UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

INVESTIGATING THE LINEAR PREDICTOR – CRITERION ASSUMPTION OF BIODATA SCALING

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

GREGORY GLENN MANLEY
Norman, Oklahoma
2003

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/684

UMI Number: 3112348

Copyright 2003 by
Manley, Gregory Glenn

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.



UMI Microform 3112348
Copyright 2004 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

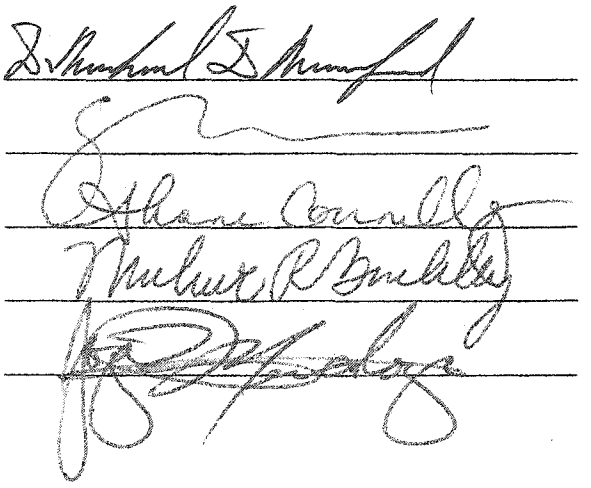
© copyright by GREGORY GLENN MANLEY 2003

All Rights Reserved.

INVESTIGATING THE LINEAR PREDICTOR – CRITERION ASSUMPTION OF BIODATA SCALING

A dissertation APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

BY



ACKNOWLEDGMENTS

I would like to gratefully acknowledge the guidance and leadership of my academic mentor, Michael Mumford, and thank him for his many hours of patience and help in my graduate studies. I would also like to recognize and give thanks to two people instrumental in my decision to undergo a Ph.D., my brother, Jeffrey Manley and my undergraduate advisor, Chuck Pierce. Without their moral and technical support, I may have never considered such an effort. I would additionally like to thank my now deceased grandparents, Fred and Virginia Browning, whom supported my efforts throughout the years and greatly wanted to witness my achievements. Last but most certainly not least, I would like to offer many thanks to two people in my life most deserving of my gratitude and most responsible for helping me obtain my degree, my parents, Joe and Carolyn Manley. They have helped me morally, spiritually, and financially through out my life, especially through my graduate and undergraduate studies.

TABLE OF CONTENTS

Acknowledgements		iv
List of Tables		vi
List of Figures		vii
List of Appendixes		viii
Abstract		ix
Chapter 1	Introduction and Literature Review	1
Chapter 2	Methodology	20
Chapter 3	Results	38
Chapter 4	Discussion	45
References		56
Tables and Figures		62
Appendixes		76

LIST OF TABLES

Table 2.1	Sample Demographics	62
Table 2.2	Descriptive Statistics of NEO sub-scales / scales	64
Table 2.3	Descriptive Statistics for Study Variable	65
Table 2.4	Sample Item Stems for Rational Scales	66
Table 2.5	Convergent/Discriminant Validity Between FFM Variables	68
Table 2.6	Fit Statistics for Confirmatory Factor Analysis of FFM scales	68
Table 3.1	Correlation Matrix of Main Study Variables	72
Table 3.2	Criterion-Related Validities for Scaling Methods	73
Table 3.3	Results of Hierarchical Quadratic Multiple Regressions (QMRs)	74
Table 3.4	Results of item Hierarchical QMRs for Scaling Stability	75

LIST OF FIGURES

Figure 2.1	Example of ABA Biodata Items	63
Figure 2.2	Latent Variable Model of Biodata Big-Five Factors and O*NET General Work Dimensions	70
Figure 2.3	Latent Variable Model of Biodata Big-Five Factors and Principle Component Sub-Factors	71

LIST OF APPENDIXES

Appendix A	Applicant Background Assessment Biodata Inventory
Appendix B	List of Items in Rational Biodata Scales
Appendix C	Permission Letter for Using FAA Data
Appendix D	Internal Review Board Approval Letter
Appendix E	Author's Biography

ABSTRACT

Study examines biographical data and Academy training performance for 4,559 Federal Aviation Administration (FAA) Air Traffic Controller Specialists collected from 1986 to 1992. Biodata measure was rationally scaled along personality factors of the Big Five and Occupational Information System (O*NET) general work activities. Study examines a linear predictor – criterion assumption made with rational and empirical scales by using evolutionary theories of personality to model a priori non-linear relationships with job and training performance among FAA candidates. Correlation and stepwise hierarchal quadratic multiple regression analyses revealed most scales had a linear relationship with Academy training performance; however, biodata scales of conscientiousness and mental processes did exhibit slight nonlinearity. High and low level of these dimensions predicted higher training and job performance than mid levels. Some sub-scales of conscientiousness showed the same pattern of results. Additionally, empirical keying of the same rational scales improved criterion-related validity over the purely rational scales and suffered less shrinkage upon cross-validation than a strict empirical key of the full biodata inventory. Implications for construct validity, theory, and practice are discussed.

INVESTIGATING THE LINEAR PREDICTOR – CRITERION ASSUMPTION OF BIODATA SCALING

Chapter 1: Introduction and Literature Review

Biodata

Many of today's organizations view people as their most valuable assets (Beatty, Schneier, & McEvoy, 1987) and, understandably, they go to great lengths to select the best performers from the available applicant pool. Organizations have at their disposal a number of techniques for selecting high performers. Of which, biographical data, or simply, "biodata" exists as a frequently used and valid predictor of human work performance (Brown, 1994; Hogan, 1994; Hough & Paullin, 1994; Mael, 1991; Mitchell & Klimoski, 1982; Mumford, Costanza, Connelly, & Johnson, 1996; Mumford & Owens, 1987; Owens, 1976; Pannone, 1984; Russell, 1994; Russell, Mattson, Devlin, & Atwater, 1990).

Biodata measures consist of items tapping developmental constructs purported to predict subsequent behavior (Brown, 1994). Mael (1991) suggested the core attribute of biodata items is pertinence to historical events that may have shaped the individual's behavior and identity. Essentially, biodata looks at past behavior or experiences in different developmental contexts or referent situations. These past behaviors can be of a developmental domain that is theoretically presumed to lead to target performance (e.g., frequent engagement in complex memory games as a teenager may predict subsequent academic performance in college) or the behaviors may simply be very similar to target performance (e.g., high school academic success predicts college academic success; Stokes, Mumford, & Owens, 1994).

There are a variety of ways to scale biodata items such as empirical keying, factorial scaling, sub-grouping, and rational scaling (Brown, 1994). Several of these methods are presented below. Essentially, all biodata scaling procedures involve a) tapping constructs thought to have a developmental antecedent to criterion performance, b) establishing an item/scale relationship with criterion performance, and c) using a criterion measure that exhibits construct validity. Validation studies typically involve unique relationships between the item pool, applicant pool, and criterion measure. Each component can contribute to the quality of the resultant performance predictor measure (e.g., Russell, 1994; Mumford & Owens, 1987). Biodata measures ideally need to a) employ items/scales that tap relevant antecedent predictor constructs for the criterion of interest, b) insure validation samples are representative of the population for which the instrument is intended, and c) employ the use of construct valid criterion measures to increase validity generalization across time and reduce shrinkage when cross-validated (Mumford & Owens, 1987).

Intent of study

Empirical keying (detailed below) has been the traditional method for scaling biodata items, however has been subject to much criticism concerning its use because of sample specific predictions and lack of demonstrated construct validity (Stokes, Mumford, & Owens, 1994). Rational scaling (also detailed below) has been offered as a viable alternative because it is construct based and has demonstrated criterion-related validity similar to empirical keys (e.g., Rieter-Palmon & Connelly, 2000). Thus the debate continues over the use of empirical keying vs. rational scaling. The intent of the

present study is to examine one issue involved in this debate: potential non-linear predictor-criterion (P-C) relationships.

One advantage of empirical keying over other methods is it can capture all linear and nonlinear item-criterion relationships. Empirical keys, when keyed at the response option (RO) level, have demonstrated increased predictive power over the same items keyed at the item level, thus providing evidence of nonlinear item-criterion relationships that RO-level keys (but not item-level keys) could capture (Kluger, Reilley, & Russell, 1991). Nonetheless, a question remains whether the empirical keys demonstrate construct validity, especially in terms of a possible non-linear P-C relationship. Non-linear item-criterion relationships do not necessarily constitute non-linear predictor construct – criterion relationships. Therefore biodata items rationally scaled along a priori predictor constructs should be examined (as opposed to empirically keyed items) for non-linear P-C relationships because they attempt to tap predictor constructs at the scale level.

The current paper examines P-C relationships for non-linearity to determine if a) hypothesized nonlinear P-C relationships at the construct level do exist, b) empirical keys are modeling hypothesized nonlinear P-C relationships at the construct level, and c) rational scales weighted nonlinearly can sufficiently model hypothesized nonlinear P-C relationships at the scale level. The implications of these three points are as follows: First, if non-linear P-C relationships are operating, we need to investigate them and integrate them in our biodata selection methodology. Second, if empirical keys capture predictor construct – criterion non-linearity then evidence for construct validity of empirical keys exist. Third, if sufficient P-C non-linearity exists, an extra step of hypothesizing and weighting the non-linearity may be necessary when constructing some rational scales.

Biodata scaling

As mentioned above, there are multiple approaches to scaling biodata items. This section will briefly review the major approaches: *empirical keying, factorial scaling, sub-grouping,* and *rational scaling* (see Stokes, Mumford, & Owens, 1994). Each method has advantages and disadvantages and the instrument's purpose may largely dictate the appropriateness for the choice of approach (Brown, 1994). For example, Brown (1994) stated as the goal of generalizability increases, the applicability of externally based measures such as empirical keying decreases because these scales are generated as a function of the development sample. Regardless of the purpose, empirical keying has traditionally been the most commonly used method and has been subject to notable criticism.

Empirical keying. Empirical keying is used to select and weight biodata items on the basis of the item's ability to predict performance. Patterns of past behavior and experiences related to specific criteria can be defined through this procedure. Using parametric or nonparametric analyses, item scores or response patterns that predict group memberships are used to assign weights. Items are assigned differential weights based on the magnitude and direction of observed relationships to the criterion. Cross-validation (the process of applying the empirically keyed items to a new sample) is performed to control for inflated correlation coefficients due to capitalization on sample-specific item - criterion relationships.

Although empirical keying procedures often result in high criterion related validity, some argue empirical keying could lead to inappropriate inferences regarding the performance of people due to various shortcomings such as lack of construct validity,

difficulty in score interpretation, validity shrinkage upon cross-validation, poor generalizability beyond the samples they are developed from, and the fact they are criterion bound (Mumford & Owens, 1987; Russell, 1994).

Factorial scaling. This procedure employs factor analytical techniques to identify meaningful psychological dimensions to predict criteria of interest. The factor analysis (typically exploratory) attempts to identify meaningful solutions that are both interpretable (have construct validity) and predictive (have criterion-related validity). Usually the solutions are rotated to an orthogonal or oblique criterion, which are then evaluated with respect to simple structure and meaningfulness. The solution that optimally satisfies structure parsimony and demonstrates psychological meaningfulness is usually retained for scaling. Items that yield factor loadings below .30 on a given dimension are typically eliminated (Brown, 1994).

The primary goal of factorial biodata scaling is to identify meaningful predictor dimensions (factors) based on a current set of items. Of course this implies the resulting factors need to be interpretable. Brown (1994) pointed out that without factor meaningfulness, this approach is just another empirical approach based on linear combinations of items or factors rather than the items themselves. It should also be noted that empirical reduction of predictor space through factor analysis might result in elimination of highly predictive, single items that do not load onto any factors. This can create a hurdle (i.e., items retained for the measure are linearly related to one another) that might interfere with finding linear or non-linear item-criterion relationships (Russell et al., 1990).

Sub-grouping. Whereas the primary goal of factorial scaling is to develop meaningful *factors* based on current set of items (reducing item data to factors), the method of sub-grouping's primary goal is to identify coherent and meaningful *groups of people* that are based on responses to biodata items (Brown, 1994). This biodata scaling technique is a statistically based classification procedure where people who display similarities in the patterns of prior experiences are sub-grouped together under the premise they will behave similarly in the future. Brown (1994) noted this procedure is useful for multi-criteria prediction and outlined the general procedure as first involving the development of an item pool of biodata experiences that cover a wide variety of situations. Items are then administered to a large sample of respondents and a principle components analysis of the subjects' item responses is conducted. Factors emerging from this analysis are used to profile individuals. Next, factor profiles are cluster analyzed to identify sub-groups of individuals with similar factor profiles. Central themes of each cluster are nominally determined and used to describe and label each cluster / sub-group. Individuals within the subgroups are described by their mean factor profile. Criterion-related validation looks at subgroup similarities and differences on the criterion of interest. Differences in behaviors and experiences that are characteristic of or related to good or bad performance can be established during validation to predict actual performance (Owens, 1976; Owens & Schoenfeldt, 1979).

The general rational behind sub-grouping is based on Owens's developmental-integrative model (Owens, 1968, 1971, 1976; Owens & Schoenfeldt, 1979) and its successor, the ecology model (Mumford, Stokes, & Owens, 1990). The ecology model views people as organisms who actively seek experiences and opportunities to maximize

long-term adaptation to their environment. Given satisfactory outcomes, individuals will actively seek out similar situations in the future – thus producing meaningful patterns of behavior (Mumford and Owens, 1987; Mumford et al., 1990; Mumford, Wesley, & Shaffer, 1987; Stokes, Mumford, & Owens, 1994).

Rational scaling. Rational scaling procedures are based on existing theory or rational reasoning and involve identification of latent predictor constructs expected to be developmental determinants of future job performance. This technique involves specification of a priori hypotheses that indicate certain items should be grouped together in a scale that theoretically or rationally measures a predictor construct. Rational scaling is helpful in emphasizing content and construct validity of the resulting scale. This makes the rational scaling method attractive for hypothesis testing, theory building, and construct validation as the method can provide considerable psychological insight (Hough & Paullin, 1994). When scaling is carefully conducted, content validity is expected to be strong because items are generated or chosen on the basis of their extent of representing or mapping the predictor domain. Further, there is real potential for finding evidence of construct validity, and thus, increased understanding of the underlying constructs and their relationship with one another because of the direct attempt to measure meaningful psychological constructs (Brown, 1994).

Criterion-related validities of rational scales reported in the biodata literature are, on average, typically .02 to .03 smaller than cross-validities reported for empirical keys (Berkley, 1953; Hough & Paullin, 1994; Mitchell & Klimoski, 1982; Mumford, Costantza, Connelly, & Johnson, 1996; Reiter-Palmon & Connelly, 2000; Schoenfeldt, 1989). This small gain in predictive power is arguably not worth the apparent lack of

construct validity of the empirical keys. Further support for the use of rational scales includes less shrinkage and greater generalization then with empirical keys. Shrinkage of predictive validities upon cross validation is usually less with rational scales than for empirical keys and the rational approach may produce inventories of considerable generalizability and may contribute to greater understanding of person-to-job matching (Berkley, 1953; Mitchell & Klimoski, 1982; Reiter-Palmon & Connelly, 2000).

Non-linear biodata P-C relationships

While the predictive advantage is generally weak, empirical keys may have distinct advantage over other methods because of, as mentioned above, their ability to capture non-linear P-C relationships. Indeed, A handful of studies employing the use of empirically keyed biodata items have shown patterns of non-linearity predicting job performance criteria at the construct level (i.e., Dean, Russell, & Muchinski, 1999; Kluger et al., 1991; Russell & Domm, 1990). Individual biodata items sometimes show non-linear P-C relationships in empirical keys for initial validation (as opposed to cross-validation); however, few studies report this to be the case for psychological constructs predicting criteria.

Exceptions to this include Russell and Domm (1990). These authors reported some biodata items showed linear relationships with a criterion when keyed at the item level, but non-linear relationships when keyed at the response option level. They also found items loading on a "negative life event" factor tended to have middle-range response options enter a vertical percent key (i.e., response options 2 – 4 on a Likert scale of 1 – 5), thus indicating a non-linear P-C relationship. Likewise, Russell et al. (1990) found similar results with a negative live events factor operating nonlinearly when

predicting U.S. Naval Officer performance in the fleet. Those officers reporting a moderate number of negative life experiences (as opposed to very few or very many) in a biodata inventory were more likely to be successful officers in the fleet (Russell et al., 1990). These findings were noted to be consistent with the goal setting literature: that moderately difficult goals tend to yield greater performance than goals too difficult or easy (Dean et al., 1999; Russell & Domm; 1990; Russell et al., 1990).

Kluger et al. (1991) conducted a study in which the same biodata instrument and empirical keys used in the Russell and Domm (1990) cross-validation sample were applied to a new sample of graduate students enrolled in business psychology courses. This study hypothesized that non-linear P-C relationships (captured by RO keys) make it difficult for respondents to guess the correct answer, thus making option-level keys more resilient to faking than item-level keys. Kluger et al. (1991) found option-level keys were not susceptible to inflated scores when subjects were instructed to "fake good;" however, item-level keys were. The more fake-resistant option-level keys presumably support an assumption of non-linear P-C relationships. Based on these studies, evidence was provided for stable nonlinear P-C relationships as the pattern of findings was replicated upon cross-validation with independent samples (Kluger et al., 1991; Russell & Domm, 1990). The following proposition is suggested from the review above:

> *Proposition*$_1$: The relationship between predictor constructs and job performance criteria is sometimes nonlinear. Some predictor constructs have nonlinear relationships with some criteria. To the extent that scales tapping those predictor constructs model the true nonlinear P-C relationship, those scales will demonstrate higher criterion-related validity in predicting job performance than the same items linearly weighted.

A priori predictor constructs: Big Five personality factors and work performance

Rational approaches to biodata scaling often involve personality constructs to base item content on (e.g., Hough & Paullin, 1994; Russell, 1994). The basic idea is items that identify past behaviors indicative of target personality dimensions can tap personality constructs believed to be related to job performance. For example, because conscientiousness has been found to predict overall job performance (e.g., Barrick & Mount, 1991) and high-conscientious people tend to behave in similar patterns (e.g., usually on time for and rarely absent from work), one might conclude biodata items such as, "How often were you late to work on your last job?" or "How many times did you miss scheduled work days for the final 4 months of your previous job?" to be tapping the conscientiousness construct domain.

