7-30-2014

The Validity of Broad and Narrow Personality Traits For Predicting Job Performance: The Differential Effects of Time

Michael B. Harari

mhara003@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

THE VALIDITY OF BROAD AND NARROW PERSONALITY TRAITS FOR

PREDICTING JOB PERFORMANCE: THE DIFFERENTIAL EFFECTS OF TIME

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Michael B. Harari

2014

To: Interim Dean Michael R. Heithaus
    College of Arts and Sciences

This dissertation, written by Michael B. Harari and entitled The Validity of Broad and Narrow Personality Traits For Predicting Job Performance: The Differential Effects of Time, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Hock-Peng Sin

_____
Stefany Coxe

_____
Valentina Bruk-Lee

_____
Chockalingam Viswesvaran, Major Professor

Date of Defense: July 30, 2014

The dissertation of Michael B. Harari is approved.

_____
Interim Dean Michael R. Heithaus
College of Arts and Sciences

_____
Dean Lakshmi N. Reddi
University Graduate School

Florida International University 2014

DEDICATION

In memory of my mother, Rina Ann Harari.

ABSTRACT OF THE DISSERTATION

THE VALIDITY OF BROAD AND NARROW PERSONALITY TRAITS FOR

PREDICTING JOB PERFORMANCE: THE DIFFERENTIAL EFFECTS OF TIME

by

Michael B. Harari

Florida International University, 2014

Miami, Florida

Professor Chockalingam Viswesvaran, Major Professor

Research into the dynamicity of job performance criteria has found evidence

suggesting the presence of rank-order changes to job performance scores across time as

well as intraindividual trajectories in job performance scores across time. These findings

have influenced a large body of research into (a) the dynamicity of validities of individual

differences predictors of job performance and (b) the relationship between individual

differences predictors of job performance and intraindividual trajectories of job

performance. In the present dissertation, I addressed these issues within the context of

the Five Factor Model of personality. The Five Factor Model is arranged hierarchically,

with five broad higher-order factors subsuming a number of more narrowly tailored

personality facets. Research has debated the relative merits of broad versus narrow traits

for predicting job performance, but the entire body of research has addressed the issue

from a static perspective -- by examining the relative magnitude of validities of global

factors versus their facets. While research along these lines has been enlightening,

theoretical perspectives suggest that the validities of global factors versus their facets

may differ in their stability across time. Thus, research is needed to not only compare the relative magnitude of validities of global factors versus their facets at a single point in time, but also to compare the relative stability of validities of global factors versus their facets across time. Also necessary to advance cumulative knowledge concerning intraindividual performance trajectories is research into broad vs. narrow traits for predicting such trajectories. In the present dissertation, I addressed these issues using a four-year longitudinal design. The results indicated that the validities of global conscientiousness were stable across time, while the validities of conscientiousness facets were more likely to fluctuate. However, the validities of emotional stability and extraversion facets were no more likely to fluctuate across time than those of the factors. Finally, while some personality factors and facets predicted performance intercepts (i.e., performance at the first measurement occasion), my results failed to indicate a significant effect of any personality variable on performance growth. Implications for research and practice are discussed.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I: INTRODUCTION

Job performance is a centrally important construct in Industrial-Organizational (I-O) Psychology (Campbell, 1990; Viswesvaran & Ones, 2000). For instance, interventions derived from I-O Psychology research and theory, such as personnel selection, are intended to influence individual job performance levels (Schmidt & Hunter, 1998). However, despite the importance of the job performance construct to the science and practice of I-O Psychology, there has been considerable trouble in its conceptualization and measurement. These issues have jointly been referred to as the *criterion problem* and have been an area of substantial research and theoretical development (Austin & Villanova, 1992).

Research into the criterion problem has considered a number of issues, such as the relative merits of different methods by which job performance can be assessed (e.g., organizational records vs. subjective judgments; Viswesvaran & Ones, 2000). Research into the criterion problem has also considered issues pertaining to the dimensionality of job performance criteria and many models that specify the lower-order structure of job performance have been proposed. A model discussed by Borman and Motowidlo (1993) distinguishes between task performance (i.e., proficiency in technical activities) and contextual performance (i.e., discretionary behaviors that contribute to organizational effectiveness). Campbell (1990) proposed an eight-dimension model, while Viswesvaran and colleagues (Viswesvaran, 1993; Viswesvaran, Ones, & Schmidt, 1996) proposed ten-dimension model. Furthermore, while research had indicated that job performance could be characterized according to a wide number of narrow dimensions, Viswesvaran,

Schmidt, and Ones (2005) reviewed evidence indicating a positive manifold across performance dimensions that could be accounted for by a general factor.

As illustrated in the preceding review, the criterion problem encompasses a number of distinct issues. However, the most relevant of these issues for purposes of the present dissertation concerns the temporal dimensionality of job performance, also referred to as dynamic criteria. I review the dynamic criteria issue in the following section.

Dynamic Criteria

Dynamic criteria refers to within-person changes in job performance across time (Ghiselli, 1956). The dynamic criteria issue has been studied in a number of different ways, for instance, by assessing group mean-level changes in criterion scores across time, by assessing rank-order changes to criterion scores across time, and by assessing intraindividual trajectories in job performance across time (Barrett, Caldwell, & Alexander, 1985; Hofmann, Jacobs, & Baratta, 1993; Hofmann, Jacobs, & Gerras, 1992). Broadly speaking, the results of the vast literature dedicated to examining the dynamic criteria issue has identified evidence suggesting that job performance changes across time, that the extent and manner in which job performance changes across time varies between people, and also suggests the need for future research in the area (Ployhart & Hakel, 1998; Sturman, Cheramie, & Cashen, 2005).

A number of theoretical perspectives have been developed to account for changes to job performance across time. The *changing subjects model* suggests that job performance changes across time because individual characteristics (e.g., knowledge) change across time. As such characteristics are conceptualized as causal antecedents of

job performance (e.g., Schmidt, Hunter, & Outerbridge, 1986), changes to these characteristics result in changes to job performance (Adams, 1957; Alvarez & Hulin, 1973). A complementary perspective is the *changing tasks model*, which suggests that job performance changes across time because job requirements are not stable. As job requirements change, so do the determinants of job performance, and rank-order changes in job performance scores would follow.

Research has also sought to account for changes to job performance across time by developing models that specify different job stages. One such model is the *maintenance-transition stage model*, which distinguishes between two job stages -- transition and maintenance stages (Murphy, 1989). These two stages can be differentiated based on the demands placed on employees in each. A *transition stage* occurs when employees experience changes to responsibilities and therefore face demands pertaining to learning and problem solving. A *maintenance stage* occurs when employees have overlearned their work roles and therefore face demands pertaining to sustained execution of well-learned behaviors. As employees progress through varying degrees of each of these stages, the determinants of job performance change, which results in changes to job performance across time.

Another job stage model is the *honeymoon-hangover model*. Similar to the maintenance-transition stage model, the honeymoon-hangover model specifies two periods. The *honeymoon period* occurs when employees are new to the job and experience heightened commitment and motivation, while a *hangover period* occurs when the novelty of the new position wears off and employees are no longer maximally motivated (Helmreich, Sawin, & Carsrud, 1986). The model proposes that determinants

of performance should vary between the two stages. Specifically, as motivation is maximized in the honeymoon period, performance should be largely determined by abilities. Conversely, when the motivation associated with the novelty of the new position begins to fade, individual differences that pertain to motivation should influence performance (Minbashian, Earl, & Bright, 2013). As the determinants of performance should vary between the two stages, job performance should vary as well.

The empirical findings and theoretical perspectives concerning performance change across time bear implications for the conceptualization of job performance and for personnel selection practices. For instance, rank-order changes in criterion scores across time may result in changes to criterion-related validity coefficients of predictors and thus, the long term utility of selection procedures may be compromised (Barrett et al., 1985; Hulin, Henry, & Noon, 1990; Keil & Cortina, 2001). Furthermore, the presence of between-person differences in performance trajectories suggests utility in assessing individual differences predictors of such performance trajectories (Deadrick, Bennett, & Russell, 1997; Ployhart & Hakel, 1998; Thoresen, Bradley, Bliese, & Thoresen, 2004). A number of studies have addressed these issues with a particular emphasis on cognitive predictors, however, as the use of personality predictors consistent with the Five Factor Model has become prevalent in the I-O literature (Hough & Ones, 2001), a number of gaps in these literatures exist pertaining to personality predictors that I address in the present dissertation.

The Five Factor Model of Personality

A major initiative faced by early personality researchers was the development of a comprehensive taxonomy of individual differences in personality (John, Angleitner, &

Ostendorf, 1988). The predominant approach used to develop such a taxonomy emerged from the *lexical hypothesis*, which stated that personality traits were encoded in the adjectives that appeared in the dictionary (Galton, 1884). Research attempted to identify the latent structure of personality by collating adjectives from the dictionary, having participants rate themselves and others on these adjectives, and using factor analysis to identify the basic dimensions underlying the adjectives (John, Naumann, & Soto, 2008). A large number of studies along these lines continuously identified five personality factors -- *extraversion*, *agreeableness*, *emotional stability*, *conscientiousness* and *openness* -- that proved robust across adjective lists, cultures, and factor analytic methods (Digman, 1972; Digman & Takemoto-Chock, 1981; Goldberg, 1981; Norman, 1963; Tupes & Christal, 1961). These five factors came to be known as the Five Factor Model of personality or simply, the Big Five (Digman, 1990).

Research suggests that individual differences in personality consistent with the Five Factor Model are arranged hierarchically. As noted by Ones, Viswesvaran, and Dilchert (2005): "When individuals respond to single items on personality inventories, they are providing self-descriptions about their own behaviors and reputations. Specific items refer to single instances of attitudes, thoughts, feelings, and behaviors. Psychologically similar items cluster together to make up specific traits. Similar traits can be grouped together to make up higher order personality dimensions" (p. 390). As can be seen through the above illustrative quote, personality can be conceptualized as narrow, homogenous dimensions known as *facets* or as broad *factors* that can account for the covariation among sets of personality facets.

In terms of the application of the Five Factor Model of personality in personnel selection contexts, a large body of research has emerged indicating that some personality factors are valid predictors of job performance (Hough & Ones, 2001). For instance, a large body of meta-analytic evidence indicates that conscientiousness predicts job performance across occupations and criteria (Barrick & Mount, 1991; Hurtz & Donovan, 2000). Furthermore, some evidence suggests that emotional stability is a valid predictor of job performance as well (Salgado, 1997). However, despite the evidence garnered in support of broad personality factors as predictors of job performance, some research suggests that the validity of personality factors might be driven by only a few of their constituent facets (Hough, 1992). Research has suggested that criterion-related validity of personality traits might be maximized by identifying those facets that best predicted job performance and incorporating those facets into predictor batteries, rather than the global factor.

The issue concerning the relative merits of broad versus narrow traits for selection purposes has come to be known as a component of the *bandwidth-fidelity dilemma* in personality assessment. Ultimately, two competing perspectives emerged. The first perspective suggested that employing broad personality measures maximizes the validity of personality traits. Proponents of the broad-trait perspective argued that criterion-related validity of personality predictors is due to variance that is shared among the facets (i.e., because of the factor), as opposed to variance that is unique to each facet (Salgado, Moscoso, & Berges, 2013). Proponents of the broad trait perspective also argued that broad traits are more practical as well because, in order to use personality scores for

decision-making purposes (e.g., selection), the decision maker would have to combine narrow traits in some way (Ones & Viswesvaran, 1996).

The second perspective suggested that employing narrow personality measures maximizes the validity of personality predictors. Proponents of the narrow trait perspective argued that the validity of personality traits are driven by variance that is unique to each personality facet as opposed to variance that is shared among facets (i.e., because of the factor; Tett, Steele, & Beauregard, 2003). Research along these lines suggested that the Big Five personality factors are too broad for selection purposes and that a more fine-grained approach would reveal meaningful predictor-criterion relationships (Hough, 1992).

While these two perspectives are predominant, it is also important to note that a third more nuanced approach also exists that attempts to seek a compromise between the two competing perspectives. Specifically, the third approach suggests that the validity of predictors is maximized when specific narrow traits are validated against conceptually relevant narrow criteria (Hogan & Holland, 2003). However, matching narrow traits to narrow criteria is not always feasible when considering the breadth of job performance criteria. For instance, research suggests that job performance dimensions are highly intercorrelated and saturated by a general factor (Viswesvaran et al., 2005). Indeed, empirical evidence in support of the narrow trait-narrow criteria matching approach is lacking, perhaps partly because of the fact the real world criteria are complex and broad (Hogan & Hogan, 1994). Finally, it is also important to note that, in order to use job performance criteria for decision-making purposes, narrow dimensions must be combined into a global measure and thus, the criterion measure would ultimately be broad.

Considering these issues, it is most useful for research to focus on the relative validity of broad versus narrow traits for predicting global criterion measures.

Purpose of Present Dissertation

At present, a large body of research has addressed issues pertaining to the bandwidth-fidelity dilemma by examining the relative magnitude of validities of global personality factors as compared to their constituent facets. However, there are two important considerations that research has virtually ignored -- (1) the relative stability of personality factor and facet validities across time and (2) the relative effects of broad vs. narrow traits on performance trends. First, the relative stability of validity coefficients of broad vs. narrow traits is a critical issue with wide implications for personnel selection practices, yet no research along these lines exists. In the present dissertation, I address issues of bandwidth-fidelity within the context of dynamic criteria research. Specifically, in subsequent sections, I propose that the validities of personality predictors would be more or less stable across time depending on the breadth of the personality predictors. I test my predictions using a four-year longitudinal design.

Second, as already indicated, research that has modeled intraindividual trajectories of job performance across time has identified between-person differences in such trajectories (Ployhart & Hakel, 1998). That is, research along these lines has found that job performance unfolds differently across time for different employees and that performance trajectories can be predicted by individual differences constructs, such as the Big Five personality factors (e.g., Minbashian et al., 2013; Thoresen et al., 2004). Research along these lines has been beneficial in advancing theories of work performance, yet gaps in this literature exist. Specifically, prior research has generally

8

assessed personality using broad measures and has not considered if narrow or broad personality measures are most useful in predicting job performance growth across time. In the present dissertation, I assess issues related to the bandwidth-fidelity dilemma by assessing broad and narrow personality traits as predictors of intraindividual trajectories of job performance scores.

Summary

In the present dissertation, I integrate the literatures pertaining to (1) the dynamicity of job performance criteria and (2) the bandwidth-fidelity dilemma in personality assessment for personnel selection purposes. In doing so, I address two issues. The first issue concerns the dynamicity of global and facet personality scores for predicting job performance across time. Indeed, the existing literature into the relative merits of global personality factors versus their facets for predicting work performance criteria has exclusively examined the relative magnitude of validities of personality factors and their facets for predicting job performance at a single point in time. In the present dissertation, I argue that another important consideration is the relative stability of criterion-related validities across time. The second issue concerns broad and narrow traits as predictors of intraindividual trajectories in job performance across time. The existing literature into this issue has focused exclusively on broad traits. Examining narrow traits as well might provide unique insights concerning determinants of performance change in organizations.

In the next chapter, I review the literature just discussed in greater depth. The chapter is organized as follows. First, I review literature concerning the dimensionality of job performance, including the temporal dimensionality (i.e., dynamic criteria). Then,

I review the literature pertaining to the Five Factor Model of personality and its use in personnel selection, including the bandwidth-fidelity dilemma. Note that, rather than focusing on the full range of personality traits consistent with the Five Factor Model, I focus exclusively on those that have demonstrated the most robust relations with job performance across occupations and criteria. I then review theoretical perspectives and prior research related to personality validity changes across time, as well as bandwidth-fidelity issues in this area. Finally, I review research and theory related to personality variables as predictors of performance trajectories across time as well as issues related to bandwidth-fidelity in personality assessment.

CHAPTER II: LITERATURE REVIEW

I begin my literature review with a discussion of the dimensionality of job performance. Following this, I discuss the Five Factor Model of personality, its use in personnel selection, and the bandwidth-fidelity dilemma in personality testing for personnel selection purposes. Finally, I discuss the Five Factor Model within the context of dynamic criteria research.

The Dimensionality of Job Performance

Job performance is defined behaviorally as the proficiency with which an employee performs the behaviors that are required as part of their job or otherwise contributes to organizational success (Campbell, 1990; Motowidlo, Borman, & Schmit, 1997; Viswesvaran & Ones, 2000). Researchers in the organizational sciences have faced considerable difficulty in adequately conceptualizing the job performance construct (Austin & Villanova, 1992). In response to these issues, in a classic paper, Ghiselli (1956) brought the following issues related to the dimensionality of job performance to light: *static dimensionality* and *temporal dimensionality*. Static dimensionality refers to the various distinct subdimensions that comprise performance on any given job and temporal dimensionality (i.e., dynamic criteria) refers to within-person changes in job performance across time. I review these dimensionality considerations in the following paragraphs, with particular emphasis placed on temporal dimensionality (i.e., dynamic criteria), as it is a key consideration for the present dissertation.

Static Dimensionality. As noted by Ghiselli (1956) and several other researchers subsequently (e.g., Cambpell, 1990, Motowidlo et al., 1997), for any job, there are a number of conceptually distinct dimensions that constitute performance. At present,

11

many investigations using a variety of distinct methods have sought to develop comprehensive typologies of performance dimensions. According to Viswesvaran and Ones (2000), these models of job performance can be categorized according to the following two dimensions: (1) *developmental context* and (2) *occupational focus*. The developmental context refers to whether the performance model explicates only a stand-alone or specific dimension of performance versus performance dimensions that are part of a larger set. Occupational focus concerns whether the performance dimension is applicable to only certain jobs versus applicable across jobs. By crossing the two dimensions of developmental context and occupational focus, we can arrive at four different types of models of job performance: (1) specific performance dimensions that are applicable to only certain jobs, (2) specific performance dimensions that are applicable across jobs, (3) a set of performance dimensions that are applicable to only certain jobs, and (4) a set of performance dimensions that are applicable across jobs. The focus of the present review will be on models of performance that are applicable across jobs.

Borman and Motowidlo (1993) proposed a model of stand-alone performance dimensions that are applicable across jobs. Their model of job performance distinguished between two dimensions: task performance and contextual performance. *Task performance* refers to proficiency in technical activities that are formally recognized as part of the job, while *contextual performance* refers to behaviors that are discretionary and are not formally recognized as part of the job, but do contribute to organizational success. Subsequent research has proposed incorporating *counterproductive work behaviors* (*CWB*) into the aforementioned model (e.g., Ones, Viswesvaran, & Schmidt,

12

1993; Robinson & Bennett, 1995). Counterproductive Work Behavior refers to behaviors that violate organizational norms and threatens the wellbeing of the organization and/or the workforce.

In terms of models that explicate a set of job performance dimensions that are applicable across jobs, Campbell (1990) proposed an eight-dimension model of job performance: job-specific task proficiency, non-job-specific task proficiency, written and oral communication, demonstrating effort, maintaining personal discipline, facilitating peer and team performance, supervision, and management or administration. Viswesvaran and colleagues (Viswesvaran, 1993; Viswesvaran et al., 1996) developed an integrative taxonomy of job performance dimensions by analyzing 486 job performance measures found in the published literature, grouping similar measures into dimensions. The effort resulted in ten performance dimensions: overall job performance, job performance or productivity, effort, job knowledge, interpersonal competence, administrative competence, quality, communication competence, leadership, and compliance with rules.

With the importance placed on explicating the latent structure of job performance, an important issue was determining the overlap among performance dimensions. For instance, Campbell (1990) proposed that there was virtually no overlap among performance dimensions and the true score correlations between dimensions could be presumed to be very small. Unfortunately, assessing this empirically posed challenges, as the intercorrelations among job performance dimensions are inflated due to *halo error* (Thorndike, 1920). Halo is a rating error that occurs when raters do not discriminate well between performance dimensions, but rather, tend to rate an employee similarly across all

dimensions (Cooper, 1981). Thus, empirical research that had identified high

correlations between distinct dimensions could simply be an artifact of halo error and

may not be due to a general factor of job performance.

Viswesvaran et al. (2005) used meta-analytic methods in order to control for the

influence of halo error on the intercorrelations among job performance dimensions. Once

accounting for halo error, the research indicated that a general factor of job performance

did indeed exist that accounted for the intercorrelations among job performance

dimensions. That is, a positive manifold across measures of job performance dimensions

was not accounted for solely due to halo error, but reflected true overlap in dimensions of

job performance and the presence of a general factor of job performance. The general

factor accounted for 60% of the total variance in supervisory ratings of job performance.

Considering the research just reviewed, it is apparent that the job performance

construct can be conceptualized at varying degrees of breadth. At the broadest level, job

performance can be conceptualized as a single higher-order factor (Viswesvaran et al.,

2005). At the narrowest level, job performance can be conceptualized as various

conceptually distinct behaviors that contribute to organizational success (Tett, Guterman,

Bleir, & Murphy, 2000). The job performance construct can also be conceptualized at

varying degrees of breadth between these two extremes (e.g., task performance,

contextual performance; Borman & Motowidlo, 1993). However, given the positive

manifold across performance dimensions and the fact that overall scores are used for

decision-making purposes in organizations, in the present dissertation, I conceptualize job

performance at the broad level, focusing on global performance scores. I now turn my

attention towards a review of the dynamic criteria (i.e., temporal dimensionality) literature.

Temporal Dimensionality. As already discussed, temporal dimensionality, also referred to as dynamic criteria, refers to within-person changes in job performance across time (Ghiselli, 1956). Dynamic criteria has been operationalized in the following three ways: (1) group mean-level changes in criterion scores across time, (2) rank-order changes in criterion scores across time, and (3) intraindividual trajectories in criterion scores across time (Barrett et al., 1985; Hofmann et al., 1992, 1993)[1]. I review each of these definitions and associated literature in the following sections.

Group Mean-Level Changes. A number of researchers have proposed that evidence for dynamic criteria would include group mean-level changes in performance scores across time (e.g., Ghiselli, 1956; Ghiselli & Brown, 1947; Ghiselli & Haire, 1960; Hanges, Schneider, & Niles, 1990). Of the various operationalizations of dynamic criteria, this one is often considered to be the most conceptually weak (Barrett et al., 1985; Sturman, 2007). For instance, the group mean-level changes definition does not address individual-level changes -- that is, there is no guarantee that changes in mean performance across time at the group level is similar to changes in mean performance across time at the individual level. Nonetheless, several studies have examined performance change using the group mean-level changes approach.

Ghiselli and Brown (1947) examined mean changes in performance (operationalized as objective accident records) in a sample of 94 streetcar and motor

---

[1] Dynamic criteria has also been defined as changes in criterion-related validity coefficients across time. I review this definition and related literature in subsequent sections.

coach operators across the first 17 months of their employment following their formal training. The results of the study indicated that accidents decreased over time at the group level in a manner that represented a positive quadratic trend, taken by the authors as evidence of dynamic criteria. However, while the study just reviewed and a handful of others have indicated that group mean-levels of job performance may change over time (e.g., Ghiselli & Haire, 1960), the implications of such findings are limited given the deficiencies associated with this definition of dynamic criteria.

Rank-Order Changes. Another method to assess dynamic criteria is examining rank-order changes in criterion scores across time (Sturman, 2007). Examining rank-order changes is accomplished by computing correlation coefficients between criterion scores assessed at different time points. At present, a number of studies have examined this issue. For instance, in a review article, Barrett et al. (1985) identified 55 studies that reported correlations between criterion scores assessed at two different time periods. These investigations identified that as criterion scores were measured further apart, their correlations became smaller (i.e., a simplex pattern; Austin, Humphreys, & Hulin, 1989).

For example, Mitchel (1975) assessed performance change across five years in a sample of managers. Performance was measured by the managers' salary histories as salary growth since time of hire. Salary growth was computed in two-year intervals -- one year, three years, and five years. The results indicated a simplex pattern. Performance at year one was more strongly correlated with performance at year two ($r = .80$) as compared to performance at year three ($r = .77$). Brenner and Lockwood (1965), using a sample of aircraft engineers, also assessed performance change by using salary data. Annual salary data was collected over a 21-year period. The results also indicated

16

that performance change across time could be characterized by a simplex pattern. Starting salary was highly correlated with early performance (e.g., $r = .73$ for 6 months, $r = .62$ for 1.5 years). However, as the measurement occasions grew further apart, the correlation became smaller. By the 21st measurement occasion, the correlation with starting salary was only $r = .15$.

