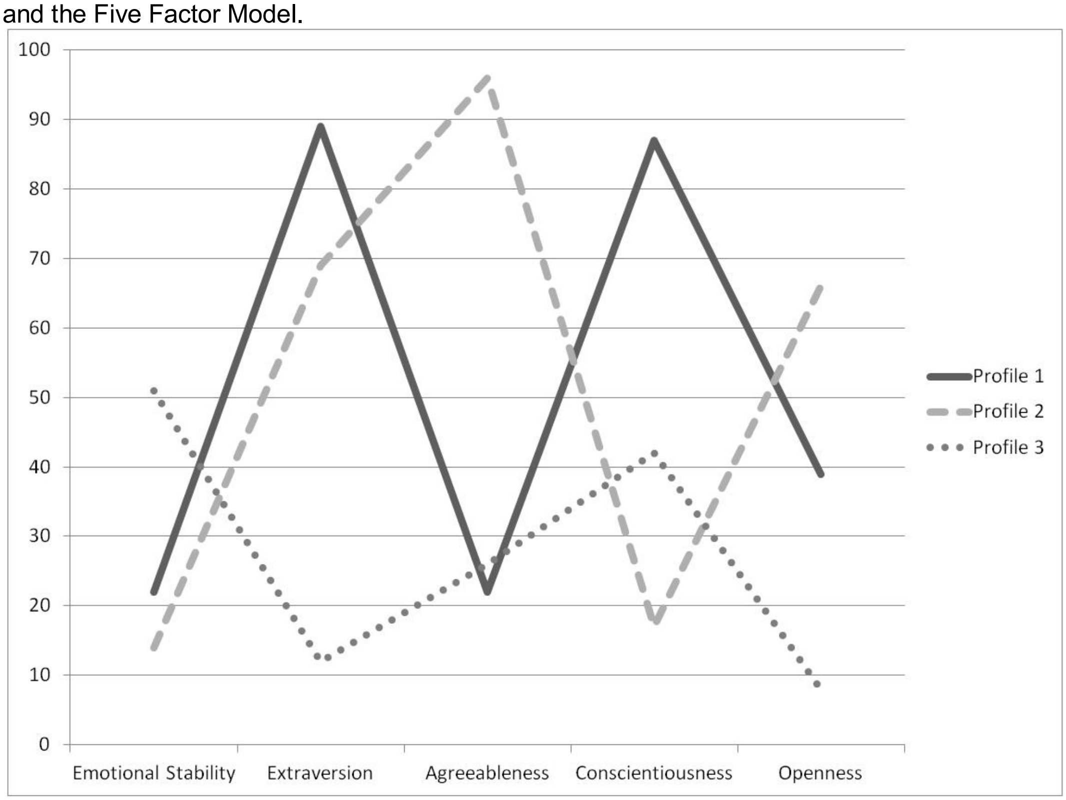
*Years in Profession __________

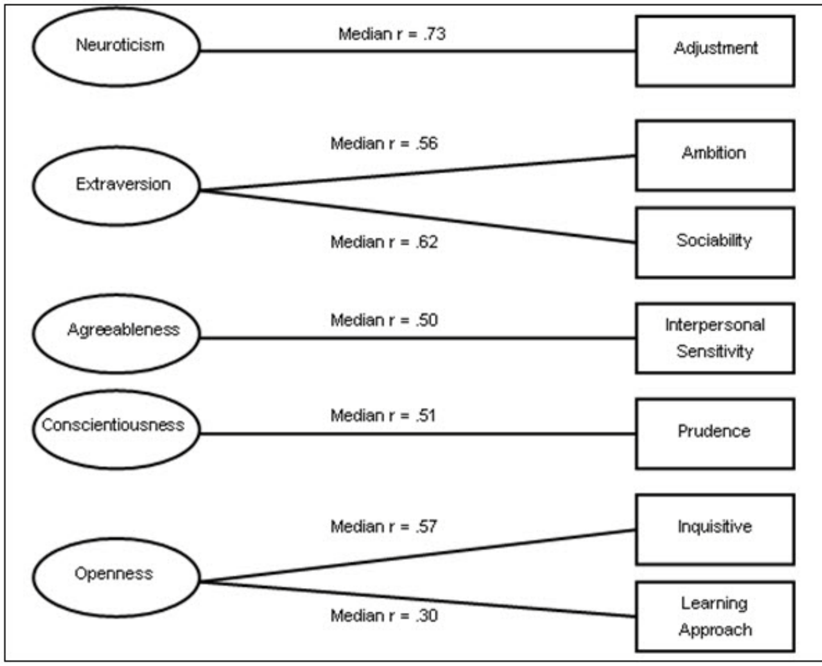
*Years experience working with FFM ____________

**Required Fields*

Rating Tasks

Review the numbered profiles shown below, along with the associated scale definitions provided. The profiles shown here are derived from HPI scores, meaning that the Extraversion score is comprised of Ambition and Sociability, and the Openness score is comprised of the Inquisitive and Learning Approach scales. As you are considering the scales and their definitions, it may also be useful to review or print the figure on the following page, which shows the relationships between the HPI scales and the Five Factor Model.





The **Adjustment** scale reflects the degree to which a person is steady in the face of pressure, or conversely, moody and self-critical (Five-Factor Model: Emotional Stability).

