


Summer 2003

# Identifying the Dimensions of Integrity: A Confirmatory and Discriminant Validity Analysis

Arlene Pace Green  
*Old Dominion University*

Follow this and additional works at: [https://digitalcommons.odu.edu/psychology\\_etds](https://digitalcommons.odu.edu/psychology_etds)

 Part of the [Industrial and Organizational Psychology Commons](#), and the [Personality and Social Contexts Commons](#)

---

## Recommended Citation

Green, Arlene P.. "Identifying the Dimensions of Integrity: A Confirmatory and Discriminant Validity Analysis" (2003). Doctor of Philosophy (PhD), dissertation, Psychology, Old Dominion University, DOI: 10.25777/zkfh-1q64  
[https://digitalcommons.odu.edu/psychology\\_etds/149](https://digitalcommons.odu.edu/psychology_etds/149)

This Dissertation is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

IDENTIFYING THE DIMENSIONS OF INTEGRITY:  
A CONFIRMATORY AND DISCRIMINANT VALIDITY ANALYSIS

By

Arlene Pace Green  
M.S. December 1996, Old Dominion University  
B.A. May 1994, James Madison University

A Dissertation submitted to the Faculty of  
Old Dominion University in Partial Fulfillment of the  
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

INDUSTRIAL/ ORGANIZATIONAL PSYCHOLOGY

OLD DOMINION UNIVERSITY  
August 2003

Approved By:

\_\_\_\_\_  
Robert M. McIntyre (Director)

\_\_\_\_\_  
Donald D. Davis (Member)

\_\_\_\_\_  
Terry L. Dickinson (Member)

\_\_\_\_\_  
Rosalyn P. Pitts (Member)

## ABSTRACT

### IDENTIFYING THE DIMENSIONS OF INTEGRITY: A CONFIRMATORY AND DISCRIMINANT VALIDITY ANALYSIS

Arlene Pace Green  
Old Dominion University, 2003  
Director: Dr. Robert M. McIntyre

While Integrity tests have demonstrated significant predictive and concurrent validity, the meaning and structure of integrity test scores are not well understood. The purpose of the present investigation was to empirically verify the results of a previous study that used an inductive method to define integrity and identify its constituent dimensions (Green, 1999). Specifically, the present investigation used item analysis, confirmatory factor analysis, discriminant validity analysis, and an analysis of social desirability to test the validity of the five integrity dimensions identified by Green (1999): Concern for Others, Conscientiousness, Emotional Control, Fairness, and Honesty. Results confirmed that Integrity acts as a second-order factor with multiple first-order dimensions. Four of the hypothesized first-order dimensions were confirmed in the study: Concern for Others, Conscientiousness, Emotional Control, and Honesty. The inadequacy of the Fairness measures made it impossible to test the relationship of this fifth dimension to the Integrity construct. A discriminant validity analysis failed to support the Integrity dimensions by indicating that Anticipated Tenure was significantly related to the Integrity construct. Also investigated was the influence of social desirability. Results indicated that social desirability influenced, but did not destroy the

factor structure of the Integrity construct. Future research into the semantic realm of integrity is suggested.

To Baby Lauren:

**“All things good and perfect come from God.”**

## ACKNOWLEDGMENTS

First I would like to acknowledge the friendship and guidance of my director, Robert M. McIntyre. Dr. McIntyre, your genuine spirit, thoughtful ideas, and willingness to help in any (and every!) way possible were invaluable and this endeavor would not have been possible without you. I would also like to thank my committee members, Dr. Don Davis and Dr. Terry Dickinson, for their most helpful thoughts and suggestions on earlier drafts – I am truly appreciative of your time and your words of encouragement along the way. I would also like to acknowledge Dr. Rosalyn Pitts for her insight on the dissertation process and uplifting phone calls at just the right time. I must also thank our department's administrative staff, Jackie, Mary, Peggy, and Dr. Justice, for assisting me at every turn in navigating this process – I would have been lost without you. I'd like to also thank my parents, sisters (Anne, Rosetta, and Shirley), brother-in-laws, nieces, nephews, and all of my loving family and friends. The “you can do its!” and “so, when will you finish?” comments, the jokes and laughs, prayers, and words of inspiration along the way will always be treasured. And finally, I'd like to thank my husband, Kelvin, for everything – supporting and encouraging me, never letting me lose my sense of humor, and always always, believing.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
INTRODUCTION .....	1
Personality and Selection .....	9
Integrity .....	13
SUMMARY .....	50
Questions Pertaining to Defining Integrity .....	51
Questions Pertaining to Integrity and Other Performance Indicators .....	51
Questions Pertaining to Understanding Integrity Instruments .....	51
Present Investigation .....	52
EXPERIMENT 1 .....	61
Method .....	61
Results .....	63
EXPERIMENT 2 .....	68
Method .....	68
Analytic Strategy and Results .....	72
CONCLUSIONS .....	104
REFERENCES .....	107
APPENDIXES .....	121
A. Critical Incident Interview .....	122
B. Concern for Others Item Selection Survey .....	125
C. Item Selection Survey Items .....	128
D. Mean Relevance Ratings for Concern for Others Items on the ISS .....	132
E. Mean Relevance Ratings for Conscientiousness Items on the ISS .....	134
F. Mean Relevance Ratings for Emotional Control Items on the ISS .....	136
G. Mean Relevance Ratings for Fairness Items on the ISS .....	138
H. Mean Relevance Ratings for Honesty Items on the ISS .....	140
E. Procedure for Creating Subscales for Maximum Likelihood Estimation In Structural Equation Modeling .....	142
VITA .....	143

## LIST OF TABLES

Table	Page
1. A Summary of Various Published Integrity Tests and Their Dimensions.....	4
2. Overt Integrity Test Factors.....	15
3. Personality-based Integrity Test Factors.....	17
4. Frequency, Percentage, and Mean goodness Ratings of Content Analysis Categories.....	20
5. Two Examples of Multi Coded Critical Incidents.....	22
6. Content Analysis Categories.....	24
7. Combination of Individual Content Categories into Five Larger Categories of Integrity.....	28
8. Correlations between Integrity Test Scores and Personality Variables.....	37
9. Correlations between Integrity Test Scores and the Big Five Dimensions of Personality.....	39
10. Items Comprising the Integrity Dimensions' Indicators.....	73
11. Correlation Matrix, Means, and Standard Deviations for the Integrity Dimension Indicators, Discriminant Construct Indicators, and Social Desirability Scale Score...79	79
12. Integrity Hypothesis Model Variants.....	82
13. Factor Correlations for the Integrity Hypothesis Measurement Model.....	88
14. Fit Statistics for the Anticipated Tenure and Social Desirability Hypotheses.....	96
15. Correlation Matrix for the Integrity Indicators with Social Desirability Partialled Out.....	99
16. Standardized Parameter Estimates for the Partialled and Non-Partialled Four Factor Models.....	100
17. Factor Correlations for Four-Factor Model with Social Desirability Partialled Out..	102



## LIST OF FIGURES

Figure	Page
1. Integrity Hypothesis Model Variants.....	76
2. Measurement Model Estimates.....	84
3. Hypothesized Model of Integrity Model Estimates.....	85
4. One Factor Model Estimates.....	86
5. Four-Factor Model Estimates.....	92

INTRODUCTION<sup>1</sup>

Personality testing has had an interesting history in Industrial and Organizational Psychology (Ghiselli & Barthol, 1953; Ghiselli, 1973; Guion & Gottier, 1965; Reilly & Chao, 1982). After examining 12 years of research pertaining to the validity of personality tests, Guion (1965) concluded "...In view of the problems, both technical and moral, one must question the wisdom and morality of using personality tests as instruments of decision in employment procedures (379)." Just 30 years later, other psychologists would review current literature and conclude that "well constructed measures of normal personality are solid predictors of performance in virtually all occupations...and (represent one) way to promote social justice and increase organizational productivity" (Hogan, Hogan, and Roberts, 1996, p. 1). The latter statement best represents the current view of personality testing in selection.

Research has demonstrated that personality tests are predictive of overall job performance (Ones, 1993; Sackett & Wanek, 1996), contextual performance (Hogan, Hogan, & Busch, 1984; Motowidlo & Van Scotter, 1994; Rosse, Miller, & Barnes, 1991), leadership (Digman, 1990; Hogan, Curphy, & Hogan, 1994), customer service (Colarelli & Dean, 1987; Williams & Sanchez, 1997), and counterproductive work behaviors (Jones & Terris, 1983; Ones, Viswesvaran, & Schmidt, 1993). Personality may be most important to selection for its ability to predict typical (i.e., 'will do') as opposed to maximal (i.e., 'can do') performance. Cronbach (1960) described typical performance as motivation-related and maximal performance as ability-related. Borman, White, Pulakos, and Oppler (1991) suggested that while maximal performance is a function of ability and

---

<sup>1</sup>The journal model format used is the *Journal of Applied Psychology*.

job knowledge, typical performance is a function of ability, job knowledge, and personality. In a test of this assertion, Driskell, Hogan, Salas, and Hoskin (1994) investigated military trainee success. The authors reported that personality was indeed more predictive of motivational performance (i.e., number of rule infractions) than academic performance or cognitive ability (as assessed by the Armed Services Vocational Aptitude Battery). As this study demonstrated, understanding both aspects of performance is critical in predicting employee success. A promising line of research in predicting the motivational aspects of performance pertains to integrity testing.

Integrity tests are paper and pencil, self-report assessments. They were initially designed to assess employee or applicant honesty (Murphy, 1993). Some of these measures rely on direct admissions. Others utilize responses to attitude or personality items to predict the likelihood that individuals will commit dishonest behaviors in the future. Since their inception, integrity tests have expanded to measure a wide variety of motivational constructs (e.g., conscientiousness, dependability, reliability) and predict a broad range of criteria (e.g., absenteeism, turnover, safety orientation, and overall performance).

Research supports the notion that integrity tests as a whole are predictive of work performance (Jones & Terris, 1983; Ones et al., 1993; Sackett & Wanek, 1996). However, only recently (Ones & Viswesvaran, 2001; Viswesvaran, 2002) has research begun to explore which dimensions of integrity (e.g., dependability, conscientiousness) are capable of predicting an individual criterion (e.g., turnover, absenteeism). That is, as a whole, we have limited knowledge to determine if one or all dimensions are necessary in predicting a single criterion (Sackett, 2002). A critical stumbling block in the

alignment of integrity dimensions and performance criteria is that the dimensions of integrity are unclear.

The test developer's definition of integrity varies significantly across individual tests and research investigations. For example, while some studies used only the honesty scales when investigating integrity (Ones, 1993; Ones et al., 1993), others used additional scales such as reliability, violence, or drug abuse (Schmidt, Viswesvaran, & Ones, 1997; Wanek, 1995). Furthermore, although different integrity tests include different dimensions (See Table 1), many tests provide the user with a single composite integrity score (O'Bannon, Goldinger, and Appleby, 1989). However, because of the differences in constituent dimensions, composite scores from different tests can have widely different theoretical meanings. While most integrity tests do focus on the prediction of counterproductive behaviors, there appears to be no accepted definition of integrity nor its underlying dimensions. As an example, O'Bannon et al. (1989) report that integrity tests tap a wide variety of dimensions including theft, opportunism, leniency, impulse control, reliable tendencies, tenure and energy level. In addition, the APA Task Force Report on integrity testing concluded that the integrity domain was "ill defined and heterogeneous" (Goldberg, Grenier, Guion, Sechrest, & Wing, 1991, p. 21).

Since this report, researchers have used exploratory and confirmatory factor analyses to delineate the dimensions of integrity (Cunningham & Ash, 1988; Harris, 1987; Hogan & Brinkmeyer, 1997; Ones, 1993; Paajanen, Hansen, & McLellan, 1993; Wanek, 1995). However, because these analyses relied on different tests or subsets of tests (which include different dimensions of integrity), research has yet to converge on a single, accepted definition of integrity and its constituent dimensions. Depending upon

Table 1

*A Summary of Various Published Integrity Tests and their Dimensions*

Test	Honesty	Theft	Substance Abuse	Violence	Socially Desirable Responding	Tenure	Overall Index	Other
Accutrac Evaluation System	X		X			X		Consistency, Mental ability
Applicant Review	X				X			Emotional stability
Employment Attitude Screening Evaluation II		X	X			X		
Employee Attitude Inventory		X	X		X		X	Job burnout, Job dissatisfaction
Employee Reliability Inventory	X		X			X		Conscientious job performance
Employee Survey		X	X					Other work violations
Hogan Reliability Inventory								Reliability
Inventory Shrinkage Evaluation		X					X	Shrinkage, Paperwork errors
Loss Prevention Analysis	X	X	X					Job satisfaction, Quality of supervision, Equity, Moral reasoning, Personal financial security, Company controls
Milby Profile	X		X	X	X	X		On-the-job performance

Table 1 Continued

Test	Honesty	Theft	Substance Abuse	Violence	Socially Desirable Responding	Tenure	Overall Index	Other
Organizational Review							X	Overall organizational effectiveness
Orion Survey			X	X	X	X		Acceptance of authority, Work attitudes, Advancement potential
Personnel Decisions Incorporated Employment Inventory						X	X	Productive behavior
Pre-employment Opinion Survey							X	Risk potential
Survey A.D.T. (Alienation, Drug & Alcohol, & Trustworthiness)	X		X				X	Alienated attitudes
Survey L.T. (Lenient Attitudes & Trustworthiness Attitudes)	X							Lenient attitudes

Table 1 Continued

Test	Honesty	Theft	Substance Abuse	Violence	Socially Desirable Responding	Tenure	Overall Index	Other
Pearson's Behavioral Construct	X							Accountability, Concern, Criticism reaction, Defensive proclivities, Empathy, Initiative, Optimism, Smoke blower tendencies, Assertiveness, Conformity, Decisiveness, Determination, Extroversion, Inquisitiveness, Organization, Rebelliousness, Concentration, Creativity, Dependability, Diplomacy, Follow-through, Meticulousness, Physical stamina, Sociability
Personal Outlook Inventory		X			X		X	
Personnel Reaction Blank							X	Dependability-conscientiousness personality factor
Personnel Selection Inventory		X	X	X	X	X	X	Employee/ customer relations, Emotional stability, Safety, Work values, Supervision attitudes
Phase II Profile	X	X			X		X	Good attitudes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1 Continued

Test	Honesty	Theft	Substance Abuse	Violence	Socially Desirable Responding	Tenure	Overall Index	Other
Reid Report	X		X				X	Social behavior, Personal achievements, Service orientation, Clerical/ math skills
Reid Survey III		X	X				X	Job satisfaction, Perceptions of company policies, Perception of supervisors, Average perception of work environment
Security Aptitude Fitness Evaluation-Resistance	X		X		X			Dependability, Socialization, Numerical ability, Reading ability
Situational Perceptions - Observations Test	X		X				X	
Sentry Survey	X				X		X	
Stanton Case Review		X					X	
Stanton Inventory	X		X					Admissions of past antisocial behavior, Information regarding other employees
Stanton Profile	X							Work motivation, Adaptability/ flexibility, Service orientation



Table 1 Continued

Test	Honesty	Theft	Substance Abuse	Violence	Socially Desirable Responding	Tenure	Overall Index	Other
Stanton Survey	X						X	
Station Employee Applicant Inventory	X		X		X	X	X	Interpersonal cooperation, Applied arithmetic, Job-specific skills and attitudes, Safety-consciousness
Station Manager Applicant Inventory	X		X		X		X	Interpersonal cooperation, Managerial arithmetic, Understanding of organization policies/ practices, Temperament
Tescor Survey		X	X	X	X			
True Test	X							
Trustworthiness Attitude Survey	X				X			
Veracity Analysis Questionnaire		X	X		X			Work attitude, Work history, Fundamental data
Wilkerson Employee Input Survey		X	X		X			Job satisfaction
Wilkerson Pre-Employment Audit		X	X		X			

the analysis, researchers have reported anywhere from four (Cunningham & Ash, 1988; Hogan & Brinkmeyer, 1997) to nineteen (Wanek, 1995) dimensions with only minimal overlap across investigations. Furthermore, the early reluctance of test publishers to release their tests for research (Camara & Schneider, 1995; Lilienfeld, 1994; Sackett & Wanek, 1996) and the large number of integrity tests available (O'Bannon et al., 1989) constrained the generalizability of research findings.

As such, the purpose of the present investigation was to overcome previous limitations by using an inductive approach to define integrity and identify its constituent dimensions. Specifically, the goals of the present study were to answer the following questions:

1. What is integrity as perceived by a diverse group of employees?
2. What are the dimensions underlying the integrity construct?