In recent decades, the factors of the "Big Five" or Five-Factor Model (FFM) have enjoyed a resurgence of interest in the personnel selection arena (Barrick & Mount, 1991; Hurtz, & Donovan, 2000; Salgado, 1997; Tett, Jackson, & Rothstein, 1991).
The Big Five or FFM taxonomy consists of five general factors: neuroticism (emotional instability), extraversion (surgency), openness to experience (open-mindedness), agreeableness (getting along with others), and conscientiousness (work orientation).
Of the five factors, conscientiousness seems to be the most robust predictor of overall job performance across a variety of occupations and is also highly predictive of training proficiency (Barrick & Mount, 1991; Barrick & Mount 1993; Hurtz, & Donovan, 2000; Salgado, 1997).

There has also been some support for the other factors of the FFM predicting performance. Agreeableness has been demonstrated to be a good predictor of overall job

performance across a variety of occupations in some studies (i.e., Tett et al., 1991). Agreeableness also seems to be related to jobs requiring interpersonal skill such as sales or management (Hurtz & Donovan, 2000). Additionally, criteria in team performance (Barrick et al., 1998) and training proficiency (Salgado, 1997) domains appear to be linked to agreeableness.

Studies have also show extraversion to be consistently related to performance in jobs requiring interpersonal skills such as sales or managerial positions (Barrick & Mount, 1991; Barrick & Mount 1993; Hurtz & Donovan, 2000; Salgado, 1997).

Openness to experience was found to be a stable and valid predictor of overall job performance across a variety of occupations in a major meta-analysis (Tett et al., 1991). Openness has also been linked to training proficiency (Barrick & Mount, 1991; Salgado, 1997) and customer service performance (Hurtz & Donovan, 2000).

Finally, neuroticism has been consistently linked to overall job performance across a variety of occupations (Salgado, 1997; Hurtz & Donovan, 2000) as well as team performance (Barrick et al., 1998). It is important to note, however, not every study examining the effectiveness of the FFM factors predicting job performance demonstrated acceptable criterion-related validity (e.g., Nikolaou & Robertson, 2001). For example, Nikolaou and Robertson (2001) found no significant relationships between any of the five factors of the FFM and overall job performance in a male Greek sample.

The factors of the FFM predicting job performance have also been examined in two domains that both lead to overall job proficiency: task performance and contextual performance (i.e., Robertson & Callinan, 1998). Borman and Motowidlo (1993) first proposed the distinction between task and contextual performance and described the

former as performance prescribed by the job, which includes behavior directly relevant to the technical core of the job. The latter, was described as behaviors that do not relate to the technical core of the job directly, but support the broader social and organizational environment in which the behaviors relevant to the technical core are conducted (Borman & Motowidlo, 1993). An example of contextual performance currently receiving much attention is organizational citizenship behavior (OCB; e.g., Borman & Motowidlo, 1993). OCB involves discretionary behavior that is not part of an employee's formal job requirements but nonetheless promotes the effective functioning of the organization (e.g., volunteering for extra duties or helping co-workers with their tasks; Organ & Ryan, 1995). Robertson and Callinan (1998) argued that the FFM is more closely linked with contextual performance than task performance. Additionally, Motowidlo and Van Scotter (1994) provided evidence that both task and contextual performance independently contribute to overall job performance; and that their results show personality variables more highly correlate with contextual than task performance.

Regardless of the criterion-related validity of the FFM and task, contextual, or overall job performance, personality dimensions (specifically the FFM) have demonstrated substantial evidence for construct validity (Digman, 1990; Digman & Inouye, 1986). Digman (1990) and Digman and Inouye (1986) provided evidence of construct validity of the FFM by demonstrating interpretable factor structures, consistency in factor structures over time, and convergent and discriminant validity. Thus, using the FFM shows great promise for construct validity of construct-based biodata scales (i.e., rational scales; Russell, 1994). Based on the review above the

following proposition is set forth with regard to the criterion-related validity of the FFM and job performance:

> *Proposition*$_2$: Conscientiousness is expected to be the strongest predictor of overall job performance, with the other factors also predicting job performance to a lesser degree. Openness to experience and conscientiousness are expected to be the strongest predictors of training proficiency.

A priori non-linear P-C relationships: Evolution of personality factors

Students of personality and laypeople alike have long asked the questions why do distinct and meaningful personality "traits" persist within persons and why do they vary across individuals (e.g., McCrae & Costa, 1999). Some authors argue convincingly for a biological basis of personality traits and that human beings are products of evolution (Adler, 1996; Buss, 1991, 1996; McCrae & Costa, 1999). Buss (1996) made a good case for personality traits pertaining to social adaptation. Individuals with different personality traits approach survival and reproduction in various ways.

In an explanation of the origins of individual differences, Buss (1991) proposed dispositions (i.e., personality traits) as evolved problem-solving strategies. Buss (1991) presented four major evolutionary routes to the emergence of consistent individual differences in dispositional strategies (the first two are discussed in detail below). The first is termed *heritable alternative strategies* – genetically based strategy differences due to selection within alternative niches or frequency-dependent selection. Second, *heritable calibration of psychological mechanisms* – cases in which adaptive optima have fluctuated or changed over time or location, thus producing heritable variation in the threshold or calibration of species-typical (e.g., psychological) mechanisms. Third, *situational contingent alternative strategies* – where different strategies are activated by

the general environmental situation, all of which constitute a species-typical repertoire inherent in each individual. The fourth route is *developmental calibration of psychological mechanisms* – individually different life experiences during development set the threshold on, or calibrate, species-typical mechanisms in a constant way, thus producing individual differences among species (Buss, 1991).

The last two evolutionary routes to dispositional individual differences (situational contingent alternative strategies and developmental calibration of psychological mechanisms) involve moderator effects of either the general environment or individual developmental experiences on the emergences of individual trait differences. The similarity between the last two evolutionary routes is strikingly similar to the ecology model of biodata and human development (Mumford et al., 1990). Recall this model views people as organisms that actively seek experiences and opportunities to maximize long-term adaptation to their environment. Given satisfactory outcomes, individuals tend to seek out similar situations in the future, producing meaningful patterns of behavior (Mumford and Owens, 1987; Mumford et al., 1987; Mumford et al, 1990; Stokes et al., 1989).

However, while identifying moderator variables is extremely important for advancing our understanding of the nomological network of personality and work performance, it is not the focus of the current paper which examines the nomological network in terms of the nature of the relationship (i.e, linear and nonlinear) between two variables (i.e, FFM and job performance). As such, the first two routes presented (heritable alternative strategies and heritable calibration of psychological mechanisms) apply directly to the current thesis and are therefore discussed in detail below.

Heritable alternative strategies. This evolutionary route maintains that effective alternative strategies for survival and reproductive success exist. Buss (1991) cited studies that identified a variety of personality characteristics that covary within female "sociosexual orientation," which encompasses an individual's tendency to form viable mateships vs. brief sexual encounters (see also Snyder, Simpson, & Gangestad, 1986; Gangestad & Simpson, 1990). One example presented was females using short-term mating strategies increased their chances of being inseminated by men of greater attractiveness while females using the long-term strategy increased the likelihood of obtaining a mate interested in substantial parental contribution. Because sociosexual orientation is clearly linked with extraversion and conscientiousness (and may be linked to a lesser degree with the other three factors), research in this area provides convincing links between evolutionary theory, personality variables, and the concept of dispositions as strategies (Buss, 1991).

McCrae and Costa (1999) also suggested individual differences in personality may reflect different but equally effective adaptation strategies. These authors offered a couple of possibilities to illustrate this point: agreeableness makes it easier to acquire allies but antagonism increases one's competitive edge and; openness leads to obtaining new resources but conventionality exploits tried-and-true methods (McCrae & Costa, 1999). Here, those very low or high in both agreeableness and openness to experience may have been at a distinct advantage in ancestral environments to those operating in the mid-regions of these dimensions. The other factors may also represent alternate strategies for survival and reproduction success. For example, individuals low in conscientiousness may employ an attraction strategy of making their mate jealous while those high in the

same factor may entice their mates by being extra considerate and thoughtful. These papers suggest that different personality characteristics may be associated with alternate strategies adapted by people for effectively performing required tasks.

This evolutionary hypothesis states a clear or extreme strategy is adaptive; however, mixes or moderation of strategies is less adaptive. The nature of a nonlinear P-C relationship under an alternate strategies hypothesis is predicted to be a U shape. The endpoints on a personality dimension represent alternate strategies; high or low on a dimension predicts better evolutionary success.

One obvious inference is these alternative adaptive strategies may be important for success and survivability at work. Operating under this inference and to the extent that extreme points on a dispositional dimension (i.e., factors of the Big Five) represent alternate effective strategies, one might expect those low or high on a factor of the Big Five to be associated with higher overall job performance than those in the mid-regions of a factor. Thus, under the theory of heritable alternative strategies for personality factors, the following proposition is made:

> *Proposition*$_{3a}$: Rational biodata scales of the FFM that are scored as a "U" (where extreme responses are weighted higher and mid-range response are weighted lower) will increase predictiveness of job performance over and above the same scales weighted linearly.
>
> *Proposition*$_{3b}$: Rational biodata scales of the FFM that are scored as a "U" will have criterion-related validities approaching the cross-validities of the same scale items empirically keyed.

Heritable calibration of psychological mechanisms. One observation that raises doubt that personality dispositions represent heritable alternative adaptations is that they seem to be continuously and normally distributed among the population (Buss, 1991;

Plomin & Nesselroade, 1990). This observation is incongruent with a bimodal distribution suggested by the alternative strategies hypothesis. On the other hand, the hypothesis of heritable calibration of psychological mechanisms suggests that adaptive optima have fluctuated over time or place. The latter hypothesis is congruent with moderate heritablity associated with these continuous distributions (Buss, 1991). For example, Zuckerman (1990) argued that heritable differences in sensation-seeking (a trait associated with extraversion) may correspond to variation in thresholds for avoiding or approaching resources (i.e., mates) – with various threshold levels possessing both benefits and costs. In terms of human ancestral environments, those with a smaller tendency to seek sensation incurred less risk but failed to accrue the reproductive benefits probabilistically associated with high amounts of sensation-seeking (i.e., vigorous approach behaviors). Conversely, those with a larger tendency to seek sensation increased the probability of reproductive success but also invited greater risk (Zuckerman, 1990).

Individual differences in sensation-seeking may signify differences in threshold setting that are products of past environments imposing different adaptive optima (Buss, 1991). For example, environments where food is in great supply may favor higher risk-taking mating behavior whereas those environments in short supply of food may favor lower risk-taking mating behavior, as risk may be incurred in the struggle for food. Ultimately, variations within a normal range of personality traits represent the normal range of historical adaptive optima for those traits (Buss, 1991).

Adler (1996) also argued the Big Five reflects fundamental traits instrumental for human adaptation in our early evolutionary past. This author noted each of the Big Five

factors was critical for survival in clan-centered, hunter-gatherer societies that characterize much of our evolutionary past. Moderate levels of neuroticism was critical to survival in the face of adversity; extraversion reflects the importance of dominance hierarchies in social organizations in which our ancestors lived; openness was required to solve newly emerging challenges to life; agreeableness promoted cooperative behavior and; conscientiousness promoted reliability and dependability in task performance. To the extent that our ancestors possessed aspects of these factors within the normal range of distribution, they were evolutionarily successful (Adler, 1996).

The theory of heritable calibration of thresholds on psychological mechanisms provides a basis for making predictions in work performance. Adler (1996) suggested the evolutionary approach may be particularly useful in establishing the links between FFM traits and adaptations required by life in organizations, especially the changing adaptations required over a career and over organizational life-cycles. Overall job performance may be partly a function of individuals operating in the normal range of personality dimensions. Under the heritable calibration theory, one might expect moderate levels (in the mid-regions) of a Big Five factor to predict better overall job performance than extreme levels (very high or low) of a factor. Thus, the following proposition is made:

> *Proposition*$_{4a}$: Rational biodata scales of the FFM that are scored as an inverted "U" (where extreme responses are weighted lower and mid-range response are weighted higher) will increase predictiveness of job performance over and above the same scales weighted linearly.
>
> *Proposition*$_{4b}$: Rational biodata scales of the FFM that are scored as an inverted "U" will have criterion-related validities approaching the cross-validities of the same scale items empirically keyed.

It is important to note Propositions 3a and 4a make exact opposite predictions and are mutually exclusive. If one is demonstrated to be the case then, by definition, the other is refuted. The alternate strategies hypothesis (Proposition 3a) predicts a U and the normal range hypothesis predicts an inverted U relationship between the FFM biodata scales and job performance.

Summary

The intent of the present study is to examine an issue involved in a debate over the use of empirical keys for scaling biodata: potential non-linear predictor-criterion (P-C) relationships. The current paper examines P-C relationships for non-linearity to determine if hypothesized nonlinear P-C relationships at the construct level do exist, empirical keys are modeling hypothesized nonlinear P-C relationships at the construct level, and rational scales weighted nonlinearly can sufficiently model hypothesized nonlinear P-C relationships at the scale level. The next chapter outlines the methodology for the current study and presents initial evidence of construct validity for the study variables.

Chapter 2: Method

Sample

The sample consisted of data from Federal Aviation Administration (FAA) air traffic controller (ATCS) applicants who subsequently had useable criterion scores (n = 4,559). The applicants were administered the Office of Personnel Management (OPM) Applicant Background Assessment (ABA) test during the period of 1986-1992. These data were from a random subset of a larger database of 11,405 ATCS candidates selected for training by the FAA for the same period from of a pool of 206,592 applicants. The ABA biodata instrument was administered for research purposes only and was not used for selection at any point.

Candidates were selected for training into the FAA Academy ATCS Nonradar Screen program with the OPM's air traffic control selection test battery. Candidates who successfully completed the multi-week training and screening program (the "Screen") were then selected as ATCSs for the FAA. Thus, the ATCS selection system was a two-stage process with the OPM test battery as the first hurdle and the Screen as the second.

The four-hour OPM test battery consisted of the Multiplex Controller Aptitude Test, the Abstract Reasoning Test, and the Occupational Knowledge Test. The selection ratio of applicants advancing to the Screen was approximately .055 (5.5%). The Screen was a nine-week initial training program that provided candidates of the FAA ATCS Academy with basic knowledge of air traffic rules and procedures then tested applicant knowledge with written exams and laboratory simulations. Due to training attrition of a number of candidates, the total number of those with usable Screen performance data was reduced to 10,014. Of those candidates advancing to the second stage of the selection

process, 54.6% passed, 33.4% failed, and 12.0% withdrew from the Screen. Those passing were ultimately selected for FAA service, with a selection ration of .546. A composite score of Screen performance was used as the criterion measure for the current study and is described in greater detail below.

The Revised NEO-PI-R personality test (Costa & McCrae, 1990) was administered to all candidates (n = 1,856) during the years 1990 to 1992. Of the candidates with usable Screen and ABA data, a total of 183 also had NEO-PI-R data. A qualitative analysis of the means of the Screen criterion and other descriptive statistics (Table 2.1) showed no real differences in Screen performance between the NEO subset (M Screen = 70.05, n = 183), the ABA subset (M Screen = 71.13, n = 4559), and the total sample who completed training (M Screen = 71.66, N = 10,014). Like the ABA, the NEO was used for research purposes by the FAA and not for air traffic controller selection at any phase.

Manning, Kegg, and Collins (1988) listed minimum initial requirements for consideration as an ATCS candidate; these included a security clearance, a medical qualification, 18-30 years of age, a high school education or equivalent, and at least three years of work experience or college.

The ABA biodata sample (N = 4,559) had an average age of 25.8 years, was 83.8% male, and predominately white (90.4%). In terms of advanced education, 11.1% had just a high school degree, 55.6% had finished some college, 31.8% had a bachelor's degree, and only 1.3% had earned an advance degree before entry to the ATCS profession. Over three-quarters (77.4%) of these applicants had no prior aviation or air traffic control experience. Table 2.1 reports descriptive statistics for the overall sample

(N = 11,405), the ABA plus Screen subset (n = 4,559), and the ABA plus Screen plus NEO subset (n = 183). As shown in the table, no meaningful differences emerged across groups.

Table 2.1 about here

Predictor

The ABA is a 142-item biodata inventory that served as the predictor in the current study. The present study examines possible non-linear P-C relationships from item response patterns measured along a continuum (i.e., Likert scale) and the ABA primarily uses Likert scale response formats. However, 15 items had nominal category response formats, which do not measure responses along a single continuum. Items with non-continuous response formats were not appropriate for the purpose of this study and were eliminated, leaving a total of 127 items analyzed here. Escape options were written for items to allow for "don't know" and not-applicable responses (Gandy, Dye, & MacLane, 1994).

Most of the items from the ABA used here were originally developed for the OPM' s Individual Achievement Record (IAR), an inventory involved in the first major application of biodata in the U.S. Civil Service. The development of an initial item inventory began with three preliminary activities: a) a review of job analysis information on federal and non-supervisory professional and administrative occupations, b) review of available taxonomies of past behavior items, and c) establishment of criteria for the acceptability of biodata for use in the public-sector (Gandy et al., 1994). Criteria of

acceptability included the following: 1) items had to deal with events that were primarily under respondent's control, 2) must have potential relevance to job performance, 3) must be verifiable in principle, 4) must be nonintrusive to personal privacy, and 5) must avoid stereotyping by race, sex, or national origin. As a result of the inclusion criteria, the domains for the acceptable items included school and educational experience, work history, skills, and interpersonal relations (Gandy et al., 1994). Figure 2.1 illustrates example items form the inventory.

Figure 2.1 about here

When developing item content, a concerted effort was made to cover life experiences, choices, and outcomes that are mediated by a wide range of cognitive, motivational, and interpersonal constructs. In general, items were developed to loosely reflect hypotheses that different levels in applicants' work, education, and interpersonal areas tapped by items would predict job performance. Item writers were aided by familiarity with job analysis information (i.e., necessary KSAOs for performance) and general characteristics of the applicant population in terms of variability in appropriate prior experiences (Gandy et al., 1994).

The FAA adopted the IAR to form the ABA, and there has been mixed evidence for construct validity of this instrument based on the ecology model framework and positive support for criterion-related validity (Dean, 1999). Likewise, Gandy et al. (1994) reported evidence for construct validity based on the four domains of acceptability (work competency, high school achievement, college achievement, and leadership skills) as

well as criterion-related validity when predicting job performance of civil service workers. The full ABA biodata inventory can be found in Appendix A.