Research has also assessed rank-order changes in performance using job performance ratings. In a longitudinal study of upward feedback, Reilly, Smither, and Vasilopoulos (1996) assessed the performance of managers (as rated by their subordinates) across four years. Their results indicated a simplex-like pattern. The time 1-time 2 performance correlation was $r = .53$, the time 1-time 3 correlation was $r = .57$, and the time 1-time 4 correlation was $r = .32$. Furthermore, the time 2-time 3 correlation was $r = .52$, while the time 2-time 4 correlation was $r = .40$. Thus, while there were some deviations in these data, overall, the results did support a simplex pattern.

Hoffman, Nathans, and Holden (1991) assessed the rank-order consistency of objective measures of productivity and ratings of quality across a three-year period in a sample of maintenance, mechanic, and field service workers. In terms of ratings of quality, a simplex pattern was observed in the data. The year 1-year 2 correlation was $r = .44$, while the year 2-year 3 correlation was $r = .26$. Similar results were observed for the objective measures of productivity. The year 1-year 2 correlation was $r = .84$, while the year 1-year 3 correlation was $r = .76$.

Overall, the pattern of results observed in these studies seemed to suggest that the rank-ordering of criterion scores does in fact change with time and that the magnitude of rank-order changes increases with time. However, subsequent analyses indicated that

while correlation coefficients did decrease as time passed, the change was often not statistically significant and furthermore, corrections for artifacts including measurement error were rarely made, further confusing the interpretation of findings (Barrett et al., 1985). For instance, interrater reliability of performance ratings is fairly low ($\rho = .52$; Viswesvaran et al., 1996), which would result in artificially deflated estimates of the correlations between criterion scores across time.

In response to these issues, Sturman et al. (2005) used a meta-analytic framework to disentangle the influence of true rank order changes in criterion scores across time from statistical artifacts on criterion-criterion correlation coefficients. Specifically, Sturman et al. estimated the effect of test retest reliability on criterion-criterion correlations across time in order to estimate the true stability, once the effect of reliability was accounted for. The results of this study indicated that a simplex pattern did in fact adequately describe the temporal decay in criterion-criterion correlation coefficients across time. They also found that performance change across time was greater for high complexity jobs as compared to low complexity jobs and when objective performance measures (e.g., sales figures) were used as compared to subjective performance measures (e.g., supervisory ratings). The results of the Sturman et al. (2005) study indicated utility in assessing rank-order changes in criterion scores across time and also suggested the need for additional primary studies into the issue.

Intraindividual Trajectories. The most recent development in assessments of dynamic criteria concerns the application of longitudinal data analysis methods in order to assess intraindividual trajectories of job performance across time (Hofmann et al., 1992, 1993). This definition is similar to the group mean-level changes definition

reviewed earlier, except the level of analysis is each individual employee as opposed to the group, and thus, the analytical framework ameliorates the limitations associated with the group mean-level changes definition (Sturman, 2007). That is, the trajectory of mean level performance across time represents the mean of each individual employee's performance trajectory across time. The longitudinal data analysis framework is also useful for assessing the variation of individual performance trajectories around the mean as well. In this way, research has been able to estimate average group-level growth parameters in job performance scores across time as well as the extent of between-person differences in such growth parameters (Ployhart & Hakel, 1998). Indeed, the longitudinal data analysis approach has been a very useful framework for explicitly assessing the manner in which individual-level job performance unfolds across time, and a number of investigations along these lines have been conducted.

The importance of studying intraindividual performance trajectories was first brought to light in Hofmann et al. (1992). Hofmann et al. reviewed several lines of evidence suggesting that systematic performance change across time, as well as between-person differences in such change, was likely to exist. In order to test the prediction, the research analyzed two datasets of baseball players. The first dataset included yearly batting averages across time and the second included yearly earned run averages across time. The results indicated systematic intraindividual change patterns, such that for some employees, performance increased across time, while for others, performance decreased across time.

In a field study, Hofmann et al. (1993) examined the intraindividual trajectory of job performance across 12 fiscal quarters (i.e., three years) using a sample of 300 sales

agents. Performance was operationalized as the value of insurance policies sold for each

quarter. Hofmann et al. used hierarchical linear modeling to assess the form of

performance change across time and to identify the extent of between-person differences

in performance change. Their results indicated significant (positive) linear and (negative)

quadratic performance trends (i.e., performance increased sharply early on and began to

level out at a later point). The results also indicated the presence of variability around the

performance trends, suggesting that performance unfolded differently across time for

different employees. Hofmann et al. used a clustering algorithm (cf. Ward, 1963) in

order to identify relatively homogenous clusters of groups whose performance unfolded

similarly across time. The results of the cluster analysis indicated three groups. The

trend of the first group's performance was largely linear and positive, the trend of the

second group's performance largely mirrored the overall results (i.e., linear then

asymptotic), and the trend of the third group's performance initially increased, but

ultimately decreased rapidly. The study provided keen insights concerning the manner in

which performance unfolds over time and the extent of systematic between-person

differences in such trends in performance.

As research has provided evidence supporting the existence of intraindividual

performance trajectories as well as between person differences in such trajectories

(Hofmann et al., 1992, 1993), research became interested in accounting for between-

person differences in performance trends using individual differences variables.

Deadrick et al. (1997) assessed performance trajectories in a sample of 408 sewing

machine operators. Performance was operationalized as average hourly earnings per

week, which was the ratio of total production earnings per week to the number of hours

worked in that week. Performance was assessed weekly across 24 weeks. The research

not only used these data to assess the form of performance change across time and

between-person differences in such change, but also to assess cognitive ability (measured

using the General Aptitude Test Battery), psychomotor ability, and sewing experience as

predictors of performance trends.

Results of the Deadrick et al. (1997) study indicated that performance across time

was characterized by a positive linear trend and that employees differed with respect to

their performance trends (the results also indicated between-person differences in the

performance intercept, or average performance at the first measurement occasion). In

terms of predictors of the performance intercept, results indicated that both experience

and psychomotor ability were positively associated with the performance intercept. In

terms of predictors of the performance trend, results indicated that experience was

negatively associated and that cognitive ability was positively associated. Thus, the

performance of employees who had sewing experience improved at a slower rater, while

the performance of employees who were high in cognitive ability improved at a faster

rate.

Ployhart and Hakel (1998) used a latent growth modeling approach to assess

intraindividual performance trajectories in a sample of 303 sales agents in a national

securities brokerage firm across eight quarters. Performance was operationalized as gross

sales commissions in dollars during each quarter. Ployhart and Hakel not only assessed

the form of intraindividual performance trajectories and between person differences in

such trajectories, but predictors of such trajectories as well. Predictors included past

sales commission and salary potential (PSCSP), self-assessed persuasive skill (persuasion), and self-assessed empathy (empathy).

Ployhart and Hakel's (1998) analyses indicated that performance change was characterized by linear (positive), quadratic (negative), and cubic (positive) growth parameters. Furthermore, results indicated significant variability around the linear and quadratic parameters, suggesting the presence of between-person differences. In terms of predictors of the latent growth parameters, results indicated that persuasion and empathy were both positively related to the linear performance trend, but no predictor was significantly related to the quadratic trend.

Overall, studies into intraindividual performance trajectories have not only shed light on the manner in which individual-level job performance unfolds over time, but have also indicated the presence of between-person differences in such trajectories that may be predicted by individual differences variables (Day, Sin, & Chen, 2004; Deadrick et al., 1997; Ployhart & Hakel, 1998; Thoresen et al., 2004). However, despite the increasingly large number of studies investigating intraindividual trajectories in job performance across time, there is a notable limitation. Specifically, the majority of research along these lines has examined only measures of effectiveness, such as sales volume, as opposed to job performance (i.e., defined behaviorally and measured via supervisory ratings). Thus, more research is needed in order to assess the extent to which the findings observed in prior research using objective measures generalize to subjective performance ratings.

As mentioned earlier, there are two important implications of the research reviewed above in relation to personnel selection and theories of work performance that I

address in the present dissertation. First, considering evidence suggesting rank-order

changes in criterion scores across time, it is possible that the validities of individual

differences variables for predicting job performance change across time as well. A

number of studies have addressed this issue in relation to cognitive ability (e.g., Barrett et

al., 1985; Hulin et al., 1990; Keil & Cortina, 2001), however, given the increasing use of

personality variables in personnel selection (Hurtz & Donovan, 2000; Ones,

Viswesvaran, & Schmidt, 1993), research is needed to examine the stability of

personality validities across time as well.

Indeed, the small number of studies that have addressed this issue are limited by a

reliance on objective measures of effectiveness as criterion scores and a general reliance

on broad measures of personality or only a select few narrow traits (cf. Helmreich et al.,

1986; Thoresen et al., 2004). Furthermore, while research into the bandwidth-fidelity

dilemma has examined the relative magnitude of validities of personality factors vs.

facets for predicting broad and narrow work performance criteria, research is needed to

address another important issue -- the relative stability of validities of personality factors

and facets for predicting broad and narrow performance criteria across time.

Second, research suggests that between-person variance in intraindividual job

performance trajectories can be predicted by individual differences variables, such as

cognitive ability and biodata (Deadrick et al., 1997; Ployhart & Hakel, 1998).

Furthermore, a few studies have assessed personality consistent with the Five Factor

Model as predictors of performance growth parameters as well (Minbashian et al., 2013;

Thoresen et al., 2004). Research along these lines has exclusively considered the broad

Big Five factors as predictors of performance change. In the present dissertation, I build

on these findings by assessing issues of bandwidth-fidelity in personality testing within the context of intraindividual trajectories of job performance across time. Bearing in mind the issues just reviewed, I now turn my attention towards the Five Factor Model of personality.

The Five-Factor Model of Personality

I begin my review of the Five Factor Model of personality by addressing historical perspectives pertaining to the development of the model and its hierarchical structure. I then review impact of the Five Factor Model on personnel selection research and practice and discuss the bandwidth-fidelity dilemma in personality testing for personnel selection purposes. Finally, I pose hypotheses and research questions concerning the validity of personality factors and facets for predicting job performance.

The Five Factor Model of personality emerged from the *lexical hypothesis*, which stated that individual differences in personality are reflected in our language (specifically, adjectives; Galton, 1884; Tellegen, 1993). Thus, the latent structure of personality could be identified by collating all of the adjectives used in the English language, scoring a sample of participants on these adjectives, and subjecting the resulting covariance matrix to factor analysis. Indeed, many researchers undertook such efforts over the years (cf. John et al., 1988). Along these lines, Allport and Odbert (1936) compiled a list of 17,953 adjectives from the dictionary. Using expert judges, they reduced the list to 4,504 words that they judged to reflect relatively stable tendencies.

Cattell (1943) drew upon the adjective list developed by Allport and Odbert (1936), grouping similar words into categories and added additional adjectives derived from the psychological literature. Cattell ultimately developed a list of 171 adjectives that

reflected the range of individual differences in personality. In subsequent work, Cattell (1945, 1946) reduced the number of traits using an empirical approach -- by having peers rate a target on the trait descriptors and using correlational analyses to determine the overlap in traits. Ultimately, Catell (1957) reduced the list of adjectives to 35, which, using factor analysis, he identified as indicators of 12 personality dimensions.

Unfortunately, Cattell's (1957) work was crude and subsequent research failed to replicate his 12 factors of personality (Digman & Takemoto-Chock, 1981). Thus, there was still little agreement regarding an adequate taxonomy of personality (e.g., Cattell, Eber, & Tatsuoka, 1970; Eysenck, 1991; Webb, 1915). However, Cattell's work did stimulate future research (cf. John et al., 1988), which soon converged on a five-factor solution to these factor analytic models (Tupes & Christal, 1992). For instance, Fiske (1949) collected ratings on 22 trait descriptors using a sample of 128 clinical psychology trainees. Across self-ratings and a variety of observer ratings, the results indicated that five factors accounted for the relationships among the observed variables. Similarly, across eight samples, Tupes and Christal (1961) found that five factors of personality emerged in their data. A number of subsequent studies replicated these five-factors (e.g., Digman, 1972; Digman & Takemoto-Chock, 1981; Goldberg, 1981; Norman, 1963).

These five-factors proved to be robust across personality scales, cultures, and factor analysis methods, and have since come to be known as the Five Factor Model of personality or simply, the Big Five (cf. Digman, 1990). The first factor, label *extraversion*, is indicated by traits such as talkative, assertive, and energetic. The second factor, *agreeableness,* is indicated by traits such as cooperative, sympathetic, and trustful. The third factor, *conscientiousness,* is indicated by traits such as dependable,

achievement-oriented, and responsible. The fourth factor, *emotional stability,* is indicated by traits such as calm and relaxed. The fifth and final factor, *openness*, is indicated by traits such as intellectual and imaginative.

The Five Factor Model of personality is hierarchical, with these five broad factors at the top, and a number of subdimensions (or *facets*) subsumed by each (Goldberg, 1981). With growing consensus around the big five factors, an important research endeavor was explicating the narrow facets subsumed by each factor. That is, having clearly defined the big five, research began to consider precisely what facets comprised each. In the following sections, I broadly review approaches used to examine the hierarchical structure of the Big Five personality factors. I then review the personality factors in greater depth and explicate the facet models used in the present study. However, it is important to note that I will not be assessing all five personality factors in the present dissertation, but rather, I will be focusing on those that are most relevant for personnel selection (as meta-analytic cumulations have indicated that they predict performance across occupations and criteria) and have generated interested within the context of the bandwidth-fidelity dilemma: conscientiousness, emotional stability, and extraversion.

First, large bodies of meta-analytic evidence cumulatively suggest that conscientiousness predicts performance across occupations and criteria (Barrick & Mount, 1991; Hurtz & Donovan, 2000). Second, while the evidence concerning emotional stability is more mixed, some meta-analytic reviews have indicated that emotional stability also predicts job performance across occupations and criteria (Salgado, 1997). Finally, very little evidence exists suggesting that extraversion predicts

job performance. However, research has argued that extraversion does not predict job performance because it is too broad, and that a focus on extraversion facets could improve the validity of this factor for predicting job performance across occupations and criteria (Hough, 1992; Minbashian, Bright, & Bird, 2009). Therefore, as extraversion is relevant within the context of the bandwidth-fidelity dilemma, I assess extraversion in the present dissertation as well. On the other hand, little evidence suggests that openness and agreeableness (nor their facets) are valid predictors of job performance across jobs and therefore, I do not address these personality traits in the present study.

Facet Models. One of the earliest attempts at identifying the facets underlying each of the Big Five factors was conducted by Norman (1963, 1967), who, using a list of almost 1,600 trait descriptor terms, replicated the five factor structure and subsequently attempted to identify the lower order facets of the Big Five using a rational method. Using the results of prior factor analytic studies, Norman assigned each trait to either the positive or negative pole of one of the Big Five factors. He then further refined each category by sorting the terms into more narrow categories within each of the Big Five factors, resulting in 75 categories. Norman still continued to combine similar terms together within each of these 75 categories, arriving at a total of 571 narrow personality facets. Thus, the result of Norman's rational analysis of the hierarchical structure of the Big Five factors resulted in a three level hierarchy in which the Big Five factors were at the top, followed by 75 mid-level traits, and finally, 571 more narrow traits.

While Norman (1963, 1967) adopted a rational approach towards deriving facets, other researchers adopted a theoretically oriented approach, consulting the psychological literature to identify the facets that comprise each of the Big Five factors (Costa, McCrae,

& Dye, 1991). Using this approach, Costa and McCrae (1992) developed a six-facet

model for each of the Big Five factors. Six facets were chosen because this is the

minimum number of facets that would allow for a replicable factor structure (Gorsuch,

1974) and because "the inclusion of more than six would soon lead to intellectual

overload" (Cost & McCrae, 1995, p. 27). Thus, the six-facet model derived by Costa and

McCrae, while grounded in practical and statistical considerations, is not necessarily

optimal (Costa & McCrae, 1998). Nonetheless, this framework has been quite influential

in the personality literature and achieves its intended purpose of broadly sampling facets

from each personality factor.

Other research has examined the lower order factors of the Five Factor Model by

using factor analysis. Saucier and Ostendorf (1999) re-analyzed two data sets -- an

American dataset that included self and peer ratings on 500 adjectives ($N = 636$;

Goldberg, 1990), and a German dataset that included self and peer ratings on 500

adjectives ($N = 775$). The researchers computed factor scores from the 100 items (20 for

each factor) with the highest loadings on each factor. Then, they correlated each of the

Big Five factor scores with the 500 adjectives and assigned each adjective to the factor

with which it correlated the highest. Finally, the researchers factor analyzed the

correlations among the adjectives grouped within each factor to assess the number and

nature of facets that represented each factor. Their results indicated a four-facet model

for extraversion (i.e., sociability, unrestraint, assertiveness, activity), agreeableness (i.e.,

warmth-affection, gentleness, generosity, modesty-humility), and conscientiousness (i.e.,

orderliness, decisiveness-consistency, reliability, industriousness), and a three-facet

model for emotional stability (i.e., low irritability, low insecurity, low emotionality) and openness (i.e., intellect, imagination-creativity, perceptiveness).

DeYoung, Quilty, and Peterson (2007) conducted a study into the facets of the Big Five factors by administering 75 facet scales from two Big Five inventories to a sample of participants and applying factor analysis to the resulting data. Their results indicated that each of the Big Five factors could be divided into two correlated facets. Conscientiousness included *industriousness* and *orderliness*, extraversion included *enthusiasm* and *assertiveness*, openness included *intellect* and *aesthetic openness*, neuroticism (reverse coded emotional stability) included *volatility* and *withdrawal*, and agreeableness included *compassion* and *politeness*.

As can be seen from the above review, there is no single model detailing the facets of each of the Big Five factors. However, while there is disagreement concerning exactly what facets comprise each of the Big Five factors, there is great consensus regarding the hierarchical structure of individual differences in personality. That is, similar to the construct of job performance, personality can be conceptualized at varying degrees of breadth, from relatively homogenous personality facets to broad personality factors that account for shared variance among sets of personality facets (Ones et al., 2005). However, it is necessary to review specific facet models for the factors assessed in the present dissertation (i.e., conscientiousness, emotional stability, extraversion) and to explicate exactly what models will serve as a basis for purposes of the present dissertation. In determining what facet models to draw upon, I considered the empirical evidence supporting the various facet models and also models commonly drawn upon in the personnel selection literature.

Conscientiousness. The literature into the facets of conscientiousness has

identified a diverse number of facets, with some studies identifying as few as two facets

(DeYoung et al., 2007) and others identifying as many as eight (Roberts, Bogg, Walton,

Chernyshenko, & Stark, 2004). MacCann, Duckworth, and Roberts (2009) assessed the

facets of conscientiousness in a sample of $N = 291$ high school students. They

administered 117 items from 12 different personality models intended to tap into

conscientiousness from the International Personality Item Pool. MacCann et al.

conducted an exploratory factor analysis and identified an eight-factor solution to the

data. These factors reflected *industriousness*, *perfectionism, tidiness, procrastination*

refrainment, *control*, *cautiousness*, *task planning*, and *perseverance*. Using a sample of

$N = 961$ participants (including university students and employees), Perugini and

Callucci (1997), using an exploratory factor analysis, found evidence for a 5 facet model

of conscientiousness, including *reliability*, *meticulousness*, *recklessness*, *inaccuracy*, and

superficiality.

In a recent review article, Roberts, Lejuez, Krueger, Richards, and Hill (in press)

noted that, while 10 facets emerged across a wide-range of studies, *order, achievement,*

self-control, and *responsibility* facets emerged most frequently. Specifically, order and

achievement were identified as conscientiousness facets in 100% of studies into the lower

order structure of conscientiousness. Furthermore, self-control and responsibility were

identified as facets of conscientiousness in 75% of studies. On the other hand, the

remaining facets emerged in only 25% of studies or less. Considering that a four-facet

model of conscientiousness that includes order, achievement, self-control, and

responsibility is the most robust across studies, these lower-order facets will be the focus

30

of the present study. The four-facet model has also been drawn upon in prior personnel selection research (e.g., Dudley, Orvis, Lebiecki, and Cortina, 2006).

The order facet reflects a tendency to impose structure on one's work environment, which would include a predisposition towards neatness, cleanliness, and planfulness (Dudley et al., 2006; Roberts et al., in press). The achievement facet is indicative of a tendency to "strive for competence in one's work" (Hough, 1992, p. 144), including a predisposition towards hard work and perseverance. The self-control facet reflects a predisposition towards cautiousness, level-headedness, and patience (Roberts et al., 2005). Finally, the responsibility facet reflects trustworthiness, respect for others, and dependability (Saucier & Ostendorf, 1999).

Emotional Stability. Most of the research into the lower order facets of emotional stability has been conducted within the context of *neuroticism* -- the low pole of emotional stability. Costa and McCrae's (1992) NEO framework proposed a six-facet model of neuroticism. The facets included *anxiety*, *angry hostility*, *depression*, *self-consciousness*, *impulsiveness*, and *vulnerability*. While this model has been influential, as noted earlier, research has argued that the NEO facets are not necessary optimal (e.g., Saucier & Ostendorf, 1999).

Perugini and Callucci (1997) conducted a factor analytic study into the facets of emotional stability (as opposed to low emotional stability -- neuroticism). Their results indicated two facets -- one pertaining to mental clarity (i.e., serenity) and the other pertaining to stability (i.e., firmness). The former pertains to calmness and the ability to relax, while the later pertains to heartiness, firmness, and psychological strength. Little research has examined the validity of emotional stability facets and therefore, there is

little guidance from the personnel selection literature for determining a model of emotional stability facets. In the present dissertation, I employ the two-facet model of emotional stability just discussed.

Extraversion. Hogan (1982) proposed a modified version of the Five Factor Model by splitting extraversion into two factors -- one representing facets that pertain to sociability and the other representing facets that pertain to ascendancy (e.g., dominance). The two-facet model of extraversion proposed by Hogan has been influential in the personnel selection literature (Hogan & Holland, 2003; Hough, 1992; Minbashian et al., 2009). DeYoung et al. (2007) assessed the lower-order facets of extraversion by administering a sample of 481 participants several personality facet scales and factor analyzing the data. Their findings indicated a two-facet model of extraversion. The first facet was *enthusiasm,* which encompassed a tendency to experience positive moods and enjoy social interaction (John, 1990), while the later encompassed agency and dominance. This two-facet model is consistent with the model proposed by Hogan (1982), which has been influential in the selection literature.

Judge et al. (2013) provided further support for this framework by meta-analytically cumulating correlations between personality facets consistent with the NEO framework (Costa & McCrae, 1992) and using the meta-analytic correlation matrix to replicate the factor structure identified by DeYoung et al. (2007) using confirmatory factor analysis. Their results supported the two-facet model of extraversion. Considering the ubiquity of the two facet model of extraversion in the literature as well as its prevalence in the personnel selection literature, I employ these two facets (i.e., assertiveness and enthusiasm) in the present study.

Having discussed the structure of individual differences in personality, I now turn my review towards the use of personality variables in personnel selection. In the following sections, I review the cumulative literature into the validity of personality variables for predicting job performance. My review includes an overview of the validity of the Big Five factors and their facets (with particular emphasis placed on the factors and facets examined in the present dissertation). I then turn my attention towards temporal issues in the validity of the Big Five personality factors and their facets.

Personality Assessment in Personnel Selection. Early research into the validity of personality variables for predicting job performance yielded bleak results. For example, Guion and Gottier (1965) reviewed personality validity studies reported in two journals -- Personnel Psychology and Journal of Applied Psychology -- from 1952 to 1963. Their results suggested that personality variables were only significantly related to job performance in 12% of the studies included in their analysis. Based on their review, Guion and Gottier (1965) noted, "taken as a whole, there is no generalizable evidence that personality measures can be recommended as good or practical tools for employee selection" (p. 159).