### Personality and Selection

Current research indicates that personality is important for predicting and understanding performance at work. In describing personality, many researchers rely on a five-factor model of personality, commonly known as "The Big Five". While each factor has been named and renamed a variety of times, the labels given by Norman (1963) are most commonly used in the literature (Digman, 1990). They are Extraversion, Emotional Stability, Agreeableness, Conscientiousness, and Openness to Experience. The Big Five were first discovered by Tupes and Christal (1992) when they used factor analysis to analyze personality data sets from Cattell (1947), Fiske (1949), and primary

data sets of their own (the research by Tupes and Christal was initially published in 1961 in a technical report, but more widely published in the *Journal of Personality* in 1992). The researchers concluded that five factors described the structure of personality in all of the data sets with substantial agreement (Tupes & Christal, 1992). The Big Five has since been replicated by means of a lexical analysis, in which participants' adjective descriptions of self and others were categorized, yielding a five-factor solution (Goldberg, 1990). The five-factor solution has also been replicated across languages (Bond, Nakazato, & Shiraishi, 1975; Borkeu & Ostendorf, 1990), instruments (Costa & McCrae, 1988; Lorr & Youniss, 1974; McCrae & Costa, 1985; McCrae & Costa, 1987), cultures (Noller, Law, & Comrey, 1987; Yoon, Schmidt, & Ilies, 2002), and populations (Digman & Takemoto-Chock, 1981; McCrae & John, 1992; Digman, 1990). The following section provides a brief definition of each factor.

Extraversion (Dimension I): This factor is commonly accepted as a representation of Eysenck's Extraversion/Introversion Dimension (Digman, 1990). Individuals on the high end of this scale are labeled extroverts and are described as assertive, talkative, active, gregarious, expressive, and sociable (Digman, 1990). Individuals on the low end of this scale are labeled introverts and are described as reserved, cautious, silent, and retiring (Ones, 1993).

Agreeableness (Dimension II): This factor has also been labeled Likeability, Friendliness, Social Conformity, Compliance versus Hostile Non-compliance, and Love (Digman, 1990). Traits associated with this high end of this dimension include courteous, flexible, trusting, good-natured, cooperative, forgiving, soft-hearted, and

tolerant (Barrick & Mount, 1991). On the low end, individuals are described as spiteful, obstructive, jealous, and suspicious (Ones, 1993).

Conscientiousness (Dimension III): This factor has also been labeled Conscience, Conformity, Dependability, Will to Achieve, Will, and Work. Traits associated with the high end of this scale include careful, thorough, responsible, organized, planful, hardworking, achievement-oriented, and persevering. On the low end, individuals are described as thrill-seeking, undependable, and frivolous (Barrick & Mount, 1991; Digman, 1990; Ones, 1993).

Emotional Stability (Dimension IV): This factor has also been named Stability, Emotionality, and Neuroticism. It is commonly defined from the low end of the scale with individuals low on Emotional Stability described as anxious, depressed, angry, emotional, worried, and insecure (Digman, 1990).

Openness to Experience (Dimension V): This factor has also been labeled Intellectance, Intellect, and Culture (Digman, 1990). Individuals on the high end of this scale are described as being imaginative, cultured, curious, original, broad-minded, intelligent, and artistically sensitive (Barrick & Mount, 1991). On the low end, individuals are described as conventional, practical, boorish, and awkward (Ones, 1993).

Not all researchers subscribe to the five factors of personality. For example, Eysenck identified three factors, Guilford identified thirteen, and Cattell identified sixteen (McCrae, 1989). In addition, there are some who support slight variations on the Big 5. For example, Hogan (1983) and Brand (1984) subscribe to six dimensions, with the primary difference being the split of Extraversion into two dimensions: Sociability and Activity (Hogan, 1983; Brand, 1984). Also, some researchers have suggested the

existence of a weak sixth factor labeled Culture or Social Class (Digman & Takemoto-Chock, 1981), while others have suggested that all factors beyond the Big Five are "error factors" (Cattell & Digman, 1962). Nonetheless, in general, most empirical research supports the five-factor model of personality (Digman, 1990).

### *The Big Five and Job Performance*

Digman (1990), based on the work of Goldberg (1981), Hogan (1983), Brand (1984), Digman (1988, as cited in Digman, 1990), and John (1989), compared various definitions of the five factors (See Digman (1990) for a full treatment of the factor comparisons). As Digman asserts, a diverse lexicon is associated with each construct. For example, the Agreeableness dimension (Dimension II) has been described with such varied words as psychoticism, likeability, love, and cortertia . This is to be expected as the Big Five theory was developed as an overarching, organizing framework for personality research. In describing the Big Five theory, Digman and Inouye (1986) concluded that there is "the intriguing implication that all linguistic conceptions of personality...may be found within this five-fold space" (p. 122). And indeed, other researchers agree (Goldberg, 1990).

Early studies on selection focused on how well each dimension of the five-factor model predicted overall job performance (Barrick & Mount, 1991; Barrick & Mount, 1993). Current research, however, focuses on only those personality constructs that can be logically linked to performance through job analysis (Hogan et al., 1996). For example, job analysis may indicate that extroversion is critical for the job of a sales

associate, but unrelated to the performance of a switchboard operator. As such, a study investigating the effectiveness of switchboard operators would not measure extroversion as a predictor of success.

Researchers have also suggested that validity studies use more specific aspects of personality and performance (Hogan et al., 1996; Hogan et al., 1994, Sackett, 2002). Instead of using a Big Five dimension to predict overall performance, researchers suggest using a facet of a Big Five dimension to predict a specific aspect of performance. In their investigation, Chidester, Helmreich, Gregorich, and Geis (1991) found that self-confidence was predictive of the number of errors made during flight crew training (a specific aspect of performance). Creating a specific link between a precisely defined predictor and criterion is believed to increase the validity of the personality dimension in predicting performance (Hogan et al., 1996). As a result, researchers are now investigating more specific aspects of personality (Tett, Jackson, & Rothstein, 1991), with integrity being one of the most promising in predicting performance.

### Integrity

Much of the research on integrity was spawned by the realization that integrity test usage was rapidly expanding in the workplace (Goldberg et al., 1991; Sackett & Harris, 1984; U.S. OTA, 1990). Instead of a theoretical definition guiding test development, it appears that the tests were developed first and researchers have had to play "catch up" in determining what these tests measure. For example, research on integrity has focused on one or more specific tests – integrity has not been defined

outside of the tests that measure the construct. Therefore, to understand the current definitions of integrity, one must understand integrity tests.

### *What are Integrity Tests?*

Integrity tests are self-report inventories designed to assess an individual's honesty and propensity toward counterproductive work behaviors. Sackett, Burris, and Callahan (1989) divide integrity testing into two categories: overt and personality-based. Overt tests use direct questioning to survey test-takers' attitudes towards counterproductive behaviors such as theft and drug abuse. They typically include a section in which test-takers are also asked about illegal behaviors they have actually committed at work. The most researched overt integrity tests are the Reid Report (Ash, 1991; Reid, 1984), Stanton Survey (O'Bannon et al., 1989), and London House Personnel Selection Inventory (PSI) (Sackett et al., 1989). Each test purports to measure a distinct set of factors (see Table 2). All of the tests offer recommendations regarding an individual's general propensity towards workplace deviance.

Personality-based integrity tests, also known as veiled purpose tests (Murphy, 1993), do not directly ask the respondent about theft or other illegal behaviors. Instead, they directly assess personality constructs that are related to a broad range of counterproductive behaviors. Commonly used personality-based tests include the Employee Reliability Inventory (ERI) (O'Bannon et al., 1989), Personnel Decisions, Inc. Employment Inventory (PDI-EI) (Sackett et al., 1989) and the Hogan Reliability Index (Hogan & Hogan, 1989). These tests purport to measure a variety of personality

Table 2

*Overt Integrity Test Factors*

Integrity Test	Theft Related Factors	Non-Theft Related Factors
Personnel Selection Inventory	Honesty	Drug Avoidance Non-Violence Customer Relations Emotional Stability Safety Work Values Supervision Tenure Employability Index Validity Scale – Distortion Validity Scale – Accuracy Detailed Personal History
Reid Report	Honesty Attitude	Social Behavior Substance Abuse Personal Achievements Service Orientation Clerical/ Math Skills
Stanton Survey	Honesty Attitude Admissions of Previous Dishonesty	-



constructs including dependability, conscientiousness, recklessness and hostility toward authority (See Table 3).

### *Reasons for the Emergence of Integrity Testing*

Integrity testing gained widespread use after the passage of the 1988 Polygraph Protection Act. This law prohibited private employers from using pre-employment polygraph testing, leading those employers concerned about theft, propensity toward deviance, and other illegal work behaviors toward an interest in integrity testing. At least 5000 companies have opted to use integrity tests and administer nearly 2.5 million of these tests annually (O'Bannon et al., 1989; Sackett & Harris, 1984). Such widespread use prompted both congressional (U.S. OTA, 1990) and psychological (Goldberg et al., 1991; Sackett & Harris, 1984) investigations into the validity and lawfulness of integrity testing (see Camara & Schneider, 1995 for a full discussion of these reports).

As a whole, the probes indicated that little was known about the construct validity of integrity measures. Under the assumption that integrity is multi-dimensional, no clear and cogent explanation of the constituent dimensions of integrity tests had been presented. The conclusions drawn spawned several studies of integrity testing. The results of these and earlier studies are summarized in the following section.

Table 3

*Personality-based Integrity Test Factors*

---

Integrity Test	Factors
Employee Reliability Inventory	Alcohol/ Substance Abuse Conscientious Job Performance Honesty Long-term Job Commitment
PDI Employment Inventory	Productive Behavior Tenure

---

### *The Validity of Integrity Tests*

The construct validity evidence for integrity tests comes from content, factor analytic, convergent and discriminant validity analyses, and criterion-related validity studies. Much of this research points to the centrality of the personality dimension conscientiousness for understanding the structure and meaning of integrity tests. Taken as a whole, the validity evidence suggests the following:

- (1) Integrity tests are multi-faceted.
- (2) Personality-based and overt integrity tests load on separate factors, with both of these factors loading on a higher-order factor.
- (3) Integrity is a significant predictor of counter-productivity and overall job performance.
- (4) Beyond these points, the meaning and structure of integrity tests is not well understood.

### *Content Analysis*

In a previous study (Green, 1999), I used a qualitative approach to identify the dimensions of the integrity construct. The goal of this inductive study was to create a model of integrity that might be used to illuminate similarities and differences among current measures of the construct. In this study, thirty-five employees, working in labor, service, and professional positions, completed a critical incident interview. The interview lasted between forty-five minutes and one hour. In the interview, participants described

incidents they witnessed or heard about – on their job – reflecting high or low integrity.

Participants answered four questions about these incidents:

- (1) What were the circumstances leading up to the incident?
- (2) What did the employee do that makes you think he or she has low or high integrity?
- (3) What were the consequences of the employee's behavior in this incident?
- (4) How good of an example do you think this is of high or low integrity?

The first three questions were open-ended and the last question was answered on a 5-point Likert-type scale.

I then content analyzed the resulting 144 critical incidents (Green, 1999). In this analysis, I coded each incident by the theme most evident in the participant's description. I selected the themes based on commonly occurring integrity test dimensions and new dimensions described by participants. Table 4 presents the range of themes for coding the critical incidents. Due to the “factorial” complexity of the incidents, many were double-or triple-coded. (For an example, See Table 5). Two hundred and ninety-nine codes were used to fully describe the 144 critical incidents.

Next, five participants served as a second cohort of coders for a portion of the content analysis. These participants were used to assess the consistency of my coding system. To accomplish this, each member of the cohort identified the most evident theme for a unique subset of the 144 critical incidents. Of the 60 incidents analyzed by these raters, 42 of them were coded identically to my analysis, indicating a 76% agreement.

Table 4 lists the frequencies of the categories observed in the critical incidents. The most commonly observed category was Honesty ( $n = 48$ ; 15.7%), followed by

Table 4

*Frequency, Percentage, and Mean Goodness Ratings of Content Analysis Categories*

Theme	Frequency	Percentage	Rating
Honesty	48	17.9	3.76
Respecting Others	34	11.2	3.94
Work Ethic	24	7.9	3.88
Taking Responsibility	24	7.9	3.79
Following Policy/ Obedience	20	6.6	3.38
Emotional Stability	20	6.6	3.88
Job Performance	16	5.2	3.88
Trustworthiness	12	3.9	4.34
Customer Service	12	3.9	4.18
Theft of Money	12	3.9	3.90
Theft of Property	10	3.3	3.90
Theft of Time	9	3.0	3.33
Moral Reasoning	9	3.0	3.63
Attendance	7	2.3	3.64
Vandalism	7	2.3	3.93
Verbal Abuse	6	2.0	4.50
Attitude	6	2.0	3.83
Motivation	5	1.6	4.00

Table 4 Continued

Theme	Frequency	Percentage	Rating
Safety	5	1.6	4.34
Timeliness	5	1.6	3.20
Stress Tolerance	3	1.0	4.00
Manipulation	3	1.0	4.17
Reliability	2	.7	3.00
Violence	2	.7	5.00
Energy Level	2	.7	4.00
Substance Abuse	1	.3	5.00
Hostility	1	.3	3.50
Job Satisfaction	0	0.00	0.00
Mental Ability	0	0.00	0.00
Tenure	0	0.00	0.00

Table 5

*Two Examples of Multi-Coded Critical Incidents*

Critical Incident	Coding Categories
<p>"I was working as a nuclear operator. This guy I knew was working on a piece of equipment. You're supposed to turn the gauge on the equipment to a certain level. The gauge controls the speed of the rotary.</p> <p>Well, it had become a game among a lot of the people to turn the gauge as far as it could go without damaging the equipment. People had left tick marks to indicate how far they had turned the gauge. Well, this guy turned the gauge way too far and the rotary tore itself apart. There was metal everywhere. Anyway, the guy just moved away from the rotary and denied that he did it, and I think that showed low integrity. It didn't matter though because the boss knew he did it and he was fired."</p>	<p>Honesty Following Policy/ Obedience Taking Responsibility</p>
<p>"I was working at an animal hospital. I thought this one employee, Anne, showed very low integrity in her job. She was a licensed veterinary technician and was a supervisor of the kennel assistants. Anne had a bad attitude and I think she brought others down. She was constantly mean and nasty to certain staff members. For example, she would talk loudly about other employees, have a temper tantrum if anyone confronted her about her behavior, talk about people's pets in their earshot, and was rough with the animals. Also, she wouldn't do things the way doctors requested and often complained about billing. She was in a supervisory position and I think people below her sometimes picked up her bad habits."</p>	<p>Attitude Emotional Stability Respecting Others</p>

Respecting Others ( $n = 34$ ; 11.2%), Working Hard ( $n = 24$ ; 7.9%), Taking Responsibility ( $n = 24$ ; 7.9%), Obedience ( $n = 20$ ; 6.6%) and Emotional Stability ( $n = 20$ ; 6.6%).

As discussed previously, the critical incident interview question 'How good of an example do you think this is of high or low integrity?', was answered on a 5-point agreement scale ranging from Poor Example (1) to Perfect Example (5). The mean ratings for each content category are listed in Table 4. As Table 4 indicates the highest mean ratings were given to violence ( $M=5.00$ ,  $SD=.00$ ) and Substance Abuse ( $M=5.00$ ,  $SD=.00$ ). However, it should be noted that each of these categories (violence and substance abuse) had only one observation.

The categories identified in the content analysis were then submitted to a rational, data reduction process. Specifically, I examined the large set of categories (See Table 6) and combined categories that described similar forms of behavior. For example, Theft of Money, Theft of Time, and Theft of Property were all placed in a single category. The purpose of this was to identify by a rational and inductive process the fundamental dimensions of the integrity construct. The result of this step was a five-facet model of integrity. The five facets were Agreeableness, Conscientiousness, Emotional Control, Honesty, and Fairness. Table 7 defines these factors and lists the individual content categories that were combined for each.