Reference FFM personality measure

A five-factor reference measure was needed to examine the extent of convergence and discrimination with the rational biodata five-factor scales. The Revised NEO-PI (1990) is a 181-item self-report measure of the five personality factors of neuroticism, extraversion, openness to experience, agreeableness, and consciousness. This inventory was administered to all candidates during the years of 1990 to 1992 and was chosen for the current study due to its widespread use in industry and accessibility for research purposes (Botwin, 1995). Forty-eight items assessed each factor of neuroticism, extraversion, and openness to experience. Each of these three factors also had six facet scales measured by eight items per facet. This version of the NEO (1990) did not yet have facet-scales for the factors of consciousness or agreeableness; these factor scales were comprised of 18 items each. The scale reliabilities were estimated with Cronbach coefficient alpha and were not unlike those typically reported in the industrial psychology literature (i.e., Botwin, 1995). All NEO scale and sub-scale means, standard deviations, and Cronbach alphas are reported in table 2.2.

Table 2.2 about here

Criterion

A nine-week initial training and assessment program referred to as the FAA Non-radar Screen was the second stage of the selection process for FAA ATCSs. This

program was administered by the FAA Academy in Oklahoma City, OK. The Screen composite score (NLCOMP) served as the primary criterion measure for the present study. Della Rocco, Manning, and Wing (1990) reported the Screen taught candidates with no prior air traffic control knowledge enough about the job to potentially advance to "full performance level" as an operating ATCS.

Screen performance assessments were categorized into three arenas: a) paper and pencil tests, b) simulations, and c) final examination. These three components were combined to form the Screen composite score. The paper and pencil tests and final exam were each weighted 20% and the simulator performance 60%. The paper and pencil measures were a series of multiple choice tests designed to measure applicant ability to learn and retain basic job knowledge. The simulations incorporated systematic evaluations of candidate performance on six 30-minute laboratory simulations of Nonradar air traffic control. The final component of the Screen composite was a multiple choice final exam assessing trainees' ability to apply rules and procedures of air traffic control.

The Screen composite was the FAA's primary performance criterion for the Academy. Traditional job performance measures were not available for this sample due, in part, to union agreements that mandated controllers be evaluated on a dichotomous, satisfactory—non-satisfactory criterion, resulting in less precise performance measures (Dean, 1999). Nonetheless, studies have shown the Screen composite serves as an adequate surrogate for job performance.

Broach and Manning (1994) studied the Screen's ability to predict future performance in subsequent on-the-job (OJT) radar training after one to two years as non-

radar ATCSs. After correcting for range restriction, Screen performance was significantly correlated ($r = .50$, $p < .01$) with OJT radar training performance. Additionally, the Screen contributed incrementally over their current selection tool, aptitude testing, in explaining OJT training variance for two types of radar training (16% & 20%; Broach & Manning, 1994).

Other criterion-related validation efforts have also provided support for the overall Screen composite as a surrogate for job performance. Researchers have found the Screen to significantly correlate with other field training success ($r = .44$ corrected for range restriction; Della Rocco et al., 1990), attrition and field training status (Manning, Della Rocco, & Bryant, 1989), and ATCS on-the-job success (Cobb, 1962, 1967; Trites, 1961). In sum, copious studies suggested Screen performance was highly predictive of on-the-job performance and should provide an adequate surrogate measure of ATCS job performance.

Rational Approach

The five factors of the FFM served as the target constructs for classifying items from the ABA questionnaire. Additionally, rational subscales were developed by classifying ABA items in terms of how they reflected general work activities (GWAs) as set forth by Occupational Information System (O*NET; 1999). The idea of this approach is to develop rational subscales that introduce work context. The reason for this rests in the notion that criterion-relatedness of rational biodata scales is somewhat dependent on performance and task domains. GWAs can be used to examine different work domains.

Manley, Halbesleben, & Mumford (2000) developed rational biodata items by targeting the constructs of the FFM within GWAs of O*NET. These authors wrote items

that asked how often respondents have engaged in behavior typical of high or low amounts of personality traits of the FFM within the context of work. For example, in tapping low conscientious (construct) behavior while interacting with customers (context), the following item resulted, "How often have you forgotten about a promise you made to a customer?" Modest evidence for the construct validity of the scales developed from these items was demonstrated with convergent and discriminant validity and factor structure goodness of fit indices (Manley et al., 2000). For the current study, ABA items were categorized along the five factors of the FFM and O*NET GWA dimensions.

*O*NET general work activities.* The GWA taxonomic structure is centered in the stimulus – organism – response (S-O-R) paradigm of behavior within environments. In terms of work behavior, S represents stimuli (information) that is received by the worker, O represents the worker who is the receptor of that stimuli or the mediation process as performed by the worker, and R represents the various responses resulting from the worker processing the stimuli (actions performed by the worker in response to the processed stimuli). The S-O-R paradigm was used in the construction of the Position Analysis Questionnaire (PAQ; McCormick, Jeanneret, & Mecham, 1969a) in an information-theory context to organize structured worker-oriented job analysis questionnaires.

One intent of the GWA model was to extend the PAQ's S-O-R paradigm to include the work context of interactions-with-others when analyzing the content of jobs (Jeanneret, Borman, Kubisiak, & Hanson, 1999). The higher order levels of the GWA taxonomy are information input (Stimulus), mental processes (Organism), work output

(Response), and interactions with others (work environment). *Information input* involves where and how information and data needed to perform the job are gained. This dimension involves a) looking for and receiving job-related information (how is the information obtained to perform this job) and b) identifying and evaluating job-relevant information (how is information interpreted to perform this job). *Mental processes* concern the processes, planning, problem-solving, decision-making, and innovating activities performed with job-relevant information. Included in this dimension are a) information and data processing and b) reasoning and decision-making. *Work output* concerns physical activities performed, equipment and vehicles operated or controlled, and complex or technical activities accomplished as job outputs. This dimension includes a) performing physical and manual work activities (activities using the body and hands) and b) performing complex or technical activities (skilled activities using coordinated movements). Finally, *interactions with others* involves interactions with other persons or supervisory activities occurring while performing the job. This last dimension concerns a) communicating and interacting, b) coordinating, developing, managing, and advising others, and c) administering (staffing, monitoring, and controlling activities).

Initial scale formation. A panel of nine advance-level (at least third year in graduate school) industrial/organizational psychology graduate students was assembled and given detailed instruction on rater error training, conceptual definitions of the FFM factors and O*NET GWAs, and classification procedure for Q-sorting items from the ABA biodata questionnaire into FFM and O*NET GWA categories. Specifically, the panel was instructed on common rater errors that occur when making subjective ratings (e.g., halo & distributional errors) and how to recognize these tendencies in an effort to

reduce the occurrence of errors made by raters. Additionally, the panel was given complete descriptions or conceptional definitions of the factors and facet traits of the FFM and of the higher order GWAs as well as lower order activities within each GWA. Finally, the Q-sort procedure (Rogers, 1951) was explained to the panel as a means of categorizing the items. Panel members reviewed each item twice, once for the FFM classification and once for the GWA determination.

Items that obtained at least a majority agreement (5/9ths or 55%) for a category were classified as such. Of the items in the ABA, 85.83% and 72.44% were classified into the FFM and O*NET GWA categories, respectfully, by initially exceeding the 55% criterion. Raters in groups subsequently discussed items that did not reach the 55% criterion until consensus was reached. All of the remaining items were subsequently classified as a result of the consensus discussions leaving a total of 127 items categorized into both FFM and GWA domains. Inter-rater agreement ranged between 55% and 100 % for the FFM and GWA scales.

Matrix subscale formation. After items from the ABA were formed into rational scales for the FFM and O*NET GWAs according to Q-sort results, a matrix approach (i.e., Manley et al., 2002) was used to combine both FFM and GWA domains. Here, factors from the FFM comprised the columns and GWAs formed the rows of the matrix. Items were placed on the matrix in their respective cells. Each cell in the matrix represented rational FFM sub-scales within GWA context. For example the following item was classified as conscientious in the first Q-sort and as mental processes in the second Q-sort:

> My high school teachers would most likely describe my self-discipline as:
> A. superior
> B. above average

C. average
D. below average
E. don't know

This item was placed in the matrix in the conscientious column and in the mental processes row. This item belongs in the conscientiousness by mental processes cell that represents a rational sub-scale of conscientiousness within the GWA context of mental processes. Likewise, any intersection between a FFM dimension and a GWA domain represented a FFM factor within a GWA context. Not every cell in the matrix had sufficient item content and some had no items at all. Of the cells containing at least three items, 7 rational FFM by GWA subscales were developed. Table 2.3 reports the descriptive statistics and reliability estimates for these variables. Alphas ranged from .26 (neuroticism) to .91 (conscientiousness) for the FFM scales, from .57 (work output) to .86 (mental processes) for the GWA scales, and from .43 (openness by interactions with others) to .90 (conscientiousness by information input) for the FFM by GWA matrix sub-scales.

Table 2.3 about here

Principle components sub-scaling. The rational scales produced from the Q-sort results described above were submitted to principle components analysis (PCA) and rotated on an oblique criterion to yield factorially derived subscales of the FFM and GWA dimensions. This provided a second set of sub-scales for examination of non-linear P-C relationships as well as procedural check for the subjective matrix approach. One might expect sub-scales resulting from either procedure (Q-sort matrix or PCA) should be somewhat similar if they are content- and construct-valid. The FFM scales yielded a total

of 7 reliable PCA sub-scales. Not unexpectedly, principle components did not emerge from the full FFM scales that exhibited low reliability (neuroticism and agreeableness). Alphas ranged from .70 to .90 for the FFM PCA sub-scales. The PCA variables, descriptive statistics, and reliability estimates are also reported in Table 2.3. Table 2.4 lists sample item stems for all rational scales and sub-scales produced from the matrix and PCA procedures. A list of all ABA biodata items within each scale and subscale can be found in Appendix B.

Table 2.4 about here

Linear item weighting. Items were linearly weighted along rational scales by giving ROs measured on a Likert scale unit values such that a response indicating a minimum amount on a dimension received a 1, moderate amounts received 2, 3, or 4, and maximum amounts received a 5. If the item had an escape option (e.g., "never had the opportunity to perform this behavior"), the option was weighted a zero, hence, dropped from the key. Next, item analysis was conducted by successively dropping an item and correlating it with the total scale. Berkeley (1953) suggested guidelines for inclusion criteria of items within homogeneous scales, which were used here; items correlating .30 or higher were retained for scaling. Those items correlating negatively with the total scale were reversed scored. After item analysis and scale trimming, internal consistencies of the rational scales were analyzed using Cronbach's coefficient alpha.

Proposition 2 hypothesized conscientiousness to be the strongest predictor of overall Screen performance, with the other Big Five factors also predicting performance

to a lesser extent. Openness to experience was also expected to be a strong predictor of training success. Because the Screen is an indicator of both training success and overall job performance, both conscientiousness and openness were expected to strongly correlate with Screen performance. Support for Proposition 2 will exist if the rational FFM biodata scales and subscales exhibit significant criterion-related validities, specifically if conscientiousness and openness have the highest values.

Construct validation. Data for a subset of ATCS candidates (described above; n = 183) in which both ABA and NEO-PI data were collected were used for convergent and discriminant validity analysis. The FFM scales from the NEO-PI test were correlated with the biodata FFM scales for initial evidence of construct validity of the FFM rational scales. The biodata rational scale of agreeableness consisted only of 3 items and had an alpha of .30, therefore this scale was not expected to exhibit a pattern of convergent and discriminant validity. Excluding the agreeableness factor, convergent validities were not unlike those typically reported in applied psychology literature. These convergent coefficients ranged from .26 (openness) to .49 (neuroticism) and are reported on the diagonal of Table 2.5. Discriminant validities were low, ranging from .01-.28 and are reported on the off-diagonals. The overall pattern (excluding agreeableness) reflected good discriminant and convergent validity. Table 2.5 reports all correlations between the biodata (ABA) FFM scales and the NEO-PI FFM scales.

Table 2.5 about here

Confirmatory factor analysis (CFA) was used to further evaluate evidence of construct validity of the FFM scales and subscales. Two models were analyzed; first a model incorporating FFM subscales of GWAs from the matrix procedure described above and second a model of the FFM with PCA sub-factors. For both models, the latent factors were permitted to correlated and were indicated by FFM scales of a) the NEO-PI and b) the ABA biodata questionnaire.

Regarding the fist model, not every GWA dimension is represented for all of the Big Five factors. The neuroticism and agreeableness scales had no items mapping onto any GWA from the matrix procedure, thus these two Big Five factors were dropped from the model. Likewise, not every FFM scale is associated with all GWAs. Extraversion only mapped onto the interactions-with-others domain and openness-to-experience was only associated with the information-input and interactions-with-others GWAs. Therefore, the first model is an incomplete representation of the Big Five factors and GWA dimensions; however, this is to be expected as the ABA items were not developed by specifically targeting the factors of the Big Five. The NEO scores and GWA by FFM subscales indicated the remaining three factors.

Because PCA yielded no reliable subscales from the neuroticism or agreeableness rational scales, the appropriate full rational scales were used as indicators for these two factors in the second model. The other three FFM factors were indicated by the PCA subscales. Figures 2.2 and 2.3 illustrate the first and second models, respectively, with factor loadings.

Figures 2.2 and 2.3 about here

Factor loadings were strong for these models, generally ranging from the .30s to .60s. The fit was modestly good for both models with the matrix GWA model fitting slightly better than the PCA model by most fit indices. Table 2.6 reports the fit statistic for both models. These preliminary results provide further initial support for construct validity of the current study's biodata scales and subscales.

Table 2.6 about here

Non-linear unit weighting. Proposition 1a hypothesized nonlinear P-C relationships to exist. This proposition will be supported when nonlinearly weighted FFM biodata scales exhibit an increase in criterion-related validity over those same scales weighted linearly. Item weighting for curvilinear P-C relationships hypothesized in Chapter 1 (i.e., "normal range" and "alternative strategies" hypotheses) was accomplished by assigning extreme values of the 5-point Likert scale (i.e., 1 and 5) a value of 1, the 3 and 4 points on the scale a value of 3, and the middle point on the scale (i.e., 3) a value of 5. With this weighting scheme an inverted U shape is produced; thus, negative correlations exceeding the value of the linear criterion-related validity of the same scale will indicate support for the "alternative strategies" hypothesis (Proposition 3). Positive correlations exceeding the linear validities for the same scales will indicate support for the "normal range" hypothesis (Proposition 4).

Non-linear polynomial weighting. Non-linear weighting was also accomplished through the use of polynomial regression. Polynomials represent general non-linear models with single curvilinear bends in the regression line. Using stepwise hierarchical quadratic multiple regression (QMR; least squares regression in which all terms were entered simultaneously), the Screen composite score was regressed on (step 1) the linear term for rational biodata FFM scales / sub-scales and (step 2) the linear and quadratic terms for each scale as shown below.

Step 1: $\text{Yhat} = \beta_0 + \beta_1 \text{FFM}$

Step 2: $\text{Yhat} = \beta_0 + \beta_1 \text{FFM} + \beta_2 \text{FFM}^2$

The F-tests (p-values) for the linear and quadratic terms were then examined for significance. The difference in multiple correlations (ΔR and p-value) was also examined for an increase in criterion-related validity with the full model. Beta weights of the quadratic terms significantly different from zero will provide support for Proposition 1a. Plotting the regression curves indicated the nature of the curves. A U shaped curve supports the "alternate strategies" hypothesis (Proposition 3). However, support for the "normal range" hypothesis (Proposition 4) exists if an inverse U shaped curve is produced.

Empirical Approach

The ABA biodata measure was empirically keyed to produce RO weights for the items. As detailed in Chapter 1, a wide variety of empirical keying methods exist. Past research suggests methods directly estimating strength of relationships between biodata ROs and a criterion provide the most stable estimates (Dean, 1999; Devlin, Abrahams, & Edwards, 1992). The point biserial correlation (r_{pb}) between each RO and the Screen

criterion were used as weights for the empirical key. The point biserial correlation is a special case of the Pearson product-moment correlation (r) applicable when correlating a truly dichotomous (e.g., a RO that is endorsed or not) with a truly continuous variable (e.g., a performance measure). Here the Pearson formula reduces to a simpler formula for r_{pb}. In contrast to other RO empirical weighting techniques, the point biserial is a more efficient estimate of the strength of this necessarily linear relationship because it uses all observations in a sample instead of discarding the middle one-third performers, which is often done in the construction of empirical keys (i.e., VPDM; England, 1971; Mumford & Owens, 1987).

The biodata sample (N = 4,559) was split into thirds for key development and cross-validation samples according to when academy training chronologically occurred (e.g., first 1/3rd to receive training, 2nd 1/3rd to receive training, etc.). A triple-cross validation design in which every possible combination of 1/3rd – 2/3rd split was used to develop three key development and three cross validation samples. The keys were developed on the 2/3rd splits and then cross-validated on the remaining 1/3rd splits. The average of the RO weights for the three development splits constituted the final key and the average of the three cross-validation splits yielded the cross-validity for each RO.

Hybrid Approach

A hybrid approach was necessary for comparisons among criterion-related validity between the linearly and nonlinearly weighted rational scales and those same items empirically keyed at the RO level. Here empirically derived RO weights were used for every item within a rational scale. The criterion-related validities of the empirically keyed – rational scales are compared to those of the linearly and non-linearly weighted

rational scales. Regarding Proposition 1b, hybrid empirical – rational scales should yield criterion-related validities similar to the hypothesized non-linearly weighted rational scales to examine evidence for the construct validity of the empirical keys. For example, if a non-linearly weighted rational scale produced a criterion-related validity coefficient of .10 and the empirically keyed item ROs for the same scale yielded a validity coefficient of .11, one may conclude the non-linear weighting to be close to the effect occurring with the RO empirical keying procedure, thus providing construct validity for non-linear P-C relationships and empirical keys.

Summary

This chapter described the sample, measures, and all methodological procedures used in this dissertation. Also covered were preliminary results for initial construct validation of the rational scales. The next chapter details the results of the analyses outlined here and provides evidence of scale-level non-linear P-C relationships between rational scales and Screen performance composite.

Chapter 3: Results

Table 3.1 reports all intercorrelations of the main study variables. These variables include the criterion composite for ATCS Academy performance (NLCOMP), the primary rational scales (GWA and FFM scales) and the subscales derived from the matrix and PCA procedures. Because the sample analyzed was previously selected by the OPM cognitive abilities test battery, severe indirect range restriction on the criterion resulted; however, all correlation coefficients reported were not corrected for this range restriction. Further, all correlations reported here are functions of linearly weighted rational scales.