However, these earlier conclusions were limited due to a lack of an adequate taxonomy for personality. That is, prior research had generally classified all personality predictors under one category of "personality" rather than distinguishing between the traits assessed by each measure (Guion & Gottier, 1965; Schmitt, Gooding, Noe, & Kirsch, 1984). As the Five Factor Model of personality emerged in the literature (Digman, 1990), research was able to address this issue by classifying personality

predictors according to the Five Factor Model, as opposed to lumping them together as measures of personality.

A comprehensive meta-analysis conducted by Barrick and Mount (1991) was the earliest such attempt. These researchers searched the literature for personality validity studies conducted between 1952 and 1988, ultimately identifying 117 studies (and 162 samples) that reported correlation coefficients that expressed the relationship between a measure of personality and a measure of job performance. Each personality variable assessed in the primary studies was classified as an indicator of one of the Big Five factors of personality. Their results indicated that across occupations and criteria, conscientiousness was a valid predictor of job performance ($\rho = .22$).

Subsequently, a number of meta-analyses were published further supporting the usefulness of the Five Factor Model as a classification scheme for personality for personnel selection purposes. For instance, Salgado (1997) extended the work of Barrick and Mount (1991) to European samples. The results were consistent with the prior work of Barrick and Mount, suggesting that conscientiousness was a valid predictor of job performance across occupations and criteria ($\rho = .25$). Salgado's meta-analysis also indicated that emotional stability was a valid predictor of job performance across occupations and criteria ($\rho = .19$).

Mount, Barrick, and Stewart (1998) conducted a meta-analysis of Big Five-job performance relationships in jobs that involved interpersonal interaction. The meta-analysis included 11 studies, each of which involved a sample of employees whose jobs either involved teamwork or dyadic service. The jobs that involved dyadic service included residential counselors, residence hall advisors, customer service representatives,

and telemarketing representatives. The results indicated that conscientiousness was a

valid predictor of performance across the jobs included in the primary studies ($\rho = .26$).

The results also indicated that emotional stability ($\rho = .27$), agreeableness ($\rho = .33$), and

extraversion ($\rho = .27$) predicted performance in jobs involving teamwork.

One limitation of the majority of meta-analyses reviewed thus far is that they

required the researchers to classify the measures used in each primary study as an

indicator of one of the Big Five personality factors. Thus, the results were potentially

influenced by the judgment of the researchers. However, these earlier meta-analyses had

a drastic influence on personnel selection research and ultimately, a large number of

primary studies that used explicit measures of the Big Five personality factors became

available in the literature. Hurtz and Donovan (2000) conducted a cumulative meta-

analytic review of this literature, including only primary studies that assessed the Big

Five using explicit measures of these personality factors. The results confirmed the

finding that conscientiousness was a valid predictor of job performance across

occupations and criteria ($\rho = .22$). Furthermore, the validity of emotional stability, while

smaller than that of conscientiousness, was also non-zero across occupations and criteria

($\rho = .14$).

Overall, by classifying personality measures according to the Five Factor Model

as opposed to treating personality predictors as a homogenous class of predictors,

research was able to identify that personality variables (and specifically,

conscientiousness) were in fact valid predictors of job performance. However, some

research began to question whether five broad categories were refined enough to truly

reveal meaningful personality-job performance relationships. That is, by refining

personality predictors further (i.e., focusing on personality facets as opposed to factors),

could the validity of personality predictors be further improved? While some researchers

suggested that this was unlikely (Ones & Viswesvaran, 1996), some research began to

examine this issue empirically. At present, a vast literature is dedicated towards

examining the relative merits of global personality factors versus their facets in a

personnel selection context and I review this literature in the following section. Apparent

in the review is that the literature into the validity of broad vs. narrow traits for predicting

performance has considered only one issue -- the relative magnitude of validities of

global personality factors versus their facets for predicting job performance at one point

in time.

   Perspectives on Broad and Narrow Personality Predictors. Considering the

hierarchical structure of personality, an important question concerns at what level of

breadth personality ought to be measured for personnel selection purposes. That is, when

attempting to use personality variables to predict work performance, should research

focus on narrow personality facets or broad personality factors? This issue, which had

come to be known as a component of the *bandwidth-fidelity dilemma*, has important

implications for applied personality assessment. Two predominant perspectives have

emerged in response to the bandwidth-fidelity dilemma -- the first favoring the use of

broad traits and the second favoring the use of narrow traits.

   The first perspective suggests that broad personality variables are better predictors

of job performance than narrow personality variables. Ones and Viswesvaran (1996)

reviewed three lines of evidence suggesting that broad traits are superior to narrow traits

for personnel selection purposes -- psychometric evidence, criterion-related validity

evidence, and practical and theoretical issues. In terms of psychometric evidence, these researchers noted that measures of personality facets are generally much shorter than measures of broad factors (as broad factors are often measured by summing scores across a range of more narrow facets) and as a result, the reliability of short measures are generally lower than that of long measures. The notion was supported by empirical evidence, as the authors found that the reliability of the conscientiousness scale from two personality measures was higher than the reliability of any of the respective facet scales.

In terms of criterion-related validity evidence, Ones and Viswesvaran (1996) indicated that the belief that specific personality facets would predict performance only in specific situations was grounded in the situational specificity hypothesis (i.e., that each job is unique and thus predictors of job performance would vary greatly across situations; Schmidt & Hunter, 1984). However, with the advent of meta-analysis, research had begun to identify that variation in criterion-related validities across validation studies was more often attributed to sampling error than to situational specificity. Ones and Viswesvaran reviewed evidence from the personality domain suggesting that the validity of personality variables for predicting job performance is a result of variance in the predictor that is accounted for by the personality factor as opposed to variance that is unique to the personality facet. Also along these lines, Ones and Viswesvaran noted that broad measures should maximize prediction because the criteria assessed in selection research are broad. Thus, broad personality predictors should be used to predict job performance because, "broad-band predictors should be used to predict broad criteria; narrow-band predictors should be used to predict narrow criteria" (Ones & Viswesvaran, 1996, p. 616). Furthermore, even though performance can be assessed at the dimension-

level, these measures are still rarely homogenous. That is, while dimension-level performance ratings may be narrower than a global performance rating, they are still relatively broad.

Finally, Ones and Viswesvaran (1996) reviewed practical and theoretical considerations suggesting that broad traits are superior to narrow traits for personnel selection purposes. In terms of practical considerations, in order to make hiring decisions, the decision maker must combine the personality data obtained in some way. Thus, even if a number of narrow facets are assessed, some composite score must ultimately be computed so that selection decisions can be determined. In terms of theoretical considerations, Ones and Viswesvaran argued that broad traits are more useful than narrow traits for developing generalizable theories of work performance. For instance, as already reviewed, broad traits have higher reliability and criterion-related validity than narrow traits and are thus more suited for explaining behaviors at work. Furthermore, broad traits would be more applicable across jobs than specific narrow traits and would thus be useful in developing comprehensive theories of individual differences in work performance. Based on these arguments just reviewed, Ones and Viswesvaran determined that broad personality traits are more useful in advancing the theory and practice of I-O Psychology than narrow traits.

While the Ones and Viswesvaran (1996) review provided compelling justification for a focus on broad traits, the second perspective on the bandwidth-fidelity dilemma suggests that narrow personality measures may be more useful than broad measures for personnel selection purposes. For instance, some researchers suggested that the Big Five are too broad to reveal meaningful relationships between personality variables and

performance criteria (Hough, 1992). Thus, the validity of personality variables could be maximized by identifying what facets of each personality factor contribute the greatest amount of variance to job performance, and using only these facets for personnel selection purposes (Hough & Ones, 2001). Other research suggested that narrow personality traits may have utility as predictors of job performance, but only when conceptually aligned with narrow performance measures (Hogan & Holland, 2003). Thus, rather than measuring broad traits, this perspective suggests that the researchers carefully consider what specific traits should be related to the specific performance criteria and these traits should be measured and used for selection purposes.

Having discussed these competing perspectives, I now turn my attention towards a review of the empirical evidence on the subject. In the following section, I review the empirical literature into the validity of personality facets for predicting job performance. I also review the empirical literature into the relative magnitude of validities of global personality factors versus their facets.

Validity of Personality Facets for Predicting Job Performance. One of the earlier investigations into this issue was a meta-analysis conducted by Hough (1992). Hough argued that, while the Big Five is a useful taxonomy for trait *description*, it is not necessarily ideal for *prediction.* That is, while the Five Factor Model can be used as a classification scheme for individual differences in personality, it may not be optimal for the specific goal of predicting work performance. Hough proposed that Hogan's (1982) personality taxonomy that split extraversion into two facets -- assertiveness and enthusiasm (DeYoung et al., 2007) -- was necessary to maximize criterion-related validity. Furthermore, Hough split conscientiousness into two facets -- achievement and

responsibility. The results of her meta-analysis indicated some utility in this approach. The results indicated that, across occupations and criteria, the mean observed validity for the assertiveness facet of extraversion ($r = .09$) was higher than the validity of the enthusiasm facet of extraversion ($r = .02$). Furthermore, the validity of the achievement facet of conscientiousness ($r = .19$) was higher than the validity of the responsibility facet ($r = .07$). However, Hough did not assess the validity of global conscientiousness and global extraversion, so her results did not provide information regarding the relative validity of broad versus narrow traits.

Hogan and Holland (2003) conducted a meta-analysis into the validity of broad and narrow personality variables for predicting broad and narrow performance criteria. Similar to Hough (1992), they split extraversion into two facets -- assertiveness and enthusiasm. They assessed the validity of these variables for predicting broad and narrow performance criteria. At the broadest level, these researchers examined overall performance criteria. They then split overall performance into two dimensions -- getting along (i.e., contextual performance) and getting ahead (i.e., task performance). Finally, moving even narrower, they classified each performance criteria used in each primary study according to their conceptual similarity to the personality variables examined in their meta-analysis (i.e., the Big Five with extraversion split into two facets).

The results of the Hogan and Holland (2003) meta-analysis were as follows. In terms of the overall performance criteria, their results were consistent with prior meta-analyses suggesting that conscientiousness was a valid predictor of performance ($\rho = .24$). Furthermore, similar to the results of Hough's (1992) meta-analysis, their results indicated that the validity of the assertiveness facet of extraversion ($\rho = .22$) was higher

than the validity of the enthusiasm facet of extraversion ($\rho = .01$). However, and again, similar to Hough's meta-analysis, Hogan and Holland did not assess the validity of global extraversion. Thus, while their meta-analysis indicated different validities for the different facets of extraversion and openness, they do not facilitate a comparison of broad versus narrow traits.

In terms of the narrow criteria measures examined by Hogan and Holland (2003), results indicated that the validity of personality predictors was enhanced when validated against conceptually relevant performance measures. For instance, the validity of agreeableness for predicting overall job performance was $\rho = .17$, but was $\rho = .34$ for predicting criteria that were classified as being similar to agreeableness. Similar conclusions were reached when considering the getting along versus getting ahead performance criteria. Specifically, the validities of emotional stability, agreeableness, and conscientiousness were higher for predicting getting along criteria as compared to overall performance criteria. Furthermore, the validity of the ambition facet of extraversion was higher for predicting getting ahead criteria than overall performance criteria. The results of this study seem to suggest that the validity of personality predictors could be improved by selecting personality predictors on the basis of their conceptual similarity to criterion measures. However, neither the Hough (1992) meta-analysis nor the Hogan and Holland (2003) meta-analysis compared the validities of broad versus narrow personality measures as predictors of various criteria measures. That is, while both meta-analyses indicated that validity might differ between facets of the Big Five factors, they did not facilitate a direct comparison of the validity of broad versus narrow traits.

A number of other studies addressed this deficit by directly comparing the validity of some broad and narrow traits for predicting broad and narrow performance criteria. Many of these initiatives focused specifically on conscientiousness. Salgado et al. (2013) addressed this issue in a field study. In a sample of 226 police officers, Salgado et al. assessed the validity of global conscientiousness and three conscientiousness facets -- order, industriousness, and self-control -- against three performance criteria that differed in breadth -- overall job performance (i.e., broad bandwidth), task performance (i.e., mid-level bandwidth), and orderliness (i.e., narrow bandwidth). The results of the study indicated that global conscientiousness was a valid predictor of all three performance criteria measures. Furthermore, none of the facets of conscientiousness contributed incremental variance to the prediction of any criteria measure beyond global conscientiousness. These results are consistent with Ones and Viswesvaran's (1996) review of the bandwidth-fidelity dilemma discussed earlier -- that it is factor variance of personality measures that contribute to the prediction of job performance as opposed to variance specific to the personality facet.

Research has examined the relative magnitude of factor and facet validities using meta-analysis as well. Mount and Barrick (1995) assessed the validity of global conscientiousness and two of its facets -- achievement and dependability -- for predicting global and specific performance criteria. A total of 173 studies were included in this meta-analysis. In terms of global criteria, the validity of global conscientiousness ($\rho =$ .31), achievement ($\rho =$ .33), and dependability ($\rho =$ .30) were similar. However, for specific criteria, the validity of global conscientiousness ($\rho =$ .40) and achievement ($\rho =$ .38) were higher than the validity of dependability ($\rho =$ .28). These results indicated that,

while the validity of personality facets may differ as a function of performance dimensions, global conscientiousness predicted both broad and narrow criteria as well as either of the two facets. These results are also consistent with Ones and Viswesvaran's (1996) assertion that broad traits are preferable to narrow traits for personnel selection purposes, regardless of the criteria employed.

Another meta-analysis, conducted by Dudley et al. (2006), addressed the same issue as Barrick and Mount (1995), but included four facets of conscientiousness -- achievement, order self-control, and responsibility -- as opposed to only two. To facilitate a comparison between global conscientiousness and these narrow facets, they used the meta-analytic estimates for global conscientiousness derived from a prior meta-analysis reported by Hurtz and Donovan (2000). Dudley et al. also compared the validities of the four facets of conscientiousness to global conscientiousness for broad and narrow performance criteria, including overall job performance, task performance, job dedication, interpersonal facilitation, and counterproductive work behaviors. Finally, they also meta-analytically cumulated the correlations between global conscientiousness and the four conscientiousness facets so that they could assess the incremental validity of the facets beyond global conscientiousness.

The results of the Dudley et al. (2006) meta-analysis indicated that global conscientiousness was a better predictor of overall job performance than any of its facets. However, responsibility contributed variance to overall job performance incremental to global conscientiousness. In terms of task performance, the validity of the achievement facet was higher than the validity of global conscientiousness ($\rho = .25$ vs. $\rho = .16$). Furthermore, achievement, order, and self-control accounted for variance in performance

ratings incremental to global conscientiousness. In terms of job dedication, both responsibility ($\rho = .46$) and achievement ($\rho = .39$) had higher validities than global conscientiousness ($\rho = .20$), and both contributed incremental variance to the prediction of job dedication beyond global conscientiousness. In terms of interpersonal facilitation, the validity of responsibility ($\rho = .23$) was higher than the validity of global conscientiousness ($\rho = .18$). Furthermore, responsibility contributed variance to interpersonal facilitation incremental to global conscientiousness. Finally, in terms of counterproductive work behaviors, the validity of responsibility ($\rho = -.34$) was higher than the validity of global conscientiousness ($\rho = -.26$). Furthermore, responsibility accounted for incremental variance in counterproductive work behaviors beyond global conscientiousness. While these results suggested that narrow traits may be better predictors of narrow performance criteria than broad traits, it is important to note that global conscientiousness was a valid predictor for each performance criteria, while the conscientiousness facets had higher validities than global conscientiousness for some criteria, but much lower validities for others. Thus, while narrow traits may predict some narrow criteria better than broad traits, broad traits still seem to predict a larger number of criteria reasonably well.

Some work has focused explicitly on the relative magnitude of validities of global extraversion vs. its facets as well. In one such study, Minbashian et al. (2009) assessed the validity of global extraversion and its assertiveness and enthusiasm facets for predicting job performance in a sample of 179 managers. The performance criteria measure was supervisory ratings. Minbashian et al. assessed the validity of extraversion and its facets for predicting overall performance criteria, getting ahead criteria, and

getting along criteria. Results were consistent with existing meta-analyses, as they failed

to detect a significant extraversion-job performance relationship for all performance

criteria. Furthermore, the results also did not indicate a significant relationship between

the enthusiasm facet of extraversion and any of the job performance criteria. However,

the assertiveness facet of extraversion was a valid predictor of overall job performance

criteria and of getting ahead criteria. Thus, counter to many of the studies reviewed into

the validity of conscientiousness and its facets, this research suggested that the validity of

extraversion might be improved by focusing only on the assertiveness facet as opposed to

the global factor.

Recently, Judge et al. (2013) conducted the most comprehensive meta-analysis to

date into the relative magnitude of validities of broad versus narrow traits for predicting

broad versus narrow performance criteria. Specifically, Judge and colleagues assessed

the validity of a three-level hierarchical model of the Big Five for predicting overall job

performance, task performance, and contextual performance. At the highest level of the

trait hierarchy were the Big Five factors. At the next level, each factor was split into two

facets consistent with the framework derived by DeYoung, et al. (2007). Finally, at the

lowest level of the hierarchy, each factor was split into six facets consistent with the

NEO-FFI framework (Costa & McCrae, 1992). The meta-analysis included 1,176 studies

derived from 410 independent samples.

The results of the vast meta-analysis indicated that, overall, optimally weighted

composites of facets resulted in greater criterion-related validity for predicting all

performance outcomes than did the Big Five factors. That is, the results indicated that

not only variance shared by the facets (i.e., variance associated with the personality

45

factor) was related to job performance, but variance unique to the facets, was related to job performance as well. However, in terms of bivariate relationships, the Big Five factors generally performed equally as well as their best predicting facets. For instance, the validity of conscientiousness for predicting overall job performance was $\rho = .26$, while the validity of the highest conscientiousness facet for predicting this criterion (i.e., industriousness) was $\rho = .24$. As another example, extraversion predicted overall job performance with a validity of $\rho = .20$, while the facet with the highest criterion-related validity (i.e., positive emotions) predicted this criterion with a validity of $\rho = .20$. As a final example, extraversion predicted task performance with a validity of $\rho = .12$, while the best predicting facet of extraversion (i.e., assertiveness) had a validity of $\rho = .15$ for predicting this criterion. Overall, this study indicated that broad measures do predict performance, and in most cases, better than any single narrow personality measure. However, validity was maximized when personality facets were optimally weighted as opposed to averaged together to form factor scores.

Criterion-Related Validity Hypotheses. As can be seen by the above review, an issue in personality testing for personnel selection purposes, which has come to be known as a component of the bandwidth-fidelity dilemma, concerns whether the best approach to maximizing the criterion-related validity of personality predictors is to assess broad (i.e., factor) versus narrow (i.e., facet) personality predictors. Two competing perspectives have emerged, one favoring broad traits and the other favoring narrow traits. In the following sections, based on the theoretical perspectives and literature just reviewed, I pose hypotheses and research questions concerning the validities of broad and narrow personality traits for predicting job performance.

Broad Personality Traits. In terms of broad personality variables, in line with prior meta-analyses and theoretical treatments, I propose that conscientiousness will be a valid predictor of composite job performance scores. I will address emotional stability- and extraversion-job performance relationships as research questions.

Hypothesis 1: Conscientiousness is a valid predictor of job performance.

Research Question 1: Is emotional stability a valid predictor of job performance?

Research Question 2: Is extraversion a valid predictor of job performance?

Narrow Personality Traits. As discussed earlier, prior reviews have indicated that the Big Five may be too broad when considering individual differences correlates of composite job performance measures, specifically with regards to conscientiousness and extraversion (e.g., Hough, 1992, Minbashian et al., 2009). Consistent with the results of prior meta-analytic reviews (Dudley et al., 2006; Hogan & Holland, 2003), I propose the following hypotheses and research questions:

Hypothesis 2: The achievement facet of conscientiousness is a valid predictor of job performance.

Research Question 3: Is the order facet of conscientiousness a valid predictor of job performance?

Research Question 4: Is the self-control facet of conscientiousness a valid predictor of job performance?

Research Question 5: Is the responsibility facet of conscientiousness a valid predictor of job performance?

Research Question 6: Is the serenity facet of emotional stability a valid predictor of job performance?

Research Question 7: Is the firmness facet of emotional stability a valid predictor of job performance?

Hypothesis 3: The assertiveness facet of extraversion is a valid predictor of job performance.

Research Question 8: Is the enthusiasm facet of extraversion a valid predictor of job performance?

The Five Factor Model and Dynamic Criteria

I now turn my attention towards issues related to the interface of the Five Factor Model of personality and dynamic criteria. These issues include (1) the stability of criterion-related validities for both broad and narrow personality measures for predicting job performance across time and (2) broad and narrow personality measures as predictors of intraindividual variation in job performance across time. Both of these issues have received scant attention in the empirical literature, but are critically important for advancing personnel selection practices and theories related to criterion conceptualization and the bandwidth-fidelity dilemma in personality assessment. I begin by reviewing relevant theories and research related to the dynamicity of criterion-related validity coefficients across time for both broad and narrow traits.

Dynamicity of Criterion-Related Validity Coefficients Across Time. As mentioned earlier, a large number of empirical investigations have examined the stability of criterion-related validity coefficients across time (and this research has largely focused on cognitive predictors). In one of the earliest studies along these lines, Ghiselli and Haire (1960) assessed the validity of a paper-and-pencil test battery (including arithmetic, speed of reactions, tapping and dotting, judgment of distance, distance discrimination,

48

mechanical principles, occupational level interests, interests outside of occupation, interest in occupations dealing with people, and interest in active occupations) for predicting the performance of 56 taxicab drivers. The performance measure used was the dollar volume of fares collected for each week. The study assessed performance across an 18-week period, but the research examined the validity of the tests for predicting performance for an average of the first three weeks and an average of the last three weeks. The results indicated that the validities of the tests changed quite drastically across the measurement occasions. For instance, the validity of the mechanical principles test was $r = .29$ for the first three weeks and was $r = .06$ for the last three weeks. As another example, the validity of the interest in active occupations predictor was $r = .09$ for the first three weeks and $r = .26$ for the last three weeks.

In another early study along these lines Bass (1962) assessed the validity of two cognitive ability tests and a job knowledge test for predicting performance in a sample of food products salesmen across 48 months. Job performance was measured using supervisory ratings. Results indicated that while validity coefficients had a tendency to fluctuate over the short term, across long spans of time, there was a clear pattern whereby validity coefficients gradually decayed across measurement occasions. Taken together, the results of these two studies seemed to indicate that the criterion-related validity of predictors might change across time.

While these studies helped to stimulate interest in dynamic criterion-related validity coefficients, they did not provide definitive evidence. In a review article, Barrett et al. (1985) noted deficits in these studies that limited the inferences that could be drawn from them concerning the dynamicity of criterion-related validity coefficients. In terms

of Ghiselli and Haire's (1960) study, Barrett et al. noted that, while the point estimates of the validity coefficients did change, it is necessary to test whether the difference between the validity coefficients is greater than would be expected by sampling error alone using the Dunn and Clark (1969) z test for dependent correlations. After conducting these analyses, Barret et al. concluded that only five out of a total of twenty three paired comparisons of validity coefficients reported in Ghiselli and Haire's study were significantly different from one another. In applying the same test to the correlations reported in Bass (1962), Barret et al. found no evidence suggesting that criterion-related validity coefficients changed significantly across time. Thus, only minimal evidence existed suggesting that criterion-related validity coefficients changed across time.

Kaufman (1972) assessed the validity of an engineering achievement test for predicting performance of 110 engineers across three time points (1955-1962, 1962-1965, and 1966-1968). Three performance criteria measures were used: supervisory ratings, number of publications, and number of patents. The results of this study suggested that the engineering achievement test predicted papers and patents and that the validities did not fluctuate significantly across time.

Deadrick and Madigan (1990) assessed the validity of cognitive ability and psychomotor ability for predicting performance across six months in a sample of 413 sewing machine operators. A measure of objective output (i.e., average hourly production earnings per week) was used as the criterion measure. The weekly measures of performance were averaged within months to form monthly performance scores. The results indicated that psychomotor ability was a valid predictor of performance across all six measurement occasions, with validity coefficients ranging from .16 to .20. Cognitive

ability was a valid predictor of job performance for four of the six measurement

occasions and the validity coefficients ranged from .09 to .17. Furthermore, the validity

of cognitive ability seemed to strengthen somewhat across time. Overall, this study

provided only limited evidence suggesting that criterion-related validity coefficients

changed across time (note that no significance tests between correlations were reported).