Noticeably missing from this inductive integrity model was the substance abuse construct, which is commonly assessed on published measures of integrity (O'Bannon et al., 1989). In the single mention of substance abuse, the participant described it as a



Table 6

*Content Analysis Categories*

Categories	Negative Expression	Positive Expression
Attendance	Absent from work	Always at work when scheduled
Attitude	Expressing negative or pessimistic feelings or thoughts	Expressing positive or can-do feelings or thoughts
Customer Service	Unfriendly, providing little help or professional aid	Providing friendly professional help or aid
Emotional Stability	Fluctuating or wildly expressing emotions	Maintain or calmly expressing emotions
Energy Level	Low level of activity	High level of activity
Honesty	Untruthful, fake	Truthful, genuine
Job Performance	Ineffective performance of job duties	Effective performance of job duties
Job Satisfaction	Displeasure, not content with the job	Pleased, content with the job
Mental Ability	Lacking intelligence or understanding	Intelligent, smart
Moral reasoning	Confusing right and wrong; acting with an unclear definition of right or wrong	Understanding and acting on clear definitions of right and wrong
Motivation	Unable to self-induce action, lacking self-motivation	Able to self-induce action, self-motivated
Obedience	Not following policy, refusing to do what is expected or instructed	Following policy, complying with commands or instructions

Table 6 Continued

Categories	Negative Expression	Positive Expression
Vandalism	Destroying property, that which is valued	Respecting property and that which is valued
Reliability	Not dependable	Dependable
Respecting Others	Treating individuals or groups poorly, violating others	Treating individuals or groups with consideration, not violating others
Taking Responsibility	Unwilling to answer to others for actions, blaming others for mistakes	Answering to others for actions, owning up to mistakes
Safety	Acting in a risky manner, lack of concern for the danger or property or others	Acting in a cautious manner, ensuring property or others are unharmed and free from danger
Stress Response	Responding poorly to pressure or urgency	Responding well to pressure or urgency
Substance Abuse	Improper use of drugs or alcohol	Not engaging in drug or alcohol abuse, even when provoked
Tenure	Short amount of time in job before leaving	Long amount of time in job before leaving
Theft of Money	Stealing money	Not stealing money when given the opportunity
Theft of Property	Stealing property	Not stealing when given the opportunity
Theft of Time	Stealing time, lying on time sheets, wasting time	Using time well, efficient
Timeliness	Tardy	On-time
Trustworthiness	Unworthy of confidence, unable to trust	Worthy of confidence, able to trust

Table 6 Continued

Categories	Negative Expression	Positive Expression
Verbal Treatment	Using harsh language with others	Not using harsh language even when provoked
Violence	Physically aggressive with others	Not physically aggressive with others, even when provoked
Work Ethic	Not working hard, "1/2 doing"	Working hard, giving it your all

precursor to an integrity violation; substance abuse was not the crux of the critical incident. It should be recalled that the purpose of this inductive model of integrity was to generate hypotheses about the meaning of integrity so that the measures of the construct could be better understood. That is, the purpose of this study was to generate rather than validate the model (Green, 1999).

### *Factor Analysis*

Researchers have also used factor analysis to delineate the dimensions of integrity. Factor analytic studies investigating a single integrity test indicate that, although overt and personality-based instruments are multi-faceted, seldom do different instruments yield identical solutions (Ash, 1991; Paajanen et al., 1993). Studies investigating the commonalties among integrity tests indicate that personality-based and overt integrity tests comprise separate factors (Ones, 1993; Woolley & Hakstian, 1992). Some research also suggests that each of these factors (i.e., personality and overt) is a component of a single, higher-order factor (Ones, 1993). The factor analytic studies for overt, personality, and mixed (that is, containing portions of overt and personality components) investigations will be discussed separately.

*Overt integrity tests.* In an attempt to draw conclusions about the structure of integrity, Ash (1991) compared the factor structures identified in three principal components analyses and one principal axis analysis of overt integrity tests. In this investigation, the Stanton Survey (Harris, 1987), PSI (Harris & Sackett, 1987) and Reid Report (Cunningham & Ash, 1988) were examined. The number of factors retained

Table 7

*Combination of Individual Content Categories into Five Larger Categories of Integrity*

Integrity Category	Individual Content Categories
<b>Honesty:</b> the tendency to act in a truthful manner at all times.	Taking Responsibility Theft of Money Theft of Time Theft of Property Trustworthiness
<b>Agreeableness:</b> the tendency to act in a courteous and cooperative manner.	Attitude Customer Service
<b>Conscientiousness:</b> the tendency to meet and exceed work expectations.	Absenteeism Following Policy/ Obedience Reliability Safety Self-Motivation Timeliness Work Ethic (working hard)
<b>Emotional Stability:</b> the tendency to control and display emotions in a professional and non-destructive manner	Stress Tolerance Vandalism Verbal Abuse Violence
<b>Fairness:</b> the tendency to use fair and consistent procedures across people and times	Moral Reasoning Respecting Others

ranged from four for the Reid report to thirteen for the PSI. All tests assessed some aspect of Theft Rumination and Projection of Dishonest Behaviors onto Others (Ash, 1991). Furthermore, of the four investigations, at least three identified factors relating to Self-Punitiveness, general Theft Punitiveness, and Admissions of Theft or other dishonest behaviors (Ash, 1991). Punitiveness was conceptualized as the degree of punishment appropriate for individuals who commit dishonest acts (Cunningham & Ash, 1988). Counter-productive applicants and employees tend to have more lenient attitudes, and thus, score lower on the integrity test. Ash's (1991) comparisons indicated that there were both similarities and differences among the factor structure of overt integrity tests.

To better understand the relationships between overt integrity tests, Ones (1993) used confirmatory factor analysis to test the hypothesis that three overt integrity tests (Personnel Selection Inventory of London House, Stanton Survey, and Reid Report) load on a single, higher-order factor. The sample size comprised 1,365 job applicants and college students, with between 300 and 500 participants taking any single instrument. The results confirmed her hypothesis with factor loadings ranging from .82 to 1.00. Ones (1993) posited that the variance shared across overt-integrity tests was due to the personality factor Conscientiousness. As a result, Ones (1993) labeled the higher-order factor Conscientiousness as Measured by Overt Integrity Tests.

Ones (1993) also investigated the true-score correlations (i.e., the correlations corrected for unreliability) among overt integrity tests. Using the same overt tests (Reid, Stanton Survey, PSI) and sample described above, Ones found an average true-score correlation of .85. Ones (1993) then used meta-analysis to investigate the same question. Ones (1993) entered 56 correlations, from a variety of studies, into a meta-analysis to

determine if various overt integrity tests were significantly related. Each correlation represented the relationship between two different overt integrity tests. In total, the correlations represented 14 different integrity instruments. Based on these correlations, Ones (1993) obtained an average true score correlation among overt integrity tests of .45. She attributed the difference between the primary data correlation (.85) and meta-analytic correlation (.45) to the large number of tests used in the meta-analysis. Based on both sets of results, Ones remarked, "to a certain extent, we can conclude that overt tests seem to share a general common core construct" (Ones, 1993, p. 71). Nonetheless, the average true score correlation of .45 also indicated that there was substantial variance among overt integrity tests that remained unshared.

*Personality-based integrity tests.* Fewer factor analytic studies of personality-based integrity tests have been conducted. Paajanen et al. (1993) used principal components analysis to investigate the factor structure of the PDI-EI. The results indicated a five-factor solution accounting for 99.8% of the common variance and 15.7% of the total variance. The five factors were labeled Irresponsibility, Sensation Seeking, Unstable Upbringing, Frankness, and Conforming Work Motivation. The remaining 84.3% of the total variance was explained by additional factors that the authors labeled Well-behaved, Unlikely Virtues, Alcohol Use, Rebelliousness, and Caution (Paajanen et al., 1993).

Ones (1993) also investigated the structure of personality-based integrity tests. She used confirmatory factor analysis to test the hypothesis that four personality-based integrity tests (PRB, PDI-EI Performance Scale, IPI critical scale, Hogan Reliability scale) loaded on a single, higher-order factor. The results were based on a primary data

set (N=1,365) and confirmed her hypothesis with loadings ranging from .75 to .90 on the single factor. Ones (1993) posited that the variance shared across integrity tests was due to the personality construct Conscientiousness, and as a result, labeled the higher-order factor Conscientiousness as Measured by Personality-based Integrity Tests. Further analysis revealed that these same tests correlated with each other with an average true score correlation of .75, further supporting Ones' hypothesis that personality-based integrity tests are highly related.

*Investigating the relationship between overt and personality-based integrity tests.*

Researchers have sought to understand the relationship between overt and personality-based integrity tests since Sackett et al., (1989) first differentiated between the two. Factor analytic research on these tests has generated what some have called, mixed results (Hogan & Brinkmeyer, 1997).

In one study, Frost and Rafilson (1989) correlated the PSI honesty scale (overt) and the PRB overall score (personality-based) for 105 participants employed various occupations. Results indicated a statistically significant correlation of .25. Based on these results, the authors concluded that the PSI honesty scale and the PRB "appear to be measuring different constructs" (Frost & Rafilson, 1989, p. 273-274).

In another study, Woolley and Hakstian (1992) investigated the factor structure of one overt integrity test (Reid Report), three personality-based integrity tests (Employee Reliability Index, PDI-EI, and the PRB), and selected personality from three personality instruments (CPI, 16PF, and the NEO-PI). The researchers submitted test sub-scores (the scores generated by the test publishers) to an exploratory factor analysis and reported a four-factor solution: Conventional Commitment, Intolerance of Dishonesty, Socialized



Control, and Active Conscientiousness. The Reid scores (Honesty and Punitive) loaded on a single factor, Intolerance of Dishonesty, with none of the personality -based scales or personality instruments loading on this factor. This suggested that the construct measured by the Reid Report was distinct from those measured by the personality and personality-based integrity instruments. All except one sub-score of the personality-based integrity instruments, along with the four scales from the CPI and one scale from the 16PF loaded on the largest factor, Socialized Control. The authors reported that they believed that this was the generalized factor which links personality-based integrity tests. They described individuals high on Socialized Control as rule-abiding individuals who were unlikely to go against societal norms. Low scorers were described as rule-breaking, less stable than high scorers and more likely to take risks.

Ones (1993) also investigated the relationship between personality-based and overt integrity tests. Analyzing four personality-based and three overt integrity tests (as previously described), she found support for the following assertions: (1) personality-based tests load on a second order factor which she labeled Conscientiousness as measured by personality-based integrity tests; (2) overt integrity tests load on a second order factor which she labeled Conscientiousness as measured by overt-based integrity tests, and (3) both the overt and personality-based Conscientiousness factors load on a third order factor which she labeled Conscientiousness as measured by integrity tests. While these results may appear contradictory to those found by Woolley and Hakstian (1992), Woolley and Hakstian (1992) did not indicate that they tested for the second order factor (generalized Conscientiousness). Therefore, it is equally possible that the

second order factor would have been found had it been tested, thereby yielding complementary, as opposed to contradictory results.

Hogan and Brinkmeyer (1997) also investigated the relationship between personality-based and overt integrity tests. The authors first submitted item-level data from the Reid Report (an overt measure) and the Hogan Employee Reliability Index (ERI, a personality-based measure) to an exploratory principal components analysis. Thirty-seven of the Reid items loaded on three factors labeled Illegal Drug Use (7 Reid items), Theft Admissions (10 Reid items), and Punitive Attitudes (20 Reid items). Punitive attitudes were defined as the extent to which an individual "expresses punitive attitudes toward theft. An example item is 'even when no one suffers, every theft should be legally charged'" (positively scored) (Hogan & Brinkmeyer, 1997, p. 593). Eighteen of the ERI items loaded on a separate factor labeled Reliability. The authors defined Reliability as "(concerning) the themes of alienation, social insensitivity, hostility to rules or authority, and impulsiveness" (Hogan & Brinkmeyer, 1997, 593). Although one item from the Reid Report loaded on the ERI Reliability factor and one item from the ERI loaded on the Drug Use factor, no other cross-loadings were evident. These results suggested that the constructs being measured by the Reid Report and the ERI were conceptually distinct, and that the constructs measured by the ERI were multi-faceted.

To further investigate the structure of integrity, the authors submitted scores on the four identified factors (Punitive Attitudes, Illegal Drug Use, Reliability, Theft Admission) to maximum-likelihood confirmatory factor analysis. The authors tested the hypothesis that the factor scores were indicators of a higher-order factor labeled Conscientiousness. Results confirmed the hypothesis with factor loadings of .70 for

Theft Admissions, .63 for Punitive Attitudes, .57 for Reliability and .28 for Illegal Drug Use. Three indices of fit (RMSEA, CFI, and NFI) indicated that the data provided a good fit to the model. Nonetheless, the range of loadings suggested that each factor was not equally important in defining the construct.

In an attempt to delineate the specific facets of integrity tests, Wanek (1995) submitted the data used by Ones (1993) to principal components factor analysis. Instead of the test level data Ones (1993) used, Wanek (1995) analyzed item-level data. One factor analysis was conducted on the items comprising the personality-based integrity tests; one was completed on the items comprising the overt integrity tests; the final analysis was completed on a combined set of items from both test types. The resultant factor solutions were complex. Factors were retained whose eigenvalues were greater than or equal to 1.0. Under this factor-retention criterion, principal components analysis extracted 127 factors for overt tests, 195 factors for personality-based tests, and 327 factors for all integrity tests combined. Although a rational investigation of the content was used to reduce the number of factors to 10, 11, and 19 respectively, many of the factors were uninterpretable or contrary to rational explanations. For example, Wanek (1995) reported that the following three questions with similar content loaded on three distinct factors: "Would you say you are too honest to steal? (PSI), Do you think you are above stealing anything at all? (Stanton), and Do you believe you are too honest too steal? (Reid)" (p. 151). Wanek (1995) concluded that these results might have been due to statistical limitations. For example, each instrument used in the analyses was completed by different subsets of participants. Of the 1,365 participants, 300 to 500 completed any given instrument. Wanek (1995) suggested that this could explain the

apparent test-specific factors that could not be explained based on content. Also, Wanek (1995) acknowledged that the reliance on more items than observations may have yielded spurious correlations.

While Wanek's (1995) study was plagued with statistical limitations, it was the first of its kind to attempt an item-level analysis of a large number of overt and personality-based integrity tests. The limitations of this study illuminate the practical difficulties in investigating the meaning of integrity tests. Using a rational inspection of the results, Wanek (1995) concluded that several themes were apparent in both personality-based and overt integrity tests. They were Theft-based factors, Trust/ Low Self-control, External Locus of Control, and Affectivity/ Locus of Control. Based on patterns of correlations, he further concluded that Self-control, not Conscientiousness (Ones, 1993; Hogan and Brinkmeyer, 1997), was the generalized factor underlying integrity tests. To further delineate the structure of integrity, Wanek (1995) stated that "the next logical step. . .call(s) for the rational formation of factors based on some judgment of item content. . .It seems reasonable to expect that cleaner composites, of at least the major factors, would result" (Wanek, 1995, p. 154). Wanek's suggestion is a driving influence on the present study.

#### *Convergent and Discriminant Validity*

Another way researchers have sought to understand the meaning of what is measured by integrity tests is through convergent and discriminant validity analyses. Numerous studies have investigated the relationship between integrity test scores and a

wide range of personality variables. Much of this research has resulted in a laundry list of relationships that lack an organizing framework for understanding or delineating the true meaning of what is measured by integrity tests (see Table 8). Researchers have also investigated the relationship between integrity test scores and other variables including cognitive ability, religiosity, and moral reasoning. Of these studies, only cognitive ability demonstrates a consistent relationship with integrity.

*Integrity and personality.* Of the Big Five dimensions, Conscientiousness, Agreeableness, and Emotional Stability have shown significant relationships with integrity test scores (see Table 9). Extroversion and Openness to Experience have also shown significant relationships, but less often and to a lesser extent (Ones, 1993) (See Table 9). Ones' (1993) study is of particular importance to the validation of integrity test scores because of its breadth and comprehensiveness. Ones (1993) meta-analyzed 1,506 correlations to test the hypothesis that the Big Five personality variables were significantly related to integrity test scores. Ones (1993) used eight integrity tests (personality-based and overt) to compute a single composite integrity score. She then used three personality instruments to compute a single composite score for each dimension of the Big Five. Three of the five personality composites regressed on the integrity composite to achieve a multiple correlation of .99. These results indicated that integrity could be almost completely explained by the personality composites. More specifically, results indicated that Conscientiousness was the best predictor of integrity ( $b=.61$ ), followed by Agreeableness ( $b=.43$ ), and Emotional Stability ( $b=.13$ ). Adding Openness to Experience and Extraversion to the equation only increased the multiple correlation by .01. Based on these results, Ones (1993) concluded that integrity testing

Table 8

*Correlations between Integrity Test Scores and Personality Variables*

Test	Author	Personality Variables							
		Achievement	Aggression	Dominance	Creativity	Depression	Inferiority	Self-Control	Socialization
ERI	Hogan & Hogan, 1989	.14*			-.08	.13*			.46*
Phase II	Logan et al., 1986			.03	.22*			-.02	
PRB	Woolley & Hakstian, 1992	.02 <sup>a</sup>						.67* <sup>a</sup>	.78* <sup>a</sup>
PDI-EI-P	Woolley & Hakstian, 1992	-.06 <sup>a</sup>						.53* <sup>a</sup>	.51* <sup>a</sup>
PDI-EI-T	Woolley & Hakstian, 1992	-.02 <sup>a</sup>						.55* <sup>a</sup>	.30* <sup>a</sup>
ERI	Woolley & Hakstian, 1992	.12 <sup>a</sup>						.81* <sup>a</sup>	.72* <sup>a</sup>
Reid Hon	Woolley & Hakstian, 1992	.16 <sup>a</sup>						.42* <sup>a</sup>	.26* <sup>a</sup>
Reid Pun	Woolley & Hakstian, 1992	.06 <sup>a</sup>						.16 <sup>a</sup>	.06 <sup>a</sup>
PSI Drug	Jones et al., 1990		.45*			.53*	.50*		
PSI Hon	Jones et al., 1990		.24*			.34*	.33*		
PSI Stab	Jones et al., 1990		.13			.32*	.33*		
PSI Viol	Jones et al., 1990		.30*			.32*	.23*		

Table 8 Continued

Test	Author	Personality Variables							
		Achievement	Aggression	Dominance	Creativity	Depression	Inferiority	Restraint	Socialization
Reid Tot	Cunningham et al., 1994	.05							
Reid Tot	Lilienfeld et al., 1994	.12	-.39*					.28*	.45*

*Note.* Achievement = achievement, achievement via independence; Aggression = aggression, hostility; Creativity = creativity, imagination, artistic; Restraint = control, impulse control, restraint, self-control; ERI = Hogan Employee Reliability Index; PDI = Personnel Description Inventory; Phase II = Phase II Profile; PRB = Personnel Reaction Blank; PSI Stab = PSI Stability; PSI Viol = PSI Violence; Reid Hon = Reid Honesty Scale; Reid Tot = Reid Total.