Proposition 2 predicted biodata scales rationally scaled along the factors of the Big Five to all predict job performance (NLCOMP) with conscientiousness and openness being the strongest predictors. A result showed the only full scale to have a significant correlation with the criterion was conscientiousness ($r = .10$, $p < .01$). Further, the matrix subscale of conscientiousness by mental processes and the PCA subscale of conscientiousness involving high school academic success had criterion-related validities of .12 and .17, respectfully. Additionally, these values all increased when weighted nonlinearly. Thus, partial support for Proposition 2 exists within the conscientiousness domain. This finding is not a surprise, as the conscientious scale comprised, by far, the most items of any FFM biodata scale. The other four scales of the FFM may have been somewhat deficient in item content. Table 3.1 reports all linearly scaled criterion-related validities of the main study variables.

Table 3.1 about here

Rational vs. empirical keying

Table 3.2 reports criterion-related validities for various scaling methods including rational linear and non-linear weighting and empirical keyed items within rational scales. Generally, empirical keys outperformed the rational scaling procedures upon cross-validation. This finding supports the predictive validity of empirical keying biodata for this sample, criterion, and job. Criterion-related validity for the full empirical key (all 142 items) was estimated at .363 (.293 when cross-validated); however, some of these items were categorical and were not used in the rational scales.

Most scales were not highly predictive of performance, however, four were moderately predictive. The four scales / sub-scales that were the most predictive for any scaling procedure were as follows, with the correlations reported in parentheses for the rational linear, rational non-linear, empirical key initial-validities, and empirical key cross-validities, respectfully: *mental processes* (GWA2; rs = .102, -.136, .266, & .241), *consciousness* (FFM5; rs = .098, -.144, .286, & .250), *consciousness by mental processes* (matrix subscale; rs = .115, -.150, .267, .244), and *consciousness- high school academic success* (PCA subscale; rs = .174, -.191, .223, & .218). All correlations for these scales, regardless of scaling procedure, were significant at the .01 alpha level (2-tailed).

The first three scales from the paragraph above show marked increases in criterion-related validity for the empirical keying procedure. This finding indicates either a) there is P-C non-linearity operating but it is operating at the item not scale level or b) there is P-C non-linearity operating but not in the hypothesized fashion (e.g., polynomial vs. quadratic). Nonetheless, this finding provides support for the empirical keying method over the rational scaling procedure for these three scales in terms or criterion-related

validity. The fourth scale mentioned in the above paragraph (high school academic success) showed only a slight increase of the empirical key over the rational scale. This indicates the P-C relationship may be non-linear as hypothesized (quadratic – a single bending curve) and provides support for the construct validity of the empirical key for this scale.

In terms of stability of the full empirical key, the criterion-related validity shrank .07 points for the full key, but decreases somewhat less with the four predictive scales of mental processes (.025), conscientiousness (.036), conscientious by mental processes (.022), and conscientiousness- high school academic success (.005). These findings support the stability and construct validity of the four predictive empirically keyed rational scales mentioned above (average shrinkage for theses scales was .025). Overall however, the empirically keyed scales shrank an average of .07. An observation that suggests combining the construct-based rational scaling procedure with empirical keying (hybrid approach) can lead to better scale stability, generalizability, and construct validity than a strict empirical approach. These validities, cross-validities, and amount of shrinkage for all scales are reported in table 3.2.

Table 3.2 about here

Linear vs. non-linear

As reported above, Table 3.2 shows the criterion-related validities for both linear and non-linear scaling methods. Three of the four predictive scales mentioned above showed marked increases in criterion-related validity when scaled non-linearly. When

compared to rational linear weighting, mental process increased (Δr) by .035 and .14 when weighted rationally non-linear and empirically keyed (cross-validity), respectfully. Likewise, conscientiousness increased .045 when weighted rational non-linear and by .15 when empirically keyed and cross-validated. The conscientiousness by mental processes subscale was increased by .035 and .13 but the conscientiousness – high school academic success subscale only increased by .02 and .04 for the rational nonlinear and cross-validated empirical key weighting schemes, respectfully.

Proposition 1 stated the relationship between predictor constructs and job performance criteria is sometimes nonlinear. The findings in Table 3.2 clearly show predictive validity can be increased for some scales. Most scales were not enhanced when scaled rational nonlinear, but interestingly, very few decreased. These findings provide initial support for Proposition 1.

Additionally, the signs of the significant correlations for the rational nonlinear scaling are mostly negative. Propositions 3a and 4a state nonlinear weighting will increase the predictiveness of rational biodata scales, but in opposite directions. Proposition 3 predicts a U relationship between the predictor and the criterion and Proposition 4 predicts an inverse U relationship. Proposition 3 (thus not Proposition 4) is primarily supported for those scales exhibiting nonlinear P-C relationships because the rational nonlinear scales were weighted as an inverse U (negative sign indicates a U relationship to be the case).

Propositions 3b and 4b state the criterion-related validities of a rational biodata scale when weighted nonlinear will approach those of the empirically keyed items within that scale. Most scales showed empirical keys (cross-validated) to outperform the rational

non-linear method with average increase in criterion-related validity about .08. Thus, support for Propositions 3b or 4b was generally not found. However, the conscientiousness-high school academic success subscale showed the rational nonlinear method to be very close to the empirical key with a difference in criterion-related validity of less then .03. Here, partial support is demonstrated for Proposition 3b.

It should be noted that for practical applications, the difficulty in using rational nonlinear weighting (assigning a priori unit values to response options) lies in predicting not only the constructs that have nonlinear P-C relationships with a particular job criterion, but also the direction (+/-) and nature of the nonlinearity (quadratic/polynomial). Hence, stepwise multiple QMR analysis was used to further examine hypothesized nonlinearity between the main study variables and job performance. Table 3.3 reports the results of QMR analysis. In terms of quadratic modeling of nonlinearity, six of the scales exhibited significant ($p < .01$) beta weights for the quadratic term. All of the significant quadratic curves were graphed and showed the P-C relationship to be consistent with Proposition 3a (U shaped curve).

Of the three scales that exhibited marked increases in criterion-related validity when scaled nonlinearly in the correlation analysis, just the conscientiousness and conscientiousness by mental processes scales exhibited significant betas weights when the criterion was regressed on the linear and quadratic terms (indicating nonlinear effect). The mental processes scale did not have significant beta weights in QMR. Additionally, the forth predictive scale discussed above (conscientiousness-high school academic success scale), which had only a marginal increase in criterion-related validity when rationally weighted, did have a significant quadratic term. These findings indicate partial

support for Propositions 1 and 3a (and refutes Proposition 4a). Not all scales showed evidence of hypothesized non-linear relationships; however, those that did consistently supported the "alternate strategies" hypothesis from evolutionary psychology and not the "normal range" hypothesis (see Chapter 1).

In terms of increasing predictive variance, none of the four predictive scales showed increases in R. Only the conscientiousness by information input and conscientiousness-college academic success scales showed significant increase in R (ΔRs = .054 & .045, respectfully). Thus for practical use, nonlinear weighting may not be warranted for the present sample-predictor-criterion.

Table 3.3 about here

Stability of predictive non-linear results of rational scales

Four predictive scales exhibiting some evidence of nonlinear P-C relationships included mental processes (GWA2), consciousness (FFM5), conscientiousness by mental processes (matrix subscale), and conscientiousness-high school academic success (PCA subscale). Not including the mental processes scale (which did not exhibit nonlinear effects in QMR), these scales were examined for item stability within the scales. Stepwise hierarchical QMR analysis was performed on individual items within each scale to examine similarity in beta weights and signs to those of the scale-level QMR analysis. Table 3.4 reports the results of this analysis for five random items within each scale. Results for all three scales generally showed a similar pattern in the beta weights and signs for the linear and quadratic terms, thus demonstrating stability and construct

validity of the non-linear effects for these scales. Appendix B lists all ABA biodata items within each scale and subscale.

Table 3.4 about here

Chapter 4: Discussion

Limitations

This study is limited in four primary ways that should be acknowledged. First, the data were severely range restricted (indirect) on the criterion. The applicant pool consisted of 209,000 applicants, of which 11,500 candidates were selected for training by a cognitive abilities test. Further, only 10,500 finished training and subsequent criterion measure. Therefore, only the top 5% of the applicant pool (in terms of cognitive ability) was analyzed for criterion-related validity. This severely reduced the total variance of criterion performance and creates a considerable muting effect on correlation coefficients.

Second, these data involve only one specific occupation with a very narrow task scope that tends to have little relationship with the personality factors of the Big Five. Post hoc analysis revealed only NEO Extraversion ($r=-.11$) and Conscientiousness ($r=.10$) scales modestly correlated with NLCOMP performance and none of the NEO scales to be significantly correlated with the performance criterion. Because the ABA biodata items were formed into rational scales judged to reflect the FFM, criterion-related validity may have decreased for these scales. Further, the present FAA sample scored higher on the conscientiousness personality dimension than the general population typically does. This atypically high conscientious sample may have affected the results.

Third, the factors of the Big Five and the four higher-order O*NET GWAs were used as target constructs for classifying preexisting biodata items and were not expressly written to tap these constructs. Therefore, construct validity of the Big Five and of the

O*NET GWAs is limited to judgments of the expert panel assembled to Q-sort the items into relevant categories.

Finally, this study had no examination of moderator effects of environment or life experiences such as might be expected with the other two evolutionary routes presented in Chapter 1. Recall, Buss (1991) proposed personality traits as evolved problem-solving strategies and presented four major evolutionary routes to the emergence of consistent individual differences in dispositional strategies. The two routes not examined here are termed *situational contingent alternative strategies* and *developmental calibration of psychological mechanisms*. The former involves differing strategies being activated by the general situation of the environment that, as a whole, constitute a species-typical repertoire inherent in each individual. The latter, proposes individually different life experiences during development that serve to calibrate, or set a threshold on, species-typical mechanisms in a systematic way – producing individual differences within species (Buss, 1991). Both routes involve moderator effects of either the general environment or individual developmental experiences on the emergence of individual trait differences. The current study is limited by lack of data for these moderator variables.

This limitation is somewhat exacerbated in light of the similarity among these proposed evolutionary routes, the current leading model of biodata construct validity – the ecology model (Mumford et al., 1990), and the method of sub-grouping for the scaling of biodata (Owens, 1976; Owens & Schoenfeldt, 1979). The ecology model maintains individuals seek out experiences and opportunities to increase long-term environmental adaptation. Upon satisfactory outcomes, individuals will seek out similar

future situations, which eventually produces meaningful behavioral patterns. The scaling method of sub-grouping attempts to identify meaningful and coherent groups of people based on biodata item responses. Individuals who display similar patterns of prior experiences are sub-grouped together under the assertion they will behave similarly in the future (Brown, 1994). The ecology model and the method of sub-grouping both involve moderator effects of the general environment and/or specific developmental life experiences. Future research should examine these environmental or experiential variables under the framework of the ecology model as it relates to evolutionary strategies of personality.

Summary of findings

Proposition 1 stated latent P-C relationships are sometimes nonlinear and that construct valid scales modeling the true form of P-C nonlinearity will increase criterion-related validity. Most study scales yielded low criterion-related validities when scaled linearly; however, the four scales of mental processes (GWA scale), conscientiousness (FFM scale), conscientiousness by mental processes (matrix subscale), and conscientiousness-high school academic success (PCA subscale) showed acceptable levels of criterion-related validity (average r = .12). Support for Proposition 1 was found when examining these four scales.

Three of these four "significantly predictive" scales exhibited marked increases in criterion-related validity when rational biodata scales were a priori weighted nonlinearly whereby mid-level responses received higher weights and extreme responses (1 and 5 on a 5-point Likert scale) received lower weights. The average increase across all four of the predictive scales was -.04 correlation points. This increase in validity was in the negative

direction indicating a U shaped single bending curve. Only the conscientiousness-high school academic success PCA subscale did not markedly increase when weighted nonlinearly.

Stepwise hierarchical QMR also revealed quadratic betas significantly different from zero ($p < .01$) for many of the study scales, including two of the four predictive scales mentioned above. The conscientiousness scale and the conscientiousness-high school academic success PCA subscale both had significant quadratic betas; however, the quadratic terms did not add significantly to the explained variance of the linear terms. Interestingly, two other subscales in the conscientious domain that were non-predictive when linearly weighted did have significant increases in explained variance when modeled with a quadratic term. The conscientiousness by information input matrix subscale ($\Delta R = .054$) and the conscientiousness-college academic success PCA subscale ($\Delta R = .045$) both increased significantly ($p > .01$) when the quadratic term was added to the linear model. Both of these scales significantly ($p > .01$) predicted the criterion when modeled with quadratic terms (average r = .09). Graphic analysis of QMR results revealed U shaped regression curves for all scales with significant quadratic betas.

Proposition 2 stated the factors of the Big Five would generally predict job performance with particular emphasis on the conscientiousness and openness to experience domains. The current study involves a job performance criterion for Academy training for FAA air traffic controller candidates. Research on the FFM suggests conscientiousness and openness to experience to be more strongly associated with this criterion measure because it is a composite of training success as well as a suitable surrogate for actual on-the-job performance (Barrick & Mount, 1991; Barrick & Mount,

1993; Hurtz & Donovan, 2000; Salgado, 1997; Tett et al., 1991). Partial support for Proposition 2 was found only with the conscientiousness domain for rational scaling. Rational biodata scales of openness to experience and the other three factors were not significantly related to this criterion.

Linearly and nonlinearly weighted rational scales of conscientious had criterion-related validities of .10 and -.14, respectively. The contentiousness by mental processes matrix subscale yielded linear and nonlinear criterion-related validities of .12 and -.15, respectively. Finally, the conscientiousness-high school academic success PCA subscale showed linear and nonlinear criterion-related validities of .17 and -.19, respectively. All other FFM scales and conscientiousness subscales had low validities when rationally scaled. It is of interest to note these two conscientiousness subscales, one from the matrix and one from the PCA procedure, yielded higher criterion-related validities than the full conscientiousness scale. These sub-scaling procedures tended to identify subsets of conscientiousness biodata items that were more predictive than the entire scale. For example, the most predictive rational scale / subscale of the entire study was the conscientiousness-high school academic success PCA subscale, which contained only eight of the 55 conscientious scale biodata items.

Proposition 3a states an inference made under the theory of heritable alternative strategies. Recall this evolutionary route states a clear or extreme strategy, such as being very low or very high on a personality dimension of the Big Five, is more adaptive than a moderated amount. Thus, rational biodata scales of the FFM scored as a U (extreme responses weighted higher than middle responses) will increase predictiveness over the same scales linearly weighted. Partial support was found for Proposition 3a as three of

the four predictive rational scales mentioned above exhibited U shape curves for the correlational analyses of the a priori nonlinearly weighted scales. Additionally, QMR analyses revealed U shaped curves when graphed for those scales exhibiting significant quadratic terms. It appears those predictor scales having nonlinear relationships with this criterion are U shaped single bending curves, supporting the heritable alternative strategies hypothesis.

Proposition 3b involves a comparison between rational and empirically keyed scales. It states nonlinearly weighted rational scales will approach the criterion-related validity of the same items empirically keyed. This proposition maintains empirical keys should operate in a predictable manner in order to show construct validity. Proposition 3b was primarily not supported. Only one scale exhibited rational validity similar to the empirically keyed validity, conscientiousness-high school academic success. However, this was a scale that was not enhanced through nonlinear weighting. This finding calls into question the construct validity of the empirically keyed scales, which will be discussed more below.

Proposition 4a states an inference made under the theory of heritable alternative strategies. This evolutionary route is termed heritable calibration of psychological mechanisms and makes exactly opposite and mutually exclusive predictions to the alternative strategies hypothesis. It suggests adaptive optima have fluctuated over time and place and, ultimately, variations within a normal range of Big Five traits represent the normal range of historical adaptive optima. Overall job performance may be associated with an individual's capacity for operating within the normal range of Big Five personality dimensions. Personality traits in the mid regions of a measure are presumed

more adaptive than extreme amounts. Thus, rational biodata scales of the FFM scored as an inverted U (extreme responses weighted lower than middle responses) will increase predictiveness over the same scales linearly weighted. Because Proposition 3a was supported, Proposition 4a (and 4b) are refuted. Future research should continue to explore the predictions made by heritable alternative strategies and heritable calibration of psychological mechanisms hypotheses in light of various moderator variables such as job and task type.

Scaling issues

Regarding the four predictive scales discussed above and illustrated in Table 3.2, the empirical key weighting method clearly outperformed the rational procedures in criterion-related validity; however, this was a highly restrictive sample in which items were not explicitly written to the constructs of the Big Five. The average validity (cross-validity) of the four scales empirically keyed was .26 (.24) whereas the average validities for the same scales linearly and nonlinearly rationally weighted were .12 and .16, respectively. Further, the average shrinkage for the empirically keys was only .02 for these four scales, indicating good stability and generalizability to others similar to the current sample.

The pattern of results shows linear weighting to be the least effective; however, once nonlinear weighting methods were used, the predictiveness of these scales usually increased. The four predictive rational scales increased an average of .04 correlation points when scaled rationally nonlinear, albeit in the negative direction (which speaks to the importance of correctly hypothesizing the direction of the curve with this procedure). When these same scales where empirically keyed and cross-validated (which allows for

nonlinear weighting when done at the RO level) the average increase was .12 correlation points. This pattern is interpreted possibly as nonlinear P-C relationships operating in which some of the nature of the nonlinearity is determined (i.e., U shape curve), but much of the nature of the curve(s) are yet to be a priori identified (i.e., additional variance vis-à-vis empirical keys). As each procedure allows for additional types of nonlinearity between predictor and criterion, the correlation increases. With this said; however, the nonlinear effects of the rational weighting procedure are generally small and may not account for the observed differences in validity between the rational and empirical approaches.

With the efficacy of empirical keys in terms of criterion-related validity duly noted, a question of construct validity for the empirical keys enters. The empirical keys for these scales were expected to follow hypothesized patterns of P-C nonlinearity; however, this was not the case. The empirical keys may likely be tapping nonlinear P-C relationships of multiple undetermined constructs.

In terms of rational scaling efficacy, the method was effective in partitioning predictive and non-predictive items into separate scales and subscales. Table 3.2 (in shading) shows the four predictive scales mentioned above as containing items that are clearly more predictive, rational or empirical, than the other scales and subscales. Further, both the matrix and PCA sub-scaling methods were effective at identifying subsets of predictive and non-predictive items within full scales. As mentioned above, smaller subsets were sometimes more predictive than the full scales from which they were extracted. Subsequent research endeavors should involve the matrix approach to rational biodata scaling and sub-scaling.