While a large number of studies have assessed the validity of cognitive predictors

across time, very few studies have examined this issue for personality predictors.

Furthermore, even fewer still have considered issues concerning the bandwidth-fidelity

dilemma. Despite the lack of empirical research into the dynamicity of broad and narrow

personality measures across time, a number of theoretical perspectives related to changes

to validity coefficients across time have emerged than can be used that provide insight.

These include (1) the changing subjects model, (2) the changing tasks model, (3) and job

stage models. I review each of these theoretical perspectives in the following sections,

followed by a review of the literature into changes in criterion-related validity

coefficients across time.

Changing Subjects Model. Consistent with the vast literature into individual

differences predictors of job performance (cf. Schmidt & Hunter, 1998), the *changing

subjects model* conceptualizes individual characteristics as antecedents of job

performance. Many constructs used in personnel selection, such as cognitive ability, are

considered to be relatively stable throughout one's job tenure, if not their entire life

(Reeve & Bonaccio, 2011). However, the mechanisms through which cognitive ability

influences job performance may be more malleable across time.

For instance, research suggests that the relationship between cognitive ability and job performance occurs through job knowledge acquisition (Schmidt et al., 1986). While cognitive ability would be expected to remain stable across one's job tenure, individual differences in job knowledge are not expected to remain stable. The changing subjects model suggests that changes to these malleable antecedents of job performance that occur across time would result in changes in both performance and criterion-related validity coefficients of individual differences predictors across time (Adams, 1957; Alvarez & Hulin, 1973). Keeping with this example, presuming that job tasks remain relatively stable across time, cognitive ability may be a robust predictor of job performance early on, as employees who are high on cognitive ability would accumulate greater amounts of job knowledge faster than employees who are lower on cognitive ability and would thus outperform them. However, as time passes, the employees who are lower on cognitive ability would continue to accumulate job knowledge and their performance would begin to catch up to the employees who are higher on cognitive ability. This suggests that the validity of cognitive ability would degrade across time for jobs that remain relatively stable (Hulin et al., 1990; Keil & Cortina, 2001).

On the other hand, the causal mechanisms through which personality influences job performance are generally considered to be motivation-related (Barrick, Stewart, & Piotrowski, 2002; Judge & Ilies, 2002). That is, personality variables influence motivation, which influences job performance. In this case, the intervening variable may be less likely to change with time and as a result, the changing subjects model would seem to predict stability in personality validities.

Changing Tasks Model. While the changing subjects model suggests that performance determinants remain stable across time but that individual differences in these determinants change, the changing tasks model takes the opposite approach. That is, the *changing tasks model* proposes that an individual's job performance and the criterion-related validity of predictors would change due to a change in job requirements that results in a change in the determinants of job performance (Alvarez & Hulin, 1972; Fleishman & Hempbel, 1954). In this case, individual characteristics remain stable, but their contribution to job performance changes as job tasks change. For instance, agreeableness may be a poor predictor of performance (Barrick & Mount, 1991), however, if the job changes in such a fashion that teamwork becomes emphasized, the validity of agreeableness may increase substantially (Bell, 2007). The changing tasks model suggests that it is not simply the passage of time that is important when considering the dynamicity of predictors across time, but what specific changes to the job take place across that time period.

Job Stage Models. A number of job stage models exist in the literature. Murphy (1989), proposed a model that distinguishes between two job stages -- the transition stage and the maintenance stage. Each stage can be defined in terms of the demands that are placed on the employee, which influences the validity of predictors. *Transition stages* occur either when the employee is new to the job or when the employee faces changes to their work roles (Thoresen et al., 2004). In this phase, the employee faces demands related to learning and problem solving. As a result, individual differences in cognitive abilities should be robust determinants of job performance. On the other hand, a *maintenance stage* occurs when work roles have become well learned and relatively

53

automated. During this phase, performance depends on the repeated execution of a set of behaviors and therefore individual differences related to motivations, such as personality variables, should be robust correlates of job performance. On the other hand, as learning and problem solving demands are lower, the validity of cognitive ability should decrease.

Another stage model relevant for the dynamicity of predictor validities across time is the honeymoon-hangover model. Similar to the maintenance-transition model, the honeymoon-hangover model distinguishes between two phases that employees move through when transitioning to a new job or to changes in responsibilities. A *honeymoon period* occurs when employees are new to a job, during which commitment and motivation are maximized due to the novelty associated with beginning a new position or performing new tasks (Helmreich et al., 1986). As motivation would be relatively homogenous across employees in a honeymoon period, individual differences in cognitive abilities would be the primary determinants of performance early in one's tenure. However, with the passage of time, employees may come to experience a *hangover period*, which occurs as work tasks that were once novel and challenging become redundant (Boswell, Boudreau, & Tichy, 2005). During this period, as motivation would no longer be maximized across employees, individual differences that relate to motivation, such as personality (Judge & Ilies, 2002), would come to influence job performance (Minbashian et al., 2013).

Empirical Evidence of Dynamic Criterion-Related Validities. Only a handful of studies have investigated the stability of personality validity coefficients across time. A longitudinal field study conducted by Helmreich et al. (1986) drew on the honeymoon effect discussed earlier, proposing that the validity of the achievement facet of

consciousness would not predict performance early on the job (because the honeymoon effect would result in homogeneity in motivation across employees). However, as the honeymoon effect wore off and performance came to be influenced by individual differences in motivations, achievement would come to predict job performance. This study employed objective criterion measures. The results of this study supported their predictions, as achievement was not a valid predictor of performance measured during the first three months of the job, but was significantly related to performance during the following two criterion measurement occasions (i.e., six months and eight months on the job). Furthermore, the validity of achievement was stable across these later two time periods.

Thoresen et al. (2004) assessed the validity of the Big Five personality factors for predicting sales performance across four quarters in two samples. One sample had been on the job for an average of 11.11 years and had experienced little changes in responsibilities, thus indicative of a maintenance stage (this sample was referred to as the *maintenance* sample). The other sample had also been on the job for some time (mean tenure was 7.92 years), but at the start of the study, this sample of employees had experienced a change in responsibilities consist with a transition job stage (this sample was referred to as the *transition* sample). This research assessed the validity of the Big Five factors for predicting objective performance outcomes (i.e., sales figures) in these two samples across four quarters of data collection. The results of this study provided support for stable personality validities across time. In the maintenance sample, only conscientiousness and extraversion were significant predictors of job performance. Furthermore, their validities were largely stable across the four measurement occasions.

In the transition sample, interestingly, agreeableness was the only valid predictor of job performance and the validity was stable across the measurement occasions.

Finally, a cross-sectional study conducted by Stewart (1999) provides some insight concerning the validity of conscientiousness and two of its facets -- achievement and order -- for predicting objective criteria (i.e., sales figures) across maintenance and transition job stages. Stewart assessed the validity of conscientiousness and the achievement and order facets in two samples drawn from the same organization. The first (i.e., transition sample) was a sample of newly hired employees, while the second (i.e., maintenance sample) was a sample of employees who had been on the job for at least a year. Drawing on the personality literature as well as the maintenance-transition stage model of job performance, Stewart proposed that conscientiousness would be a valid predictor of performance in both samples. However, he proposed that the validities of order and achievement would vary between the samples. Specifically, during a transition stage when employees are learning their roles, behaviors and motivations related to the order facet of conscientiousness, such as an orientation towards structure and time management, would facilitate performance. However, in a maintenance stage, once employees are familiar with work processes, behavioral tendencies and motivations associated with order would bear less of an influence on job performance. On the other hand, behavioral tendencies and motivations related to the achievement facet of conscientiousness, such as goal commitment and persistence, would not be expected to influence performance in a transition stage, because motivation to perform well would be maximized and thus homogenous across employees (see the discussion above in relation to the *honeymoon effect*). However, in the maintenance stage, individual differences in

motivations to excel become critical, as employees must be motivated to repeat well-learned behaviors. Thus, Stewart proposed that achievement would be a valid predictor in the maintenance sample as opposed to the transition sample. The results of this study supported the predictions. Specifically, order was a valid predictor of performance in the transition sample but not in the maintenance sample. On the other hand, achievement was a valid predictor of performance in the maintenance sample but not in the transition sample. Finally, conscientiousness was a valid predictor of performance in both samples.

Cumulatively, the results of these studies suggest that the validities of the broad personality factors may be more stable across time than those of personality facets. That is, facets of the Big Five factors may be valid predictors of job performance at certain points in time, for instance, during certain job stages, when the motivations and behavioral tendencies associated with that narrow trait are aligned with job requirements (e.g., Helmreich et al., 1986; Stewart, 1999). However, as the Big Five factors encompass a number of motivations and behavioral propensities that would be relevant for performance across job stages, the validity of the Big Five factors for predicting job performance should be more stable across time.

> **Hypothesis 4(a):** The validity of global conscientiousness for predicting job performance is stable across time.
>
> **Hypothesis 4(b):** The validity of conscientiousness facets for predicting job performance change across time.
>
> **Hypothesis 5(a):** The validity of global emotional stability for predicting job performance is stable across time.
>
> **Hypothesis 5(b):** The validity of emotional stability facets for predicting job

57

performance change across time.

Hypothesis 6(a): The validity of global extraversion for predicting job performance is stable across time.

Hypothesis 6(b): The validity of extraversion facets for predicting job performance change across time.

Predictors of Intraindividual Trajectories. Recall that dynamic criteria can be conceptualized as within-person trajectories in job performance scores across time (Hofmann et al., 1992, 1993). Indeed, this is the most explicit way to examine the manner in which performance unfolds across time at the individual-level (Sturman, 2007). As reviewed earlier, research into this phenomenon has indicated that job performance change across time can be defined in terms of linear, quadratic, and cubic slopes (Day et al., 2004; Thoresen et al., 2004). Furthermore, research has indicated considerable amounts of between-person variance in performance growth parameters, suggesting that performance unfolds over time differently for different people. This notion has received notable empirical support, as research has indicated that a number of individual differences constructs, such as cognitive ability, job experience, and biodata are useful predictors of growth parameters (Deadrick et al., 1997; Ployhart & Hakel, 1998). However, while a number of investigations along these lines have been conducted, only two have considered personality predictors.

First, a study by Thoresen et al. (2004), which I discussed earlier in relation to dynamic personality validities, examined the Big Five factors as predictors of performance trajectories in a sample of employees that was said to be in a maintenance job stage. The researchers proposed that conscientiousness should be positively related

58

to performance growth, because employees who are high on this trait would have a strong desire for continuous improvement. This prediction was supported, as the results of this study indicated that conscientiousness was positively related to the cubic performance trend.

Another study examined openness as a predictor of performance growth in a sample of newly hired financial consultants across four years (Minbashian et al., 2013). This study was unique in that job performance was measured by annual supervisory ratings as opposed to objective outcomes. These researchers drew on the honeymoon-hangover effect to propose that openness should be positively related to the performance trend. This was based off the notion that employees who are high on openness would be less susceptible to performance detriments associated with the hangover effect, as openness is indicative of a greater motivation to continue to learn and develop. The results supported their conclusions.

A limitation of both of these studies is a reliance on broad personality measures. That is, issues of bandwidth-fidelity in personality testing have not been considered in relation to intraindividual trajectories in job performance. I address this issue in the present dissertation.

Predictors of Performance Growth. Job performance growth may be dependent upon a motivation to achieve and self-directed goal setting (Murphy, 1989; Stewart, 1999). As a result, I propose that conscientiousness is a valid predictor of performance growth. Additionally, the facet of conscientiousness most relevant in accounting for performance in later stages of employees is that of achievement (Helmreich et al., 1986).

As a result, I propose that achievement will predict performance growth. I assess the remaining facets of conscientiousness as predictors of growth as research questions.

Hypothesis 7: Conscientiousness is positively related to performance growth.

Hypothesis 8: The achievement facet of conscientiousness is positively related to performance growth.

Research Question 9: Is the order facet of conscientiousness related to performance growth?

Research Question 10: Is the self-control facet of conscientiousness related to performance growth?

Research Question 11: Is the responsibility facet of conscientiousness related to performance growth?

In terms of the remaining factors of personality and their facets, there is little theoretical rationale for predicting relationships with performance growth. Therefore, I assess these effects as research questions.

Research Question 12: Is emotional stability related to performance growth?

Research Question 13: Is the serenity facet of emotional stability related to performance growth?

Research Question 14: Is the firmness facet of emotional stability related to performance growth?

Research Question 15: Is extraversion related to performance growth?

Research Question 16: Is the enthusiasm facet of extraversion related to performance growth?

Research Question 17: Is the assertiveness facet of extraversion related to

performance growth?

CHAPTER III: METHODS

Database

Data were collected from incumbents of a large business services firm located in the United States as part of a concurrent local validation study. A total of $N = 1,113$ incumbents were administered a personality assessment for research purposes. Item-level supervisory ratings of job performance were collected from the organization's records for four years prior to the personality test administration. The data files containing the personality scores and the supervisory ratings were merged together. The resulting data set included $N = 173$ employees with scores on the personality measure and performance scores for the full four year period.

Participants. Participants were $N = 173$ employees. The employees occupied one of the following three jobs: Consultant (16.2%), Customer Service (48.6%), and Information Technology (32.9%). These data were missing for the remaining 2.3% of employees. Note that these three positions were determined to be similar enough to be included in the same local validation study and to be rated on the same job performance dimensions. Demographic information was not available for any employees.

Measures.

Personality. Personality was assessed using a forced-choice adaptive assessment of personality designed for personnel selection purposes. The measure was explicitly developed around the facets of the Five Factor Model of personality. The development procedure followed the methods outlined by Stark, Chernyshenko, Drasgow, and White (2012). Note that because of the proprietary nature of the scale, the sample items discussed in the present section do not reflect the actual content of the measure. In terms

of the conscientiousness facets, achievement was assessed by the *industriousness* subscale (e.g., "I do my best to succeed at anything I do."), order was assessed by the *orderliness* subscale (e.g., "I always clean up my workspace without being asked"), self-control was assessed by the *restraint* subscale (e.g., "I always think things through before acting."), and responsibility was assessed by the *dutifulness* subscale (e.g., "I always do what I have promised."). In terms of emotional stability, serenity was measured by the *worry free* subscale (e.g., "I usually have no problem relaxing.") and firmness was measured by the hardiness subscale (e.g., "I never get depressed."). Finally, in terms of extraversion, assertiveness was measured by the *assertiveness* subscale (e.g., "I often assume leadership roles.") and enthusiasm was measured by the *gregariousness* subscale (e.g., "I enjoy meeting new people.").

Job Performance. Job performance was assessed using supervisory ratings. Each year, each employee's direct supervisor evaluated the employee's job performance for developmental purposes. Performance was assessed using the following five items: drive and persistence, quality, personal development, facilitating teamwork, and proactivity. Each item was defined behaviorally on the rating scale. The description of each performance dimension was broad, encompassing a range of behaviors that spanned performance dimensions. Therefore, it was not possible to group items together into more narrow performance dimensions. Each item was rated on a four-point Likert-type scale, with 4 being the highest rating and 1 being the lowest.

Because I am examining issues pertaining to performance change and validity change across time, it was necessary to establish that the construct validity of the performance measure did not vary across time. To do so, I first established that a unitary

factor structure adequately accounted for the data across the measurement occasions. Therefore, I specified a four-factor Confirmatory Factor Analysis (CFA) model where performance at each time point were the four factors, each indicated by the five items assessed at that time period. Residuals for each respective item were allowed to correlate across measurement occasions. Missing data was handled using Full Imputation Maximum Likelihood (FIML).

In order to test the consistency of the factor structure across time, I tested for metric invariance (i.e., equivalence of factor loadings across time points; Vandenberg & Lance, 2000). To do so, I first added the constraint to the CFA model that the factor loadings associated with each item were equal across measurement occasions. To establish metric invariance, it is necessary to demonstrate that adding the equality constraint did not significantly reduce the fit of the model (Horn & McArdle, 1992). Therefore, I tested the fit of the constrained CFA model against the unconstrained CFA model using a nested chi-square difference test. A significant chi-square difference would suggest that the factor loadings (and thus the construct validity of the performance measure) did in fact change across time. Conversely, a non-significant chi-square difference test would indicate failure to detect significant changes to the factor loadings across time and I would therefore be able to proceed with my analyses. I computed overall job performance scores by averaging across items within each measurement occasion.

Analyses

Before testing my substantive hypotheses, I conducted a series of analyses pertaining to criterion dynamicity. First, I assessed rank-order changes to criterion scores

across time. Second, I assessed intraindividual trajectories in job performance scores across time. In terms of rank-order changes, I first computed observed correlations between criterion scores across measurement occasions. However, these correlations would be biased downward due to measurement error. In order to account for the influence of measurement error on observed correlations, I used the following approaches (a) I corrected the criterion-criterion correlations for unreliability using the interrater reliability estimated derived by Viswesvaran et al. (1996; i.e., .52) and (b) I corrected the criterion-criterion correlations using coefficient alpha reliability estimates derived in my data.

Note that each of these approaches have unique strengths and weaknesses. The first approach is useful because it accounts for all sources of measurement error that influence the observed criterion-criterion correlations, but it relies on published estimates of interrater reliability rather than those that are local to my sample. Furthermore, the first approach does not account for the effect of differences in reliability across measurement occasions, but rather, assumes that measurement error influences all estimates equally. On the other hand, the coefficient alpha approach does use estimates of reliability derived in my sample, but only accounts for the influence of internal consistency on observed correlations as opposed to interrater reliability. However, using both of these approaches provides a more comprehensive assessment of true performance change across time than relying on only a single approach in isolation.

In terms of assessing intraindividual trajectories in job performance across time, I used a Structural Equation Modeling (SEM) framework (i.e., latent growth modeling) in order to assess individual-level growth parameters and between-person variation among

the growth parameters. I conducted these analyses in M Plus (Muthén & Muthén, 2012) following a model building approach whereby, using criterion scores across the four time points, I fit an intercept only model. Following this, I incorporated a linear trend to the model and tested the fit of the linear trend model against the intercept only model using a -2LL difference test. Finally, I fit a quadratic trend to the model and tested the quadratic trend model against the best fitting model from the prior step using a -2LL difference test. In each of these models, the residual vairances were constrained to be equal. In order to approximate the range in which 95% of participants' growth parameters fell, I took the square root of the variance associated with each growth parameter, multiplied it by two, and added and subtracted the resulting value from the mean of the respective growth parameter. The resulting values reflected the range in which approximately 95% of participants' growth parameters fell.

Hypotheses 1-3 and research questions 1-8 concerned the criterion-related validity of broad and narrow personality traits for predicting job performance criteria. To address these hypotheses, I first formed composite job performance scores by averaging composite scores across each of the four criterion measurement occasions. To test these hypotheses and research questions, I computed zero-order criterion-related validity coefficients by correlating the predictor and criterion scores. A significant correlation would indicate that the individual difference variable was a valid predictor of job performance.

In addition to assessing bivariate predictor-criterion relationships, it is useful to assess how predictor variables operate as a set in accounting for job performance criterion. Therefore, I followed up on the correlational analyses with multiple regression

66

analyses. I specified five regression models. For the first model, I regressed overall composite job performance scores onto the global conscientiousness, emotional stability, and extraversion scores. In the second model, I regressed overall job performance scores onto all of the personality facets examined in my dissertation. For the third regression model, I regressed overall job performance onto the conscientiousness facets. For the fourth regression model, I regressed overall job performance onto the emotional stability facets. Finally, for the fifth regression model, I regressed overall job performance onto the extraversion facets. These analyses provided me some insight concerning how these predictors operated as a set in predicting the criterion. Specifically, the multiple correlation indicates that amount of variance accounted for in job performance by all personality variables included in the regression model, while the beta-weights provide some insight concerning the contribution of each predictor to prediction of the criterion.

However, it is important to note that interpreting the beta-weights alone provides an incomplete picture of the relative contribution of each predictor in a regression model towards prediction of the criterion. As a supplement to a traditional regression approach, research suggests incorporating relative importance analysis in order to accurately determine the proportion of predicted variance accounted for by each predictor included in a regression model (LeBreton, Hargis, Griepentrog, Oswald, & Ployhart, 2007; LeBreton, Ployhart, & Ladd, 2004; Tonidandel & LeBreton, 2011). I conducted these analyses using a relative weight analysis approach, consistent with the methods outlined by Tonidandel and LeBreton (in press). Relative weight analysis produces two types of coefficients -- relative weights and rescaled-relative weights. The former indicates the proportion of variance in the criterion variable accounted for by each predictor variable,

while the later indicates the percentage of predicted variance accounts for by each predictor variable (i.e., the relative weight divided by the model R^2).

The statistical significance of relative weights is estimated by bootstrapping 95% confidence intervals around each relative weight. However, as relative weights reflect a proportion of the model R^2, they can never be less than or equal to zero. Therefore, it is not possible to assess the overlap of the lower bound of the bootstrapped confidence interval with zero. As an alternate approach, research suggests that the statistical significance of relative weights can be ascertained by incorporating a random variable in the model. As the random variable has a population relative weight of zero, evidence of a statistically significant relative weight is observed when the value is significantly different than the relative weight of the random variable (Tonidandel, LeBreton, & Johnson, 2009). I assessed the statistical significance of the relative weights using this procedure, again, consistent with the methods outlined by Tonidnadel and LeBreton (in press).

Hypotheses 4-6 concern the stability of criterion-related validity coefficients across time for both broad and narrow personality variables. I used two approaches to test these hypotheses. First, I computed zero-order correlation coefficients between each predictor and job performance measured at each of the four measurement occasions. Then, consistent with the methods outlined by Barrett et al. (1985), I tested for significant differences between all pairs of validity coefficients using the Dunn and Clark (1969) z test. A significant z difference score would indicate that the two dependent validity coefficients differ from one another outside of what would be expected by chance. While this approach has been widely used in the dynamic validity literature, there is an

important shortcoming associated with this approach. Observed validity coefficients may vary from one another due to changes in criterion reliability across time and thus, a significant z difference test could be driven by changes in reliability alone, as opposed to true changes to criterion-related validity coefficients. In other words, validity change may be confounded with changes to criterion reliability across time. Considering this, it was also important to test these hypotheses using a framework that accounted for any changes to criterion reliability across time.

In response to this issue, I used a second approach to testing the dynamic validity hypotheses and research questions using a SEM framework. Specifically, I began with the job performance CFA model reviewed earlier (with the factor loading constraints in place). I estimated the validity of each personality variable for predicting job performance at each time point by regressing the four criterion factor scores onto each predictor (note that I did not use a multiple regression framework here, but rather, a separate model was specified for each personality trait). In these analyses, the disturbance terms associated with each performance factor were allowed to intercorrelate. The path coefficients provided an estimate of the validity of the personality variable for predicting job performance at each measurement occasion. Note that, as the job performance scores were modeled as latent variables, the path coefficients were not influenced by measurement error (as indexed by internal consistency -- note that these path coefficients could still be biased downward due to idiosyncratic rating errors as would be indexed by interrater reliability). Thus, as compared to observed validity coefficients, these path coefficients would be less likely to vary from one another due to variation in the reliability of criterion scores across time.

In order to test for significant differences between path coefficients across measurement occasions (i.e., dynamic validities), I used logic similar to the metric invariance test discussed earlier. Specifically, I added a constraint to the model specifying that the path coefficients were equal across the measurement occasions. If imposing the constraint to this model made the model fit the data significantly worse (as evidenced by a statistically significant nested chi-square difference test), it would suggest that the path coefficients were not equal and thus, I could conclude that the validity coefficients changed significantly across time.

Hypotheses 7 and 8 and research questions 9-15 concern broad and narrow personality variables as predictors of performance growth across time. In order to address these hypotheses, I returned to the best fitting of the latent growth models as discussed earlier. I then specified a model for each personality trait whereby the latent growth parameters were regressed onto that trait. A significant relationship between the personality trait and the performance intercept would suggest that the personality trait was associated with performance at the first measurement occasion. A significant relationship between the personality trait and a slope parameter would suggest that the personality trait was associated with the trajectory of performance across time.