\*Value is statistically significant ( $p < .05$ )

<sup>a</sup>Value represents the correlation coefficient for male respondents only.

Table 9

*Correlations between Integrity Test Scores and the Big Five Dimensions of Personality*

Integrity Test	Author	Big Five Dimensions of Personality				
		Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Openness to Experience
Phase II	Logan et al., 1986			.03	.28*	
PRB	Woolley & Hakstian, 1992	.01 <sup>a</sup>	.53* <sup>a</sup>	.15 <sup>a</sup>	-.34* <sup>a</sup>	-.02 <sup>a</sup>
PDI-EI-P	Woolley & Hakstian, 1992	-.26* <sup>a</sup>	.36* <sup>a</sup>	.05 <sup>a</sup>	.08 <sup>a</sup>	.10 <sup>a</sup>
PDI-EI-T	Woolley & Hakstian, 1992	-.35* <sup>a</sup>	.25* <sup>a</sup>	.22* <sup>a</sup>	.11 <sup>a</sup>	-.17* <sup>a</sup>
ERI	Woolley & Hakstian, 1992	-.05 <sup>a</sup>	.48* <sup>a</sup>	.24* <sup>a</sup>	-.32* <sup>a</sup>	-.13 <sup>a</sup>
RR-Ho	Woolley & Hakstian, 1992	.08 <sup>a</sup>	.25* <sup>a</sup>	.22* <sup>a</sup>	-.21* <sup>a</sup>	-.06 <sup>a</sup>
RR-Pu	Woolley & Hakstian, 1992	.02 <sup>a</sup>	.08* <sup>a</sup>	.06 <sup>a</sup>	-.08 <sup>a</sup>	.06 <sup>a</sup>
Reid Total	Cunningham et al., 1994	-.07			-.19*	
Reid Pun	Hogan & Brinkmeyer 1997	-.13*	.20*	.42*	.35*	.08*
Reid Drug	Hogan & Brinkmeyer 1997	-.02	.04	.07*	.10*	.01
Reid Theft	Hogan & Brinkmeyer 1997	-.09*	.22*	.38*	.41*	.05*



Table 9 Continued

Integrity Test	Author	Big Five Dimensions of Personality				
		Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Openness to Experience
Reid Total	Hogan & Brinkmeyer 1997	-.15*	.26*	.50*	.47*	.10*
Composite <sup>a</sup>	Ones, 1993	-.05	.26	.28	.22	.08

*Note.* Agreeableness = agreeableness, warm, likeability; Extraversion = extraversion, sociability; Openness to Experience = openness to experience, intellectance; Emotional Stability = emotional stability, adjustment; Conscientiousness = conscientiousness, prudence.

<sup>a</sup>Meta-analytic composite of 1,506 correlations.

was simply an extension of personality testing from the Big Five perspective with Conscientiousness, Agreeableness, and Emotional Stability accounting for the most variance in integrity test scores. Ones' (1993) results provide important information regarding the relationship of overarching personality constructs to integrity.

In a review of the literature, Sackett et al. (1989) described evidence supporting the notion that different personality-based integrity tests are similar in meaning. Sackett et al. (1989) demonstrated that the pattern of correlations between two personality-based integrity tests and the California Psychological Inventory (CPI; Gough, 1972) were highly similar. The authors reported that both the Personnel Reaction Blank and the Hogan Reliability Scale correlated highest with the Socialization, Self-control, and Good Impression scales of the CPI. Furthermore, "the next three highest correlations were with the (CPI) Conformance, Sense of Well-being, and Responsibility scales" (Sackett et al., 1989, p. 515). The authors used these correlation pattern similarities to support their hypothesis that there is a high degree of similarity among personality-based integrity tests (Sackett et al., 1989).

*Integrity and other variables.* Cognitive ability (g) is the best single predictor of overall job performance (Hunter & Hunter, 1984). As such, the extent to which integrity tests explain variance in job performance beyond that explained by cognitive ability is critical in establishing the construct validity and utility of the measures. Hogan and Hogan (1989) correlated participant ASVAB scores (quantitative and verbal) with a personality-based integrity measure (overall score for the Hogan Reliability scale). The authors reported correlations of .07 between integrity and the ASVAB quantitative intelligence score and -.09 between integrity and the ASVAB verbal intelligence score.

Others (Gough, 1972; Hunter, 1980; Jones & Terris, 1983; Ones et al., 1993) have reported similar, near zero correlations. These results provided discriminant validity evidence by indicating that integrity tests are not significantly related to cognitive ability. Based on the assumption that the correlation between integrity and cognitive ability was zero, Ones et al. (1993) used meta-analytic correlations to compute the validity coefficient when both cognitive ability and integrity were used to predict job performance. Results indicated that adding integrity to the cognitive ability equation significantly increased the predictability of job performance (e.g., 27% increase for medium complexity jobs) (Ones et al., 1993).

Along with cognitive ability, researchers have also investigated integrity's relationship with religiosity and morality. Due to the similarity of the topics and scarcity of studies, religiosity and morality will be considered together. Sackett and Wanek (1996) reason that religiosity could be related to integrity in two ways. Religious individuals may be more likely to tell the truth on integrity inventories resulting in decreased integrity scores. Or, religious individuals may have greater integrity in general, resulting in increased integrity scores. Research has yet to demonstrate a consistent relationship between integrity and religiosity. For example, Andrews and Lilienfeld (1993) found no differences in integrity scores when they compared a group of nuns and monks (religious) with a group of college students. Lasson (1992), however, found a significant correlation of .24 between Reid scores and level of religiousness, and a correlation of -.16 between religiousness and theft admissions in a college student sample. In a third study, Cochran (1991) found no significant relationships between

morality, as measured by the Defining Issues Test, and integrity, as measured by the Reid Report. More research is needed to clarify this relationship.

### *Criterion-Related Validity*

A large body of research indicates that integrity tests are significantly related to counterproductive work behaviors (Jones, Joy, & Martin, 1990; Jones, Joy, Werner, & Orban, 1991; Ones, 1993; Ones & Viswesvaran, 2001; Terris & Jones, 1982, Viswesvaran, 2002) and overall job performance (Ones, 1993; Ones et al., 1993). The finding that integrity tests significantly predict counterproductive work behaviors is expected in that integrity tests are designed for this purpose—to identify individuals with a propensity to steal, use drugs, or commit other illegal behaviors on the job. However, the finding that integrity tests also relate to positive work behaviors represents a serendipitous, yet easily interpretable finding. Individuals high on integrity are described as dependable, planful, achievement oriented, and having a will to work. It makes sense that along with avoiding negative behaviors, these individuals would also be more likely to exhibit positive work behaviors (Sackett & Wanek, 1996). In the following sections, research is discussed regarding the integrity-job performance and integrity-counterproductivity relationships.

*Integrity and job performance.* In the most comprehensive meta-analysis to date, Ones et al. (1993) used 665 validity coefficients over 576,460 respondents to investigate the relationship between integrity and overall job performance. The authors reported a true score validity of .34 (SD=.13) for the integrity-job performance relationship. The

true score validity represents the true, theoretical correlation between various measures of integrity and job performance over repeated independent testings (Allen & Yen, 1979).

That is, the true score validity depicts the correlation between two variables when there is no error in measurement. Further analyses indicated that job complexity moderated this integrity-job performance relationship. The highest validity was obtained for high complexity jobs (.46), followed by low complexity (.45) and medium complexity (.32), respectively. Because not all studies described the complexity of the positions investigated, Ones et al. used 110 studies in this part of the investigation (17% classified as low, 73% classified as medium, and 10% classified as high).

For selection purposes, validity can be estimated by means of a predictive validation strategy. In 23 predictive validation studies where supervisory ratings of performance served as the criterion, Ones et al. (1993) found a true score validity coefficient of .41 for integrity tests. Moderator analyses indicated that overt and personality-based tests yielded equivalent validities in predicting supervisor ratings of performance (Ones et al., 1993).

A recent study investigated the impact of conscientiousness on the integrity-job performance relationship. Ones (1993) used true score correlations from her meta-analysis and true score correlations from a meta-analysis completed by Barrick and Mount (1991). One must first understand the conclusions presented by Barrick and Mount. They investigated the relationship between the Big Five dimensions of personality and job performance and concluded that conscientiousness was the only dimension consistently predictive of job performance across five occupational groups (Barrick & Mount, 1991). This conclusion along with findings that conscientiousness

and integrity are highly correlated (Collins & Schmidt, 1993; Nolan, 1991), prompted Ones (1993) to investigate the relative impact of conscientiousness and integrity on job performance.

Ones (1993) used partial correlations first to hold constant the effect of conscientiousness on the integrity-job performance relationship, and next to hold constant the effect of integrity on the conscientiousness-job performance relationship. Partialing out conscientiousness from integrity and job performance yielded a decrease in true score validity for the integrity-job performance relationship (from .46 to .33). Murphy and Lee (1994) reported similar findings. These findings indicated that conscientiousness was a significant factor in integrity's ability to predict job performance. However, it should be noted that even when the effect of conscientiousness was removed, integrity maintained a significant positive correlation with job performance (.33). Partialing the effect of integrity from the conscientiousness-job performance relationship also resulted in a decrease in true score validity (from .23 to .05, a near-zero correlation). After the effect of integrity was removed from this relationship, conscientiousness did not maintain a significant correlation with job performance.

Ones' (1993) results indicated two things. First, by comparing the two partial correlations, it appeared that integrity predicted job performance (.33) better than did conscientiousness (.05). Second, as discussed above, when the effect of integrity was removed, conscientiousness did not maintain a significant relationship with job performance. That is, while conscientiousness could only partially explain the relationship between integrity and job performance, integrity could completely explain the relationship between conscientiousness and job performance. Therefore, taken as a

whole, these results support the idea that conscientiousness is best understood as a facet of integrity.

*Integrity and counterproductivity.* To investigate the relationship between integrity and counterproductivity, researchers have used a variety of validation strategies including predictive, concurrent, and group differentiation (Cascio, 1991). The body of evidence indicates that integrity tests are effective predictors of a variety of counterproductive work behaviors, although several moderators affect this relationship.

Group differentiation studies use integrity scores to differentiate known offenders from a comparison group in which the level of offense is unknown. The term "offenders" has been defined in a variety of ways including employees fired for gross misconduct (Jones et al., 1991), incarcerated felons (Ash, 1974), employees identified as having mishandled employer's cash (Terris & Jones, 1982), or known child abusers (Jones et al., 1991). In one study, Jones et al. (1991) compared a group of 100,000 job applicants from a normative database to a comparison group of 1,073 employees who had been terminated for gross misconduct on the job. Possible offenses included theft, drug abuse, vandalism, policy violations, poor job performance, or chronic absenteeism. Results indicated that the counterproductive group had significantly lower integrity scores than the normative group, as measured by the honesty scale of the PSI (Jones et al., 1991). Overall, research indicates that both overt and personality-based integrity tests are effective in differentiating between offender and comparison groups.

Concurrent and predictive validation studies also support the assertion that integrity tests are important predictors of counterproductive work behaviors. In their meta-analysis, Ones et al. (1993) reported a true score validity coefficient of .47 between

integrity and counterproductivity. Counterproductivity was defined in terms of a person's engaging in disruptive behaviors such as actual theft, admitted theft, dismissals for theft, illegal activities, absenteeism, tardiness, and violence. These behaviors were measured by means of both external and self-report measures. The results of the study by Ones et al. (1993) and the results of others also indicated that several moderators affect the relationship between integrity and counterproductivity. Some of these include validation strategy and sample, test type, the criterion's definition, and criterion measurement. Each of these is discussed below.

*Validation strategy and sample.* Many researchers believe that the optimal strategy for estimating the predictive capability of a pre-hire test is the predictive validity paradigm as opposed to the concurrent paradigm. With respect to counterproductive work behaviors, research has indicated that the concurrent strategy (which requires samples of current employees) tends to overestimate validity (Ones et al., 1993; Sackett & Wanek, 1996). This apparent overestimation may be due to the fact that research on current employees is likely to provide anonymity, and therefore, employees are more likely than applicants to admit to counterproductive work behaviors. Regardless of the reason, this difference in admission of counterproductive behavior implies that the variance of the applicant sample would be more restricted than that of the employee sample (Sackett & Wanek, 1996).

*Test type.* Overt and personality-based integrity tests are important predictors of counterproductive behaviors. However, which is the better predictor is unknown. Although some analyses have demonstrated greater validity for overt integrity tests (Frost & Rafilson, 1989; Ones et al., 1993), others have demonstrated greater validity for



personality-based tests (Ones et al., 1993). In their meta-analysis, Ones et al. (1993) reported greater validity for overt tests in predicting counterproductivity. However, the tendency of overt studies to use different criteria from studies on personality-based tests makes the research difficult to interpret and makes definitive conclusions impossible. Thus, to date, research does not indicate whether personality-based or overt tests are better predictors of counterproductive behaviors.

*Criterion-definition.* Counterproductive behaviors can be defined narrowly or broadly (Sackett et al., 1989). Narrow definitions typically focus on theft behaviors or absenteeism. Broad definitions can include a range of behaviors including theft, drug abuse, absenteeism, turnover, disciplinary problems, and violence. Sackett et al. (1989) hypothesized that overt tests would predict narrow criteria better than personality-based tests. Overt-tests typically focus on specific behavior, and, as a result, are directly linked to the narrow criteria on which they are focused. Alternately, the authors hypothesized that personality-based tests would out-perform overt tests in predicting broad criteria because personality-based integrity tests focus on broad personality constructs that are better suited for predicting broad criteria. In a partial test of the hypothesis, Ones et al. (1993) found evidence that countered the hypothesis of Sackett et al. (1989). Ones and her colleagues reported that overall, the pattern of their study's relationships indicated that overt tests predicted broad criteria better than narrow criteria. More research is needed to determine the robustness of these findings.

*Criterion-measurement.* Counterproductive behaviors have been most often measured by means of self-report or external criteria (e.g., performance write-ups, terminations, and absentee records). Research indicates that validity coefficients are

lower with external criteria (McDaniel & Jones, 1988; Ones et al., 1993; Sackett & Wanek, 1996). This is no doubt due to the difficulty of objectively detecting counterproductive work behaviors. Murphy (1990) pointed out that because of this difficulty, validity coefficients computed on the basis of external criteria tend to be lower than those based on self-report data.

### *Summary of Validation Evidence*

The content and factor analytic studies have resulted in multiple descriptions of integrity and its dimensions. Nonetheless, research indicates that integrity tests as a whole are effective predictors of job performance and counterproductive work behaviors. Although a variety of factors may moderate the integrity—counterproductivity relationship, the validity coefficient is consistently greater than zero. Based on their meta-analysis, Ones et al. (1993) reported true score validity coefficients of .34 (SD=.13) for integrity-job performance and .47 (SD=.37) for integrity-counterproductivity. Their results and the results of others support the continued use of integrity testing in predicting job performance. However, their research only partially addresses the focal questions in the current study pertaining to the definition of integrity and a parsimonious identification of its facets.

## SUMMARY

The meaning and structure of integrity test scores are not well understood. Early applied researchers developed integrity tests to measure applicant and employee honesty. Now integrity testing has expanded to assess a variety of concepts from conscientiousness to safety (Sackett & Wanek, 1996). Indeed, integrity tests have become a catchall category of tests presumed to measure a number of constructs related to employee performance.