Theoretical implications

The heritable alternate strategies evolutionary route received some support for those scales displaying a non-linear P-C relationship with the criterion. This evolutionary route maintains that effective alternative strategies for survival and reproductive success do exist and that clear or extreme strategies are adaptive; however, mixes or moderate strategies are less adaptive. The primary targeted construct showing this effect was conscientiousness. Indeed, conscientiousness has been noted to have a bimodal relationship with job performance (Robertson & Callinan, 1998; Rust, 1999).

In a review examining the use of personality at work settings, Robertson and Callinan (1998) argued that the relationship between conscientiousness and job performance should not be treated in a linear fashion. These authors suggested people high or low in conscientiousness might perform at the same level of competence by caring out their work in alternative ways (Robertson & Callinan, 1998). Similarly, Rust (1999) found individuals who score very low on the conscientiousness scale are often fully functioning and successful individuals; in spite of the fact that those low in this factor are often described in pejorative terms such as hedonistic, unreliable, lazy, and lacking in self-discipline. Rust (1999) maintained relatively little attention has been paid to low levels of conscientiousness and that overall job performance is associated with both high and low levels of conscientiousness. These studies coupled with the current set of findings suggest that different personality characteristics in conscientiousness may be associated with alternate strategies adapted by people for effectively performing required tasks.

Other studies have also demonstrated nonlinear P-C relationships between personality measures of conscientiousness and other factors of the Big Five with performance; however, it is not always a U shaped relationship (Cucina & Vasilopoulos, 2003; Vasilopoulos, Cucina, & Goldenberg, 2002). Cucina and Vasilopoulos (2003) found that conscientiousness and openness to experience had nonlinear relationships with GPA among college students. These authors found moderate levels conscientiousness (inverted U curve) and very low or high levels of openness (U curve) predicted the highest performance (Cucina & Vasilopoulos, 2003). Similarly, Vasilopoulos et al. (2002) found nonlinear relationships between conscientiousness and emotional stability with training performance in a sample of Federal law enforcement trainees. The P-C relationships in this study were found to be inverted U shaped – moderated amounts predicted higher performance (Vasilopoulos et al., 2002).

It appears the nature of the nonlinearity may depend on moderator variables such as sample demographics (e.g., student vs. worker) and criterion type (e.g., air traffic controller performance vs. college GPA). Perhaps several of the evolutionary routes proposed by Buss (1991) are valid – depending on the situation. Future research in this area should seek to more fully examine the four evolutionary routes proposed for the Big Five in light of various work and school settings.

Practical implications

There are three primary practical implications offered from the results of the current study. First, predictability of FAA ATCS Academy performance can be enhanced through the use of rational biodata scales, especially when empirically keyed; however, it is not much enhanced through rational non-linear weighting. When weighting rational

scales with this biodata inventory-sample-criterion, linear weighting usually suffices. With this said, some of the scales were enhanced when weighted rationally non-linear. Thus, this method may be warranted in some situations. Again, the difficulty of this approach is in determining the direction (+/-) and nature of the nonlinearity (single or multiple bending curves). Future researchers should further explore these relationships before using in applied situations.

The second practical implication involves the use of top-down selection practices. The Society for Industrial and Organizational Psychologists Principles (1987) state the use of top-down selection practices is based on the assumption of a linear P-C relationship. The top-down approach may not always be recommended for measures of conscientiousness. If in some situations the very high or low conscientious individuals are the best performers and in other cases the moderately conscientious employees are the best, a top-down strategy will not always be selecting the most qualified applicants. The P-C relationship between selection tests and job performance should be validated in light of possible nonlinearity before employing top-down selection practices.

The third implication speaks to the potential of using O*NET GWAs for constructing rational biodata scales. GWAs may be a viable means for forming rational subscales to introduce work context in the predictor. The current study incorporated the four higher-level GWAs of the O*NET taxonomic structure, which was effective in introducing work context in the classification of biodata items. This method yielded subscales of items that were differently related to the criterion. Further studies may seek to use these as well as the lower level GWAs in similar scaling and item generation procedures.

References

Adler, S. (1996). Personality and work behaviour: Exploring the linkages. *Applied Psychology: An International Review, 45*, 207-224.

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

Barrick, M.R., & Mount, M.K. (1993). Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology, 78*, 111-118.

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

Berkley, M.H. (1953, July). *A comparison between the empirical and rational approaches for keying a heterogeneous test* (Research Bulletin 53-24). Lackland Air Force Base, Texas: 6570th Personnel Research Laboratory, Human Resources Research Center.

Borman, W.C., & Motowidlo, S.J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W.C. Borman (Eds.), *Personnel selection in organizations*. San Francisco: Jossey-Bass

Botwin, M.D. (1995). Review of the Revised NEO Personality Inventory. In J.C. Conuley & J.C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 862-863). Lincoln, NB: University of Nebraska Press.

Broach, D., & Manning, C.A. (1994). *Validity of the air traffic control specialist nonradar screen as a predictor of performance in radar-based air traffic control training* (report No. FAA-AM-94-9). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Brown, S.H. (1994). Validating biodata. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 197-236). Palo Alto, CA: Consulting Psychologists Press.

Buss, D.M. (1991). Evolutionary personality psychology. *Annual Review of Psychology, 42*, 459-491.

Buss, D.M. (1996). Social adaptation and the five major factors of personality. In J.S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180-207). New York: Guilford Press.

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

Digman, J.M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*, 116-123.

Cobb, B.B. (1962). *Problems in air traffic management: II. Prediction of success in air traffic controller school* (Civil Aeromedical Research Institute Report No. 62-2). Washington, DC: Federal Aviation Administration.

Cobb, B.B. (1967). *Problems in air traffic management: V. Identification and potential of aptitude test measures for selection of tower air traffic controller trainees* (Report No. FAA-AM-67-1). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Costa, P.T. Jr., & McCrae, R.R. (1990). *NEO PI-R Professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Cucina, J.M., & Vasilopoulos, N.L. (2003). *Nonlinear personality-performance relationships and the spurious moderating effects of traitedness*. Paper presented at the 18th annual meeting of the Society for Industrial and Organizational Psychology, Orlando, Florida.

Dean, M.A. (1999). *On biodata construct validity, criterion-related validity, and adverse impact*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge, LA.

Dean, M.A., Russell, C.J., & Muchinski, P.M. (1999). Life experiences and performance prediction: Toward a theory of biodata. In G. Ferris (Ed.). *Research in personnel and human resource management* (Vol 18, pp. 231-358). New York: JAI Press.

Della Rocco, P.S., Manning, C.A., & Wing, H. (1990). *Selection of controllers for automated systems: Applications from current research* (Report No. AM-90-13). Washington DC: Federal Aviation Administration, Office of Aviation Medicine.

Devlin, S.E., Abrahams, N.M., & Edwards, J.E. (1992). Empirical keying of biographical data: Cross-validity as a function of scaling procedures and sample size. *Military Psychology, 4*, 119-136.

England, G.W. (1971). *Development and use of weighted application blanks* (Rev. ed.). Minneapolis: University of Minnesota, Industrial Relations Center.

Gandy, J.A., Dye, D.A., & MacLane, C.N. (1994). Federal government selection: The Individual Achievement Record. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information*

in selection and performance prediction (pp. 109-146). Palo Alto, CA: Consulting Psychologists Press.

Gangestad, S.W., & Simpson, J.A. (1990). Toward an evolutionary history of female sociosexual variation. *Journal of Personality, 00*, 00-00.

Hogan, J.B. (1994). Empirical keying of background data measures. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 69-108). Palo Alto, CA: Consulting Psychologists Press.

Hough, L., & Paullin, C. (1994). Construct-oriented scale construction: The rational approach. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 109-146). Palo Alto, CA: Consulting Psychologists Press.

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

Jeanneret, P.R., Borman, W.C, Kubisiak, U.C., & Hanson, M.A. (1999). Generalized work activities. In N.G. Peterson, M.D. Mumford, W.C. Borman, P.R. Jeanneret, & E.A. Fleishman (Eds), *An occupational information system for the 21st century: The development of O*NET* (pp. 105-145). Washington, DC: American Psychological Association.

Kluger, A.N., Reilly, R.R., & Russell, C.J. (1991). Faking biodata tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology, 76*, 889-896.

Mael, F.A. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology, 44*, 763-792.

Manley, G. G., Halbesleben, J., & Mumford, M. D. (2002, April), *Differences in Biodata and Personality Tests in Measuring the Five-Factor Model*. Paper presented at the 17th annual meeting of the Society for Industrial and Organizational Psychology, Toronto, Canada.

Manning, C.A., Del Rocco, P.S., & Bryant, K. (1989). *Prediction of success in air traffic control field training as a function of selection and screening performance* (Report No. DOT-FAA-AM-89-6). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Manning, C.A., Kegg, P.S., & Collins, W.E. (1988). *Studies of post-strike air traffic control specialist trainees: II. Selection and screening* (Report No. FAA-AM-88-3). Washington, DC: Department of Transportation/Federal Aviation Administration.

McCormick, E.J., Jeanneret, P.R., & Mecham, R.C. (1969a). The development and background of the position analysis questionnaire (PAQ) (Tech. Rep. No. 5). Lafayette, IN: Purdue University, Occupational Research Center.

McCrae, R. R., & Costa, P. T., Jr. (1999) A five-factor theory of personality. In L. A. Pervin and O. P. John (Eds.). *Handbook of personality: Theory and research.* (2nd Ed., pp. 139-154). New York: Guilford.

Mitchell, T.W., & Klimoski, R.J. (1982). Is it rational to be empirical? A test of methods for scoring biographical data. *Journal of Applied Psychology, 67*, 411-418.

Motowidlo, S.J., & Van Scotter, J.R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology, 79*, 475-480.

Mumford, M.D., & Owens, W.A. (1987). Methodological review: Principles, procedures, and findings in the application of background data measures. *Applied Psychological Measurement, 11*, 1-31.

Mumford, M.D., Stokes, G.S., & Owens, W.A. (1990). *Patterns of life adaptation: The ecology of human individuality.* Hillsdale, NJ: Erlbaum.

Mumford, M.D., Wesley, S.S., & Shaffer, G.S. (1987). Individuality in a developmental context II: The crystallization of developmental trajectories. *Human Development, 30*, 291-321.

Mumford, M.D., Costantza, D.P., Connelly, M.S., & Johnson, J.F. (1996). Item generation procedures and background data scales: Implications for construct and criterion-related validity. *Personnel Psychology, 49*, 361-398.

Nikolaou, I., & Robertson, I.T. (2001). The five-factor model of personality and work behavior in Greece. *European Journal of Work and Organization Psychology, 10*, 161-186.

Organ, D.W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology, 48*, 775-802.

Owens, W.A. (1968). Toward one discipline of scientific psychology. *American Psychologist, 23*, 782-785.

Owens, W.A. (1971). A quasi-actuarial basis for individual assessment. *American Psychologist, 26*, 992-999.

Owens, W.A. (1976). Background data. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 609-644). Chicago: Rand McNally.

Owens, W.A., & Schoenfeldt, L.F. (1979). Towards a classification of persons. *Journal of Applied Psychology, 53*, 569-607.

Pannone, R.D. (1984). Predicting test performance: A content valid approach to screening applicants. *Personnel Psychology, 64*, 563-607.

Plomin, R., & Nesselroade, J.R. (1990). Behavioral genetics and personality change. *Journal of Personality, 58*, 191-220.

Reiter-Palmon, R., & Connelly, M.S. (2000). Item selection counts: A comparison of empirical key and rational scale validities in theory-based and non-theory-based item pools. *Journal of Applied Psychology, 85*, 143-151.

Robertson, I.T., and Callinan, M. (1998). Personality and work behavior. *European Journal of Work and Organization Psychology, 7*, 321-340.

Rogers, C.R. (1951). *Client-centered therapy: Its current practice, implications, and theory*. Boston: Houghton Mifflin.

Russell, C.J. (1994). Generation procedures for biodata items: A point of departure. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 17-38). Palo Alto, CA: Consulting Psychologists Press.

Russell, C.J., & Domm, D.R. (1990). *On the construct validity of biographical information: Evaluation of a theory based method of item generation.* Unpublished manuscript, Department of Industrial Relations and human Resources, Rutgers, New Brunswick, NJ.

Russell, C.J., Mattson, J., Devlin, S.E., & Atwater, D. (1990). Predictive validity of biodata items generated from retrospective life experience essays. *Journal of Applied Psychology, 75*, 569-580.

Rust, J. (1999). Discriminant validity of the big five personality traits in employment settings. *Social Behavior and Personality, 27*, 99-108.

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

Schoenfeldt, L.F. (1989). *Biographical data as the new frontier in employee selection research*. Paper presented at the annual meeting of the American Psychological Association, New Orleans.

Snyder, M., Simpson, J.A., & Gangestad, S.W. (1986). Personality and sexual relations. *Journal of Personality and Social Psychology*, 51, 181-190.

Society for Industrial and Organizational Psychology (1887). *Principles for the validation and use of personnel selection procedures*. (3rd Ed.). College Park, MD: University of Maryland Press.

Stokes, G.S., Mumford, M.D., & Owens, W.A. (1994). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: Consulting Psychologists Press.

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

Trites, D.K. (1961). *Problems in air traffic management: I. Longitudinal prediction of effectiveness in air traffic controllers* (Civil Aeromedical Research Institute report No. 61-1). Washington, DC: Federal Aviation Administration.

Vasilopoulos, N.L., Cucina, J.M., & Goldenberg, R. (2002). *Personality and training proficiency: Issues of Bandwidth-Fidelity, validity, and curvilinearity*. Paper presented at the 17th annual meeting of the Society for Industrial and Organizational Psychology, Toronto, Canada.

Zuckerman, M. (1990). The psychophysiology of sensation seeking. *Journal of Personality, 58*, 313-345.

Table 2.1. Sample Demographics

Variable	Total sample (N = 10014 to 11405)	ABA + Screen sample (n = 4559)	Screen + ABA + NEO sample (n = 183)
Ave. age	25.91	25.75	26.06
% male	82.3	83.8	66.1
Minority: % Native American	.6	.5	.5
% Asian	1.4	1.1	2.2
% Black	3.4	2.9	6.6
% Hispanic	3.0	2.5	5.5
% White	89.1	90.4	82.5
% from "comfortable" home life	49.8	50.6	45.9
% from suburban home life	33.4	34.2	37.7
Education: % finished high school only	11.0	11.1	12.6
% with some college	55.7	55.6	61.7
% finished college	31.9	31.8	24.0
% advanced degree	1.2	1.3	1.6
% with no prior experience in field	74.5	77.4	73.2
Academy: % passed	54.6	59.8	57.4
% failed	33.4	40.2	42.6
% withdrawn	12.0	———	———
ave. NLCOMP score	71.66	71.13	70.05

* Note: Only those who completed the Screen were included in the subset analyses.

Figure 2.1. Examples of ABA / IAR Biodata Items

1. My high school teachers would most likely describe my self-discipline as:
   a. superior
   b. above average
   c. average
   d. below average
   e. don't know

2. The number of high school clubs and organized activities (such as band, sports, newspaper, etc.) in which I participated was:
   a. 4 or more
   b. 3
   c. 2
   d. 1
   e. didn't participate

3. My grade point average in my college major was:
   a. I did not go to college or I went less than 2 years
   b. less than 2.90
   c. 2.90 - 3.19
   d. 3.20 - 3.49
   e. 3.5 or higher

4. In the past three years, the number of different paying jobs I have held for more than two weeks is:
   a. 7 or more
   b. 5 – 6
   c. 3 – 4
   d. 1 – 2
   e. none

5. My previous supervisors (or teachers if not previously employed) would most likely describe my problem-solving skills as:
   a. superior
   b. above average
   c. average
   d. below average
   e. don't know

Table 2.2 Descriptive statistics of NEO sub-scales / scales (n = 183)

	Mean	Std. Deviation	Alpha
ANXIETY	12.7158	4.53389	.71
HOSTILITY	9.1366	4.24496	.72
DEPRESSION	9.0820	4.28374	.73
SELF-CONSCIOUSNESS	12.6011	4.08016	.66
IMPULSIVENESS	14.1311	4.28827	.69
VULERNABILITY	6.9071	3.35567	.79
WARMTH	23.8743	4.04243	.80
GREGARIOUSNESS	17.4317	4.15952	.61
ASSERTIVENESS	18.8962	4.33807	.70
ACTIVITY	19.4918	3.91653	.63
EXCITEMENT-SEEKING	20.9290	4.29674	.62
POSITIVE EMOTIONS	22.1475	3.94538	.69
FANTASY	17.3989	4.74971	.76
AESTHETICS	16.1257	5.75035	.80
FEELINGS	21.1858	3.95836	.68
ACTIONS	16.6284	3.62435	.57
IDEAS	21.5137	4.28370	.75
VALUES	21.9836	3.74822	.64
CONSCIENTIOUSNESS	53.5410	8.36904	.87
AGREEABLENESS	49.2842	7.29640	.80
NUEROTICISM	64.5738	18.14059	.90
EXTROVERSION	122.7705	16.72998	.87
OPENNESS TO EXPERIENCE	114.8361	15.16595	.83

Table 2.3. Descriptive statistics for study variables (n = 4,559)

	Mean	Std. Deviation	alpha	# of items
NLCOMP (criterion composite)	71.1346	11.45943	---	---
information input (gwa1) scale	3.6789	.48080	.61	14
mental processes (gwa2) scale	3.6277	.39741	.86	36
work output (gwa3) scale	3.4107	.52828	.57	12
interactions w/ others (gwa4) scale	3.4033	.42253	.85	30
neuroticism (ffm1) scale	2.2450	.34651	.26	8
extraversion (ffm2) scale	3.3107	.48848	.81	16
openness (ffm3) scale	2.7227	.51977	.68	14
agreeableness (ffm4) scale	3.8307	.64653	.30	3
consciousness (ffm5) scale	3.4840	.43456	.91	55
matrix: extraversion by interactions w/ others	3.2432	.50321	.80	15
matrix: openness by information input	2.9937	.62063	.60	8
matrix: openness by interactions w/ others	2.8488	.54350	.43	6
matrix: consciousness by information input	3.4276	.88108	.90	11
matrix: consciousness by mental processes	3.6494	.43653	.86	29
matrix: consciousness by work output	3.4892	.42518	.58	17
matrix: consciousness by interactions w/ others	4.0434	.50187	.56	7
PCA: extraversion-oral communication / persuasiveness	3.7440	.56194	.71	7
PCA: extraversion-leadership experience	3.0358	.70451	.79	5
PCA: openness-college experience / science orientation	2.4932	.87484	.70	4
PCA: conscientiousness-college academic success	2.8551	.93086	.90	8
PCA: conscientiousness-work orientation	4.2333	.51856	.83	8
PCA: conscientiousness-high school academic success	3.4600	.58743	.76	8
PCA: conscientiousness-written communication / comprehension	2.3299	.75290	.83	4