CHAPTER IV: RESULTS

My results section is organized as follows. I first review results pertaining to the test of metric invariance of the job performance measure across measurement occasions. Following this, I review results pertaining to performance change across time. My next set of results concern cross-sectional personality-job performance relationships. Finally, I review the results of my longitudinal analyses.

Metric Invariance

As many of my analyses involve assessing changes to job performance scores across time as well as changes to validity coefficients across time, it is necessary to assess the equivalence of the construct validity of performance scores across time. To do so, I assessed the equivalence of factor loadings across time (i.e., metric invariance; Vandenberg & Lance, 2000). I first tested whether a unitary factor structure adequately accounted for the data across measurement occasions. Specifically, I specified a four-factor CFA model where performance at each time point were the four factors, each indicated by the five items assessed at that time point. Residuals for each respective item were allowed to correlate across measurement occasions. Missing data was handled using Full Imputation Maximum Likelihood (FIML). The model fit the data well [χ^2 (134) = 139.64, *n.s.*, CFI = .99, RMSEA = .02]. The factor loadings associated with the model are reported in Table 1. All factor loadings were statistically significant at *p* < .05.

Table 1

Standardized factor loadings for the job performance CFA model

Item	Time 1	Time 2	Time 3	Time 4
Drive and Persistence	.55	.65	.56	.52
Quality	.76	.52	.61	.56
Personal Development	.52	.60	.50	.27
Facilitating Teamwork	.71	.50	.52	.37
Proactivity	.49	.57	.62	.37

Note. All factor loadings are statistically significant at *p* < .05.

In order to test the equivalence of factor loadings across measurement occasions (i.e., metric invariance), I first added the constraint to the CFA model that the factor loadings associated with each item were equal across measurement occasions. This model fit the data well [χ^2 (146) = 159.14, *n.s.*, CFI = .98, RMSEA = .02]. To establish metric invariance, it is necessary to establish that adding the equality constraint did not significantly reduce the fit of the model (Horn & McArdle, 1992). Therefore, I tested the fit of the constrained CFA model against the unconstrained CFA model using a nested chi-square difference test. The results indicated no significant difference [$\Delta\chi^2$(12) = 19.50, *n.s.*], and thus, I did not observe evidence suggesting that the factor structure of performance ratings differed between the measurement occasions. The factor loadings for the constrained model are reported in Table 2.

Table 2

Standardized factor loadings for the constrained job performance CFA model

Item	Time 1	Time 2	Time 3	Time 4
Drive and Persistence	.61	.59	.58	.45
Quality	.71	.62	.63	.49
Personal Development	.57	.51	.49	.35
Facilitating Teamwork	.64	.52	.57	.44
Proactivity	.57	.56	.53	.39

Note. Factor loadings were constrained to be equal across measurement occasions. All factor loadings are statistically significant at *p* < .05.

Performance Change

My analyses pertaining to performance change included an assessment of rank-order changes to criterion scores across time as well as intraindividual trajectories in performance scores across time. I first review results pertaining to rank-order changes. Specifically, I review observed criterion-criterion correlations followed by the correlations corrected for unreliability using the following indices of criterion reliability: (1) internal consistency and (2) inter rater. Following this, I review results pertaining to intraindividual performance trajectories.

Rank-Order Changes. My first set of results pertains to rank-order changes to job performance across the four measurement occasions. Along these lines, I first conducted zero-order correlations between criterion scores across the measurement occasions. The results of this analysis are reported in Table 3.

Table 3

Observed zero-order correlations between criterion scores across time

		1	2	3	4
1	Time 1	(.74)			
2	Time 2	.29*	(.68)		
3	Time 3	.32*	.42*	(.70)	
4	Time 4	.30*	.26**	.39*	(.50)

Note. *p < .05. Coefficient alpha reliability estimates are listed in the diagonal.

The results of this analysis indicated that criterion scores were moderately intercorrelated across time. An interesting finding is that the familiar simplex pattern was not observed. Rather, the relationships between performance at time one and performance at the remaining measurement occasions were similar, as opposed to decreasing across time. The relationship between performance at time 1 and time 2 was r = .29, while the time 1-time 3 relationship was r = .32, and the time 1-time 4 relationship was r = .30.

In terms of the generally small magnitude of these correlations, it is important to consider that they are biased downward due to measurement error. In order to estimate the correlations between latent performance constructs, I conducted two additional analyses. First, I corrected the observed correlations for unreliability using the interrater reliability estimate reported in Viswesvaran et al. (1996) of .52. Second, I corrected the observed correlations for criterion unreliability using the internal consistency estimates (i.e., coefficient alpha) derived in the present study. These results are reported in Tables 4 and 5.

Table 4

Zero-order correlations between criterion scores across time corrected for interrater

reliability

		1	2	3
1	Time 1			
2	Time 2	.56*		
3	Time 3	.62*	.81*	
4	Time 4	.58*	.50*	.75*

Note. **p < .05.*

Table 5

Zero-order correlations between criterion scores across time corrected using coefficient

alpha reliability estimates

		1	2	3
1	Time 1			
2	Time 2	.41*		
3	Time 3	.44*	.61*	
4	Time 4	.49*	.45*	.65*

Note. **p < .05.*

The results reported in Table 4 pertaining to the corrections corrected for
interrater reliability are useful for illustrating the problems associated with interpreting
observed criterion-criterion correlations as indicators of rank-order performance change
across time. As interrater reliability of supervisory ratings is only .52 (Viswesvaran et
al., 1996), observed correlations would underestimate these relationships. In reviewing

the corrected correlations in Table 4, it is apparent that, while rank-order changes did occur, there is nonetheless a large stable component of job performance across times. As these analyses used the same reliability estimate to correct all correlations, these analyses do not help me to determine if the lack of a simplex patterns was potentially due to differences in reliability of criterion scores across time points.

Table 5 reports the correlations between criterion scores across time corrected for unreliability using local internal consistency estimates. While these estimates are not fully corrected, as they are still biased downward due to idiosyncratic rater error (e.g., halo error), they do allow me to account for the influences of differences in internal consistency across measurement occasions on observed correlations. Thus, these analyses are useful for understanding if a simplex pattern simply does not exist in these data, or if it has been obscured by variation in reliability across measurement occasions.

Inspecting the values reported in Table 5 provides support for the former explanation. It appears that, once correcting the observed values for unreliability using the coefficient alpha estimates derived in the present study, a simplex pattern still does not emerge in these data. The time 1-time 2 correlation was $r = .41$, the time 1-time 2 correlation was $r = .44$, and the time 1-time 3 correlation was $r = .49$. Thus, the correlation actually strengthened somewhat across time. Overall, despite the often-regarded ubiquity of a simplex pattern in criterion-criterion relationships across time (Austin et al., 1989), such a pattern of correlations was not observed in these data.

Intraindividual Performance Trajectories. My next set of results pertains to the form of performance change across the four measurement occasions. I conducted these analyses using a latent growth modeling approach in SEM. I followed a model building

approach in order to identify the manner in which job performance unfolded across time. I began with an intercept only model whereby I estimated the average intercept (i.e., performance at time 1) and the amount of variation around that intercept. The model fit the data poorly [$\chi^2(11) = 81.20$, $p < .05$, CFI $= .21$, RMSEA $= .19$]. The -2LL for the model was 300.02 with 3 degrees of freedom. Due to the poor fit of the model, I did not interpret the latent intercept parameter.

Next, I added a linear trend parameter to the model. The model fit the data poorly [$\chi^2(8) = 34.62$, $p < .05$, CFI $= .70$, RMSEA $= .14$]. The -2LL for the model was 253.50 with 6 degrees off freedom. Comparing the intercept only model to the intercept and linear trend model using a -2LL difference test indicated that the model that included a linear trend fit the data better than the model including only the intercept parameter [$\Delta\chi^2(3) = 46.52$, $p < .05$].

Finally, I specified a model that included an intercept, linear trend, and quadratic trend. This model fit the data well [$\chi^2(4) = 9.10$, $n.s.$, CFI $= .94$, RMSEA $= .09$]. The -2LL for the model was 227.98 with 10 degrees of freedom. Comparing the intercept and linear trend model to the current model using a -2LL difference test indicated that the current model fit the data better than the model including only the intercept and linear trend parameters [$\Delta\chi^2(4) = 25.52$, $p < .05$]. The parameter estimates associated with the model are reported in Table 6 and the average performance trend is depicted graphically in Figure 1.

Table 6

Latent growth parameters

		M	SD	1	2
1	Int	3.31*	.25*		
2	Lin	.18*	.23*	-.03*	
3	Quad	-.04*	.06	.01	-.01

Note. *p < .05. Int = intercept, Lin = linear trend, Quad = quadratic trend. Covariances between growth parameters are indicated in the cells.

Figure 1

Graphical depiction of mean performance growth
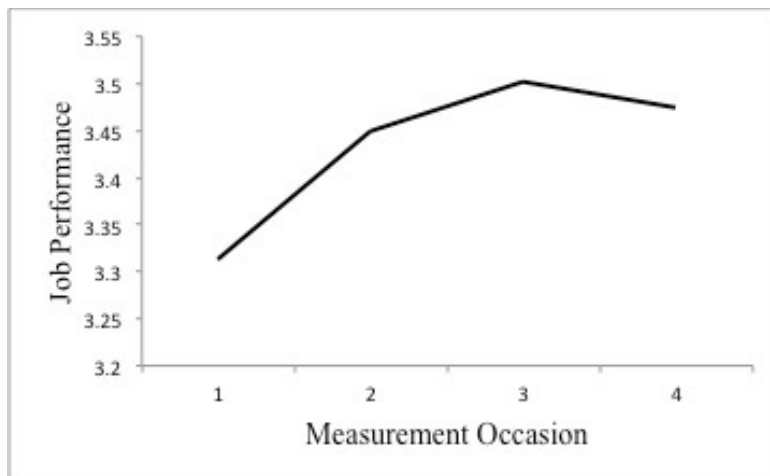
Overall, the results of the latent growth modeling analysis are consistent with the vast majority of research into performance trends, suggesting that performance tends to increase earlier on and ultimately levels off. The results indicated an overall mean intercept of 3.31 ($\sigma_{intercept} = .25$). The variation around the intercept is depicted graphically in Figure 2. The intercept was negatively correlated with the linear trend, suggesting that employees with higher performance at time 1 tended to improve at a

slower rate as compared to employees with lower performance at time 1. In terms of the linear trend, the results indicated a mean of .18 ($\sigma_{intercept}$ = .23). The variation around the linear trend is depicted graphically in Figure 3. Finally, the results indicated a mean quadratic performance trend of -.04 and failed to detect significant variability around the mean trend. Thus, the results failed to detect any evidence suggesting that the quadratic trend varied between employees.

Figure 2

Graphical depiction of mean performance growth and variability around intercept
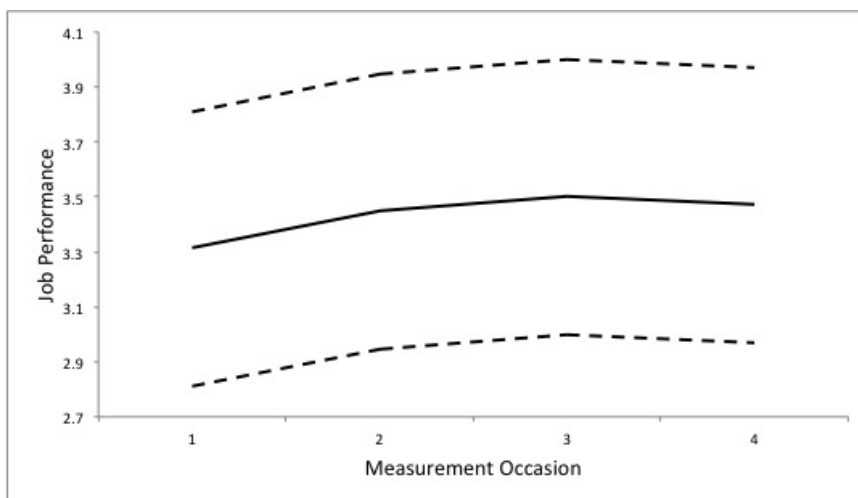
Figure 3

Graphical depiction of mean performance growth and variation around linear trend
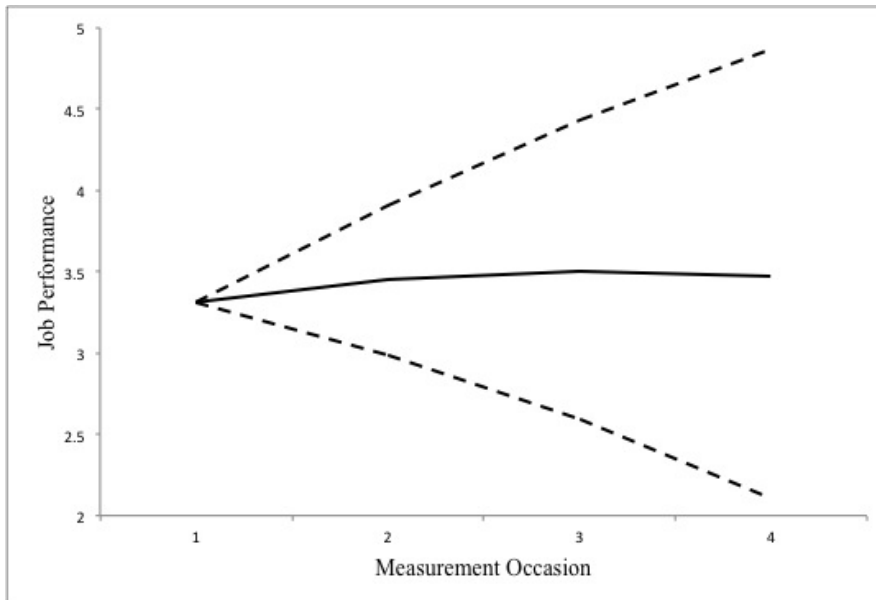
Criterion-Related Validity

In this section, I begin testing my substantive hypotheses pertaining to criterion-related validity. First, I test my cross-section hypotheses and research questions pertaining to the validity of broad and narrow personality traits as predictors of job performance. Following this, I test my longitudinal hypotheses pertaining to changes to criterion-related validity across time.

Cross-Sectional Analyses. My first set of hypotheses and research questions concerned the validity of conscientiousness, emotional stability, and extraversion for predicting overall job performance criteria (i.e., averaged across measurement occasions). To test these hypotheses, I first computed zero-order conscientiousness-, emotional stability-, and extraversion-job performance correlations. Following this, I specified a multiple regression model whereby job performance was regressed onto

conscientiousness, emotional stability, and extraversion. Finally, I conducted a relative

weight analysis in order to estimate the relative contribution of each of the personality

traits to the prediction of job performance. The results of the correlational analyses are

reported in Table 7, the results of the multiple regression model are reported in Table 8,

and the results of the relative weight analyses are reported in Table 9.

Table 7

Observed validity coefficients for conscientiousness, emotional stability, and extraversion

		M	SD	1	2	3
1	Conscientiousness	.11	.37			
2	Emotional Stability	-.08	.50	.28*		
3	Extraversion	-.22	.56	.23*	.39*	
4	Performance	3.44	.22	.26*	.03	.17*

Note. *p < .05.

Table 8

Results of personality factor multiple regression analysis

	B/R	SE	β	t	p
Conscientiousness	.15	.05	.25	3.29	.001
Emotional Stability	-.04	.04	-.10	-1.26	.210
Extraversion	.06	.03	.15	1.89	.060
$R(R^2)$	.30(.09)				

Table 9

Results of personality factor relative weight analysis

	RW	Rescaled-RW
Conscientiousness	.06*	69.79%
Emotional Stability	.00	3.69%
Extraversion	.02	26.52%
R^2	.09	

***Note.** *p < .05*

Hypothesis 1 predicted that conscientiousness was a valid predictor of job performance. The zero-order correlations reported in Table 7 indicated that conscientiousness was a valid predictor of job performance ($r = .26$, $p < .05$). In terms of the regression analysis (see Table 8), overall, the three personality factors accounted for 9% of the variance in job performance. The results of the regression analysis indicated a statistically significant effect of conscientiousness on job performance ($B = .15$, $\beta = .25$, $p < .05$). Furthermore, in terms of the relative weight analysis (see Table 9), the results indicated that conscientiousness accounted for 69.97% of the variance in job performance explained by the three personality factors (i.e., 9%) and the relative weight associated with conscientiousness was statistically significant. Overall, these results provided support for hypothesis 1.

Research question 1 asked if emotional stability was a valid predictor of job performance. The zero-order correlations reported in Table 7 failed to support an emotional stability-job performance relationship ($r = .03$, $p > .05$). The results of the regression analysis (see Table 8) failed to detect a statistically significant effect of

emotional stability on job performance (B = -.04, β = -.10, p > .05). Finally, in terms of

the relative weight analysis (see Table 9), emotional stability accounted for very little

explained variance in performance ratings (Rescaled-RW = 3.69%). Overall, I failed to

detect any evidence suggesting that emotional stability was a valid predictor of job

performance.

Research question 2 asked if extraversion was a valid predictor of job

performance. The zero-order correlations reported in Table 7 indicated that extraversion

was a valid predictor of job performance (r = .17, p < .05). In terms of the regression

analysis (see Table 8), the results failed to detect a statistically significant effect of

extraversion on job performance (B = .06, β = .15, p > .05). However, the effect of

extraversion on job performance did approach the threshold for statistical significance, as

the p value was equal to .06. In terms of the relative weights (see Table 9), extraversion

accounted for 26.52% of explained variance in performance ratings, though its relative

weight was not statistically significant. Cumulatively, these results were somewhat

mixed regarding the validity of extraversion. In terms of bivariate predictor-criterion

relationships, my results indicated that extraversion did predict job performance.

However, once accounting for the effect of conscientiousness, the effect of extraversion

was much smaller. Nonetheless, overall, the results did suggest that extraversion was a

valid predictor of job performance.

In terms of narrow traits, I specified a series of hypotheses and research questions

concerning the relationship between personality facets and job performance.

In order to test these hypotheses and research questions, I computed zero-order

correlations between the narrow personality traits and job performance. In order to

estimate how these traits predicted performance as a set, I specified a multiple regression model whereby job performance was regressed onto the narrow traits. In order to estimate the relative contribution of each predictor included in the regression model to predicted variance, I supplemented the multiple regression model with relative weight analysis. I conducted the multiple regression and relative weight analysis including all narrow traits across personality factors and each narrow trait within personality factors. The correlations are reported in Table 10. The results of the multiple regression and relative importance analyses including all narrow traits across personality factors are reported in Tables 11 and 12. Note that the results of the multiple regression analysis indicated that the personality facets accounted for 13% of the variance in job performance.

Table 10

Zero-order correlations between personality facets and job performance

	M	SD	1	2	3	4	5	6	7	8
1 Achievement	.17	.66								
2 Order	-.05	.66	.14							
3 Self-Control	.17	.60	.26*	.05						
4 Responsibility	.15	.45	.37*	.02	.30*					
5 Serenity	-.14	.58	.15*	-.02	.15*	.25*				
6 Firmness	-.03	.61	.21*	.05	.12	.34*	.41*			
7 Assertiveness	-.11	.72	.42*	.09	-.02	.23*	.28*	.33*		
8 Enthusiasm	-.32	.66	.20*	.04	-.17*	.14	.12	.32*	.33*	
9 Performance	3.44	.22	.30*	.16*	.08	.09	.02	.02	.25*	.02

Note. * $p < .05$.

Table 11

Results of personality facet multiple regression analysis across personality factors

	B/R	*SE*	β	*t*	*p*
Achievement	.07	.03	.23	2.59	.010
Order	.04	.02	.11	1.51	.133
Self-Control	.01	.03	.02	.25	.807
Responsibility	-.00	.04	-.01	-.08	.935
Serenity	-.01	.03	-.03	-.41	.679
Firmness	-.02	.03	-.06	-.70	.483
Assertiveness	.06	.03	.20	2.25	.025
Enthusiasm	-.02	.03	-.07	-.80	.428
$R(R^2)$	.36(.13)				

Table 12

Results of personality facet relative weight analysis across personality factors

	RW	Rescaled-RW
Achievement	.06*	45.02%
Order	.02	13.40%
Self-Control	.00	2.26%
Responsibility	.00	2.55%
Serenity	.00	0.80%
Firmness	.00	1.50%
Assertiveness	.04*	32.87%
Enthusiasm	.00	1.61%
R^2	.13	

Hypothesis 2 predicted that the achievement facet of conscientiousness was a valid predictor of job performance. In terms of the zero-order correlations (see Table 10), the results indicated that achievement was a statistically significant predictor of job performance ($r = .30$, $p < .05$). The results of the first multiple regression model (i.e., including all personality facets; see Table 11) indicated a statistically significant effect of achievement on job performance ($B = .07$, $\beta = .23$, $p < .05$). Furthermore, the results of the first relative weight analysis (i.e., including all personality facets; see Table 12) indicated that achievement accounted for 45.02% of the variance in job performance explained by the personality facets. The results of the multiple regression and relative weight analyses including the conscientiousness facets only are reported in Tables 13 and

14. The results indicated that the conscientiousness facets as a set accounted for 10% of the variance in job performance. The results of the multiple regression analysis indicated a statistically significant effect of achievement on job performance ($B = .09$, $\beta = .29$, $p < .05$) and the results of the relative weight analysis indicated that achievement accounted for 75.23% of the variance in job performance explained by the conscientiousness facets. Overall, these results provided strong support for hypothesis 2.

Table 13

Results of multiple regression analysis for conscientiousness facets

	B/R	SE	β	t	p
Achievement	.09	.03	.29	3.53	.001
Order	.04	.02	.12	1.57	.118
Self-Control	.00	.03	.00	.03	.975
Responsibility	-.01	.04	-.02	-.20	.845
$R(R^2)$	.32(.10)				

Table 14

Results of relative weight analysis for conscientiousness facets

	RW	Rescaled-RW
Achievement	.08*	75.23%
Order	.02	18.40%
Self-Control	.00	2.37%
Responsibility	.00	4.00%
R^2	.10	

Note. *$p < .05$.

Research question 3 asked if the order facet of conscientiousness was a valid

predictor of job performance. In terms of the zero-order correlations (see Table 10), the

results indicated a statistically significant validity coefficient of $r = .16$ ($p < .05$). The

results of the first set of multiple regression and relative weight analyses (i.e., including

all personality facets) failed to detect a significant effect of order on job performance (see

Tables 11 and 12). Specifically, the results of the multiple regression analysis failed to

detect a significant effect of order ($B = .04$, $\beta = .11$, $p > .05$), while the results of the

relative weight analysis indicated that order accounted for 13.40% of explained variance

in job performance. While 13.40% is a nontrivial amount, the results failed to detect a

statistically significant relative weight. The results of the regression and relative weight

analyses including only conscientiousness facets (see Tables 13 and 14) were parallel to

those just reviewed. The results of the regression analysis failed to detect a significant

effect of order on job performance once accounting for the other facets of

conscientiousness ($B = .04$, $\beta = .12$, $p > .05$), while the relative weight analysis failed to

detect a significant relative weight for order (RW = .02, Rescaled-RW = 18.40%, *p* > .05). Overall, my results indicated that, in isolation, order was a significant predictor of job performance, but once accounting for the effect of other personality facets, the effect was not significant.

Research question 4 asked if the self-control facet of conscientiousness was a valid predictor of job performance. The zero-order correlation (see Table 10) failed to detect a significant relationship between self-control and job performance (*r* = .08, *p* > .05). In terms of the first regression model (see Table 11), the results failed to detect a significant effect of self-control on job performance (*B* = .01, β = .02, *p* > .05). The results of the first relative weight analysis (see Table 12) indicated that self-control accounted for a trivial portion of explained variance in job performance (RW = .00, Rescaled-RW = 2.26%, *p* > .05). The results of the regression (*B* = .00, β = .00, *p* > .05) and relative weight analyses (RW = .00, Rescaled-RW = 2.37%, *p* > .05) involving conscientiousness facets only were parallel (see Tables 13 and 14), failing to detect a significant effect of self-control on job performance. Thus, my results failed to detect any evidence suggesting that the self-control facet of conscientiousness was a valid predictor of job performance.