This state of affairs makes theory building with regard to integrity in the work place difficult. As Landy (1986) explained, "For some time...*validity* was considered a correlation between a predictor and a criterion...Such a positivist view was (and remains) only minimally helpful in developing...a basic understanding of what was being measured... (p. 1183)." Landy's remarks demonstrate that the fact that one construct correlates with another measured construct does not imply that the two constructs are essentially equivalent. Most researchers engaged in employee selection research would not state that intelligence and performance are equivalent on the basis of the established fact that the two constructs covary. Yet, in the field of integrity measurement, there may be just such a tendency. It is time for integrity test research to expand upon the prediction studies and focus on the basic understanding that Landy (1986) described. Recent research has made significant strides in understanding the meaning of integrity (e.g., Ones, 1993; Ones & Viswesvaran, 2001). However, there are important questions that research has yet to answer. These questions can be divided into three areas: defining integrity, understanding integrity's relationship to other performance indicators, and understanding integrity instruments.

### Questions Pertaining to Defining Integrity

1. What is integrity? One might conclude that given the many perspectives on the construct, there is no one definition.
2. What dimensions accurately and completely describe the integrity domain?

### Questions Pertaining to Integrity and Other Performance Indicators

3. How does the integrity construct relate to the five dimensions of personality (Agreeableness, Emotional Stability, Conscientiousness, Extraversion, Openness to Experience) *and* why are these relationships evident? The questions here pertain to logical and empirical distinctions and similarities between known personality dimensions and integrity.
4. How are the dimensions of integrity related to dimensions of employee performance (e.g., theft, customer service, turnover) (Ones & Viswesvaran, 2001; Sackett et al., 1989)? And specifically what dimensions of employee performance are NOT related (e.g., tenure) or are negatively related (e.g., creativity) to integrity?

### Questions Pertaining to Understanding Integrity Instruments

5. What is the meaning of the dimensions commonly assessed by integrity instruments?
6. Do integrity instruments assess the negative and positive aspects of the construct (Hogan & Ones, 1997)?

7. What factors cause the variation in individual tests' abilities to predict performance criteria (Sackett & Wanek, 1996)?

### Present Investigation

The purpose of the present investigation was to develop a basic understanding of integrity by beginning to answer the first two questions: What is the definition of integrity and what dimensions accurately and completely describe the integrity domain? Specifically, this study expanded upon the research conducted by Green (1999).

As discussed previously, in an earlier study (Green, 1999), I used a qualitative, critical incident approach to identify integrity dimensions. This approach had two advantages. First, research suggests that the interpretation of integrity as measured by different tests varies across these tests. Therefore, an arbitrary choice of one integrity measure over another as though they are equivalent may reduce the interpretability of the research findings. The use of a qualitative research approach and inductive reasoning avoided reliance on a single theory or model of integrity (Green, 1999). Second, qualitative approaches are helpful in the exploratory stages of research. While significant research had been conducted on the predictive nature of integrity tests, only minimal research had been conducted on the psychological meaning of integrity. Therefore, this study was an important step in identifying the ways in which high and low integrity is demonstrated at work.

The present investigation expanded upon my earlier work (Green, 1999) by empirically testing the qualitative results. Specifically, this study used item analysis,

confirmatory factor analysis, and discriminant validity analysis to test the validity of the five integrity dimensions previously identified: Concern for Others, Conscientiousness, Emotional Stability, Fairness, and Honesty (See Table 7). A detailed explanation of each dimension follows. The following discussion goes beyond Table 7 in explaining the nuances of the five dimensions identified in the qualitative study. This discussion is based on the 144 critical incident analyzed in my earlier study (Green, 1999).

The Honesty dimension is defined as the tendency to act in a truthful manner at all times. This category is divided into two components that are best described by means of the negative pole of the dimension: dishonesty. The two components are active dishonesty and passive dishonesty.

Active dishonesty involves acts of commission. This component is defined as the tendency to engage in dishonest behaviors and includes three content categories from the critical incident (CI) analysis: Theft of Time, Theft of Money and Theft of Property.

Passive dishonesty involves acts of omission. This component is defined as the tendency to omit the truth, especially when such behavior results in personal gain. This component contains two content categories from the CI analysis: Taking Responsibility (for one's action) and Trustworthiness.

The Concern for Others dimension is defined as the tendency to act in a manner that promotes cooperation and displays concern for others (e.g., customers, co-workers, and work group). Specifically, this category focuses on an individual's ability to display positive (can do), courteous, and helpful behaviors, especially when such behaviors contribute to work goal accomplishment. This category is a combination of two content categories from the CI analysis: Customer Service and Attitude. The Customer Service

category included several critical incidents in which participants described employees behaving particularly courteously (or discourteously) and being helpful (or unhelpful) to customers and co-workers. The Attitude category contained numerous examples of individuals' displaying negative, pessimistic, and uncooperative attitudes. As an example, one critical incident described a veterinary technician who complained often about her duties and her co-workers (pessimistic/ can't do attitude). The participant described the technician as rude to clients (discourteous) and uncooperative with doctors and other technicians (unhelpful).

The Emotional Control dimension is defined as the ability to display emotions in a professional and non-destructive manner. This category is a combination of four content categories from the CI analysis: Stress Tolerance, Vandalism, Verbal Abuse, and Violence. Vandalism, Verbal Abuse, and Violence are all considered outcomes of a lack of emotional control. In the critical incidents, individuals who vandalized, physically or verbally abused others, did so out of anger, frustration or retribution. As an example, in one incident, a participant described a co-worker who threw office supplies at another co-worker for repeatedly parking in her designated parking space. Stress tolerance is defined as how well individuals respond to stressful situations and is conceived as a reflection of their emotional control. As an example, one critical incident in the Stress Tolerance category described a physician who discovered fifty dollars worth of supplies were mistakenly thrown away (the stressful situation). The doctor's response to this situation was to verbally abuse the person he believed was responsible, humiliate the person with verbal attacks in front of patients, and require this person to look for the

supplies in the city trash dump. The doctor's response demonstrated that in this instance he lacked emotional control and displayed his anger in an unprofessional manner.

The Conscientiousness dimension is defined as the tendency to meet and exceed work expectations. This category is a combination of seven content categories from the CI analysis: Following Policy/ Obedience, Absenteeism, Timeliness, Safety, Working Hard, Reliability and Self-Motivation. Following Policy/ Obedience, Absenteeism, Timeliness, and Safety are common employer expectations. Employers expect their employees to follow policy and instruction, be at work when scheduled, arrive to work and meetings on time, and act in a safe manner. Therefore, an employee who meets these expectations is acting in a conscientious manner. Working Hard and Self-Motivation are necessary for employees seeking to exceed work expectations. Reliability is an outcome of conscientious behavior. That is, employees who meet expectations were often labeled reliable in the critical incidents. As an example, in one incident in the Reliable category, the participant described an individual as "unreliable". When asked to explain further, she indicated that the individual was "often late to work, unclean when he was there, and disappeared from the floor at various times...". Clearly, this individual did not have a tendency to meet or exceed work expectations.

The Fairness category is defined as the ability to act consistently across people and times. This category is a combination of two content categories from the CI analysis: Moral Reasoning and Respecting Others. In the critical incidents, individuals described as moral and respectful were often described as fair or fair-minded as well. As an example, one incident in the Moral Reasoning category described a supervisor who failed to punish a work group that admittedly broke company policy, because the leader of this



work group was the supervisor's friend. The participant indicated that he questioned the supervisor's "moral reasoning." In a similar situation, with a different work group, the same supervisor punished the group and leader. As a result, the supervisor's behavior clearly displayed a lack of consistency across people.

### *Model of Construct Validation*

To empirically assess the accuracy of the integrity dimensions previously identified (Green, 1999), a variety of validation strategies were used. Specifically, in examining the construct validity of the integrity dimensions, I carried out item analyses, hierarchical confirmatory factor analysis, a discriminant validity analysis, and a factor structure analysis.

### *Hypotheses*

Following are the working hypotheses of this study.

*Hypothesis 1: Support for hierarchical model of integrity.* My previously developed model of integrity (Green, 1999) is an empirically supportable conception of integrity.

*Rationale.* According to this model, integrity comprises five factors—Concern for Others, Conscientiousness, Emotional Control, Fairness, and Honesty. Byrne (1998) asserts that when a theory identifies a construct as a general, higher-order factor indicated by lower-order latent constructs, hierarchical confirmatory factor analysis (HCFA)

provides the appropriate type of empirical support. Hence, this type of evidence will be used to support the hypothesis.

*Hypothesis 2: Discriminant validity.* There is reasonable empirical support for the discriminant validity of Integrity—conceived of as a higher-order construct—and its constituent dimensions.

*Rationale.* Discriminant validity is established when scores on the construct of interest (i.e., integrity) are unrelated to scores on other distinct constructs. The selection of these “distinct constructs” involves a certain degree of strategy. It makes little sense to compare the constructs of interest with constructs that are conceptually and semantically unrelated. For example, it would make little sense to compare a new test of intelligence with measures of height because the constructs fall into different semantic domains. For sure, intelligence and height are distinct constructs. But having empirical evidence for that fact does little to provide support for the instrument designed to measure intelligence or height. The goal in establishing discriminant validity of some focal measure is to show that this measure is able to distinguish its associated construct with other measured constructs that are semantically similar. This means that the analysis must choose comparison constructs that are logically and conceptually related—semantically similar—but not semantically identical to the construct of interest. These comparison constructs are referred to as discriminant constructs in this study, in that a comparison with them provides evidence of discriminant validity. It should be noted that when a construct (such as integrity) is multifaceted, the different facets could also serve as discriminant constructs for each other.

For the present investigation, Integrity and its five constituent dimensions represented the measured construct of interest (Green, 1999). Two discriminant constructs—substance abuse and tenure—were chosen for the reasons presented below. Statistical comparisons among the five integrity dimensions and between the set of constructs of interest and the discriminant constructs were used to assess the discriminant validity of the construct of interest.

Substance abuse was chosen as a discriminant construct because it appears on 24 published integrity tests (O'Bannon et al., 1989), but was not identified as an integrity dimension by Green (1999). Research is unclear regarding the relationship of substance abuse to integrity. As an example, Jones (1980) demonstrated a positive correlation between substance abuse and the Personnel Security Index integrity test (PSI, 1977), whereas Hogan and Brinkmeyer (1997) demonstrated no relationship between substance abuse and the Employee Reliability integrity test (Hogan & Hogan, 1989). Nonetheless, many consider substance abuse an important marker of integrity as evidenced by its widespread use on integrity tests (See Table 1). Based on my qualitative, critical incident analysis (Green, 1999), it was hypothesized that substance abuse was a discriminant construct and would not correlate with the identified dimensions of Integrity.

Anticipated tenure, defined as the likelihood of remaining on the job, was selected as a discriminant construct primarily because it appears on seven published integrity tests (O'Bannon et al., 1989), but again was not identified as an integrity dimension by Green (1999). The Personnel Decisions, Inc. Employment Inventory (PDI-EI) is an integrity test that provides two composite scores: (1) a performance score (P) that assesses the likelihood that the applicant will engage in counterproductive work behaviors and (2) a

tenure score (T) that assesses the likelihood that the applicant will remain on the job longer than three months (PDI, 1985). The PDI-EI tenure score has demonstrated positive correlations with the PDI-EI performance score (.49 for females, .67 for males, Woolley & Hakstian, 1992) and several other integrity tests including the Personnel Reaction Blank (.45, PDI, 1985), the Employee Reliability Inventory (.44), the Reid Honesty scale (.35), and the Reid Punitive scale (.12). Based on Green's (1999) qualitative, critical incident analysis, it was hypothesized that anticipated tenure was a reasonable discriminant construct and would not correlate with the hypothesized dimensions of Integrity.

*Hypothesis 3: Social desirability.* Social desirability and integrity are conceptually and empirically distinct.

*Rationale.* Considerable research has been conducted investigating the relationship of integrity tests to social desirability, or deception. In a review of the research, Sackett et al. (1989) concluded that there is a significant relationship between tests of social desirability and measures of integrity (McFarland & Ryan, 2000). However, some research indicates that these correlations do not damage the factor structure or predictive validity of many of these tests (Ellingson, Smith, & Sackett, 2001; Lilienfeld, Andrews, & Stone-Romero, 1994; Ones et al., 1993; Ones, Viswesvaran, & Reiss, 1996; Zickar & Robie, 1999).

As honesty is an integrity dimension identified by Green (1999), a relationship between the honesty construct and a test of social desirability or deception would be expected. Nonetheless, whether or not this relationship significantly affects the test's factor structure is yet to be determined. In the present study, a test of social desirability

was included in the survey. In short, I hypothesized that although integrity and social desirability are significantly, positively related, social desirability would not significantly impact the factor structure of the integrity construct.

### *Summary of Present Investigation*

In the present study I used item analysis with confirmatory and discriminant validity analyses to assess the validity of the integrity dimensions previously identified (Green, 1999). The procedures and statistical analyses for this study were conducted in two phases: Experiment One and Experiment Two. Experiment One involved a conceptual grouping of the items. This was conducted to confirm the link between individual items and the facets of integrity identified in the critical incidents. Experiment Two involved a survey administration, confirmatory factor analyses, analyses of discriminant validity evidence, and an analysis of social desirability.

## EXPERIMENT 1

### Method

#### *Participants*

Eighty-five individuals, familiar with personality assessment literature and methodology, were contacted via mail about study participation. Eighteen individuals responded and participated in Study One. Participants included 11 university faculty, 3 personality researchers, 3 professional psychologists and 1 graduate student. Participant time in current profession ranged from 1 to 50 years with a mean time of 13 years in their current profession. Participant age ranged from 27 to 78 years with a mean age of 40.5.

#### *Measures*

The Item Selection Survey (ISS) was used to assess the perceived relevance of individual items to the five personality constructs under investigation (See Appendix B). The constructs were the integrity dimensions identified by Green (1999) (Agreeableness, Conscientiousness, Emotional Stability, Fairness, Honesty). A pool of 17 to 28 items per construct was developed (See Appendix C for individual items).

Some of the items were generated by me, some were selected from the International Personality Item Pool (International Personality Item Pool, 2001), and some were selected from the Comparative Emphasis Scale (Meglino, Ravlin, & Adkins, 1989). The International Personality Item Pool (IPIP) is a set of 1,412 items measuring the

lower-level facets of the five-factor model of personality (International Personality Item Pool, 2001). The items were located on the Internet and the web site listed each item as well as each item's correlation with published and well known personality scales and items. The Comparative Emphasis Scale (CES) contains items that assess three constructs investigated in the present study: concern for others, fairness and honesty. The items for the CES are designed to be used in a forced choice format.

On the Item Selection Survey (ISS) participants used a 5-point scale to rate the relevance of an individual item to a given construct. The scale ranged from slightly relevant (1) to extremely relevant (5). Participants could also choose "not at all relevant" if they believed the item was unrelated to the given construct.

There were five versions of the survey. Each version defined one of the five constructs and included a different set of 17 to 28 items (See Appendices B and C).

It should be noted that in the ISS instructions (See Appendix B) participants were told that the items they were rating, when presented on future personality assessments, might vary in response format. That is, the items being rated might require objective, Likert-type, or forced choice responses on future personality assessments. Participants were instructed that this information may or may not assist them in rating the relevance of a given item. The majority of the items on the survey would require Likert-type responses.

### *Procedure*

The fifty potential participants were mailed a cover letter explaining the study and one version of the ISS. Individuals were asked to return the survey in a pre-stamped and addressed envelope. Thirteen participants responded to the first mailing.

In a second mailing, 35 additional people were contacted about study participation. Each person was mailed all five versions of the ISS survey. Five participants responded to the second mailing. Combined, these two mailings yielded five to eight respondents for each version of the survey.

### *Results*

Participant ratings on the ISS were used to select the final items for Study Two. Only items with a mean relevance rating of 3.0 (i.e., moderately relevant) or higher were used in the final survey. If more than 12 items met this criterion, the mean relevance rating for each item was reviewed. Items with the lowest ratings were dropped until there were a more reasonable number of items (i.e., 10 to 12). It should be noted that I retained any items measuring a unique portion of the construct domain based on my own reasoning and judgment, even if that left more than 12 items for a given construct. An overview of the results for each survey is provided below.



### *Concern for Others*

Based on the ISS ratings, 14 of 23 items were selected for inclusion in the final version of the Concern for Others survey. The mean relevance rating for the selected items ranged from 5.00 to 3.50, with an overall mean of 3.96 for all selected items. The most relevant item on the final survey was “Can’t be bothered with others’ needs” ( $M=5.00$ ,  $SD=.00$ ). The least relevant item on the final survey was “Acknowledge others’ accomplishments” ( $M=3.5$ ,  $SD=1.31$ ). (See Appendix D for a list of the mean relevance ratings for all items assessed on the Concern for Others version of the ISS). In the second experiment, all of the selected items were scored on a Likert-type scale.

### *Conscientiousness*

Based on the ISS ratings, 13 items were selected for inclusion in the final version of the Conscientiousness survey. The mean relevance rating for the selected items ranged from 4.83 to 3.83, with an overall mean of 4.21 for all selected items. The most relevant item on the final survey was “Work hard” ( $M=4.83$ ,  $SD=.41$ ). The least relevant item on the final survey was “Pay my bills on time” ( $M=3.83$ ,  $SD=.98$ ). (See Appendix E for a list of mean relevance ratings for all items assessed on the Conscientiousness version of the ISS). In the second experiment, all of the selected items were scored on a Likert-type scale.