(Note: descriptive statistics a function of linearly weighted rational scales)

Table 2.4. Sample item stems for rational scales

Rational biodata scale	Sample ABA items
information input (gwa1) scale	• Prior to accepting my present job, I last attended college as a full-time student: • Prior to accepting my present job, the number of different federal agencies I worked for was:
mental processes (gwa2) scale	• My previous supervisor (or teachers if not previously employed) would most likely describe my skill at thinking on my feet as: • My previous supervisor (or teachers if not previously employed) would most likely describe the speed at which I work as:
work output (gwa3) scale	• The age at which I first started to earn money (other than an allowance) was: • The amount of time I have been out of work between jobs has been:
interactions w/ others (gwa4) scale	• Relative to the other high school students in my major field of study, my classmates would most likely describe my interpersonal skills as: • Relative to the other high school students in my major field of study, my classmates would most likely describe my leadership skills as:
neuroticism (ffm1) scale	• My previous supervisor (or teachers if not previously employed) would most likely describe my self-control as: • During my years in high school, I was singled out for discipline reasons:
extraversion (ffm2) scale	• My high school classmates would most likely describe the amount of my participation in extracurricular activities as: • The number of elected offices I held in high school was:
openness (ffm3) scale	• The number of times I changed my college major before I selected the one in which I graduated was: • The number of times I elected non-required college science courses was:
agreeableness (ffm4) scale	• In the three years prior to accepting my present job, the number of formal suggestions I have submitted to my former employer(s) was: • My previous supervisor (or teachers if not previously employed) would most likely describe my skill at getting along with others as:
consciousness (ffm5) scale	• Relative to the other high school students in my major field of study, my most demanding teacher would most likely describe my academic work as: • My previous supervisor (or teachers if not previously employed) would most likely describe my dependability as:
matrix: extraversion by interactions w/ others	• My high school classmates would most likely describe my leadership in extracurricular activities as: • My previous supervisor (or teachers if not previously employed) would most likely describe my oral communication skills as:
matrix: openness by information input	• The number of times I elected non-required college English courses was: • The number of times I elected non-required college math courses was:
matrix: openness by interactions w/ others	• The number of high school clubs and organized activities (such as band, newspaper, etc.) in which I participated was: • My previous supervisor (or teachers if not previously employed) would most likely describe my responsiveness to other persons' viewpoints as:
matrix: consciousness by information input	• My final year in high school, I was absent: • My highest education level is:
matrix: consciousness by mental processes	• The number of college courses in which I received a failing grade was: • The number of national scholastic honor societies I belonged to in college was:

matrix: consciousness by work output	• During my last year in college, my average number of hours of paid employment per week was: • The proportion of my college expenses that I earned was:
matrix: consciousness by interactions w/ others	• Prior to accepting my present job, I was asked to serve as supervisor in my boss's absence: • Prior to accepting my present job, I was chosen to serve on special task forces or committees at work:
PCA: extraversion-oral communication / persuasiveness	• My previous supervisor (or teachers if not previously employed) would most likely describe my oral communication skills as: • My peers would likely rate my skill in influencing people to my point of view as:
PCA: extraversion-leadership experience	• The number of years of leadership experience I have had (such as work supervisor, commissioned or non-commissioned officer, scout patrol leader, school or social club president, athletic captain, etc.) is: • The number of elected offices I held in high school was:
PCA: openness-college experience / science orientation	• Prior to accepting my present job, I last attended college as a full-time student: • The number of times I elected non-required college science courses was:
PCA: conscientiousness-college academic success	• The college grade I most often received was: • My class standing in college put me in the: top:
PCA: conscientiousness-work orientation	• My previous supervisor (or teachers if not previously employed) would most likely describe my planning and organizing skills as: • My previous supervisor (or teachers if not previously employed) would most likely describe my dependability as:
PCA: conscientiousness-high school academic success	• The high school grade I most often received was: • My class standing in high school put me in the top:
PCA: conscientiousness-written communication / comprehension	• My previous supervisor (or teachers if not previously employed) would most likely rate my writing skills as: • My previous supervisor (or teachers if not previously employed) would most likely rate my speed of reading skill as:

Table 2.5. Convergent / discriminant validity correlations between FFM variables

	NEO neuroticism	NEO extraversion	NEO openness	NEO agreeableness	NEO conscientious -ness
Biodata-ABA neuroticism	.492 (.000)	-.159 (.032)	-.013 (.861)	-.243 (.001)	-.284 (.000)
Biodata-ABA extraversion	-.287 (.000)	.477 (.000)	.097 (.193)	.135 (.069)	.202 (.006)
Biodata-ABA openness	-.111 (.134)	.069 (.355)	.260 (.000)	.199 (.006)	.107 (.151)
Biodata-ABA agreeableness	-.181 (.014)	.311 (.000)	.089 (.233)	.160 (.031)	.058 (.432)
Biodata-ABA conscientiousness	-.272 (.000)	.170 (.021)	.166 (.024)	.163 (.027)	.300 (.000)

(P-values for 2-tailed test in parentheses. N = 183.)

Table 2.6. Fit statistics for confirmatory factor analyses of biodata FFM scales and subscales (N = 183)

Goodness-of-Fit Statistic	Matrix FFM * GWA model	FFM principle components model
GFI	.88	.91
AGFI	.81	.84
RMR	.08	.09
Chi Square (p-value)	<.0001	<.0001
RMSEA	.08	.09
AIC	26.7	28.0

[This page intentionally left blank.]

Figure 2.2. Latent variable model of biodata constructs of 3 Big Five factors and 4 O*NET general work dimensions.

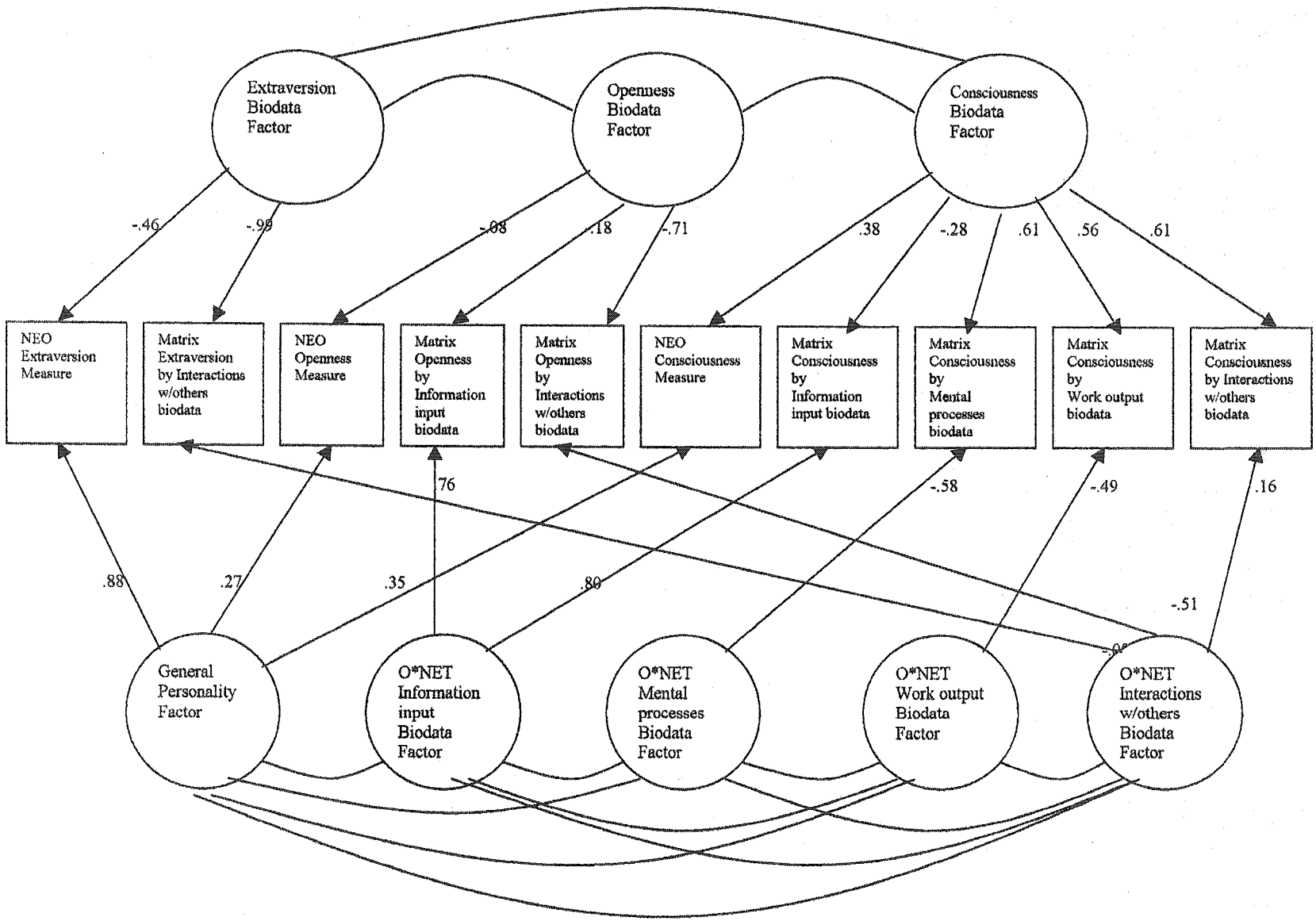


Figure 2.3. Latent variable model of biodata constructs of Big Five factors and principal component sub-factors.

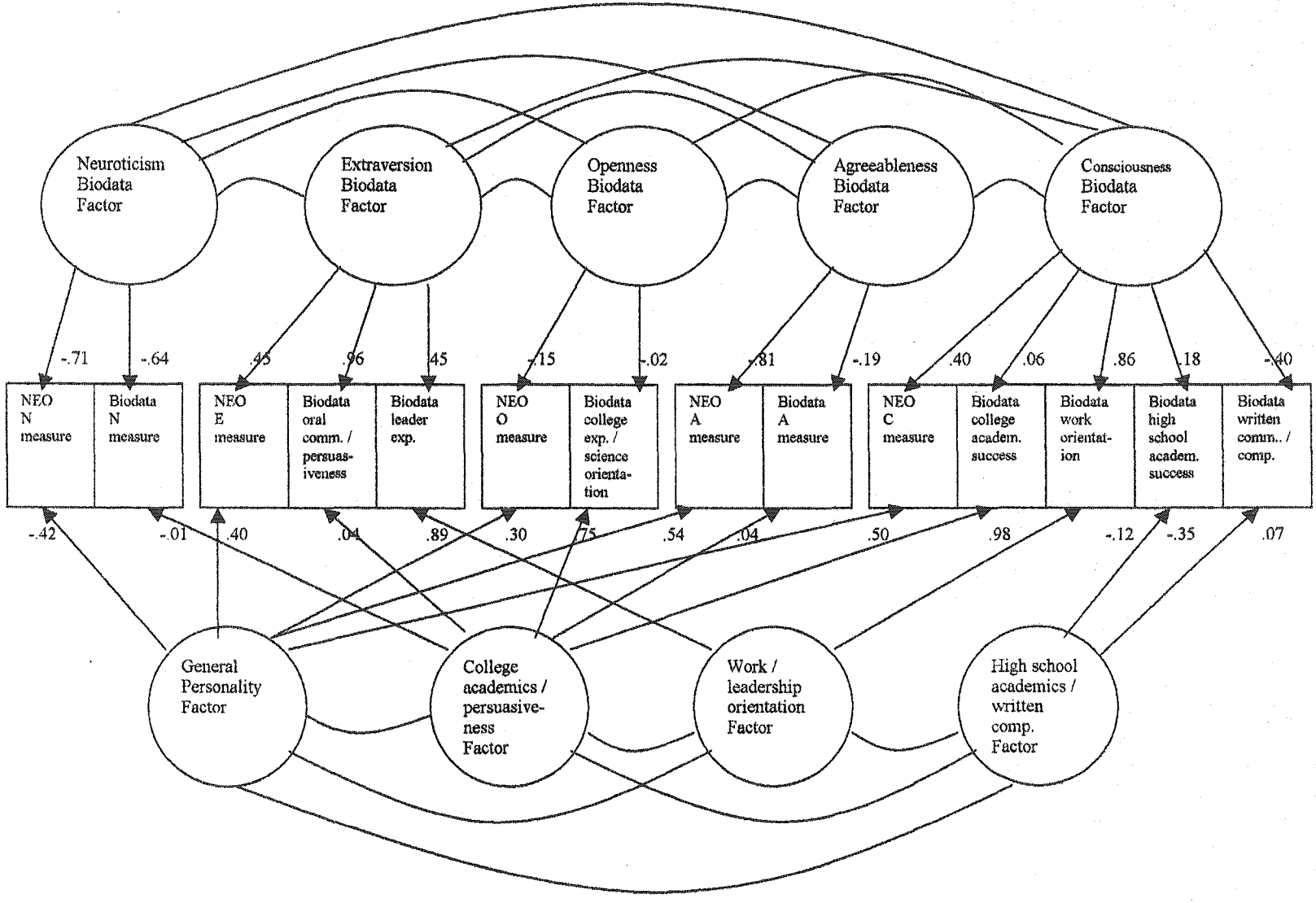


Table 3.1. Correlation matrix of main study variables (N = 4,559)

(Note: correlations > (+/-).05 significant at the 0.01 level (2-tailed). Values a function of linearly weighted rational scales. All correlations uncorrected for indirect range restriction)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. criterion composite NLCOMP																							
2. information input (gwa1)	-0.07																						
3. mental processes (gwa2)	0.10	-0.32																					
4. work output (gwa3)	0.02	-0.18	0.41																				
5. interactions w/ others (gwa4)	-0.01	-0.18	0.55	0.36																			
6. neuroticism (ffm1)	0.04	0.09	-0.35	-0.15	-0.40																		
7. extraversion (ffm2)	-0.02	-0.17	0.50	0.31	0.95	-0.34																	
8. openness (ffm3)	-0.03	-0.09	0.49	0.37	0.30	-0.10	0.24																
9. agreeableness (ffm4)	-0.04	-0.05	0.37	0.27	0.55	-0.26	0.48	0.08															
10. consciousness (ffm5)	0.10	-0.54	0.88	0.60	0.49	-0.26	0.42	0.63	0.25														
11. matrix: extravert-interact w/ others	-0.03	-0.17	0.48	0.30	0.95	-0.33	1.00	0.24	0.46	0.41													
12. matrix: openness-information input	-0.02	0.00	0.42	0.38	0.14	-0.03	0.11	0.87	0.02	0.57	0.11												
13. matrix: openness-interact w/ others	-0.01	-0.20	0.48	0.29	0.74	-0.22	0.62	0.54	0.31	0.48	0.62	0.30											
14. matrix: conscious-information input	-0.04	0.72	-0.50	-0.34	-0.20	0.10	-0.17	-0.61	-0.02	-0.79	-0.17	-0.55	-0.34										
15. matrix: conscious-mental processes	0.12	-0.32	0.98	0.37	0.51	-0.31	0.46	0.51	0.26	0.88	0.44	0.44	0.46	-0.52									
16. matrix: conscious-work output	0.03	-0.21	0.52	0.95	0.39	-0.19	0.34	0.45	0.27	0.69	0.33	0.44	0.33	-0.42	0.48								
17. matrix: conscious-interact w/ others	0.07	-0.09	0.37	0.34	0.54	-0.24	0.41	-0.06	0.42	0.33	0.39	-0.07	0.20	0.00	0.31	0.34							
18. PCA: extravert-oral comm./persuasive	0.01	-0.14	0.47	0.28	0.81	-0.34	0.84	0.06	0.48	0.35	0.81	0.01	0.41	-0.06	0.41	0.30	0.52						
19. PCA: extravert-leadership experience	-0.03	-0.11	0.35	0.19	0.73	-0.22	0.78	0.24	0.29	0.30	0.79	0.10	0.57	-0.14	0.34	0.21	0.20	0.43					
20. PCA: openness-college exp./science	0.01	-0.26	0.46	0.47	0.20	-0.07	0.17	0.78	0.05	0.64	0.17	0.84	0.33	-0.61	0.47	0.51	0.00	0.08	0.14				
21. PCA: conscientious-college success	0.04	-0.57	0.58	0.50	0.23	-0.10	0.20	0.70	0.05	0.82	0.20	0.65	0.37	-0.85	0.60	0.57	0.01	0.08	0.16	0.73			
22. PCA: conscientious-work orientation	0.08	-0.09	0.66	0.26	0.54	-0.37	0.47	0.06	0.41	0.46	0.43	0.03	0.28	-0.07	0.60	0.33	0.60	0.59	0.24	0.06	0.08		
23. PCA: conscientious-H.S. success	0.17	-0.15	0.64	0.15	0.26	-0.14	0.24	0.32	0.06	0.57	0.24	0.25	0.30	-0.32	0.69	0.21	0.10	0.13	0.28	0.23	0.32	0.18	
24. PCA: conscientious-written comm.	0.03	0.14	-0.62	-0.12	-0.37	0.20	-0.35	-0.14	-0.24	-0.44	-0.34	-0.13	-0.25	0.14	-0.62	-0.17	-0.24	-0.39	-0.19	-0.12	-0.17	-0.42	-0.23

Table 3.2. Criterion-Related Validities (w/ NLCOMP) for Various Scaling Methods (N = 4,559)