Research question 5 asked if the responsibility facet of conscientiousness was a valid predictor of job performance. The correlation analyses (see Table 10) failed to detect a significant responsibility-job performance relationship (*r* = .09, *p* > .05). The results of the regression (*B* = -.00, β = -.01, *p* > .05) and relative weight analyses (RW = .00, Rescaled-RW = 2.55%, *p* > .05) involving all personality facets (see Tables 11 and 12) failed to detect a significant effect of responsibility on job performance.

Furthermore, the results of the regression ($B = -.01$, $\beta = -.02$, $p > .05$) and relative weight analyses (RW = .00, Rescaled-RW = 4.00%, $p > .05$) involving only conscientiousness facets (see Tables 13 and 14) also failed to detect a significant effect of responsibility on job performance. Thus, I failed to detect any evidence indicating that responsibility was a valid predictor of job performance.

Research question 6 asked if the serenity facet of emotional stability was a valid predictor of job performance. In terms of the zero-order correlation (see Table 10), the results failed to detect a significant effect ($r = .02$, $p > .05$). The results of the regression model including all personality facets (see Table 11) failed to detect a significant effect of serenity on job performance ($B = -.01$, $\beta = -.03$, $p > .05$) and the results of the relative weight analysis (see Table 12) also failed to detect a significant effect and indicated that serenity accounted for only a small proportion of variance in job performance explained by personality facets (RW = .00, Rescaled-RW = .80%, $p > .05$). The results of the multiple regression and relative weight analyses involving only emotional stability facets are reported in Tables 15 and 16. The results indicated that emotional stability facets accounted for 0% of the variance in job performance. Overall, my results failed to detect a significant effect of the serenity facet of emotional stability on job performance.

Table 15

Results of multiple regression analysis for emotional stability facets

	B/R	SE	β	t	p
Serenity	.01	.03	.01	.16	.873
Firmness	.01	.03	.02	.22	.826
$R(R^2)$	.03(.00)				

Table 16

Results of relative weight analysis for emotional stability facets

	RW	Rescaled-RW
Serenity	.00	40.89%
Firmness	.00	59.11%
R^2	.00	

Research question 7 asked if the firmness facet of emotional stability was a valid predictor of job performance. As indicated in Table 10, the firmness-job performance correlation was $r = .02$ and was not statistically significant ($p > .05$). The results of the regression analysis involving all personality facets (see Table 11) failed to detect a significant effect of firmness on job performance ($B = -.02$, $\beta = -.06$, $p > .05$) and the relative weight analysis (see Table 12) indicated that firmness accounted for a trivial proportion of variance in job performance explained by the personality facets (RW = .00, Rescaled-RW = 1.50%, $p > .05$). The results of the regression model involving only emotional stability facets (see Table 15) also failed to detect an effect of firmness on job

performance, as both emotional stability facets as a set accounted for 0% of the variance in job performance. Thus, these results failed to detect a significant effect of the firmness facet of emotional stability on job performance.

Hypothesis 3 predicted that the assertiveness facet of extraversion was a valid predictor of job performance. In terms of the zero-order correlations (see Table 10), the results provided initial support for hypothesis 3, as the results indicated a significant assertiveness-job performance correlation ($r = .25$, $p < .05$). The results of the regression analysis involving all personality facets (see Table 11) indicated a significant effect of assertiveness on job performance ($B = .06$, $\beta = .20$, $p < .05$). Furthermore, the associated relative weight analysis (see Table 12) indicated that assertiveness accounted for 32.87% of the variance in job performance explained by personality facets. The results of the multiple regression and relative weight analyses involving only extraversion facets appear in Tables 17 and 18 and indicate that the extraversion facets accounted for 7% of the variance in job performance. The regression analysis indicated a significant effect of assertiveness on job performance ($B = .08$, $\beta = .27$, $p < .05$). Furthermore, the results of the relative weight analysis indicated that the variance in job performance explained by extraversion facets was predominantly driven by assertiveness (RW = .06, Rescaled-RW = 96.49%), though the relative weight was not statistically significant. Overall, the pattern of results observed here suggested that the assertiveness facet of extraversion was a valid predictor of job performance, providing support for hypothesis 3.

Table 17

Results of multiple regression analysis for extraversion facets

	B/R	*SE*	*β*	*t*	*p*
Assertiveness	.08	.02	.27	3.44	.001
Enthusiasm	-.02	.03	-.07	-.87	.388
$R(R^2)$	.26(.07)				

Table 18

Results of relative weight analysis for extraversion facets

	RW	Rescaled-RW
Assertiveness	.06	96.49%
Enthusiasm	.00	3.51%
R^2	.07	

Research question 8 asked if the enthusiasm facet of extraversion was a valid predictor of job performance. The results of the correlation analysis (see Table 10) indicated a correlation of $r = .02$ ($p > .05$). The results of the regression analysis that included all personality facets (see Table 11) failed to detect a significant effect of the enthusiasm facet of extraversion on job performance ($B = -.02$, $β = -.07$, $p > .05$) and the results of the associated relative weight analysis (see Table 12) indicated that enthusiasm accounted for a trivial amount of the variance in job performance explained by personality facets (RW = .00, Rescaled-RW = 1.61%, $p > .05$). The results of the regression ($B = -.02$, $β = -.07$, $p > .05$) and relative weight analyses (RW = .00, Rescaled-

RW = 3.51%) including only extraversion facets (see Tables 17 and 18) were parallel. Overall, these results failed to detect an effect of the enthusiasm facet of extraversion on job performance.

Longitudinal Analyses. The first set of longitudinal hypotheses and research questions concerned the stability of criterion-related validity coefficients across time. To test these hypotheses and research questions, I used the following two approaches. First, I computed zero-order correlations between the predictors and criterion scores at each measurement occasion and tested significant differences between validity coefficients across time using the Dunn and Clark (1969) z test. A significant z difference score would provide evidence suggesting that validity coefficients were dynamic across time. Second, I used a SEM approach whereby I modeled job performance as latent factors and regressed each latent job performance factor onto each personality predictor. I then tested for significant differences between path coefficients by adding a constraint to the model that path coefficients had to be equal. Evidence for dynamic validities would be observed if the constrained model fit the data significantly worse than did the unconstrained model.

Hypothesis 4a predicted that the validity of conscientiousness is stable across time. The zero-order validity coefficients for conscientiousness and its facets across the four time points are reported in Table 19. The results indicated that global conscientiousness was a valid predictor of job performance for three of the four time points when employing two-tailed significance testing criteria, and for all four time points when employing one-tailed significance testing criteria. The observed correlations ranged from $r = .14$ to $r = .24$. The results of the Dunn and Clark (1969) z tests (see

Table 20) failed to indicate any significant variation in conscientiousness validities across

time. In terms of the SEM model where I regressed the latent job performance scores

onto conscientiousness, results indicated good model fit [$\chi^2(162) = 174.92$, *n.s.*, CFI =

.98, RMSEA = .02; see Table 21]. Furthermore, the path coefficients reported in Table

22 indicated that conscientiousness was a valid predictor of latent job performance scores

across all four measurement occasions. As indicated in Table 21, constraining the path

coefficients to be equal across measurement occasions did not significantly reduce the fit

of the model [$\Delta\chi^2(3) = 1.31$, *n.s.*]. The path coefficients from the constrained model are

reported in Table 23. Overall, the pattern of results observed here provided support for

hypothesis 4a, as the results failed to detect any evidence suggesting that the validity of

conscientiousness was dynamic across time.

Table 19

Zero-order validity coefficients for conscientiousness and facets

	Time 1	Time 2	Time 3	Time 4
Conscientiousness	.19*	.14$^+$	.24*	.17*
Achievement	.24*	.05	.30*	.26*
Order	.08	.15*	.15*	.05
Self-Control	.06	.05	.05	.06
Responsibility	.07	.10	.06	.02

Note. *$p < .05$. $^+p < .05$ (one-tailed).

Table 20

Results of Dunn and Clark (1969) z tests comparing global conscientiousness and facet validities across time

	$r_{t1,t2}$	$r_{t1,t3}$	$r_{t1,t4}$	$r_{t2,t3}$	$r_{t2,t4}$	$r_{t3,t4}$
Conscientiousness	.56	-.58	.23	-1.24	-.33	.85
Achievement	2.12*	-.71	-.23	-3.12*	-2.31*	.50
Order	-.77	-.79	.33	.00	.08	1.19
Self-Control	.11	.11	.00	.00	-.11	-.12
Responsibility	-.33	.11	.55	.49	.86	.47

Note. *$p < .05$.

Table 21

Fit of constrained and unconstrained models involving conscientiousness and facets

	Unconstrained			Constrained			
	χ^2(162)	CFI	RMSEA	χ^2(165)	CFI	RMSEA	$\Delta\chi^2$(3)
Conscientiousness	174.92	.98	.02	176.23	.98	.02	1.31
Achievement	179.76	.97	.03	188.04	.96	.03	8.28*
Order	169.69	.99	.02	171.82	.99	.02	2.13
Self-Control	171.16	.99	.02	171.29	.99	.02	0.13
Responsibility	167.79	.99	.01	168.85	.99	.01	1.06

Note. *$p < .05$. $\Delta\chi^2$(3) = results of a nested chi-square difference test with three degrees of freedom comparing the constrained to the unconstrained model.

Table 22

Path coefficients from unconstrained conscientiousness and facet SEM models

	Time 1	Time 2	Time 3	Time 4
Conscientiousness	.19(.22)*	.14(.19)*	.19(.28)*	.13(.24)*
Achievement	.13(.27)*	.03(.07)	.14(.35)*	.11(.36)*
Order	.05(.10)	.08(.19)*	.07(.17)*	.02(.08)
Self-Control	.03(.06)	.04(.09)	.03(.06)	.03(.08)
Responsibility	.06(.08)	.08(.13)	.04(.07)	.02(.03)

Note. *$p < .05$. Standardized coefficients appear in cells with unstandardized coefficients following in parentheses.

Table 23

Path coefficients from constrained conscientiousness and facet SEM models

	Time 1	Time 2	Time 3	Time 4
Conscientiousness	.15(.18)*	.15(.21)*	.15(.23)*	.15(.28)*
Achievement	.10(.22)*	.10(.24)*	.10(.28)*	.10(.35)*
Order	.05(.10)	.05(.12)	.05(.13)	.05(.16)
Self-Control	.03(.06)	.03(.07)	.03(.07)	.03(.09)
Responsibility	.04(.06)	.04(.07)	.04(.07)	.04(.09)

Note. *$p < .05$.

Hypothesis 4b predicted that the validities of conscientiousness facets are dynamic across time. In terms of achievement, the zero-order correlations (see Table 19) indicated a significant relationship between achievement and job performance for three of the four measurement occasions. The correlations ranged from $r = .05$ to $r = .30$. The

results of the Dunn and Clark (1969) z tests (see Table 20) indicated that the time 2

correlation differed significantly from the time 1, time 3 and time 4 correlations. In terms

of the SEM model, the results indicated good model fit [$\chi^2(162) = 179.76$, $n.s.$, CFI $= .97$,

RMSEA $= .03$; see Table 21] and the path coefficients (see Table 22) indicated that

achievement was a valid predictor of job performance for three of the four measurement

occasions. The results of the nested chi-square difference test comparing the

unconstrained model to constrained model indicated that constraining the path

coefficients to be equal significantly reduced model fit [$\Delta\chi^2(3) = 8.28$, $p < .05$]. The path

coefficients from the constrained model are reported in Table 23. Overall, these results

provided evidence suggesting that the validity of the achievement facet of

conscientiousness was dynamic across time.

In terms of the order facet of conscientiousness, the results of the correlational

analyses (see Table 19) indicated that order was a valid predictor of job performance

during two of the four measurement occasions. The observed validity coefficients ranged

from $r = .05$ to $r = .15$. The results of the Dunn and Clark (1969) z tests (see Table 20)

failed to detect any evidence suggesting that the validity of order changed significantly

across time. In terms of the SEM analysis, the results indicated good model fit [$\chi^2(162) =$

169.69, $n.s.$, CFI $= .99$, RMSEA $= .02$] and the path coefficients (see Table 22) indicated

that order was a valid predictor of job performance during only two of the four

measurement occasions. The results of the nested chi-square difference test comparing

the constrained to the unconstrained model [$\Delta\chi^2(3) = 2.13$, $n.s.$] failed to indicate that the

constrained model fit the data significantly worse than the unconstrained model. The

path coefficients from the constrained model are reported in Table 23. The results of

these analyses provided some support for dynamic validities of order, as the results indicated a significant order-job performance relationship during only two of the four measurement occasions, however, the fluctuation in validity coefficients was not greater than what would be expected due to sampling error alone.

In terms of the self-control facet of conscientiousness, the results of the correlational analyses (see Table 19) failed to detect a statistically significant self-control-job performance relationship during any of the four measurement occasions. The observed correlations ranged from $r = .05$ to $r = .06$. The results of the Dunn and Clark (1969) z tests (see Table 20) failed to detect evidence of dynamic validity for self-control. The SEM model whereby the four latent job performance scores were regressed onto self-control fit the data well [$\chi^2(162) = 171.16$, CFI = .99, RMSEA = .02; see Table 21] and the results failed to detect a significant effect of self-control on job performance (see Table 22). The results of the nested chi-square difference test comparing the fit of the unconstrained to the constrained model [$\Delta\chi^2(3) = 0.13$, $n.s.$] indicated that fixing the path coefficients to be equal did not significantly reduce model fit. Path coefficients from the constrained model are reported in Table 23. Overall, these results did not provide any evidence suggesting that the validity of the self-control facet of conscientiousness was dynamic across time.

In terms of the responsibility facet of conscientiousness, the results of the correlational analyses (see Table 19) failed to detect a significant responsibility-job performance relationship during any of the four measurement occasions. The observed correlations ranged from $r = .02$ to $r = .10$. The results of the Dunn and Clark (1969) z tests (see Table 20) failed to detect any evidence suggesting that these validity

coefficients changed significantly across time. The SEM model in which the four latent job performance factors were regressed onto the responsibility facet fit the data well $[\Delta\chi^2(162) = 167.79, CFI = .99, RMSEA = .01;$ see Table 21]. The results failed to detect a significant effect of responsibility on job performance at any of the measurement occasions (see Table 22). The results of the nested chi-square difference test comparing the constrained model to the unconstrained model $[\Delta\chi^2(3) = 1.06, n.s.]$ was not statistically significant, and thus failed to indicate that the validity of responsibility changed significantly across time. Path coefficients from the constrained model are reported in Table 23. Overall, the results observed in these analyses failed to indicate that the validity of responsibility was dynamic across time.

When considering all of the dynamic validity analyses pertaining to conscientiousness facets as a whole, I did observe some support for hypothesis 4b. Specifically, in terms of achievement, the results indicated that the validity coefficients varied significantly across time. Furthermore, while I did not observe evidence of significant validity change with relation to order, the results of the correlational and SEM analyses both indicated a significant effect of order on job performance for only two of the four measurement occasions. Overall, hypothesis 4b was partially supported.

Hypothesis 5a predicted that the validity of emotional stability is stable across time. The zero-order emotional stability-job performance correlations across the four measurement occasions are reported in Table 24. The results indicated that validities range from $r = -.01$ to $r = .08$ and none were statistically significant. The results of the Dunn and Clark (1969) z tests comparing these correlations are reported in Table 25 and failed to detect any evidence suggesting that the validities changed across time. The

SEM model in which the four latent job performance scores were regressed onto emotional stability fit the data well [χ^2(162) = 178.64, CFI = .97, RMSEA = .02; see Table 26] and the results failed to detect a significant effect of emotional stability on job performance for any of the four measurement occasions (see Table 27). The results of the chi-square difference test comparing the constrained to the unconstrained model [$\Delta\chi^2$(3) = 2.01, *n.s.*] indicated that constraining the path coefficients to be equal across the measurement occasions did not significantly reduce model fit. Path coefficients from the constrained model are reported in Table 28. The results reviewed here provided support for hypothesis 5a.

Table 24

Zero-order validity coefficients for emotional stability and facets

	Time 1	Time 2	Time 3	Time 4
Emotional Stability	-.01	-.00	.08	.02
Serenity	.01	-.02	.04	.04
Firmness	-.02	.01	.10	-.01

Table 25

Results of differences tests between global emotional stability and facet validities across time

	$r_{t1,t2}$	$r_{t1,t3}$	$r_{t1,t4}$	$r_{t2,t3}$	$r_{t2,t4}$	$r_{t3,t4}$
Emotional Stability	-.11	-1.01	-.33	-.97	-.21	.71
Serenity	.33	-.34	-.33	-.73	-.64	.00
Firmness	-.33	-1.35	-.11	-1.09	.21	1.30

Table 26

Fit of constrained and unconstrained models involving emotional stability and facets

	Unconstrained			Constrained			
	$\chi^2(162)$	CFI	RMSEA	$\chi^2(165)$	CFI	RMSEA	$\Delta\chi^2(3)$
Emotional Stability	178.64	.97	.02	180.65	.97	.02	2.01
Serenity	183.52	.96	.03	184.45	.97	.03	0.93
Firmness	179.54	.97	.03	182.55	.97	.03	3.01

Note. $\Delta\chi^2(3)$ = results of a nested chi-square difference test with three degrees of freedom comparing the constrained to the unconstrained model.

Table 27

Path coefficients from unconstrained emotional stability and facet SEM models

	Time 1	Time 2	Time 3	Time 4
Emotional Stability	-.01(-.01)	-.00(-.01)	.06(.11)	.02(.05)
Serenity	.01(.02)	-.01(-.03)	.03(.06)	.03(.08)
Firmness	-.02(-.04)	.01(.01)	.05(.13)	.00(.01)

Table 28

Path coefficients from constrained emotional stability and facet SEM models

	Time 1	Time 2	Time 3	Time 4
Emotional Stability	.02(.03)	.02(.03)	.02(.04)	.02(.05)
Serenity	.02(.03)	.02(.03)	.02(.04)	.02(.05)
Firmness	.01(.02)	.01(.02)	.01(.03)	.01(.03)

Hypothesis 5b predicted that the validities of the emotional stability facets will change across time. In terms of serenity, the correlational analyses (see Table 24) indicated observed validity coefficients ranging from $r = -.02$ to $r = .04$. None of the validity coefficients were statistically significant. Results of the Dunn and Clark (1969) z tests comparing these validity coefficients (see Table 25) failed to detect any evidence suggesting that the validity coefficients differed significantly across time. The SEM model whereby the four latent job performance scores were regressed onto serenity fit the data well [$\chi^2(162) = 183.52$, CFI = .96, RMSEA = .02; see Table 26]. The path coefficients (reported in Table 27) failed to indicate a significant effect of serenity on job

performance for any of the measurement occasions. The results of the chi-square difference test comparing the constrained model to the unconstrained model [$\Delta\chi^2(3) =$ 0.93, *n.s.*] was not statistically significant and thus, my results failed to detect any evidence suggesting that the validity of the serenity facet of emotional stability changed significantly across time. Path coefficients from the constrained model are reported in Table 28. Overall, the results observed in these analyses failed to indicate that the validity of the serenity facet of emotional stability was dynamic across time.

In terms of firmness, the correlational analyses indicated observed validity coefficients ranging from $r = -.02$ to .10 (see Table 24). None of the validity coefficients were statistically significant. Results of the Dunn and Clark (1969) *z* tests (See Table 25) failed to detect any evidence suggesting that these validity coefficients changed significantly across time. The SEM model in which the four latent job performance factors were regressed onto the firmness facet fit the data well [$\chi^2(162) = 179.54$, CFI = .97, RMSEA = .03; see Table 26]. Consistent with the results of the correlation analyses, the results of the SEM analyses failed to detect a significant effect of firmness on job performance at any of the measurement occasions (see Table 27). The results of the nested chi-square difference test comparing the constrained model to the unconstrained model [$\Delta\chi^2(3) = 3.01$, *n.s.*] was not statistically significant, and thus, failed to indicate that the validity of firmness changed significantly across time. Path coefficients from the constrained model are reported in Table 28. The pattern of results observed for the serenity and firmness facets of emotional stability failed to support hypothesis 5b.

Taken together, the results of the analyses just reviewed pertaining to the validities of emotional stability and its facets across time failed to detect any evidence

supporting dynamic validities. However, this finding must be tempered by the absence of

validity evidence observed for the emotional stability factor and its facets. That is, our

results indicated that emotional stability and its facets were overall poor predictors of job

performance and that they remained poor predictors of job performance across time.

Hypothesis 6a predicted that the validity of extraversion was stable across time.

The zero-order extraversion-job performance correlations (see Table 29) ranged from $r =$

.09 to $r = .17$. Only one of the validity coefficients was statistically significant. The

results of the Dunn and Clark (1969) z tests (see Table 30) failed to detect any evidence

suggesting that the validities were dynamic across time. The SEM model whereby the

four latent job performance factors were regressed onto extraversion fit the data well

$[\chi^2(162) = 3.01, .98, \text{RMSEA} = .02;$ see Table 31]. The results of the SEM model were

consistent with the zero-order correlations, indicating that extraversion was a valid

predictor of job performance during only one of the four measurement occasions (see

Table 32). Furthermore, constraining the path coefficients to be equal did not

significantly reduce the fit of the model $[\Delta\chi^2(3) = 0.40, n.s.]$. Path coefficients from the

constrained model are reported in Table 33. Overall, the results of these analyses failed

to indicate that the validity of extraversion changed significantly across time, providing

support for hypothesis 6a.

Table 29

Zero-order validity coefficients for extraversion and facets

	Time 1	Time 2	Time 3	Time 4
Extraversion	.09	.10	.13	.17*
Assertiveness	.18*	.13	.21*	.19*
Enthusiasm	-.04	.04	-.01	.09

Note. *$p < .05$.

Table 30

Results of differences tests between global extraversion and facet validities across time

	$r_{t1,t2}$	$r_{t1,t3}$	$r_{t1,t4}$	$r_{t2,t3}$	$r_{t2,t4}$	$r_{t3,t4}$
Extraversion	-.11	-.45	-.89	-.37	-.76	-.48
Assertiveness	.56	-.34	-.11	-.99	-.66	.24
Enthusiasm	-.88	-.34	-1.44	.61	-.54	-1.18

Table 31

Fit of constrained and unconstrained models involving extraversion and facets

	Unconstrained			Constrained			
	χ^2(162)	CFI	RMSEA	χ^2(165)	CFI	RMSEA	$\Delta\chi^2$(3)
Extraversion	174.55	.98	.02	174.95	.98	.02	0.40
Assertiveness	184.07	.96	.03	186.19	.97	.03	2.13
Enthusiasm	168.21	.99	.02	170.34	.99	.01	2.12

Note. $\Delta\chi^2$(3) = results of a nested chi-square difference test with three degrees of freedom comparing the constrained to the unconstrained model.

Table 32

Path coefficients from unconstrained extraversion and facet SEM models

	Time 1	Time 2	Time 3	Time 4
Extraversion	.07(.12)	.05(.11)	.08(.17)	.08(.23)*
Assertiveness	.10(.22)*	.05(.13)	.10(.27)*	.07(.25)*
Enthusiasm	-.02(-.04)	.02(.04)	-.00(-.00)	.04(.12)

Note. *$p < .05$.

Table 33

Path coefficients from constrained extraversion and facet SEM models

	Time 1	Time 2	Time 3	Time 4
Extraversion	.07(.13)*	.07(.15)*	.07(.16)*	.07(.20)*
Assertiveness	.08(.17)*	.08(.20)*	.08(.22)*	.08(.27)*
Enthusiasm	.02(.04)	.02(.04)	.02(.04)	.02(.06)

Note. *$p < .05$.

  Hypothesis 6b predicted that the validities of the extraversion facets will change across time. In terms of assertiveness, the correlational analyses indicated observed validity coefficients ranging from $r = .13$ to .21 (see Table 29). Three out of four of these validity coefficients were statistically significant. Results of the Dunn and Clark (1969) z tests (see Table 30) comparing these validity coefficients failed to detect any evidence suggesting that the validity coefficients changed significantly across time. The SEM model whereby the four latent job performance scores were regressed onto assertiveness fit the data well [$\chi^2(162) = 184.07$, CFI = .96, RMSEA = .03]. These results also indicated a significant effect of assertiveness on job performance for three out of the four

measurement occasions (see Table 32). The results of the chi-square difference test

comparing the constrained model to the unconstrained model [$\Delta\chi^2(3) = 2.13$, *n.s.*] was

not statistically significant and thus, my results failed to detect any evidence suggesting

that the validity of the assertiveness facet of extraversion changed significantly across

time. Path coefficients from the constrained model are reported in Table 33.