### *Emotional Control*

Based on the ISS ratings, 13 of 18 items were selected for inclusion in the final version of the Emotional Control survey. The mean relevance rating for the selected items ranged from 5.00 to 3.43, with an overall mean of 4.30 for all selected items. The most relevant item on the final survey was “Keep my emotions under control” ( $M=5.00$ ,  $SD=.00$ ). The least relevant item on the final survey was “Act quickly without thinking” ( $M=3.83$ ,  $SD=1.27$ ). (Appendix F lists the mean relevance ratings for all items on the Emotional Control version of the Item Selection Survey). In the second experiment, all of the selected items were scored on a Likert-type scale.

### *Fairness*

The items on the Fairness version of the Item Selection Survey were chosen from three sources: the Comparative Emphasis Scale, the IPIP, and items that I created. (See Appendix G for a list of the mean relevance rating for each Fairness item on the ISS). Respondents to the ISS rated 12 of the 13 items 3.0 or greater. The single item rated less than 3.0 was from the pool of items created by the author. Of the 3 sources of items, only the Comparative Emphasis Scale has previous and significant studies of reliability and validity in a work context (Meglino et al., 1989; Meglino, Ravlin, & Adkins, 1992; Ravlin & Meglino, 1987). Therefore, the fairness items from the Comparative Emphasis Scale were used to measure fairness in the second experiment. The mean relevance rating for the selected items ranged from 4.88 to 3.50, with an overall mean of 4.15 for all

selected items. In the second experiment, all of the selected Fairness items were scored on a forced-choice scale. For example, for the item with the highest mean relevance rating, respondents indicated which of the following two behaviors they valued most: (a) Being impartial in judging disagreements or (b) helping others on difficult projects or assignments. In this example, item "a" is the fairness item.

It should be noted that six Comparative Emphasis Scale (CES) items not included on the ISS were included in Study Two. The CES utilizes a forced choice format. Similarly worded items are matched with items assessing different work values that are equal in social desirability. On the ISS, similarly worded items were not included to limit the participant burden of rating duplicate items. For example, the ISS included the item "Being impartial in judging disagreements", but did not include the item "Trying to bring about a fair solution to a dispute". Nonetheless, the author of the Comparative Emphasis Scale notes the importance of using all items to accurately measure even a single value on the scale (Meglino, 2000). Therefore, all 12 of the items related to fairness were included in Study Two.

### *Honesty*

Based on the ISS ratings, 14 of 25 items were selected for inclusion in the final version of the Honesty survey. The relevance ratings for the selected items ranged from 5.00 to 4.00, with an overall mean of 4.57 for all selected items. The most relevant item on the final survey was "Tell the truth" ( $M=5.00$ ,  $SD=.00$ ). The least relevant item on the final survey was "Break my promises" ( $M=4.00$ ,  $SD=1.22$ ). (Appendix H lists the mean

relevance ratings for the 25 items assessed on the Honesty version of the Item Selection Survey). In the second experiment, eleven of the fourteen items were scored on a Likert-type scale. The remaining three items were open-ended questions.

## EXPERIMENT 2

### Method

#### *Participants*

Two hundred and fifty-four individuals, currently employed in a job they have held for at least 30 days, participated in study two. There were 199 females and 53 males, 2 participants did not respond to the gender item. One hundred and twenty-nine of the participants were White, 92 were Black, 15 were Asian, 6 were Hispanic, 4 were other, and 8 participants did not respond to the ethnicity item. Participants worked an average of 32.34 hours a week ( $SD=12.07$ ). The average age of the participants was 27 ( $SD=9.4$ ). Many of the participants were recruited from a mid-sized university. These participants received research credit in a psychology course for their participation. Other participants in the study were not compensated.

#### *Measures*

##### *Integrity Measures*

The five integrity constructs (Concern for Others, Conscientiousness, Emotional Control, Fairness, and Honesty) were assessed by means of the scales developed during Study One. For the majority of the items, participants were asked to use a Likert-type scale to indicate how accurately each item described their personality and behavior as

they "generally are now", not as they "wish to be in the future". The 5-point scale ranged from Very Inaccurate to Very Accurate. The Fairness items were answered using a forced choice format and two of the Honesty items were open ended questions.

### *Discriminant Measures*

*Anticipated tenure.* To measure anticipated tenure, participants were asked about their intentions to turnover or leave their current position. Intention to turnover was measured by means of a 3-item scale developed by Camman, Fichman, Jenkins, and Klesh (1979): "(1) I often think of leaving the organization, (2) it is very possible that I will look for a new job next year, and (3) if I may choose again, I will choose to work for the current organization (reverse coded)". All items were answered on a 7-point scale. These items have demonstrated acceptable internal consistency in previous studies (Chen, Hui, & Segó, 1998; George & Jones, 1996).

*Substance abuse.* The Substance Abuse Scale was developed specifically for this research and was based on items contained in the Drug Abuse Screening Test (Skinner, 1982) and research on the effects of alcohol at work (Hollinger, 1988). The Drug Abuse Screening Test (DAST) is a 28-item instrument that assesses the extent of an individual's drug-related problems. The instrument assesses drug-related problems with work, family, spouse, and health. Only the three items assessing work-related problems were utilized in the present study. DAST items are answered on a yes/ no scale. It should be noted that the DAST items refer only to drug abuse (Skinner, 1982). As a result, three additional items were included in the present study to assess the effects of alcohol abuse.

These items were identical to the three drug abuse items, except, the word "drug" was replaced by "alcohol." This allowed for the assessment of both alcohol and drug abuse in the Substance Abuse Scale.

A single-item measure of working under the influence was also included in the Substance Abuse Scale: "How often have you come to work during the past year under the influence of alcohol or drugs?". The question was answered on a five-point scale (1 = almost daily, 2 = about weekly, 3 = four to twelve times per year, 4 = one to three times per year, and 5 = never) (Hollinger, 1988).

#### *Social Desirability Measure*

Social desirability was measured with a short-form of the Marlowe-Crowne Social Desirability (SD) Scale (Crowne & Marlowe, 1960; Reynolds, 1982). When first published, the Marlowe-Crowne SD scale was the first scale of social desirability that was independent of psychopathology. Since its first appearance, the Marlowe-Crowne SD scale has been used in various experimental and applied research studies. In 1982, Reynolds (1982) published three short forms of the original Marlowe-Crowne SD Scale. While the original scale had 33 items, Reynolds developed three short forms with 11, 12, and 13 items respectively. Psychometric analyses by Reynolds (1982) and others Ballard (1992) indicate that the 13-item short form is a reliable and valid alternative to the original Marlowe-Crowne SD Scale. The 13-item short form was administered in the present investigation.

*Procedure*

Participants were administered all of the measures described above (Integrity scales, Discriminant scales, and Social Desirability measure). To encourage honest and careful responses, participants were told that their responses would be anonymous and used for experimental purposes only. Specifically, experimental use was explained to participants in the following manner:

The materials in this study are extremely important to businesses and their employees. Businesses all over the country use tests to learn about the characteristics of their applicants and employees. However, some people question how well these businesses are measuring these characteristics.

This is where you come into play. I am investigating the statistical properties of the items on this questionnaire so tests of the highest quality can be developed. This is why it is extremely important to answer the items CAREFULLY and TRUTHFULLY. In doing so, you will be providing the information needed to ensure that personality characteristics are measured accurately.

Neither your name nor any identifying information will be recorded. However, only those questionnaires with every item answered will be included in the study. If you are willing to be truthful and careful with your responses, please turn the page and begin the questionnaire.

Old Dominion University's Institutional Review Board for the Department of Psychology approved the administration of the survey.



## Analytic Strategy and Results

Confirmatory and hierarchical confirmatory factor analyses (HCFA) were used to test the study hypotheses. The computer program, LISREL Version 8.52 (Joreskog & Sorbom, 2001), provided the basis for estimating model parameters through the maximum likelihood estimation strategy. Subscales were created to serve as indicators for the constructs of interest (i.e., latent variables). In contrast to using the item-level responses themselves as multiple indicators, subscales serve as useful multiple indicators in structural equation modeling. They avoid the problems associated with the polytomous scale items comprising the original scale and legitimize the use of the maximum likelihood estimation strategy. Of course, multiple indicators are desirable because they allow the software to estimate the measurement error variance of the sample variance covariance matrix (Bollen, 1989). I used the procedure described in Appendix I to create two indicators per construct (see Table 10 for a listing of the items that comprised each indicator). For each model tested, the latent variables were standardized to create a scale for estimate interpretation. All estimates are presented in their completely standardized form.

Three hypotheses were investigated in this study: the Integrity hypothesis, the discriminant hypothesis, and the social desirability hypothesis. The model established while testing the Integrity hypothesis was used as a baseline model for the discriminant and social desirability analyses. Therefore, the analytic strategy and results for the Integrity hypothesis are presented before the analytic strategy and results for the discriminant and social desirability hypotheses.

Table 10

*Items Comprising the Integrity Dimensions' Indicators*

Indicator	Response Format	Items
Concern for Others 1	Likert-type	<ol style="list-style-type: none"> <li>1. Believe there's not always time to be kind to others<sup>f</sup></li> <li>2. Cut others to pieces<sup>f</sup></li> <li>3. Acknowledge other's accomplishments</li> <li>4. Am able to cooperate with others</li> </ol>
Concern for Others 2	Likert-type	<ol style="list-style-type: none"> <li>1. Look down on others<sup>f</sup></li> <li>2. Approach others in a positive manner</li> <li>3. Like to be of service to others</li> <li>4. Am polite to strangers</li> </ol>
Conscientiousness 1	Likert-type	<ol style="list-style-type: none"> <li>1. Pay attention to details</li> <li>2. Do just enough work to get by<sup>f</sup></li> <li>3. Set high standards for myself and others</li> <li>4. Follow through on my commitments</li> <li>5. Do more than what's expected of me</li> </ol>
Conscientiousness 2	Likert-type	<ol style="list-style-type: none"> <li>1. Try to follow the rules</li> <li>2. Accomplish my work on time</li> <li>3. Check over my work</li> <li>4. Work hard</li> <li>5. Work on improving myself</li> </ol>
Emotional Control 1	Likert-type	<ol style="list-style-type: none"> <li>1. Am calm even in tense situations</li> <li>2. Snap at people<sup>f</sup></li> <li>3. Panic easily<sup>f</sup></li> <li>4. Keep my emotions under control</li> <li>5. Get stressed out easily<sup>f</sup></li> </ol>
Emotional Control 2	Likert-type	<ol style="list-style-type: none"> <li>1. Get angry easily<sup>f</sup></li> <li>2. Shoot my mouth off<sup>f</sup></li> <li>3. Keep my cool</li> <li>4. Get irritated easily<sup>f</sup></li> <li>5. Take offense easily<sup>f</sup></li> </ol>

Table 10 Continued

Indicator	Response Format	Items
Fairness 1	Forced-Choice	<ol style="list-style-type: none"> <li>1. Encouraging someone who is having a difficult day Considering different points of view before Taking action<sup>f</sup></li> <li>2. Being impartial in judging disagreements<sup>f</sup> Helping others on difficult projects or Assignments</li> <li>3. Making sure each person has an equal chance to get rewards or credit<sup>f</sup> Taking on more responsibility to advance in Your career</li> </ol>
Fairness 2	Forced-Choice	<ol style="list-style-type: none"> <li>1. Judging people fairly based on their abilities rather than only on their personalities<sup>f</sup> Seeking out all opportunities to learn new Skills</li> <li>2. Trying to be helpful to a friend Being sure that any assignments you make Are fair to everyone<sup>f</sup></li> <li>3. Providing fair treatment for each person<sup>f</sup> Lending a helping hand to someone having Difficulty</li> </ol>
Honesty 1	Likert-type	<ol style="list-style-type: none"> <li>1. Truthful in dealing with others</li> <li>2. Lie to make myself look good<sup>f</sup></li> <li>3. Break my promises<sup>f</sup></li> </ol>
Honesty 2	Likert-type	<ol style="list-style-type: none"> <li>1. Tell the truth</li> <li>2. Cheat to get ahead<sup>f</sup></li> <li>3. Do things behind other people's backs<sup>f</sup></li> <li>4. Take credit for others ideas<sup>f</sup></li> </ol>

<sup>f</sup> Reverse scored items. <sup>f</sup> Fairness option within the Forced Choice Items

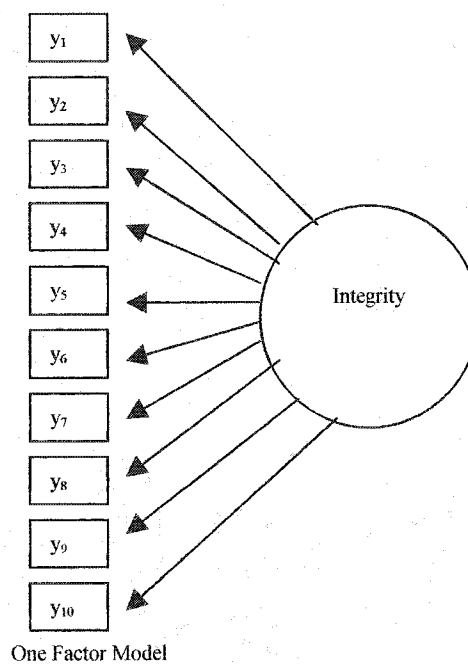
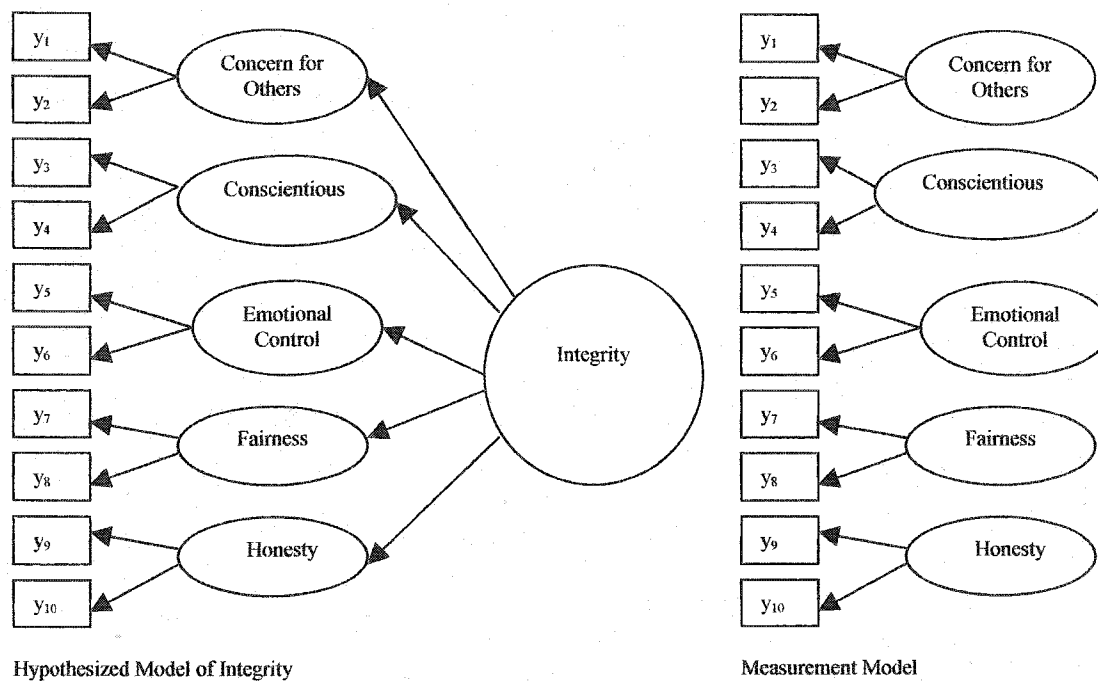
### *Integrity Hypothesis*

The Integrity hypothesis asserted that my previously developed model of Integrity (Green, 1999) was an empirically supportable conception of the Integrity construct. This model contained five first-order factors (Concern for Others, Conscientiousness, Emotional Control, Fairness, and Honesty) that loaded on a single, higher-order factor, Integrity. This model will be referred to as the hypothesized model of Integrity (See Figure 1).

### *Analytic Strategy*

To test the Integrity hypothesis, I used confirmatory and hierarchical confirmatory factor analysis to assess the adequacy of the hypothesized model and also to compare the hypothesized model with other, plausible models of the Integrity construct. To the extent that the hypothesized model described the sample data better than the remaining, plausible models, construct validity was considered strengthened (Bentler and Bonett, 1980). Based on a strategy developed by Rindskopf and Rose (1988), four models were investigated: the hypothesized model of Integrity and three alternative models. The three alternative models were the null model, the one factor model, and the measurement model.