Scale / Scaling method ==>	rational linear	rational non-linear	empirical key - dev.	empirical key - cross-val.	empirical key - shrinkage
Full ABA (142 items)	----	----	0.363*	0.293*	0.070*
GWA1: information input	-0.069*	-0.068*	0.160*	0.131*	0.029
GWA2: mental processes	0.102*	-0.136*	0.266*	0.241*	0.025
GWA3: work output	0.020	-0.026	0.141*	0.069*	0.072*
GWA4: interactions w/others	-0.014	0.022	0.206*	0.131*	0.075*
FFM1: neuroticism	0.042	0.001	0.104*	0.024	0.079*
FFM2: extraversion	-0.022	0.039	0.173*	0.124*	0.049*
FFM3: openness to experience	-0.035	-0.022	0.156*	0.069*	0.086*
FFM4: agreeableness	-0.040	0.055*	0.068*	0.058*	0.010
FFM5: consciousness	0.098*	-0.144*	0.286*	0.250*	0.036
matrix: extraversion by interactions w/ others	-0.030	0.051*	0.142*	0.096*	0.046*
matrix: openness by information input	-0.019	-0.019	0.132*	0.058*	0.074*
matrix: openness by interactions w/ others	-0.014	-0.007	0.089*	0.008	0.081*
matrix: consciousness by information input	-0.037	-0.059*	0.143*	0.119*	0.025
matrix: consciousness by mental processes	0.115*	-0.150*	0.267*	0.244*	0.022
matrix: consciousness by work output	0.030	-0.044	0.154*	0.075*	0.078*
matrix: consciousness by interactions w/ others	0.073*	-0.071*	0.122*	0.083*	0.039
PCA: extraversion-oral communication / persuasiveness	0.006	0.008	0.145*	0.109*	0.036
PCA: extraversion-leadership experience	-0.027	0.033	0.071*	0.054*	0.017
PCA: openness-college experience / science orientation	0.008	-0.014	0.084*	0.028	0.056*
PCA: conscientiousness-college academic success	0.036	-0.057*	0.142*	0.110*	0.033
PCA: conscientiousness-work orientation	0.076*	-0.088*	0.118*	0.101*	0.017
PCA: conscientiousness-high school academic success	0.174*	-0.191*	0.223*	0.218*	0.005
PCA: conscientiousness-written comm. / comp.	0.031	0.004	0.070*	0.006	0.064*

(Note: values with * are significant at .01 level [2-tailed]. All correlations uncorrected for indirect range restriction. Shaded rows are the four predictive scales discussed in text)

Table 3.3. Results of Stepwise Hierarchical Quadratic Multiple Regression (QMR) Analysis (N = 4,559)

	Step 1 – Linear term		Step 2 – Linear + Quadratic terms			
Scale	R	F-test	R	Beta-linear	Beta-quadratic	ΔR
GWA1 information input	.070*	22.08*	.074*	-.338	.270	.004
GWA2 mental processes	.102*	47.85*	.105*	1.176	.279	.003
GWA3 work output	.020	1.80	.021	.071	-.051	.001
GWA4 interactions w/others	.014	.88	.018	.111	-.125	.003
FFM1 neuroticism	.042	8.00*	.043	.126	-.085	.001
FFM2 extraversion	.022	2025	.022	.005	-.027	.000
FFM3 openness	.035	5.58	.052*	-2.89*	2.62*	.017
FFM4 agreeableness	.040	7.20*	.040	-.065	.025	.000
FFM5 conscientiousness	.098*	43.77*	.105*	-.386	.485*	.007
matrix: extraversion by interactions w/ others	.030	4.02	.030	-.051	.021	.000
matrix: openness by information input	.019	1.34	.050	-.351*	.335*	.031
matrix: openness by interactions w/ others	.014	.850	.014	-.047	.033	.000
matrix: consciousness by information input	.037	6.18	.091*	-.622*	.591*	.054*
matrix: consciousness by mental processes	.115*	61.52*	.120*	-.180	.297	.005
matrix: consciousness by work output	.031	4.24	.032	-.080	.111	.001
matrix: consciousness by interactions w/others	.073*	24.44*	.073*	.085	-.012	.000
PCA: extraversion:- oral comm. / persuasiveness	.006	.183	.028	.268	-.263	.022
PCA: extraversion-leadership experience	.026	3.25	.032	-.142	.118	.006
PCA: openness-college exp./science orient.	.008	.270	.012	-.034	.043	.004
PCA: conscientiousness-college success	.036	5.88	.081*	-.337*	.380*	.045*
PCA: conscientiousness-work orientation	.077*	26.70*	.080*	-.060	.138	.003
PCA: conscientiousness-H.S. success	.174*	142.96*	.182*	-.310*	.486*	.008
PCA: conscientiousness-written comm. / comp.	.031	4.28	.033	.078	-.049	.002

(Note: values with * are significant at .01 level [2-tailed]. Shaded rows are the four predictive scales discussed in text)

Table 3.4. Results of Stepwise Hierarchical QMR Item Analysis for Scaling Stability (N = 4,559)

	Step 1 – Linear term		Step 2 – Linear + Quadratic terms			
Scale	R	F-test	R	Beta-linear	Beta-quadratic	ΔR
FFM5 conscientiousness scale items:	.098*	43.77*	.105*	-.386	.485*	.007
ABA25	.057*	14.93*	.070*	-.248*	.195*	.013
ABA 30	.062*	17.15*	.109*	-.507*	.454*	.047*
ABA 31	.056*	14.31	.114*	-.540*	.494*	.058*
ABA R33	.031	4.33	.097*	-560*	.537*	.066*
ABA 68	.046*	9.54	.081*	-.280*	.243*	.035
matrix: consciousness by mental processes scale items:	.115*	61.52*	.120*	-.180*	.297*	.005
ABA 8	.096*	42.47*	.135*	-.414*	.331*	.039
ABA 41	.081*	29.72*	.166*	-.850*	.783*	.085*
ABA 42	.036	5.98*	.084*	-.479*	.449*	.048*
ABA 84	.093*	39.48*	.107*	-.258*	.173*	.014
ABA 102	.132*	80.28*	.171*	-.515*	.398*	.039
PCA: conscientiousness-H.S. academic success scale items:	.174*	142.96*	.182*	-.310*	.486*	.008
ABA 3	.104*	49.36*	.119*	-.330*	.234*	.015
ABA 16	.128*	75.78*	.140*	-.346*	.225*	.012
ABA 17	.155*	111.41*	.183*	-.480*	.340*	.028
ABA 20	.243*	285.32*	.254*	-.520*	.286*	.011
ABA 21	.125	72.25	.137*	-.321*	.204*	.012

(Note: values with * are significant at .01 level [2-tailed]. See Appendix A for full inventory of ABA biodata items.)

Appendix A:

Office of Personnel Management

APPLICANT BACKGROUND ASSESSMENT

APPLICANT BACKGROUND ASSESSMENT

Please answer all the questions on this biographical questionnaire to the best of your ability. **Your answers, which will be used for research purposes only and remain confidential,** will assist the Federal Aviation Administration Civil Aeromedical Institute (CAMI) in its longitudinal study of the Air Traffic selection process.

ALL the questions, which follow, are in a multiple-choice format. Answer each one by blackening the oval in the appropriate column of your choice. Choose the response that best fits you and **MAKE ONLY ONE RESPONSE PER QUESTION.**

ACADEMIC EXPERIENCE: HIGH SCHOOL

1. During *high school* (grades 9-12) I made the semester honor roll:
   - O never
   - O once or twice
   - O three or four times
   - O five or six times
   - O seven or eight times

2. When I graduated from *high school* I was:
   - O 16 years old or younger
   - O 17 years old
   - O 18 years old
   - O 19 years old
   - O 20 years old or older

3. Relative to the other *high school* students in my major field of study, my most demanding teacher would most likely describe my *academic work* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

4. During my last year in *high school*, my average number of hours of *paid employment* per week was:
   - O more than 20
   - O 16 to 20 hours
   - O 10 to 15 hours
   - O fewer than 10 hours
   - O none

5. Relative to the other *high school* students in my major field of study, my classmates would most likely describe my *interpersonal skills* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

Social Security Number

			-			-				

Example: 999 - 99 - 9999

51787

6. Relative to the other *high school* students in my major field of study, my classmates would most likely describe my *leadership skills* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

7. My *high school* teachers would most likely describe my *self discipline* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

8. My *high school* teachers would most likely describe my *academic potential* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

9. My *high school* classmates would most likely describe the amount of my *participation* in extracurricular activities as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

10. My *high school* classmates would most likely describe my *leadership* in extracurricular activities as:
    - O superior
    - O above average
    - O average
    - O below average
    - O don't know

11. The number of different *high school* sports I participated in was:
    - O 4 or more
    - O 3
    - O 2
    - O 1
    - O didn't play sports

12. The number of letters I received in *high school* sports was:
    - O 4 or more
    - O 3
    - O 2
    - O 1
    - O 0

13. The number of *high school* clubs and organized activities (such as band, newspaper, etc.) in which I participated was:
    - O 4 or more
    - O 3
    - O 2
    - O 1
    - O *didn't participate*

14. My final year in *high school*, I was absent:
    - O more than 15 days
    - O 10 to 14 days
    - O 5 to 9 days
    - O fewer than five days
    - O never

15. During my years in *high school*, I was singled out for disciplinary reasons:
    - O 5 or more times
    - O 3 or 4 times
    - O twice
    - O once
    - O never

16. My class standing in *high school* put me in the:
    - O top 10%
    - O top 33%
    - O top 50%
    - O top 90%
    - O did not graduate from high school

17. The *high school* grade I most often received was:
    - O A
    - O B
    - O C
    - O D or lower
    - O don't remember

18. The number of *high school* courses which I failed was:
    - O 5 or more
    - O 3 or 4
    - O 2
    - O 1
    - O none

19. The *high school English* grade I most often received was:
    - O A
    - O B
    - O C
    - O D or lower
    - O don't remember or didn't take English

20. The *high school math* grade I most often received was:
    - O A
    - O B
    - O C
    - O D or lower
    - O don't remember or didn't take math

21. The *high school science* grade I most often received was:
    - O A
    - O B
    - O C
    - O D or lower
    - O don't remember or didn't take science

22. The *high school* subject in which I received my lowest grades was:
    - O science
    - O math
    - O English
    - O history/social sciences
    - O physical education

23. The number of elected offices I held in *high school* was:
    - O 5 or more
    - O 3 to 4
    - O 2
    - O 1
    - O none

ACADEMIC EXPERIENCE: UNDERGRADUATE COLLEGE

24. My highest education level is:
    - O no college
    - O 1 to 2 years of college or associate degree
    - O 3 to 4 years of college, no degree
    - O Bachelor's degree
    - O advanced degree

25. During *college* the number of times I made the *Dean's List* was:
    - O 5 or more times
    - O 3 to 4 times
    - O 1 to 2 times
    - O never
    - O didn't go to college

26. Prior to accepting my present job, I last attended college as a full-time student:
    - O did not attend college
    - O less than a year prior to accepting my first job in my present series
    - O one year prior to accepting my first job in my present series
    - O 2 to 3 years prior to accepting my first job in my present series
    - O over 3 years prior to accepting my first job in my present series

27. During my last year in *college*, my average number of hours of paid employment per week was:
- O more than 20 hours
- O 10 to 20 hours
- O fewer than 10 hours
- O none
- O didn't go to college

28. The number of different *undergraduate colleges* I attended prior to graduation was:
- O 4 or more
- O 3
- O 2
- O didn't change colleges
- O didn't go to college

29. The number of times I changed my *college* major before I selected the one in which I graduated was:
- O 3 times or more
- O 2 times
- O 1 time
- O didn't change majors
- O didn't go to college

30. My class standing in *college* put me in the:
- O top 10%
- O top 33%
- O top 50%
- O bottom 50%
- O didn't go to college

31. The *college* grade I most often received was:
- O A
- O B
- O C
- O D or lower
- O didn't go to college

32. On a 4 point scale where A=4, my grade point average the *first two years of college* was:
- O I did not go to college or went less than two years
- O less than 2.90
- O 2.90 to 3.19
- O 3.20 to 3.49
- O 3.50 or higher

33. My grade point average *after the first two years of college* was:
- O I did not go to college or went less than two years
- O less than 2.90
- O 2.90 to 3.19
- O 3.20 to 3.49
- O 3.50 or higher

34. My grade point average in my *college major* was:
- I did not go to college or went less than two years
- less than 2.90
- 2.90 to 3.19
- 3.20 to 3.49
- 3.50 or higher

35. My *overall grade point average in college* was:
- I did not go to college or went less than two years
- less than 2.90
- 2.90 to 3.19
- 3.20 to 3.49
- 3.50 or higher

36. Of the following, the *college* subject in which I received my lowest grades was:
- science
- English
- math
- history/political science
- didn't go to college

37. The number of *college* courses in which I received a failing grade was:
- 3 or more
- 2
- 1
- none
- didn't go to college

38. At the time I applied for my present job, my *undergraduate* education consisted of having completed:
- less than 30 semester hours (45 quarter hours)
- 30 to 59 semester hours (45 to 89 quarter hours)
- 60 to 90 semester hours (90 to 134 quarter hours)
- more than 90 semester hours (135 quarter hours) but no degree
- Bachelor's Degree

39. At the time I applied for my present job, my *graduate* education consisted of having completed:
- 0 to 5 graduate semester hours (0 to 8 quarter hours)
- 6 to 11 graduate semester hours (9 to 17 quarter hours)
- 12 to 23 graduate semester hours (18 to 35 quarter hours)
- 24 graduate semester hours or more (36 quarter hours)
- Master's Degree, Ph.D. Degree, or other graduate degree

40. The *college English* grade I most often received was:
- A
- B
- C
- D or lower
- didn't take English or didn't go to college

41. The *college math* grade I most often received was:
- A
- B
- C
- D or lower
- didn't take math or didn't go to college

42. The *college science* grade I most often received was:
- A
- B
- C
- D or lower
- didn't take science or didn't go to college

43. The number of times I elected non-required *college English* courses was:
- 3 or more
- 2
- 1
- never
- didn't go to college

44. The number of times I elected non-required *college math* courses was:
- 3 or more
- 2
- 1
- never
- didn't go to college

45. The number of times I elected non-required *college science* courses was:
- 3 or more
- 2
- 1
- never
- didn't go to college

46. The proportion of my *college* expenses that I earned was:
- more than 50%
- 25% to 50%
- some but less than 25%
- none
- didn't go to college

47. The amount of my *college* expenses covered by scholastic scholarships was:
- more than 50%
- 25% to 50%
- some but less than 25%
- none
- didn't go to college

48. The amount of my *college* expenses covered by athletic scholarships was:
- more than 50%
- 25% to 50%
- some but less than 25%
- none
- didn't go to college

49. I received my *college* degree:
- prior to obtaining full-time employment
- while working on a full-time job
- while in the armed forces
- didn't graduate
- didn't go to college

50. Prior to accepting my present job series, I had been out of college for:
   - O 5 or more years
   - O 3 to 4 years
   - O 1 to 2 years
   - O less than one year
   - O didn't go to college or didn't graduate

51. The number of *college* clubs and organized activities (band, newspaper, etc.) in which I participated was:
   - O 3 or more
   - O 2
   - O 1
   - O didn't participate
   - O didn't go to college

52. The number of letters I received in *college* sports was:
   - O 3 or more
   - O 2
   - O 1
   - O 0
   - O didn't go to college

53. The number of student offices to which I was elected in *college* was:
   - O 3 or more
   - O 2
   - O 1
   - O 0
   - O didn't go to college

54. The number of national scholastic honor societies I belonged to in *college* was:
   - O 3 or more
   - O 2
   - O 1
   - O 0
   - O didn't go to college

WORK EXPERIENCE

55. In the three years prior to accepting my present job, the number of different paying jobs I *held* for more than two weeks was:
   - O 7 or more
   - O 5 to 6
   - O 3 to 4
   - O 1 to 2
   - O none

56. In the three years immediately before accepting my present job, the number of different full or part-time jobs I *applied* for was:
   - O none
   - O 1 to 2
   - O 3 to 4
   - O 5 to 6
   - O 7 or more

57. Prior to accepting my present job, I had been employed in work similar to that of my present job for:
- O never employed in a similar job
- O less than 1 year
- O 1 to 2 years
- O 3 to 4 years
- O over 5 years

58. In the three years before accepting my present job, the number of promotions I received in all previous jobs was:
- O not employed
- O 0
- O 1
- O 2
- O 3 or more

59. I left my last full-time job because:
- O I was laid off or discharged
- O there was little chance for advancement or increase in pay
- O important personal reasons - such as moving or pregnancy
- O something else
- O have never had a full time job

60. Prior to accepting my present job, I worked on my last full-time job:
- O have not held full-time job
- O less than six months
- O 6 months up to a year
- O one to two years
- O more than two years

61. Prior to accepting my present job, the number of different federal agencies I worked for (not including military service) was:
- O 0
- O 1
- O 2
- O 3
- O 4 or more

62. I learned about the opportunity to apply for my present job through:
- O a public notice or media advertisement
- O a friend or relative
- O college recruitment
- O working in some other capacity for the agency
- O some other way

63. My military service was:
- O none
- O non-career enlisted
- O non-career officer
- O career enlisted
- O career officer

64. My employment status prior to accepting my present job was:
- O employed full-time
- O employed part-time
- O student, not employed
- O self-employed
- O unemployed

65. The number of months I was unemployed during the three years immediately before accepting my present job was:
- O 0
- O 1 to 2
- O 3 to 4
- O 5 to 6
- O 7 or more

66. Prior to accepting my present job, I worked extra hours during evenings and on weekends:
- O much more often than most persons in the job
- O somewhat more often than most persons in the job
- O about the same as most persons in the job
- O somewhat less often than most persons in the job
- O not employed prior to present job

67. In the three years immediately before accepting my present job, my work experience (military or civilian) was in:
- O professional or administrative occupations
- O clerical or sales occupations
- O service occupations
- O trades or labor occupations
- O not employed during the three years immediately before accepting my present job

68. On my last job (prior to accepting my present job), my supervisor rated me as:
- O outstanding
- O above average
- O average
- O below average
- O not employed or received no rating

69. Prior to accepting my present job, I was late (tardy) for work:
- O once or twice a year or less
- O once or twice in a six month period
- O once or twice a month
- O once or twice a week
- O not employed prior to present job

70. In the three years prior to accepting my present job, the number of formal awards I received for my job performance was:
- O not employed prior to present job
- O 0
- O 1
- O 2
- O 3 or more

71. The amount of time I have been out of work between jobs usually has been:
- never out of work
- less than one month
- 1 to 2 months
- 3 to 4 months
- 5 or more months

72. In the three years prior to accepting my present job, the number of formal suggestions I have submitted to my former employer(s) was:
- not employed prior to present job
- 0
- 1
- 2
- 3 or more