In terms of enthusiasm, the correlational analyses indicated observed validity

coefficients ranging from $r = -.04$ to .09 (see Table 29). None of the validity coefficients

were statistically significant. Results of the Dunn and Clark (1969) z tests (see Table 30)

failed to detect any evidence suggesting that these validity coefficients changed

significantly across time. The SEM model in which the four latent job performance

factors were regressed onto the enthusiasm facet fit the data well [$\chi^2(162) = 168.21$, CFI

= .99, RMSEA = .02]. The results failed to detect a significant effect of enthusiasm on

job performance at any of the measurement occasions (see Table 32). The results of the

nested chi-square difference test comparing the constrained model to the unconstrained

model [$\Delta\chi^2(3) = 2.12$, *n.s.*] was not statistically significant, and thus, failed to indicate

that the validity of the enthusiasm facet changed significantly across time. The path

coefficients from the constrained model are reported in Table 33. Overall, the results

observed in these analyses failed to indicate that the validity of the enthusiasm facet of

extraversion was dynamic across time. The pattern of results observed for the

assertiveness and enthusiasm facets of extraversion failed to support hypothesis 5b.

The next series of hypotheses and research questions pertain to personality factors

and facets as predictors of performance growth. In order to test these hypotheses and

research questions, I began with the latent growth model estimated earlier. Then, I

regressed the latent growth parameters onto the personality factor or facet specified in the hypothesis or research question. The results of these analyses are reported in Table 34.

Hypothesis 7 predicted that conscientiousness was positively related to performance growth. To test hypothesis 7, I regressed the latent performance growth parameters onto conscientiousness. The model fit the data well [$\chi^2(5) = 10.29$, *n.s.*, CFI = .95, RMSEA = .08]. Results indicated that conscientiousness was significantly related to the performance intercept (i.e., performance at time 1; see Table 34), but failed to detect a significant relationship between conscientiousness and either the linear or quadratic growth parameters. Therefore, hypothesis 7 was not supported. Hypothesis 8 predicted that the achievement facet of conscientiousness was positively related to performance growth. To test this hypothesis, I regressed the latent performance growth parameters onto achievement. The model fit the data well [$\chi^2(5) = 15.73$, $p < .05$, CFI = .90, RMSEA = .11]. Results indicated that achievement was significantly related to the performance intercept (see Table 34), but failed to detect a significant relationship between achievement and either the linear or quadratic growth parameters. Thus, hypothesis 8 was not supported.

Table 34

Effect of personality factors and facets on latent growth parameters

	Intercept	Linear Trend	Quadratic Trend
Conscientiousness	.16*	.00	-.00
Achievement	.11*	-.05	.02
Order	.04	.05	-.02
Self-Control	.03	-.00	.00
Responsibility	.06	.02	-.01
Emotional Stability	-.01	.04	-.01
Serenity	.00	-.00	.00
Firmness	-.02	.06	-.02
Extraversion	.06	-.00	.00
Assertiveness	.08*	-.02	.00
Enthusiasm	-.01	.01	.00

Notes. *$p < .05$.

Research question 9 asked if the order facet of conscientiousness was a predictor of performance growth. To test this research question, I regressed the latent growth parameters onto order. The model fit the data well [$\chi^2(5) = 9.14$, *n.s.*, CFI = .95, RMSEA = .07]. Results failed to detect a significant effect of order on any of the latent growth parameters. Research question 10 asked if the self-control facet of conscientiousness was related to performance growth. To test research question 10, I regressed the latent growth parameters onto self-control. The model fit the data well [$\chi^2(5) = 9.11$, *n.s.*, CFI = .95, RMSEA = .07]. Results failed to detect a significant effect of self-control on any of the

latent growth parameters. Research question 11 asked if the responsibility facet of conscientiousness was a predictor of performance growth. To test research question 11, I regressed the latent performance growth parameters onto responsibility. The model fit the data well [$\chi^2(5) = 9.18$, *n.s.*, CFI = .95, RMSEA = .07]. Results failed to detect a significant effect of responsibility on the latent growth parameters.

Research Question 12 asked if emotional stability was a predictor of performance growth. To test research question 12, I regressed the latent performance growth parameters onto emotional stability. The model fit the data well [$\chi^2(5) = 10.00$, *n.s.*, CFI = .94, RMSEA = .08]. Results failed to detect a significant effect of emotional stability on the performance growth parameters. Research question 13 asked if the serenity facet of emotional stability was a predictor of performance growth. To test research question 13, I regressed the latent performance growth parameters onto the serenity facet. The model fit the data well [$\chi^2(5) = 9.29$, *n.s.*, CFI = .95, RMSEA = .07]. Results failed to indicate a significant effect of serenity on any growth parameters. Research question 14 asked if the firmness facet of emotional stability predicted performance growth. To test research question 14, I regressed the latent performance growth parameters onto firmness. The model fit the data well [$\chi^2(5) = 10.39$, *n.s.*, CFI = .94, RMSEA = .08]. Results failed to detect a significant effect of firmness on performance growth.

Research question 15 asked if extraversion predicted performance growth. To test research question 15, I regressed the latent performance growth parameters onto extraversion. The model fit the data well [$\chi^2(5) = 9.14$, *n.s.*, CFI = .95, RMSEA = .07]. Results failed to detect a significant effect of extraversion on performance growth. Research question 16 asked if the assertiveness facet of extraversion was a valid predictor

of performance growth. To test research question 16, I regressed the latent performance growth parameters onto assertiveness. The model fit the data well [$\chi^2(5) = 9.60$, *n.s.*, CFI = .95, RMSEA = .07]. Results indicated that assertiveness was significantly positively related to the performance intercept (see Table 34). However, I did not observe any evidence suggesting that assertiveness was related to the performance trend. Finally, research question 17 asked if the enthusiasm facet of extraversion was related to the performance trend. To test research question 17, I regressed the latent performance growth parameters onto enthusiasm [$\chi^2(5) = 10.27$, *n.s.*, CFI = .94, RMSEA = .08]. Results failed to detect a significant effect of enthusiasm on the performance trend. Thus, overall, I did not observe any evidence suggesting that any personality traits assessed in the present study were related to performance trend.

Ancillary Analyses

My dissertation focused on the validity of personality factors and facets for predicting global performance criteria. However, some research has argued that the validity of personality facets should be highest when validated against narrow performance criteria. In the following sections, I review the results of analyses pertaining to the validity of personality factors and facets for predicting the five dimensions of performance assessed in the present study. Note that each dimension was assessed with only a single item. Therefore, in order to obtain a more stable (i.e., reliable) estimate of dimension-level criteria scores, I averaged each item across time points. As a result, each performance dimension was a composite of four items rather than only one item. The results of these analyses pertaining to conscientiousness are reported in Table 35.

Table 35

Validity of conscientiousness and facets for predicting narrow performance criteria

	Persistence	Quality	Development	Teamwork	Proactivity
Conscientiousness	.19*	.24*	.09	.12	.28*
Achievement	.24*	.24*	.12	.18*	.28*
Order	.15[a]	.18*	.00	.05	.16*
Self-Control	-.02	.07	.06	.04	.14
Responsibility	.09	.10	.04	.00	.08

Note. *$p < .05$. Persistence = drive and persistence, Development = personal development, Teamwork = facilitating teamwork. [a]rounded up from .145.

First, consistent with the analyses pertaining to overall job performance scores, I failed to detect significant self-control- and responsibility-job performance relationships. In terms of global conscientiousness, achievement, and order, a visual inspection of the correlations indicates that the validity of these traits for predicting performance was somewhat lower for personal development and facilitating teamwork dimensions as compared to the remaining three performance dimensions. Global conscientiousness was a significant predictor of drive and persistence, quality, and proactivity performance dimensions. Achievement was a significant predictor of drive and persistence, quality, facilitating teamwork, and proactivity performance dimensions. Achievement was the only significant correlate of facilitating teamwork identified in these analyses. Finally, order was a significant predictor of quality and proactivity performance dimensions.

The results of the analyses pertaining to global emotional stability and its facets are reported in Table 36. In terms of global emotional stability, results failed to indicate a

significant relationship with any of the performance criteria. I observed the same results

for the serenity and firmness facets of emotional stability. Cumulatively, the results of

these analyses failed to indicate a significant relationship between emotional stability or

its facets with any of the performance dimensions.

Table 36

Validity of emotional stability and facets for predicting narrow performance criteria

	Persistence	Quality	Development	Teamwork	Proactivity
Emotional Stability	-.05	.05	-.10	.05	.12
Serenity	.01	.05	-.10	.01	.10
Firmness	-.09	.03	-.06	.08	.10

Note. *$p < .05$. Persistence = drive and persistence, Development = personal

development, Teamwork = facilitating teamwork.

Finally, the results of the analyses pertaining to global extraversion and its facets

are reported in Table 37. Global extraversion predicted job performance on the drive and

persistence and proactivity dimensions. The assertiveness facet predicted performance

on the drive and persistence, quality, and proactivity dimensions. Finally, the results

failed to detect a significant effect of the enthusiasm facet on any performance

dimension.

Table 37

Validity of extraversion and facets for predicting narrow performance criteria

	Persistence	Quality	Development	Teamwork	Proactivity
Extraversion	.15*	.15	.03	.10	.19*
Assertiveness	.22*	.21*	.12	.10	.25*
Enthusiasm	.02	.03	-.08	.05	.06

Note. *p* < .05. Persistence = drive and persistence, Development = personal development, Teamwork = facilitating teamwork.

CHAPTER V: DISCUSSION

The purpose of the present dissertation was to integrate two literatures that are both critically important to the science and practice of I-O Psychology. The first was the literature pertaining to the dynamicity of job performance. The dynamic criteria literature has been vitally important to research and practice in I-O Psychology (Austin & Villanova, 1992; Ghiselli, 1956; Sturman, 2007). Indeed, such research has vast implications for conceptualizations of job performance as well as for personnel selection practices. Research into the dynamicity of job performance has indicated that (a) rank-order changes in job performance occur across time and (b) intraindividual trajectories in job performance exist, as do between-person differences in such trajectories. At present, the dynamic nature of job performance criteria has been examined in a large number of studies (e.g., Sturman et al., 2005; Hofmann et al., 1993).

The second of these literatures was that pertaining to the bandwidth-fidelity dilemma in personality testing. The emergence and popularization of the Five Factor Model of personality has had a dramatic influence on personnel selection research and practice (Hough & Ones, 2001). However, while research had indicated generalizable validity coefficients for some of the Big Five personality factors (Barrick & Mount, 1991; Hurtz & Donovan, 2000; Salgado, 1997), subsequent research had argued that validity gains could be realized by examining personality facets rather than the broad factors (Hogan & Holland, 2003; Hough, 1992). A sizeable body of literature has been dedicated to examining the relative merits of personality factors and facets in selection contexts by comparing the magnitude of validity coefficients of broad vs. narrow traits (Dudley et al., 2006; Judge et al., 2013; Mount & Barrick, 1995; Salgado et al., 2013).

116

In the present dissertation, I argued that considering issues pertaining to dynamic criteria and the bandwidth-fidelity dilemma in concert could advance our knowledge concerning the relative merits of broad vs. narrow traits in selection and concerning the conceptualization of job performance criteria. Specifically, insights into the relative merits of broad versus narrow traits in selection could be gained by examining the stability of validities of broad vs. narrow traits across time. Furthermore, insights concerning predictors of intraindividual performance trajectories could be gained by examining broad and narrow traits as predictors of such trajectories.

My analyses first considered issues of dynamic criteria and bandwidth-fidelity separately. My analyses pertaining to the dynamicity of job performance across time began with an assessment of rank-order changes in job performance scores across time. The results of these analyses indicated that such changes did not follow the simplex pattern. Note that, while a simplex pattern is often discussed as being ubiquitous in the literature (Austin et al., 1989), some studies have failed to detect such a pattern of correlations (e.g., Hofmann et al., 1993).

It is also important to note that, across a four year period, corrected validity coefficients remained large in magnitude. Viswesvaran et al. (2005) found evidence suggesting that a general factor of job performance exists that accounts for the shared variance among distinct dimensions of job performance. A positive manifold among job performance scores across time may also reflect a general factor of job performance that remains stable across time, accounting for the positive manifold across measurement occasions.

My second set of dynamic criteria analyses pertained to intraindividual trajectories in job performance across time. The results of these analyses were consistent with the existing literature (cf. Hofmann et al., 1992, 1993; Minbashian et al., 2013; Ployhart & Hakel, 1998; Thoresen et al., 2004). Specifically, the overall job performance trajectory was characterized by a positive linear trend and a negative quadratic trend, whereby performance improved early on, and then began to level off. Furthermore, and also consistent with existing research, my analyses indicated significant between-person variability in growth parameters (excluding the quadratic trend).

In terms of my analyses pertaining exclusively to the bandwidth-fidelity dilemma (i.e., my cross-sectional analyses concerning the validity of personality factors and facets), my findings were also largely consistent with the existing literature. In terms of conscientiousness and its facets, my results indicated that global conscientiousness was a valid predictor of job performance. Furthermore, my results also indicated that the validity of the achievement facet of conscientiousness was somewhat larger than that of global conscientiousness. This was consistent with the results of Hough's (1992) meta-analysis. In terms of my cross-sectional analyses pertaining to emotional stability and its facets, I observed no evidence suggesting that these traits predicted job performance. This finding was counter to meta-analytic evidence suggesting that emotional stability is a valid predictor of job performance (Salgado, 1997). Finally, my results concerning the validity of extraversion and its facets were also consistent with Hough's (1992) meta-analysis. Specifically, while extraversion was a significant predictor of job performance, the validity of the assertiveness facet was somewhat larger.

While my cross-sectional results mirrored findings from the existing literature, my longitudinal analyses provided new insights. The existing literature into the relative merits of personality factors vs. facets for predicting job performance has exclusively addressed the issue by examining the relative magnitude of validities of global personality factors versus their facets for predicting job performance at a single point in time (e.g., Dudley et al., 2006; Judge et al., 2013; Hogan & Holland, 2003; Hough, 1992; Mount & Barrick, 1995; Saglado et al., 2013). Although this literature has been influential and has lead to many important insights, some research and theoretical perspectives suggest that the validity of personality facets may be more likely to change across time than the validity of personality factors (Murphy, 1989; Stewart, 1999). Thus, while personality facets may have higher validity than their factors for predicting performance at a single point in time, it is possible that the somewhat lower validity of personality factors might be counterbalanced by greater stability in validity coefficients across time. To the best of my knowledge, my longitudinal analyses were the first to empirically address this issue.

Overall, I observed evidence suggesting that the validities of personality factors were stable across time. Furthermore, in some cases, the validities of the global factors were more stable across time than those of their facets. Conscientiousness predicted performance across the four measurement occasions and I observed no evidence suggesting that the validities fluctuated significantly across time. On the other hand, the achievement and order facets predicted performance during only certain measurement occasions and the validities of achievement fluctuated significantly across time. Thus, the somewhat higher validity associated with achievement observed in the present study

was offset by lower temporal stability across time. These results supported existing cross-sectional research suggesting that conscientiousness facet validities may be more likely to change across job stages as compared to conscientiousness factor validities (Stewart, 1999). However, my dissertation represented the first empirical test of the prediction.

In terms of emotional stability, I did not observe any evidence suggesting that the global personality factor or its facets predicted performance at any point in time. Furthermore, I did not observe any evidence suggesting that their validities changed significantly across time. Thus, the results ultimately indicated that global emotional stability and its facets were poor predictors of job performance and that they remained poor predictors of performance across time.

The results concerning extraversion and its facets also yielded novel and interesting conclusions. Specifically, the results failed to detect any evidence suggesting that the validities of extraversion or its facets changed significantly across time. Furthermore, at each measurement occasion, the validity of the assertiveness facet of extraversion was higher than that of global extraversion. Thus, while my results pertaining to conscientiousness indicated that the factor was more stable than the facet, this was not the case for extraversion, as the factor and facets were equally as stable across time. Considering that the factor and facets did not differ in their dynamicity across time, and the assertiveness facet had higher validity than the global factor at each measurement occasion, my findings supported the notion expressed by Hough (1992) -- that extraversion is too broad for personnel selection purposes and that validity gains could be achieved by examining its facets rather than the broad factor. My results

provided further support for this notion as I found that the assertiveness facet of extraversion is not only higher than that of extraversion at one point in time, but that it remained higher for four years.

It is interesting to note that, despite the evidence suggesting that job performance changed across time, the criterion-related validity of many personality traits assessed were stable. The finding highlights that performance change and validity change are two different phenomena and while performance change may allow for validity change, it does not necessarily mean that validity change would occur. The finding is consistent with the sentiment reflected in Ackerman (1989): "The only clear inference from the simplex-like pattern of intercorrelations in skilled performance is that individuals continuously change their rank order on performance over practice. As such, the underlying determinants of performance cannot be fixed over practice. However, this fact does not require that validity coefficients drop with time-on-task" (p. 361).

The results of my longitudinal criterion-related validity analyses bear implications for selection practices. Selection research has noted the importance of designing selection systems in such a fashion that utility is maximized across time (Hulin et al., 1990). Therefore, in determining the relative merits of global personality traits versus their facets for predicting job performance, it is necessary to not only examine the relative magnitude of their validities, but the relative stability of their validities as well.

Some cross-sectional research has indicated that the achievement facet of conscientiousness may be a better predictor of job performance than global conscientiousness (Hough, 1992). However, this conclusion must be tempered by the findings observed in my dissertation indicating that, while achievement was a superior

predictor of job performance than global conscientiousness at certain points in time, at other points in time its validity was much lower than that of global conscientiousness. My findings indicated that global conscientiousness seems to reliably predict performance across time and thus incorporating global conscientiousness into predictor batteries would result in tangible performance improvement across time. On the other hand, when incorporating the achievement facet into predictor batteries, the validity of the selection system is likely to vary across employee tenure.

In terms of extraversion, research has indicated that, while global extraversion might not be a valid predictor of job performance, its assertiveness facet might be (Hough, 1992). However, no research has provided any guidance concerning the stability of the validity of the assertiveness facet. My findings indicated that assertiveness was not only a superior predictor of job performance as compared to extraversion at one point in time, but that it remained a superior predictor of job performance across time. Thus, my findings bolster support for the notion that the assertiveness facet of extraversion is a better predictor of job performance than the global extraversion factor. My findings suggest that organizations could benefit from incorporating the assertiveness facet of extraversion into predictor batteries.

In discussing the implications of dynamic criterion-related validity coefficients, an implicit assumption is that stability is good. However, that may not always be the case. For example, in the present study, I found that the validity of emotional stability for predicting job performance was low and remained low across the four measurement occasions. In this case, the long-term utility of the selection system would have been improved if the validity of emotional stability were dynamic (i.e., if the validity of

emotional stability improved during at least one of the four measurement occasions). Therefore, stable validity coefficients might not always be preferred in operational contexts. However, validity coefficients that are practically meaningful in magnitude and remain stable across time would result in greater utility over the long term as compared to validity coefficients that are practically meaningful in magnitude at only certain points in time.

In terms of predictors of performance growth, none of my hypotheses were supported as my results failed to detect a significant effect of any personality variable on performance growth. Global conscientiousness and the achievement facet were both positively related to the performance intercept, suggesting that these traits were positively related to mean performance at the first measurement occasion. However, these traits did not predict performance growth. Furthermore, the only other trait to predict the performance intercept was the assertiveness facet of extraversion. No broad or narrow traits examined in the present study predicted performance trends. The null findings concerning personality predictors of performance growth are somewhat counter to the existing literature that has indicated significant relationships between some of the Big Five personality factors and performance trends (Thoresen et al., 2004). However, much of this work was conducted using objective performance measures over shorter periods of time. For example, Thoresen et al. found that global conscientiousness predicted growth of objective sales criteria across four quarters. In a subsequent study, Minbashian et al. (2013) examined performance trends using subjective ratings (i.e., annual performance evaluations) across four years. The results of the Minbashian et al. study were somewhat counter to the results of Thoresen et al., as the research failed to detect a significant effect

of conscientiousness on performance trends. Research does indicate that rank-order

consistency of job performance across time is higher when ratings of job performance are

employed as compared to objective performance measures (Sturman et al., 2005).

Therefore, there may have been less variance in performance change in the present study

as compared to those using objective measures. This could have potentially accounted

for my null findings. Another consideration is that annual performance evaluations may

be too long of a time period to assess performance trajectories. In order to fully

understand the manner in which performance unfolds across time and predictors of such

change patterns, a more fine-grained approach may be necessary whereby performance is

measured at more closely spaced intervals.

It is also important to note that my sample was slightly unique. Specifically, my

sample was comprised of employees who worked for the organization for a full four-year

period. The attraction-selection-attrition framework would suggest that this sample of

employees should be more similar in personality to one another as compared to the

population at large (Schneider, 1987). Therefore, as is the case in any concurrent validity

study, there was indirect range restriction on the predictor. Also, as my sample included

only employees who remained with the organization for such a long period, there was

also likely range restriction on the criteria. Specifically, poor performing employees may

have been terminated during such a prolonged period (Sturman & Trevor, 2001). It is

also possible that superior performing employees were promoted or otherwise left for

other positions. Thus, the variation in criterion scores for my final sample was likely

restricted. The range restriction observed here might have contributed to my null

findings with respect to my latent growth modeling analyses as well as some of the null

findings (and small effect sizes) observed earlier.

Limitations and Future Directions

While my dissertation is the first study into the relative stability of validities of

global personality factors versus their facets across time, and I found evidence suggesting

that the validities of global conscientiousness were more stable across time than those of

its facets, I did not test any theories concerning how or why this effect occurred. For

instance, I have no data to suggest what, if any, changes in job demands had occurred

over this period or any changes that took place in the jobs under consideration that would

signify different job stages (Murphy, 1989). Thus, future research is needed to advance

theories that can account for dynamic validities of global conscientiousness and its facets.

However, I note that some research along these lines has taken place (Helmreich et al.,

1986; Stewart, 1999; Thoresen et al., 2004), while my study was the first to empirically

examine the issue of differential dynamicity between global conscientiousness and its

facets.

Nonetheless, future research can build on my findings by, for example, examining

the stability of validities for predicting job performance using a cohort sample of

employees from their first criterion measurement occasion and onward. A cohort sample

may appear desirable, as such a sample of employees may be likely to experience a

transition stage early on and settle into a maintenance stage thereafter (Murphy, 1989),

and this would allow research to assess changes to conscientiousness factor and facet

validities across time and across job stages. However, when considering the changing

nature of work in many jobs, regardless of an employee's tenure, there is often a need to

develop new skills and face new demands (Howard, 1995; Patterson, 2001). Thus, when employees will be in different job stages may not be easily predictable, especially when considering that job stages occur at the individual-level rather than group-level (Murphy, 1989). As a result, identifying exactly what facets of conscientiousness predict performance during what specific job stage may be of limited utility, as the distinction between different job stages becomes less clear as the work place becomes less static and less predictable. Thus, while testing theories that account for dynamic validities is important, empirical research into the relative stability of validities of broad and narrow traits is equally important, has immediate relevance for personnel selection practices, and has been the focus of less research than have theories of dynamic validities. Nevertheless, future research should explore the possibility personality facets can be a better predictor of job performance than personality factors at specific points in an employee's career.

While the above discussion noted that usefulness of using a cohort sample of employees for assessing validity change across time, such a sample is also useful for assessing intraindividual performance trajectories. In the present dissertation, my sample was not a cohort of employees and employees differed in their tenure at the first measurement occasion. As I did not examine a cohort sample, I was assessing trajectories in performance that occurred at different points in each employee's tenure. While my results did conform to existing research in that I identified a negative quadratic performance trend, I was unable to account for any variance in the trend parameters using the personality predictors. It is possible that if I assessed a single cohort of employees across times, the findings would have been different. I was able to test this assertion by

limited my sample to only those employees for whom time 1 represented their first performance measurement occasion with the organization (i.e., the first measurement occasion represented a rating of their performance for their first year on the job). However, when rerunning my latent growth modeling analyses using this subset of employees, I still failed to detect a significant effect of any personality trait on performance growth.