The null model assumed no correlations among the ten indicator variables. This model was a baseline model in that a fit of this model signified that no relationships existed among the variables, and further statistical analyses were unwarranted. The one



*Figure 1. Integrity Hypothesis Model Variants*

factor model assumed ten indicator variables that loaded on a single factor, Integrity (See Figure 1). This model hypothesized that distinctions among the first-order factors were unnecessary. Instead, this model asserted that the variables represented multiple indicators of a single construct, Integrity. The measurement model assumed five, correlated, first-order factors (See Figure 1). This model hypothesized that the indicators were adequate measures of five distinct factors, and that there were significant relationships among the factors. The fit of this model was an upper-bound limit for the fit of the hypothesized model of Integrity. In hierarchical factor analysis, a second-order model can only fit as well as the associated first-order model. Therefore, if there was inadequate support for the measurement model, an analysis of the hypothesized, second-order model would be unnecessary. The adequacy of each of the four models was assessed via measures of overall fit and a logical review of the factor loadings, measurement error variances, error variances, and factor correlations.

To further investigate the validity of the hypothesized model of Integrity, the fit of each alternative model was compared to the fit of the hypothesized model. The hypothesized model was considered supported if it described the sample data significantly better than the alternative models. As the four models investigated were nested models, a statistical comparison was employed. A model is nested within another if it has the same observed variables (indicators), but greater restrictions than the comparison model. Restrictions can be imposed in a variety of ways including constraining relationships to zero, requiring equivalence among parameter estimates, or setting parameter estimates to a pre-determined value. In the present investigation, the

null model represented the most restricted model, followed by the one factor model, the hypothesized model of Integrity, and the measurement model. Nested models are advantageous in that they allow for statistical, rather than descriptive, model comparisons. In the present investigation, the hypothesized model of Integrity was compared to each alternative model with the Chi-Square Difference Statistic. Specifically, the arithmetic difference between the Chi-Square for the hypothesized model and the Chi-Square for the alternate models was computed. The hypothesized model was considered supported if the Chi-Square Difference statistic was significant, indicating that the hypothesized model provided the better fit to the data.

### *Results*

The Integrity hypothesis asserted that the hypothesized model of Integrity (see Figure 1) was an empirically supportable conception of the Integrity construct. This hypothesis was considered supported if the fit statistics, model estimates, and difference statistic indicated that the hypothesized model described the sample data better than each of the alternative models. A correlation matrix with means and standard deviations of the Integrity dimensions' indicators is provided in Table 11.

*Fit statistics.* Measures of overall fit indicate the degree to which the sample data fit the hypothesized factor structure. In the present investigation, the Non-Normed Fit Index (NNFI), Comparative Fit Index (CFI), and Standardized Root Mean Square Residual (SRMSR) were used to assess the overall 'goodness' of the models. The Chi-Square statistic was also computed for model comparison purposes. The Chi-Square was

Table 11

*Correlation Matrix, Means, and Standard Deviations for the Integrity Dimension Indicators, Discriminant Construct Indicators, and Social Desirability Scale Score*

Scale	CFO1	CFO2	CNS1	CNS2	ECO1	ECO2	FAR1	FAR2	HON1	HON2	SUB1	SUB2	TEN1	TEN2	SOCD
CFO1	1.00														
CFO2	.64*	1.00													
CNS1	.39*	.45*	1.00												
CNS2	.51*	.46*	.62*	1.00											
ECO1	.37*	.48*	.27*	.19*	1.00										
ECO2	.44*	.54*	.25*	.30*	.73*	1.00									
FAR1	-.05	-.03	.07	-.01	.03	.03	1.00								
FAR2	.05	.02	.14	.07	.07	.06	.36*	1.00							
HON1	.31*	.34*	.49*	.44*	.24*	.25*	.06	.06	1.00						
HON2	.46*	.44*	.43*	.45*	.31*	.36*	.07	.05	.61*	1.00					
SUB1	-.18*	-.14	-.18*	-.32*	.03	-.05	-.05	-.13	-.12	-.17*	1.00				
SUB2	-.12	-.03	-.12	-.16	-.04	-.05	-.16	-.16	-.17*	-.11	.41*	1.00			



Table 11 Continued

Scale	CFO1	CFO2	CNS1	CNS2	ECO1	ECO2	FAR1	FAR2	HON1	HON2	SUB1	SUB2	TEN1	TEN2	SOCD
TEN1	-.05	-.15	-.13	-.07	-.17*	-.03	.03	-.06	-.15*	-.05	-.07	.10	1.00		
TEN2	-.09	-.20*	-.15	-.06	-.16	-.03	-.07	-.11	-.22*	-.15	-.02	.12	.58*	1.00	
SOCD	.41*	.48*	.32*	.31*	.41*	.44*	.01	.07	.35*	.38*	-.08	-.10	-.13	-.20*	1.00
Mean	4.39	4.33	4.27	4.40	3.49	3.61	1.43	1.35	4.30	4.50	1.10	1.03	3.37	4.30	1.54
SD	.61	.53	.54	.44	.82	.80	.32	.30	.62	.50	.34	.10	1.76	2.46	.23

Note. CFO = Concern for Others; CNS = Conscientiousness; ECO = Emotional Control; FAR = Fairness; HON = Honesty; SUB = Substance Abuse; TEN =

Anticipated Tenure; SOCD = Social Desirability

\*Value is statistically significant ( $p < .001$ ).

the only statistic generated for the null model. Table 12 provides the fit statistics for the four models.

The NNFI indicates the percentage of variance in the covariance matrix accounted for by the hypothesized factor structure and is adjusted for degrees of freedom (i.e., model complexity) (Bentler & Bonett, 1980). The comparative fit index is interpreted in the same manner as the NNFI, but is thought to be less biased by sample size (Bentler, 1990). A value of .90 for the NNFI and CFI statistics represents a reasonable fit of the data to the model. A value  $>.95$  for the two indicates a superior fitting model. As Table 12 indicates, the NNFI and CFI for the one-factor model were .74 and .80 respectively, indicating a poor fitting model. The NNFI and CFI for the hypothesized model of Integrity were .96 and .97 respectively, indicating a superior fitting model. The NNFI and CFI for the measurement model also indicated a superior fitting model with values of .98 and .99 respectively.

The SRMSR represents the difference between the observed covariances and the predicted covariances – the smaller values represent better model fit (Kline, 1998). For the SRMR, a value between .08 and .05 represents a reasonably well fitting model (Marsh, Balla, & McDonald, 1988). A value less than .05 represents an excellent fitting model. The SRMSR for the one factor model was .10, again, indicating a poor fitting model. The SRMSR for the hypothesized model was .05 and the SRMSR for the measurement model was .03. This indicated that both the hypothesized model and the measurement model provided an excellent fit to the data.

Table 12

*Integrity Hypothesis Model Variants*

Model	$\chi^2$	df	NNFI	CFI	SRMSR	$\chi^2$ Diff <sup>^</sup>	Diff df <sup>^</sup>
Null	1411.46	45	-	-	-	1342.00*	15
One Factor	309.48	35	.74	.80	.10	240.02*	5
Hypothesized	69.46	30	.96	.97	.05	-	-
Measurement	40.93	25	.98	.99	.03	28.53*	5
Four Factor	59.14	16	.94	.97	.06	-	-

*Note.* Dashes indicate the fit statistic was not computed. NNFI = Non-Normed Fit Index; CFI = Comparative Fit Index; SRMSR = Standardized Root Mean Square Residual

<sup>^</sup> The  $\chi^2$  Diff statistic compares each model  $\chi^2$  to the  $\chi^2$  for the Hypothesized Model of Integrity.

\* Value is statistically significant ( $p < .05$ ).

*Model comparisons.* The hypothesized model of Integrity was compared to each of the three alternate models by means of the Chi-Square Difference Statistic. As Table 12 indicates, the hypothesized model explained the sample data significantly better than the null or one factor models. However, the measurement model explained the data significantly better than the hypothesized model.

While fit statistics assess the adequacy of a model overall, model estimates provide detailed information about specific model elements. A thorough review of model estimates can provide important information regarding components and relationships of the overall model. Figures 2 through 4 provide the factor loadings, measurement error, and error variances for the one factor, hypothesized, and measurement models (The factor loadings for the null model were set to zero). All estimates are provided in their completely standardized form. T-values were used to determine if factor loadings and error variances were significantly different from zero.

*Measurement model.* The overall fit statistics indicated that the measurement model provided an excellent fit to the sample data. The obtained factor loadings and measurement error variances supported this contention (See Figure 2).

Factor loadings are similar in conception to regression coefficients in multiple regression analyses. That is, each factor loading represents the amount of change in an observed variable for a unit change in the latent variable. For the measurement model, all of the estimated factor loadings were significant. Measurement error variance is that unique portion of the indicator not caused by the latent variable (Bollen, 1989). One measurement error variance is associated with each indicator in the model. The smaller the error variance, the better the indicator is explained by the latent variable. Eight of the

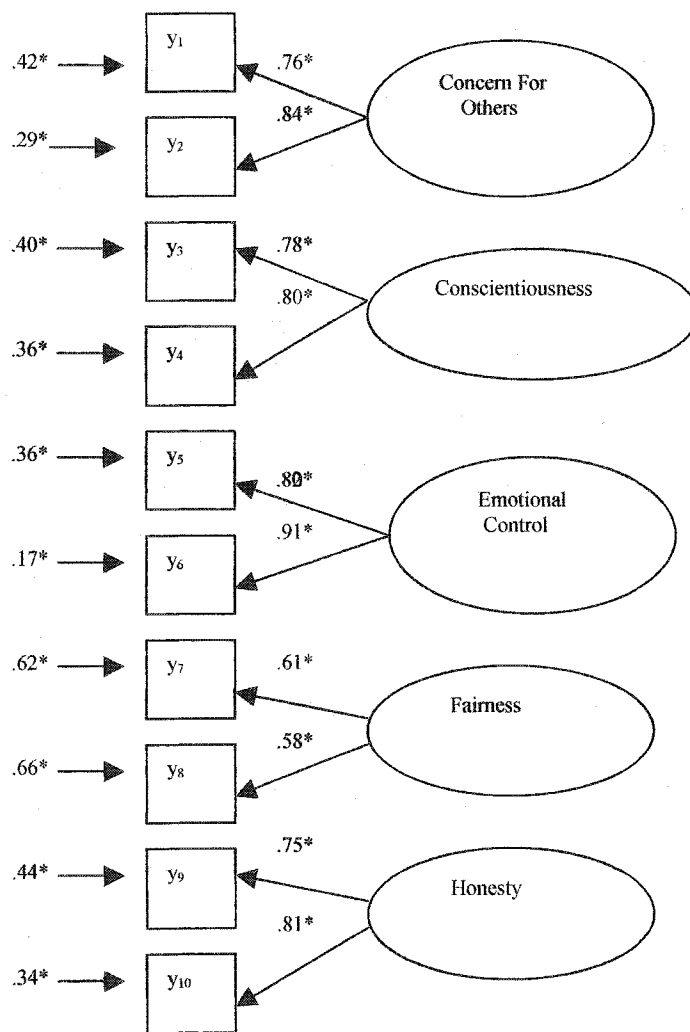


Figure 2. Measurement Model Estimates

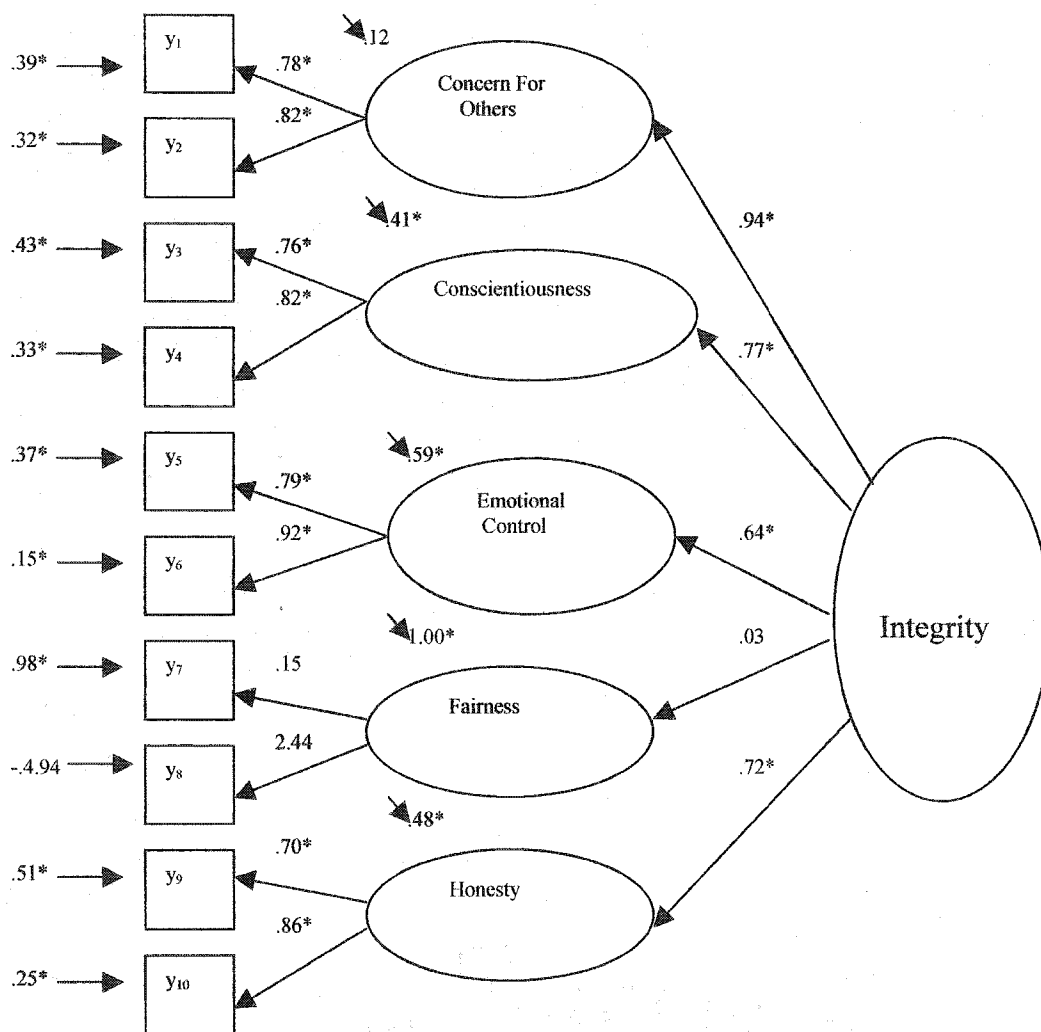


Figure 3. Hypothesized Model of Integrity Model Estimates

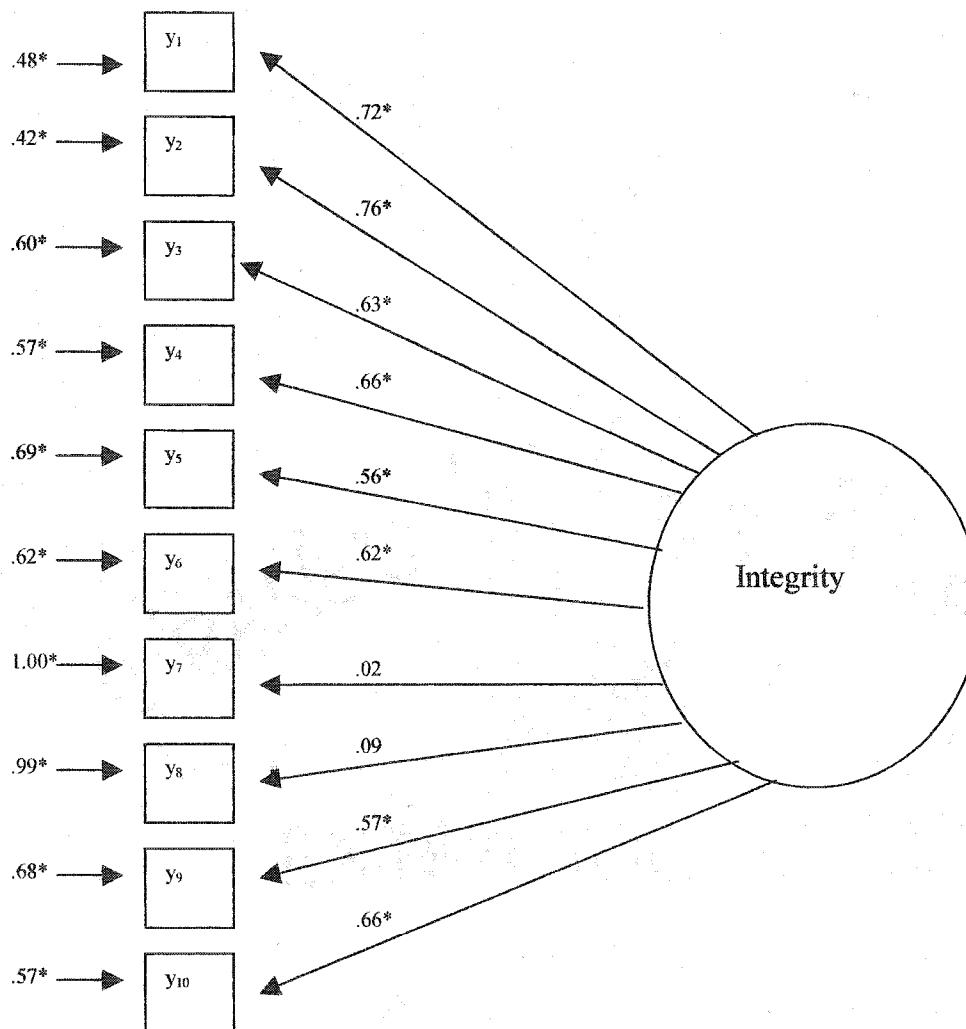


Figure 4. One Factor Model Estimates

ten indicators demonstrated statistically significant amounts of measurement error variance. The measurement error variances for the indicators ranged from .17 for the Emotional Control indicator two to .66 for the Fairness indicator two. The high measurement error variance for the Fairness indicator two suggested that this was not an adequate indicator of the latent factor, Fairness.