The **Ambition** scale evaluates the degree to which a person seems leader like, seeks status, and values achievement (Five-Factor Model: Extraversion).

The **Sociability** scale assesses the degree to which a person needs and/or enjoys social interaction (Five-Factor Model: Extraversion).

The **Interpersonal Sensitivity** scale reflects social sensitivity, tact, and perceptiveness (Five- Factor Model: Agreeableness).

The **Prudence** scale concerns self-control and conscientiousness (Five-Factor Model: Conscientiousness).

The **Inquisitive** scale reflects the degree to which a person seems imaginative, adventurous, and analytical (Five-Factor Model: Intellect/Openness).

The **Learning Approach** scale reflects the degree to which a person enjoys academic activities and values education as an end in itself (Five-Factor Model: Intellect/Openness).

Next, for each profile shown, think about how a person matching this profile might behave at work. Think about their everyday reputation, habits, strengths, and weaknesses.

Now, for each profile, develop a descriptive phrase or label of your own choosing. Try to think of a label that will be meaningful to yourself and others. In other words, develop a label that when used, would immediately call to mind a very specific type of person who habitually displays a circumscribed set of behaviors.

For example, Burke and Witt (2004) note that repeated demonstrations of the behaviors associated with individuals who are high on Conscientiousness and low on Agreeableness “will likely evoke undesirable labels from peers and supervisors (i.e., "the supreme whiner," "the thorn in our side," "the one who's never pleased," "the person who gets under everyone's skin," etc.) “ (p. 359). Labels need not be this colorful if not warranted though; “Talkative” or “Delinquent” may suffice.

Please bear the following guidelines in mind as you develop your labels:

- Above all, labels should be free from vulgarity, slurs, or stereotypes aimed at any protected group or class.
- Labels should be succinct. A few words should suffice.
- Labels need not be negative. Profiles may describe problem employees like Burke and Witt note, but may also describe particular types of high performers or strong organizational citizens, or individuals with a unique set of behaviors that may be independent of job performance altogether.

Please enter your profile labels in the spaces below.

1. __________
2. __________
3. __________
4. __________

Next, please assign the profile that you believe would best predict performance to each combination of industry and job level below. Simply fill in each blank with the number from a profile above.

Manager: Sales	_______	Manager: Service	_________
Individual Contributor: Sales	_______	Individual Contributor: Service	_________

Finally, please take a moment to think about working with personality profiles.

If asked to hypothesize about the validity of using personality profiles for making employment decisions (i.e., using the extent to which an individual fits a given profile via a fit score or similar indices), I think the validity would be:	○ Better ○ Worse ○ About the same
When thinking about how to make sense out of personality data, I think that profiles would be:	○ More Useful ○ Less Useful ○ About the same
In comparison to working with individual scales of the FFM, I found thinking about personality data from a profile perspective to be:	○ Easier ○ Harder ○ About the same

Thank you very much for your assistance!

APPENDIX B

MANUAL GENERALIZABILITY TEST

The significance of the differences between the cross-validated classification percentages, per sample pairing, may be evaluated via chi-square analysis in terms of both specific profiles and overall classification rate; Chi square is defined as follows:

$$\chi^2 = \Sigma(O - E)^2 / E$$

where O, in the present case, = observed frequency of classification in cross-validation, and E = the expected frequency of classification in cross-validation. The expected values are taken as the proportion of cases classified in the derivation sample (e.g., Sample 1) applied to the cross-validation sample (e.g., Sample 2). For example, if 80 of 300 cases are classified to Profile 1 in Sample 1, this represents a 26.7% classification rate. Applying this percentage to $N = 200$ in Sample 2 would yield an expected classification of 26.7% x 200 = 53.33 cases into Profile 1 (derived from Sample 1). As classification is dichotomous (i.e., classified versus not classified), $df = 2 - 1 = 1$, yielding a critical value of 3.84 ($p < .05$). A significant chi-square would indicate poor generalizability of Profile 1 from Sample 1 to Sample 2. The same procedure can be applied to overall classification rate (i.e., into any profile, not just a specific profile).

However, chi square is known to be sensitive to sample size (favoring lack of generalizability of classification at the sample level, where N is larger). Accordingly, fit will also be assessed using the Incremental Fit Index (IFI; Bollen, 1989; 1990), defined as

$$IFI = (\chi^2_{null} - \chi^2_{model}) / (\chi^2_{null} - df_{model}),$$

where χ^2_{null} is chi-square with 0 cases classified in cross-validation, χ^2_{model} is the chi-square with the observed cases classified, and $df_{model} = 2 - 1 = 1$. The advantage of the IFI and similar fit indices is that they control for sample size, allowing fit at the profile level to be compared directly to fit at the overall sample level (within samples), and across samples, regardless of differences in *N*. Also, they produce fit indices approximating a 0 to 1 range.[6] The value marking acceptable fit with the IFI is subjective, but a cutpoint of .90 is often used; some argue for .95. Regardless, the IFI permits use as a relative effect size with which to evaluate the independent and joint effects of within versus between level and industry on profile generalizability. The higher the IFI, the higher the generalizability.

[6] The IFI can be greater than 1, but such cases are rare and discrepancies from 1 are small in magnitude.