Another limitation of the present dissertation was the exclusive focus on conscientiousness, emotional stability, and extraversion. I chose to examine conscientiousness and emotional stability due to the meta-analytic evidence suggesting that these two traits (but specifically conscientiousness) are valid predictors of job performance across occupations and criteria (e.g., Salgado, 1997). Furthermore, I included extraversion because of its significance within the context of the bandwidth-fidelity dilemma. Specifically, research has long since argued that extraversion was too broad for selection purposes and that validity gains could be realized by utilizing specific facets as opposed to the broad factor (e.g., Hough, 1992). However, more recently, some research has argued that openness might also be too broad for selection purposes and that certain facets may be valid predictors of job performance. Research suggests that openness can be split into two facets -- aesthetic openness and intellect (DeYoung et al., 2007, Judge et al., 2013). Some research suggests that, while global openness is not a valid predictor of job performance, at the facet level, intellect is a valid predictor of job performance while aesthetic openness is not (Griffin & Hesketh, 2004). More research into this issue is needed, but especially longitudinal research. Indeed, if the validity of the intellect facet is in fact higher than that of global openness, another question emerges

127

-- is the validity of the intellect equally as stable across time as that of openness as well? For example, if the openness factor is more stable across time than that of the intellect facet, it is possible that the lower validity at a single point in time is counterbalanced by greater stability across time. Research is needed to address this issue.

An interesting finding of the present dissertation was that emotional stability and its facets did not predict performance at any measurement occasion. Similarly, in my ancillary analyses, I failed to detect any evidence suggesting that emotional stability or its facets predicted performance on any dimension. This finding was somewhat contradictory to meta-analytic evidence. For instance, in a meta-analytic review, Salgado (1997) found evidence suggesting that emotional stability was a valid predictor of job performance across occupations and criteria. These findings could potentially be due to the model of emotional stability employed in the present dissertation. Specifically, I employed the two-facet model reviewed in Perugini and Callucci (1997) that included serenity and firmness facets.

More recently, DeYoung et al. (2007) proposed a two-facet model of emotional stability that included the facets of low withdrawal and low volatility. In terms of the later model, both of Perugini and Callucci's (1997) emotional stability facets would be encompassed by low withdrawal. Therefore, the model employed in the present study may be somewhat deficient, as I did not assess emotional stability facets pertaining to low volatility. In a recent meta-analytic review, Judge et al. (2013) found evidence suggesting that the low volatility facet as opposed to the low withdrawal facet largely drove the validity of emotional stability. Specifically, the validity of low volatility for predicting overall job performance was $\rho = .12$, while the validity of low withdrawal was

$\rho = .05$. Therefore, the null results concerning the validity of emotional stability or its facets for predicting job performance could have been due to my choice of a facet model of emotional stability that did not include low volatility facets.

As reviewed earlier in my literature review, research has debated the relative merits of broad vs. narrow traits for personnel selection purposes. However, some research has sought a compromise between these competing perspectives by proposing that trait-criteria bandwidth matching is a key concern for the validity of personality traits. Specifically, research along these lines has suggested that narrow personality traits should best predict conceptually related narrow criterion measures, while broad personality factors should best predict broad performance criteria (Hogan & Holland, 2003). In the present study, because of the broad bandwidth of the performance rating items, I was unable to assess the validity of personality factors and facets for predicting narrow performance criteria across time. It is possible that when narrow traits are conceptually aligned to narrow criteria measures, the validity of personality facets may be stable across time. It would be useful for future research to address this issue. However, it is important to note that most real world criteria are broad and therefore, assessing the validity of narrow traits for predicting narrow performance criteria may be of limited utility (Hogan & Hogan, 1994; Ones & Viswesvaran, 1996). For instance, research indicates a positive manifold across job performance dimensions that is accounted for by a higher-order general factor (Viswesvaran et al., 2005). Therefore, if the bandwidth matching perspective is correct, this would suggest that broad traits should be preferred over narrow, as most criteria measures are in fact broad.

A final limitation of the present dissertation concerns an aspect of the data collection. Specifically, personality scores were collected after all of the criterion scores were collected. The timing of the personality data collection is problematic in that, as a four-year longitudinal study, it is unclear as to whether or not personality scores collected after time 4 were the same as personality scores would have been if collected prior to time 1. That is, if personality traits changed across time, the results of the present dissertation may not generalize to practice where personality scores are collected prior to job performance scores. However, research has indicated high rank-order consistency in personality across one's adult years. For example, in a cumulative meta-analytic review, Roberts and DelVecchio (2000) identified estimates of trait consistency between the ages of 22-29 of $\rho = .57$, 30-39 of $\rho = .62$, 40-49 of $\rho = .59$, 50-59 of $\rho = .75$, and 60-63 of $\rho = .72$. In a diverse sample of middle-aged employees, Sutin and Costa (2010) found evidence of relative rank-order stability in personality scores across 10 years. Coefficients of stability were large in magnitude, ranging from $r = .66$ to $r = .76$. Considering the evidence just reviewed, I believe that it is likely that the findings of the present study would generalize to a pre-employment selection context where personality scores were measured before criterion scores were available as opposed to afterwards.

Conclusions

In my dissertation, I considered issues pertaining to the dynamicity of job performance along with issues pertaining to the bandwidth-fidelity dilemma in personality testing for personnel selection purposes. Both of these issues have been critically important for conceptualizations of criteria and for personnel selection research and practice and, by examining these issues in concert, my results provided new insights.

In terms of conscientiousness, my results indicated that, while the achievement facet might have higher validity than the global factor at certain points in time, the validity of the global factor is more stable. Thus, research that has suggested that incorporating the achievement facet of conscientiousness into selection batteries as opposed to the global factor would yield improved validity (e.g., Hough, 1992) must be tempered by the conclusions reached in my study.

On the other hand, my results pertaining to extraversion were consistent with the notion that this personality factor is too broad for selection purposes. My results indicated that the validity of the assertiveness facet of extraversion was higher than that of global extraversion across all four measurement occasions and I observed no evidence suggesting that its validity fluctuated across time. Overall, my findings supported the use of global conscientiousness and of the assertiveness facet of extraversion for selection purposes. The study reported in my dissertation is the first to apply a longitudinal perspective to the bandwidth-fidelity dilemma in personality testing for personnel selection purposes. Given the implications of such a perspective for selection research and practice, more research along these lines is encouraged.

LIST OF REFERENCES

Ackerman, P. L. (1989). Within-task intercorrelations of skilled performance: Implications for predicting individual differences? (a commentary on Henry & Hulin, 1987). *Journal of Applied Psychology, 74*(2), 360-364.

Adams, J. A. (1957). The relationship between certain measures of ability and the acquisition of a psychomotor response. *Journal of General Psychology, 56,* 121-134.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs, 47,* i-171

Alvarez, K. M., & Hulin, C. L. (1972). Two explanations of temporal changes in ability-skill relationships: A literature review and theoretical analysis. *Human Factors, 14,* 295-308.

Alvares, K. M., & Hulin, C. L. (1973). An experimental evaluation of a temporal decay in the prediction of job performance. *Organizational Behavior and Human Performance, 9,* 169-185.

Austin, J. T., Humphreys, L. G., & Hulin, C. L. (1989). Another view of dynamic criteria: A critical reanalysis of Barrett, Caldwell, and Alexander. *Personnel Psychology, 42,* 583-596.

Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917-1992. *Journal of Applied Psychology, 77,* 836-874.

Barrett, G. V., Caldwell, M. S., & Alexander, R. A. (1985). The concept of dynamic criteria: A critical reanalysis. *Personnel Psychology, 38,* 41-56.

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., Stewart, G. L., & Piotrowski, M. (2002). Personality and job performance: Test of the mediating effects of motivation among sales representatives. *Journal of Applied Psychology, 87*(1), 43-51.

Bass, B. M. (1962). Further evidence on the dynamic character of criteria. *Personnel Psychology, 15,* 93-97.

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology, 92*(3), 595-615.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt and W. C. Borman (eds.), *Personnel Selection in Organizations* (pp. 71-98). San Francisco, CA: Jossey Bass.

Boswell, W. R., Boudreau, J. W., & Tichy, J. (2005). The relationship between employee job change and job satisfaction: The honeymoon-hangover effect. *Journal of Applied Psychology, 90,* 882-892.

Brenner, M. H., & Lockwood, H. C. (1965). Salary as a predictor of salary: A 20-year study. *Journal of Applied Psychology, 49*(4), 295-298.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette, 7 L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687-782). Palo Alto, CA: Consulting Psychologists Press.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38,* 476-506.

Cattell, R. B. (1945). The principal trait clusters for describing personality. *Psychological Bulletin, 42*, 129-161.

Cattell, R. B. (1946). *Description and Measurement of Personality*. New York, NY: World Book.

Cattell, R. B. (1957). *Personality and Motivation: Structure and Measurement.* New York, NY: World Book.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF).* Champaign, IL: Institute for Personality and Ability Testing.

Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90,* 218-244.

Costa, P. T. Jr., & McCrae, R. R. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences, 12*(9), 887-898.

Costa, P. T. Jr. & McCrae, R. R. (1992). *Revised NEO personality inventory and NEO five-factor inventory: Professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.

Costa P. T. Jr. & McCrae, R. R. (1995). Primary traits of Eysenck's P-E-N-system: Three- and five-factor solutions. *Journal of Personality and Social Psychology, 69,* 308-317.

Costa, P. T. Jr., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: Examples from conscientiousness. *European Journal of Personality, 12,* 117-134.

Costa, P. T. Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences, 12*(9), 887-898.

Day, D. V., Sin, H., & Chen, T. T. (2004). Assessing the burdens of leadership: Effects of formal leadership roles on individual performance over time. *Personnel Psychology, 57,* 573-605.

Deadrick, D. L., Bennett, N., & Russell, C. J. (1997). Using hierarchical linear modeling to examine dynamic performance criteria over time. *Journal of Management, 23*(6), 745-757.

Deadrick, D. L., & Madigan, R. M. (1990). Dynamic criteria revisited: A longitudinal study of performance stability and predictive validity. *Personnel Psychology, 43,* 717-744.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93,* 880-896.

Digman, J. M. (1972). The structure of child personality as seen in behavior ratings. In R. M. Dreger (Ed.), *Multivariate Personality Research* (pp. 587-611). Baton Rouge, LA: Claitor's.

Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology, 41,* 417-440.

Digman J. M., & Takemoto-Chock, M. K. (1981). Factors in the natural language of personality: Re-analysis and comparison of six major studies. *Multivariate Behavioral Research, 16,* 149-170.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*(1), 40-57.

Dunn, O. J., & Clark, V. A. (1969). Correlation coefficients measured on the same individuals. *Journal of American Statistical Association, 64*, 366-377.

Eysenck, H. J. (1991). Dimensions of personality: 16, 5, or 3? – Criteria for a taxonomic paradigm. *Personality and Individual Differences, 12,* 773-790.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology, 44,* 329-344.

Fleishman, E. A., & Hempbel, W. E. Jr. (1954). Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika, 19,* 239-252.

Galton, F. (1884). Measurement of character. *Fortnightly Review, 36,* 179-185.

Ghiselli, E. E. (1956). Dimensional problems of criteria. *Journal of Applied Psychology, 1,* 46-60.

Ghiselli, E. E., & Brown, C. W. (1947). Learning in accident reduction. *Journal of Applied Psychology, 31, 580-582.*

Ghiselli, E. E., & Haire, M. (1960). The validation of selection tests in the light of the dynamic character of criteria. *Personnel Psychology, 13*, 225–231.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of Personality and Social Psychology* (Vol. 2, pp. 141-165). Beverly Hills, CA: Sage.

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

Gorsuch, R. L. (1974). *Factor analysis.* Philadelphia, PA: Saunders.

Griffin, B., & Hesketh, B. (2004). Why openness to experience is not a good predictor of job performance. *International Journal of Selection and Assessment, 12*(3), 243-251.

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18,* 135-164.

Hanges, P. J., Schneider, B., & Niles, K. (1990). Stability of performance: An interactionist perspective. *Journal of Applied Psychology, 75,* 658-667.

Helmreich, R. L., Sawin, L. L., & Carsrud, A. L. (1986). The honeymoon effect in job performance: Temporal increases in the predictive power of achievement motivation. *Journal of Applied Psychology, 71,* 185-188.

Hoffman, C. C., Nathan, B. R., & Holden, L. M. (1991). A comparison of validation criteria: Objective versus subjective performance measures and self-versus supervisor ratings. *Personnel Psychology, 44,* 601-619.

Hofmann D. A., Jacobs R., & Baratta J. E. (1993) Dynamic criteria and the measurement of change. *Journal of Applied Psychology*, 78(2), 194-204.

Hofmann, D. A., Jacobs, R., & Gerras, S. J. (1992). Mapping individual performance over time. *Journal of Applied Psychology, 77,* 185-195.

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

Hogan, J., & Hogan, R. (1994, April). Fidelity and bandwidth in assessment. In *Personality and Job Performance: Big Five versus Specific Traits*, Symposium conducted at the meeting of the Society for Industrial and Organizational Psychology, Nashville, TN.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*(1), 100-112.

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

Hough, L. M. (1992). The "Big Five" personality variables-construct confusion: Description versus prediction. *Human Performance, 5*(1 & 2), 139-155.

Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. R. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *International Handbook of Work and Organizational Psychology* (pp. 233-377). Thousand Oaks, CA: Sage.

Howard, A. (1995). *The changing nature of work.* San Francisco, CA: Jossey-Bass.

Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin, 107,* 328-340.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology, 85*(6), 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, NY: Guilford Press.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2,* 171-203.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (pp. 114-158). New York, NY: Guilford Press.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*(4), 765-780.

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology, 87*(4), 797-807.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representation of the Five-Factor Model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology, 98*(6), 875-925.

Katerberg, R., & Blau, G. J. (1983). An examination of level and direction of effort and job performance. *Academy of Management Journal, 26*(2), 249-257.

Kaufman, H. G. (1972). Relations of ability and interest to currency of professional knowledge among engineers. *Journal of Applied Psychology, 56*(6), 495-499.

Keil, C. T., & Cortina, J. M. (2001). Degradation of validity over time: A test and extension of Ackerman's model. *Psychological Bulletin, 127,* 673-697.

LeBreton, J. M., Hargis, M. B., Griepentrog, B., Oswald, F. L., & Ployhart, R. E. (2007). A multidimensional approach for evaluating variables in organizational research and practice. *Personnel Psychology, 60,* 475-498.

LeBreton, J. M., Ployhart, R. E., & Ladd, R. T. (2004). A Monte Carlo comparison of relative importance methodologies. *Organizational Research Methods, 7,* 258-282.

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19,* 451-458.

Minbashian, A., Bright, J. E. H., & Bird, K. D. (2009). Complexity in the relationships among the subdimensions of extraversion and job performance in managerial occupations. *Journal of Occupational and Organizational Psychology, 82,* 537-549.

Minbashian, A., Earl, J., & Bright, J. E. H. (2013). Openness to experience as a predictor of job performance trajectories. *Applied Psychology: An International Review, 62*(1): 1-12.

Mitchel, J. O. (1975). Assessment center validity: A longitudinal study. *Journal of Applied Psychology, 60*(5), 573-579.

Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance, 10,* 71-83.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resource Management, 13*(3), 153-200.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-Factor Model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*(2/3), 145-165.

Murphy, K. R. (1989). Is the relationship between cognitive ability and job performance stable over time? *Human Performance, 2,* 183-200.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

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.

Norman, W. T. (1967). *2,800 personality trait descriptors: Normative operating characteristics from a university population.* Department of Psychology, University of Michigan.

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609-626.

Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Personality at work: Raising awareness and correcting misconceptions. *Human Performance, 18*(4), 389-404.

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

Patterson, F. (2001). Developments in work psychology: Emerging issues and future trends. *Journal of Occupational and Organizational Psychology, 74,* 381-390.

Perugini, M., & Gallucci, M. (1997). A hierarchical faceted model of the Big Five. *European Journal of Personality, 11,* 279-301.

Ployhart, R. E., & Hakel, M. D. (1998). The substantive nature of performance variability: Predicting interindividual differences in intraindividuial performance. *Personnel Psychology,* 51, 859-901.

Reeve, C. L., & Bonaccio, S. (2011). On the myth and reality of the temporal validity degradation of general mental ability across time. *Intelligence, 39,* 255-272.

Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. (1996). A longitudinal study of upward feedback. *Personnel Psychology, 49,* 599-612.

Roberts, B. W., Bogg, T., Walton, K., Chernyshenko, O., & Stark, S. (2004). A lexical approach to identifying the lower-order structure of conscientiousness. *Journal of Research in Personality, 38,* 164-178.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*(1), 3-25.

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (in press). What is conscientiousness and how can it be assessed? *Developmental Psychology.*

Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal, 38,* 555-572.

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

Salgado, J. F., Moscoso, S., & Berges, A. (2013). Conscientiousness, its facets, and the prediction of job performance ratings: Evidence against narrow measures. *International Journal of Selection and Assessment, 21*(1), 74-84.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology, 76*(4), 613-627.

Schmidt, F. L., & Hunter, J. E. (1984). A within setting empirical test of the situational specificity hypothesis in personnel selection. *Personnel Psychology, 37*(2), 317-326.

Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71*, 432-439.

Schmitt, N., Gooding, R. Z., Noe, R. D., & Kirsch, M. (1984). Meta-analysis 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.

Stark, S., Chernyshenko, O. S., Drasgow, F., & White, L. A. (2012). Adaptive testing with multidimensional pairwise preference items: Improving the efficiency of personality and other noncognitive assessments. *Organizational Research Methods, 15,* 463-487.

Stewart, G. L. (1999). Trait bandwidth and stages of job performance: Assessing differential effects for conscientiousness and its substraits. *Journal of Applied Psychology, 84*, 959-968.

Sturman, M. C. (2007). The past, present, and future of dynamic performance research. *Research in Personnel and Human Resources Management, 26*, 49-110.

Sturman, M. C., Cheramie, R. A., & Cashen, L. H. (2005). The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *Journal of Applied Psychology, 90,* 269-283.

Sturman, M. C., & Trevor, C. O. (2001). The implications of linking the dynamic performance and turnover literatures. *Journal of Applied Psychology, 80*(4), 684-696.

Sutin, A. R., & Costa, P. T. Jr. (2010). Reciprocal influence of personality and job characteristics across middle adulthood. *Journal of Personality, 78*(1), 257-288.

Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry, 4,* 122-130.

Tett, R. P., Guterman, H. A., Bleir, A., & Murphy, P. J. (2000). Development and content validation of a "hyperdimensional" taxonomy of managerial competence. *Human Performance, 13*(3), 205-251.

Tett,R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality-job performance relationship. *Journal of Organizational Behavior, 24,* 335-356.

Thoresen, C. J., Bradley, J. C., Bliese, P. B., & Thoresen, J. D. (2004). The Big Five personality traits and individual job performance growth trajectories in maintenance and transitional job stages. *Journal of Applied Psychology, 89*, 835-853.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4,* 25-29.

Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology, 26,* 1-9.

Tonidandel, S., & LeBreton, J. M. (in press). RAW-Web – A free, comprehensive, web-based, user-friendly tool for relative weight analysis. *Journal of Business and Psychology.*

Tonidandel, S., LeBreton, J. M., & Johnson, J. W. (2009). Determining the statistical significance of relative weights. *Psychological Methods, 14,* 387-399.

Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings.* Technical Report, USAF, Lackland Air Force Base, TX.

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on traits. *Journal of Personality, 60,* 225-251.

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

Viswesvaran, C. (1993). *Is there a general factor in job performance?* Unpublished doctoral dissertation, University of Iowa.

Viswesvaran, C., & Ones, D.S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment, 8*, 216-226.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*, 557-574.

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology, 90*(1), 108-131.

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association, 58*, 236-244.

Webb, E. (1915). Character and intelligence. *British Journal of Psychology Monographs, 1*(3), 1-99.

VITA

MICHAEL BENJAMIN HARARI

Born, Miami, Florida

M.S.   Florida International University
             Miami, FL

B.S.   University of North Florida
             Jacksonville, FL

PUBLICATIONS AND PRESENTATIONS

Harari, M. B., Naemi, B., Viswesvaran, C., Roberts, R. D., & Rodriguez, J. (In Press).
    The validity of conscientiousness and its facets: Stable or dynamic across time?
    Academy of Management Proceedings.

Rudolph, C. W., Michel, J. S., Harari, M. B., & Stout, T. (2014). Perceived social
    support and work-family conflict: A comparison of Hispanic immigrants and non-
    immigrants. *Cross Cultural Management: An International Journal, 21*(3), 306-
    325.

Harari, M. B., Viswesvaran, C., & O'Rourke, R. (2014). Gender differences in work
    sample assessments: Not all tests are created equal. *Journal of Work and
    Organizational Psychology, 30*(1), 29-34.

Harari, M. B., Laginess, A. J., & Viswesvaran, C. (2014, May). Examining
    dynamic validities across maintenance and transition job stages. Poster session
    presented at the annual meeting of the Society for Industrial Organizational
    Psychology, Honolulu, HI.

Harari, M. B., Laginess, A. J., & Rudolph, C. W. (2014, May). A meta-analysis of
    rater individual differences-rating level relationships. Poster session
    presented at the annual meeting of the Society for Industrial Organizational
    Psychology, Honolulu, HI.

Harari M. B., Rudolph, C. W., Viswesvaran, C., & Woodward, M. M. (2013, August).
    Structured interview validities across time: An analysis of stable and dynamic
    relationships. Paper presented at the annual meeting of the Academy of
    Management, Orlando, FL.

Harari M. B., Allen, J. W., & Rudolph, C. W. (2013, August). Clarifying the role of observer attributions in ratings of dynamic performance. Paper presented at the annual meeting of the Academy of Management, Orlando, FL.

Harari, M. B., & Rudolph, C. W. (2013, April). Dynamic criteria and within-person performance variability: Where are we now? Panel discussion held at the annual meeting of the Society for Industrial Organizational Psychology, Houston, TX.

Rutherford, K., Harari, M. B., & Rudolph, C. W. (2013, April). A multidimensional measure of observer attributions for performance. Poster session presented at the annual meeting of the Society for Industrial Organizational Psychology, Houston, TX.

Allen, J., Wittgenstein J., Harari, M. B., & Bruk-Lee, V. (2012, November). A qualitative review of workplace aggression: Current trends and future directions. In P. Spector (Chair), *Employee responses to workplace abuse.* Symposium conducted at the Southern Management Association Meeting, Ft. Lauderdale, FL.

Wittgenstein, J., Harari, M. B., Allen, J., & Rudolph, C. W. (2012, November). When does performance variability matter? The effect of dynamic performance characteristics on rater attributions and performance ratings. Paper presented at the annual Southern Management Association Conference, Ft. Lauderdale, FL.

Rudolph, C. W., Harari, M. B., & Wittgenstein, J. (2012, August). The influence of positive and negative performance trends on performance ratings of older and younger sales people. In T. McCausland (Chair), *New views of age stereotypes: Form, prevalence, and consequences*. Symposium conducted at the annual meeting of the Academy of Management, Boston, Massachusetts.

Wittgenstein, J., Harari, M. B., & Allen J. (2012, March). Organizational justice's role as a stressor in the workplace. Poster session presented at the annual Industrial-Organizational/Organizational Behavior Conference, Orlando, FL.

Harari, M. B., Wittgenstein, J., & Allen, J. (2012, February). Distributive justice as an organizational stressor. Poster session presented at the annual Sunshine Educational Research Center, Tampa, FL.

Leone, C., Harari, M. B., Smith, R. (2011, January). Self-monitoring and self-selection to employment opportunities. Poster session presented at the annual meeting of the Society for Personality and Social Psychology, San Antonio, TX.

Davis, C., Harari, M. B., Hawkins, L., Leone, C. (2009, November). Behavioral identification and responsibility attributions: Interpreting and explaining the physical abuse of children. Poster session presented at the annual Meeting of the Society for Southeastern Social Psychologists. Fort Myers, FL.