73. The age at which I first started to earn money (other than an allowance) was:
- less than 12 years old
- 12 to 13 years old
- 14 to 15 years old
- 16 to 17 years old
- 18 years or older

74. In the year before accepting my present job, the number of time I had been late for work (or class) was:
- more than 14 times
- 10 to 14 times
- 5 to 9 times
- fewer than five times
- never

75. In the three years prior to accepting my present job, the number of jobs I had been fired from was:
- 5 or more
- 3 to 4
- 2
- 1
- none

76. Prior to accepting my present job, I was asked to serve as supervisor in my boss' absence:
- somewhat more often than most
- about the same as most others
- somewhat less often than most
- much less than most
- not employed prior to present job

77. Prior to accepting my present job, I was selected to attend training:
- somewhat more often than most
- about the same as most others
- somewhat less often than most
- much less than most
- not employed prior to present job

78. Prior to accepting my present job, I was chosen to serve on special task forces or committees at work:
- O somewhat more often than most
- O about the same as most others
- O somewhat less often than most
- O much less than most
- O not employed prior to present job

SKILLS

79. The *number of civic organizations of special organizations* (which have regular meetings and a defined membership) that I ***belonged*** to prior to accepting my present job is:
- O none
- O 1
- O 2 or 3
- O 4 or 6
- O 7 or more

80. Which one of the following have you ever *organized or assisted* in organizing? ***If you organized more than one, mark the one most important to you.***
- O Athletic team or sport competition
- O Financial or charity campaign to raise funds
- O Literary, debating, choral, religious, or social club
- O Some other civic, social, work related, or professional organization
- O Have never organized or assisted in organizing any club or group

81. The number of *elective offices* (other than in high school or college organizations) I have held in the last five years is:
- O none
- O 1 or 2
- O 3 or 4
- O 5 or 6
- O 7 or more

82. In organizations to which I belong, my *participation* is best described as:
- O do not belong to any organizations
- O not very active
- O a regular member but not an office holder
- O have held at least one important office
- O have held several important offices

83. My previous supervisor (or teachers if not previously employed) would probably describe my *attendance record* as:
- O much worse than my peers
- O somewhat worse than my peers
- O about the same as my peers
- O somewhat better than my peers
- O much better than my peers

84. My previous supervisor (or teachers if not previously employed) would most likely describe my *problem solving skills* as:
- O superior
- O above average
- O average
- O below average
- O don't know

85. My previous supervisor (or teachers if not previously employed) would most likely describe my skill at *thinking on my feet* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

86. My previous supervisor (or teachers if not previously employed) would most likely describe the *amount of supervision* that I need as:
   - O more than average
   - O average
   - O less than average
   - O very little
   - O don't know

87. My previous supervisor (or teachers if not previously employed) would most likely describe my *dependability* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

88. My previous supervisor (or teachers if not previously employed) would most likely describe the *speed at which I work* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

89. My previous supervisor (or teachers if not previously employed) would most likely describe the amount of *time I needed to complete assignments* as:
   - O a great deal
   - O more than average
   - O average
   - O less than average
   - O don't know

90. My previous supervisor (or teachers if not previously employed) would most likely describe my skill at *meeting deadlines under pressure* as:
   - O superior
   - O above average
   - O average
   - O below average
   - O don't know

91. My previous supervisor (or teachers if not previously employed) would most likely describe me as *taking on more than I can handle*:
   - O most of the time
   - O a great deal of the time
   - O sometimes
   - O infrequently
   - O don't know

92. My previous supervisor (or teachers if not previously employed) would most likely describe *me* as *mastering my assignments*:
- O most of the time
- O a great deal of the time
- O sometimes
- O infrequently
- O don't know

93. My previous supervisor *(or teachers if not previously employed) would most likely describe my* supervisory potential as:
- O superior
- O above average
- O average
- O below average
- O don't know

94. My previous supervisor (or teachers if not previously employed) would most likely describe my skill at *getting along with others* as:
- O superior
- O above average
- O average
- O below average
- O don't know

95. My previous supervisor (or teachers if not previously employed) would likely describe my *oral communication* skills as:
- O superior
- O above average
- O average
- O below average
- O don't know

96. My previous supervisor (or teachers if not previously employed) would likely describe my *self-control* as:
- O superior
- O above average
- O average
- O below average
- O don't know

97. My previous supervisor (or teachers if not previously employed) would likely describe my *responsiveness to other person's viewpoints* as:
- O superior
- O above average
- O average
- O below average
- O don't know

98. My previous supervisor (or teachers if not previously employed) would most likely describe my skill at *speaking before a group* as:
- O superior
- O above average
- O average
- O below average
- O don't know

99. My previous supervisor (or teachers if not previously employed) would most likely describe my *logical reasoning* skills as:
- O superior
- O above average
- O average
- O below average
- O don't know

100. My previous supervisor (or teachers if not previously employed) would most likely describe my *planning and organizing* skills as:
- O superior
- O above average
- O average
- O below average
- O don't know

101. My previous supervisor (or teachers if not previously employed) would most likely describe my *analytical* skills as:
- O superior
- O above average
- O average
- O below average
- O don't know

102. My previous supervisor (or teachers if not previously employed) would most likely describe my basic *math* skills as:
- O superior
- O above average
- O average
- O below average
- O don't know

103. My previous supervisor (or teachers if not previously employed) would most likely describe my *vocabulary* as:
- O superior
- O above average
- O average
- O below average
- O don't know

104. My previous supervisor (or teachers if not previously employed) would most likely rate my *writing* skills as:
- O superior
- O above average
- O average
- O below average
- O don't know

105. My previous supervisor (or teachers if not previously employed) would most likely rate my speed of *reading* skill as:
- O superior
- O above average
- O average
- O below average
- O don't know

106. My previous supervisor (or teachers if not previously employed) would most likely rate my *reading comprehension* skill as:
- O superior
- O above average
- O average
- O below average
- O don't know

107. My previous supervisor (or teachers if not previously employed) would most likely rate my skill at *doing several different jobs at the same time* as:
- O superior
- O above average
- O average
- O below average
- O don't know

108. My previous supervisor (or teachers if not previously employed) would most likely describe my attention to *detail* as:
- O superior
- O above average
- O average
- O below average
- O don't know

109. My previous supervisor (or teachers if not previously employed) would most likely describe my ability to *recall facts and details of information* as:
- O superior
- O above average
- O average
- O below average
- O don't know

110. My previous supervisor (or teachers if not previously employed) would most likely describe my skill at *getting work done on time* as:
- O superior
- O above average
- O average
- O below average
- O don't know

111. The number of years of *leadership experience* I have had (such as work supervisor, commissioned or non-commissioned officer, scout patrol leader, school or social club president, athletic captain, etc.) is:
- O 5 or more years
- O 3 or 4 years
- O 2 years
- O 1 year
- O less than one year

112. In the past six months, the average number of hours per week I spent *reading* newspapers, books, magazines, etc. outside of work is:
- O 5 or more hours per week
- O 3 to 4 hours per week
- O 2 hours per week
- O 1 hour per week
- O less than 1 hour per week

113. My peers would likely rate my *interpersonal skills* as:
- O superior
- O above average
- O average
- O below average
- O don't know

114. On a list of 100 typical people in the kind of job I can do best, my peers would probably place me in the:
- O top 10%
- O top 25%
- O top 50%
- O top 75%
- O top 90%

115. In terms of punctuality, my peers would probably say that I usually arrive:
- O much later than most
- O later than most
- O on time
- O earlier than most
- O much earlier than most

116. If you were to ask my peers, they would probably say that the amount of recognition I receive relative to my accomplishments is:
- O a great deal less than deserved
- O somewhat less than deserved
- O as much as is deserved
- O somewhat more than deserved
- O much more than deserved

117. My peers would probably say that the highest level I could reach if I chose a career in a major corporation would be:
- O a top level executive (e.g. vice president)
- O a middle manager
- O a first level supervisor
- O a professional or technical expert
- O other non-supervisory technical or administrative position

118. My peers would probably describe me as a person who:
- O never takes chances
- O hardly ever takes chances
- O sometimes takes chances
- O often takes chances
- O very often takes chances

119. My peers would probably describe me as:
- O much more aggressive than most of my peers
- O somewhat more aggressive than most of my peers
- O about as aggressive as most of my peers
- O somewhat less aggressive than most of my peers
- O much less aggressive than most of my peers

120. My peers would probably say that getting me to change once I have made up my mind is:
- O much harder than most
- O somewhat harder than most
- O about the same as most
- O somewhat easier than most
- O much easier than most

121. Which of the following communication situations would your peers say you would handle best?
- writing a lengthy report
- giving a lecture or speech to a large group
- mixing and conversing with a room full of strangers
- discussing a topic with another individual
- don't know

122. Which of the following would your peers say describes your behavior in a group situation?
- you freely express your views, and sway the group considerably
- you freely express your views, but the group does not always share them
- you are reluctant to express your views, but when you do they are usually well received
- you usually don't express your views
- don't know

123. Which of the following would your peers say describes your behavior in a social situation?
- always at ease in a social situation
- almost always at ease in a social situation
- generally at ease in a social situation
- occasionally at ease in a social situation
- don't know

124. My peers would probably say that having someone criticize my performance (i.e. point out a mistake) bothers me:
- much less than most
- somewhat less than most
- about the same as most
- somewhat more than most
- much more than most

125. My peers would probably describe me as being:
- much more confident than most
- somewhat more confident than most
- about as confident as anyone else
- somewhat less confident than most
- much less confident than most

126. Which of the following would your peers consider your *weakest* trait?
- learning new things quickly
- composing effective written reports
- working with and getting along with other people
- speaking and expressing yourself effectively to others
- working well under pressure

127. Which of the following would your peers consider your *strongest* trait?
- learning new things quickly
- composing effective written reports
- working with and getting along with other people
- speaking and expressing yourself effectively to others
- working well under pressure

128. My peers would likely rate my skill in *influencing people* to my point of view as:
- superior
- above average
- average
- below average
- don't know

129. Compared to others in my unit, my rate of *promotion in the military* was:
- much faster than most
- somewhat faster than most
- about the same as most
- somewhat slower than most
- never served in the military

130. Compared to others on my last *full-time* job, my rate of *promotion* was:
- much faster than most
- somewhat faster than most
- about the same as most
- somewhat slower than most
- not employed full-time prior to present job

131. Prior to accepting my present job I:
- never worked for this agency
- worked part-time for this agency while in college
- worked for this agency during summer vacations while in college
- worked full time for this agency for a period of time but then resigned
- was employed full-time with agency immediately prior to accepting my present job

132. Before I joined the government, the information I had about the type of work that air traffic controllers are expected to do was:
- none
- practically no information
- some information
- quite a bit
- knew in considerable detail

133. Prior to accepting my present job, the amount of formal training that I had (other than college) related directly to my present job was:
- less than 6 months
- 6 months to a year
- 1 to 2 years
- 3 to 4 years
- 5 or more years

134. During my teens, I usually spent most of my summers (*choose one*):
- taking life easy
- attending summer school
- attending honors classes
- working part-time
- working full-time

135. Before accepting my present job, the length of time I had worked shift work was:
- never worked shift work
- less than 6 months
- 6 to 12 months
- 13 months to 2 years
- more than 2 years

136. The number of times in the past five years I was denied an award I deserved is:
- O never
- O once or twice
- O three or four times
- O five or six times
- O seven or more times

137. In the past year, I have been annoyed by my coworkers:
- O never
- O rarely
- O occasionally
- O frequently
- O constantly

138. Compared to my peers, I find myself leading others:
- O much less often than most
- O somewhat less often than most
- O about the same as most
- O somewhat more than most
- O much more often than most

139. Compared to my coworkers, people come to me for advice:
- O much more often than most
- O somewhat more often than most
- O about the same as most
- O somewhat less often than most
- O much less often than most

140. If I could have any full-time job I wanted, the reason I would pick the job which I would finally choose is that:
- O I would be recognized for the work I do
- O I would be with people I really like
- O I would have the freedom to be creative
- O I would have great possibilities for monetary rewards
- O I could do the kind of work that I find very interesting

141. When I think about being an air traffic controller, the first thing that turns me off most about the job is that:
- O achieving anything of significance might be difficult
- O doing the same things over and over might be boring
- O lacking control over my work activities would be frustrating
- O having little prestige as a controller would be unsatisfying
- O working under constant pressure could be very hard

142. The aspect of being an air traffic controller that appeals to me most is that:
- O my job is secure in the future
- O I'm responsible for the safety of many others
- O I'll receive a good salary, which will grow
- O I'll be constantly challenged to resolve situation, which arise
- O the work will always be interesting

Appendix B: ABA Items Within Study Rational Scales

GWA: information input	GWA: mental processes	GWA: work output	GWA: interactions w/ others	FFM: neuroticism	FFM: extraversion	FFM: openness	FFM: agreeableness	FFM: conscientiousness	Matrix: extraversion by interactions w/others	Matrix: openness by info input	Matrix: openness by interactions w/others	Matrix: conscientiousness by info input	Matrix: conscientiousness by mental processes	Matrix: conscientiousness by work output	Matrix: conscientiousness by interactions w/others	PCA: extraversion-oral comm / persuasivness	PCA: extraversion-leadership experience	PCA: openness-college exp./ science orientation	PCA: conscientiousness-college academic success	PCA: conscientiousness-work orientation	PCA: conscientiousness-high school academic success	PCA: conscientiousness-written comm. / comp.
33	3	27	5	96	5	13	94	3	5	43	11	14	3	27	76	85	5	43	30	84	3	103
34	7	46	6	107	6	29	113	7	6	44	13	24	7	28	78	95	6	44	31	87	16	104
35	8	50	9	119	9	43	72r	8	9	45	51	32	8	46	87	98	9	45	40	88	17	105
57	16	52	10	124	10	44		16	10	57	97	33	16	50	93	111	10	26r	42	90	18	106
61	19	65	11	136	23	45		17	23	112	139	34	19	52	117	125	23		46	92	19	
63	20	66	13	137	53	51		19	53	132	79r	35	20	65	60r	128			47	93	20	
131	21	68	23	15r	85	57		20	95	26r		38	21	66	115r	138r			54	99	21	
132	28	71	51	116r	95	97		21	98	56r		39	37	68					24r	100	1r	
133	37	73	53		98	132		25	111			25r	40	71								
25r	40	130	76		111	135		27	122			30r	41	73								
26r	41	58r	78		122	26r		28	123			31r	42	110								
30r	42	129r	87		123	56r		30	125				47	114								
31r	47		93		125	79r		31	128				54	130								
56r	54		94		128	120r		37	82r				84	58r								
	84		95		82r			40	138r				88	70r								
	85		97		138r			41					90	83r								
	88		98					42					92	129r								
	90		111					46					99									
	91		113					47					100									
	92		117					50					101									
	96		119					52					102									
	99		122					54					103									
	100		123					66					104									
	101		124					68					105									
	102		125					71					106									
	103		128					73					108									
	104		139					76					109									
	105		79r					77					1r									
	106		82r					78					89r									
	107		138r					84														
	108							87														
	109							88														
	1r							90														
	72r							92														
	86r							93														
	89r							99														
								100														
								101														
								102														
								103														
								104														
								105														
								106														
								1r														
								18r														
								24r														
								32r														
								33r														
								34r														
								35r														
								38r														
								58r														
								83r														
								86r														
								89r														

U.S. Department
of Transportation
**Federal Aviation
Administration**

800 Independence Ave., S.W.
Washington, D.C. 20591

February 22, 2002

Greg Manley
Department of Psychology
University of Oklahoma
Norman, OK 73019

Dear Mr. Manley,

In support of agency research objectives on alternative selection measures for the air traffic control specialist (FV-2152) occupation, you are granted permission to use archival biographical, demographic, cognitive aptitude test, and training performance data and measures in your dissertation on construct validity of existing biodata inventories. The data are provided for research purposes only, and may not be used for any commercial purpose. You agree to acknowledge the FAA as the source for your research data, and provide a bound copy of your doctoral dissertation to the FAA.

Edna Fiedler, PhD

Edna Fiedler, Ph.D.
Manager, Training and Organizational Research Laboratory
Civil Aerospace Medical Institute

The University of Oklahoma

OFFICE OF RESEARCH ADMINISTRATION

March 18, 2002

Mr. Gregory G. Manley
1538 Pecan Ave.
Norman, OK 73072

Dear Mr. Manley:

Your research application, "Biodata Scaling and Prediction: a Question of Linearity," has been reviewed according to the policies of the Institutional Review Board chaired by Dr. E. Laurette Taylor, and found to be exempt from the requirements for full board review. Your project is approved under the regulations of the University of Oklahoma - Norman Campus Policies and Procedures for the Protection of Human Subjects in Research Activities.

Should you wish to deviate from the described protocol or the research is to extend beyond 12 months, you must notify this office, in writing, noting any changes or revisions in the protocol and/or informed consent document, and obtain prior approval or request an extension of this ruling. A copy of the approved informed consent document is attached.

Should you have any questions, please contact me at irb@ou.edu.

Sincerely,

Susan Wyatt Sedwick

Susan Wyatt Sedwick, Ph.D.
Director of the Office of Research Administration and
Administrative Officer for the
Institutional Review Board – Norman Campus (MPA #1146)

SWS:lk
FY2002-310

cc: Dr. E. Laurette Taylor, Chair, Institutional Review Board
Dr. Michael D. Mumford, Psychology

Appendix E

AUTHOR'S BIOGRAPHY

Gregory G. Manley is an assistant professor at the University of Texas at San Antonio in the Department of Psychology. He received his MS in 1999 and PhD in 2003 in the area of industrial-organizational psychology at the University of Oklahoma. Dr. Manley teaches undergraduate and graduate courses in I/O psychology, psychological measurements, research methods, and introductory psychology. His research interests include background data and personality measures for employment selection, leader assessment, business ethics, and multivariate modeling. Dr. Manley has published his research in the *Journal of Education for Business* and has presented research at the annual meetings of the Society for Industrial and Organizational Psychology, Industrial Organizational – Organizational Behavior, and the Southern Management Association. Prior to graduate studies, he received his BS in psychology with honors at Montana State University in 1996 and served in the US Navy as a missile systems analyst aboard the aircraft carrier USS Constellation (CV-64) from 1984 to 1990.

Dr. Manley has a lovely wife, Pamela, two boys, Adrian and Jordan, and a stepdaughter, Cami, with whom he enjoys spending much of his free time. When not teaching and researching, he enjoys mountain biking, skiing, motorcycling, reading, and visiting family.