The measurement model yielded six, significant correlations among the latent factors. Except for Fairness, all of the first-order factors demonstrated significant correlations among themselves (See Table 13). Two correlations were particularly high: Concern for Others-Conscientiousness (.71) and Conscientiousness-Honesty (.72). As a whole, the correlation values indicated significant relationships among the first order factors.

*Hypothesized model of Integrity.* The fit statistics indicated that the hypothesized model of Integrity also provided an excellent fit to the data. A review of the factor loadings, measurement error variances, and factor correlations supported this contention for four of the five factors. For the hypothesized model, the Fairness indicators and construct failed to maintain significant relationships within the model.

*First order estimates.* For the hypothesized model, the Fairness construct failed to explain a significant amount of variance in the Fairness indicators. The factor loading for the Fairness indicator one was low (.15) and the factor loading for the Fairness Indicator 2 was improper (2.44). Factor loadings for the eight remaining indicators on the four associated factors (Concern for Others, Conscientiousness, Emotional Control, and Honesty) had significant T values and ranged from .70 for the Honesty indicator one to .92 for the Emotional Control indicator two (See Figure 3).



Table 13

*Factor Correlations for the Integrity Hypothesis Measurement Model*

Factor	CFO	CNS	ECO	FAR	HON
CFO	1.00				
CNS	.71*	1.00			
ECO	.68*	.37*	1.00		
FAR	-.02	.14	.08	1.00	
HON	.62*	.72*	.44*	.13*	1.00

*Note.* CFO = Concern for Others; CNS = Conscientiousness; ECO = Emotional Control; FAR = Fairness;  
HON = Honesty

\* Value is statistically significant ( $p < .001$ ).

Not surprisingly, a review of the measurement error variances demonstrated that the Fairness indicators were weak measures of their first-order factor, Fairness. As was expected, the measurement error variance for Fairness indicator one was high (.98) and the measurement error variance for Fairness indicator two was improper (-4.94). The eight remaining indicators also demonstrated significant measurement error variances. Measurement error variances for the remaining indicators ranged from .15 for the Emotional Control indicator two to .51 for the Honesty indicator one.

*Second order estimates.* Based on the first-order factor estimates, it was not surprising that the Integrity construct was unable to explain a significant amount of variance in the Fairness construct (.03). The factor loadings for the remaining factors on Integrity were statistically significant: .94 for Concern for Others, .77 for Conscientiousness, .64 for Emotional Control, and .72 for Honesty. Error variance for the Fairness factor was 1.00. A factor's error variance represents the amount of variance in the factor that cannot be explained by its relationship to the higher-order construct, in this case, Integrity. Error variances for the Conscientiousness, Honesty, and Emotional Control factors were also significant: .41 for Conscientiousness, .48 for Honesty, and .59 for Emotional Control.

*One factor model.* Fit statistics indicated that the one factor model failed to accurately describe the sample data. The factor loadings and measurement error variances supported this contention (See Figure 4).

All of the measurement error variances for the one factor model were significant and relatively high, ranging from .42 for the Concern for Others indicator two to 1.00 for the Fairness indicator one. While eight of the ten factor loadings were significant, the

magnitudes of the factor loadings were lower than those of the other models (See Figure 4).

### *Conclusions*

Fit statistics indicated that both the hypothesized and measurement models provided an excellent overall fit to the data. However, model comparisons indicated that the measurement model explained the sample data significantly better than the hypothesized model. These results were not surprising as the measurement model acts as an upper bound limit for the fit of the hypothesized model. Nonetheless, the hypothesized model was the most theoretically valid and parsimonious model. It was based on considerable research, fit statistics indicated an excellent fitting model, and the model accounted for the data with a single latent variable, Integrity. As such, I concluded that the hypothesized model of Integrity best described the sample data.

A review of the model estimates indicated that within the hypothesized model, the Fairness indicators were inadequate measures of the Fairness construct, and that Integrity was not accounting for a significant amount of variance in the Fairness factor. As a whole, the fit statistics and estimates provided partial support for the Integrity hypothesis. Of the five Integrity dimensions, Concern for Others, Conscientiousness, Emotional Control, and Honesty were confirmed as first order factors of a single, higher-order factor.

Based on the results, I carried out a second HCFA without the Fairness construct. Of course, post hoc analysis such as this may be criticized as taking advantage of chance.

Technically, a new set of data would be preferred. However, it appeared reasonable to carry out this next step provided that all conclusions were presented with caution. This revised model (without the Fairness construct) was used in the follow-up discriminant and social desirability analyses.

#### *Higher-Order Integrity Model without Fairness*

HCFA was used to assess the validity of the hypothesized model of Integrity without the Fairness construct, hereafter referred to as the four-factor model. This model contained four first order factors (i.e., Concern for Others, Conscientiousness, Emotional Control, and Honesty) loading on a single, higher-order factor Integrity. The indicators utilized in the previous analyses were again used as indicators of the first-order factors. The CFI, NNFI, and SRMSR measures of overall fit, along with factor loadings, measurement error variances, and factor correlations were used to assess the adequacy of the model.

For the fit statistics, a value of .97 was obtained for the CFI and a value of .94 was obtained for the NNFI, indicating a reasonable to excellent fit of the data to the model. The SRMSR for the model was .06, also representing a reasonably well fitting model (See Figure 5). The factor loadings for the first-order factors on Integrity were all significant: .94 for Concern for Others, .77 for Conscientiousness, .64 for Emotional Control, and .72 for Honesty. The error variances for the first-order factors were also all significant, ranging from .12 for Concern for Others to .59 for Emotional Control (See Figure 5).

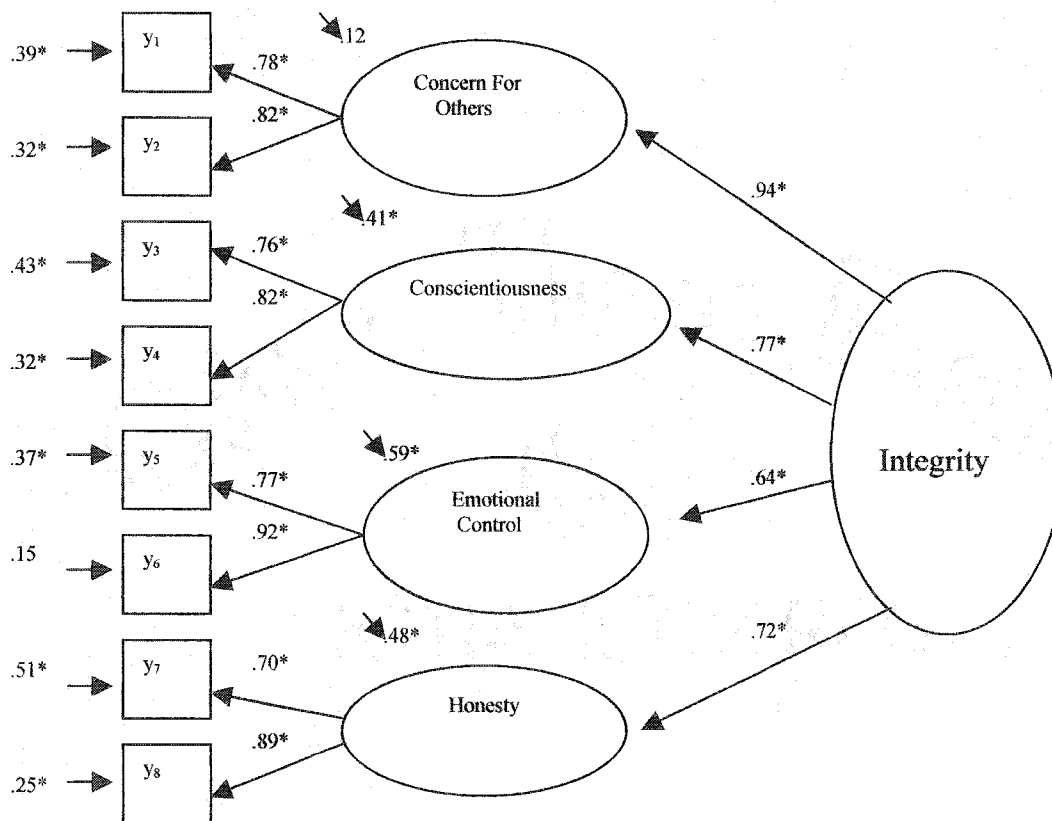


Figure 5. Four-Factor Model Estimates

Results for the second HCFA were similar to those found for the initial, hypothesized model. Overall fit statistics indicated a well fitting model. The four factors maintained significant and stable relationships with Integrity. As a result, the four-factor Integrity model was used as the comparison model for the discriminant validity and social desirability analyses.

### *Discriminant Hypothesis*

The discriminant hypothesis asserted that there was reasonable support for Substance Abuse and Anticipated Tenure as discriminant constructs of Integrity. The analytic strategy and results for this hypothesis follow.

### *Analytic Strategy*

To test the discriminant hypothesis, two models were to be compared for each discriminant construct. The baseline model was the four-factor Integrity model. The second, alternate model included five first-order factors (the four Integrity dimensions and a single discriminant construct) that loaded on a single higher-order factor, Integrity. As the alternate models contained two more indicators than the baseline, four-factor model, the models could not be compared statistically. Instead, a relative comparison was employed. Specifically, three overall fit statistics were computed for the two models: the NNFI, CFI, and SMRSR. As research indicates that many fit statistics are biased toward model complexity, statistics that control for degrees of freedom were

preferred (Bollen, 1989). Each of these statistics controlled for degrees of freedom, and thus, attenuated the potential bias toward model complexity (Bollen, 1989). The discriminant hypotheses were considered supported if the four-factor Integrity model provided a 'better' fit to the data than the alternate models.

### *Results*

A correlation matrix of the Integrity and discriminant indicators is provided in Table 12. A preliminary review of this matrix indicated that the Integrity dimensions were more highly correlated with each other than with the two discriminant constructs.

*Substance abuse.* In the current sample, the reported incidence of substance abuse was very low. Among the 254 participants, only 18 participants, or 7% of the sample reported any substance abuse behavior. While HCFA with Maximum Likelihood Estimation is highly robust to non-continuous data, it was determined that these data deviated too far from normality to make conclusions based on the analyses valid. Therefore, the discriminant hypothesis regarding substance abuse was not tested.

*Anticipated Tenure.* Using the 3-item scale, two indicators were created for the Anticipated Tenure construct. The mean of the first two items served as the first indicator. The third item was used as a second, single-item indicator of the construct. It should be noted that the latter, single-item indicator did not approach the 15 categories recommended for a variable to be classified as continuous (Joreskog & Sorbom, 1989). However, several researchers have verified the robustness of Maximum Likelihood Estimation procedures to data that deviate far from normality (Joreskog & Sorbom,

1989). Based on this research, the single item indicator, along with the indicator created from two items, was deemed appropriate for use in the analyses.

In HCFA, the validity of the first-order indicators is an upper-bound limit for any higher-order relationships (Marsh & Hovecar, 1988). Therefore, I first reviewed the ability of the indicators to measure the Anticipated Tenure construct. The factor loadings for the two indicators were both significant: .66 for the first indicator and .89 for the second. The first Anticipated Tenure indicator demonstrated significant measurement error variance (.57), while the other did not (.21). Overall, the Anticipated Tenure indicators were deemed adequate measures of the Anticipated Tenure construct.

As previously described the fit statistics for the four-factor model of Integrity indicated a reasonably well to excellent fitting model. For the Anticipated Tenure model, the fit statistics were as follows: .94 for the NNFI, .96 for the CFI, .06 for the SRMSR, and 87.54 (df=30) for the Chi-Square statistic. These statistics indicated a reasonably well fitting model. As Table 14 demonstrates, the fit statistics for this alternate model were very similar to those obtained for the four-factor model. The error variance for the Anticipated Tenure factor was statistically significant and high at .95. However, the construct maintained a significant loading on the Integrity construct (.22). Anticipated Tenure was correlated with the Integrity dimensions as follows: Concern for Others (.20), Conscientiousness (.17), Emotional Control (.14), and Honesty (.16). Only the Anticipated Tenure-Concern for Others correlation was statistically significant.

Results for the Anticipated Tenure analyses do not provide support for the discriminant validity hypothesis. Adding Anticipated Tenure as a first-order factor of Integrity only slightly reduced model fit. In addition, Integrity explained a small, yet



Table 14

*Fit Statistics for the Anticipated Tenure and Social Desirability Hypotheses*

Model	$\chi^2$	Df	NNFI	CFI	SRMSR
Four Factor	59.14	16	.94	.97	.06
Anticipated Tenure	87.54	30	.94	.96	.06
Social Desirability	60.91	16	.91	.95	.07

*Note.* The Social Desirability Model represents the Four-Factor Model analyzed with a covariance matrix in which Social Desirability has been partialled out. NNFI = Non-Normed Fit Index; CFI = Comparative Fit Index; SRMSR = Standardized Root Mean Square Residual

significant amount of variance in the Anticipated Tenure construct, indicating that the two constructs were indeed related.

### *Social Desirability Hypothesis*

The social desirability hypothesis asserted that social desirability and Integrity were conceptually and empirically distinct constructs. Specifically, the hypothesis asserted that social desirability did not significantly impact the factor structure of the Integrity construct.

### *Analytic Strategy*

To test this hypothesis, the four-factor model of Integrity was tested with a covariance matrix from which participants' social desirability scores were partialled. This model was then compared to the results obtained for the non-partialled four-factor model of Integrity. Participants' mean social desirability scale scores were partialled from the Concern for Others, Conscientiousness, Emotional Control, and Honesty indicators. The partialled and non-partialled four-factor models contained equal degrees of freedom and, thus, could not be compared statistically. Therefore, a relative comparison was employed. The partialled covariance matrix was compared to the non-partialled matrix on the previously described fit statistics: NNFI, CFI, and SRMSR. Factor loadings and measurement error variances were also investigated and compared.

## Results

The coefficient alpha for the social desirability scale was .71. The correlation matrix of the partialled indicators is provided in Table 15. As previously described, the fit statistics for the four-factor model of Integrity indicated a reasonably well fitting model. For the partialled four-factor model, the fit statistics were as follows: .91 for the NNFI, .95 for the CFI, .07 for the SRMSR, and 60.91 (df=16) for the Chi-Square statistic. These statistics also indicated a reasonably well fitting model. A comparison of model estimates follows.

*First-order model estimates.* For the partialled model, the factor loadings for the eight indicators were statistically significant and relatively high (See Table 16). Loadings ranged from .65 for the Honesty indicator one to .95 for the Emotional Control indicator two. As a whole, these loadings were slightly lower in value than those obtained for the non-partialled model. However, the ordering of the indicators for the partialled and non-partialled models was similar. For example, for both models, the two highest loadings were for the Emotional Control indicator two and Honesty indicator two. Likewise for both models, the two lowest loadings were for the Honesty indicator one and Conscientiousness indicator one.

The indicator measurement error variances for the partialled model were all statistically significant and ranged from .10 for the Emotional indicator two to .58 for the Honesty indicator one. These variances were slightly higher than those obtained for the

Table 15

*Correlation Matrix for the Integrity Indicators with Social Desirability Partialled Out*

Scale	Mean	SD	CFO1	CFO2	CNS1	CNS2	ECO1	ECO2	FAR1	FAR2	HON1	HON2
CFO1	4.38	.61	1.00									
CFO2	4.33	.53	.56*	1.00								
CNS1	4.27	.54	.30*	.35*	1.00							
CNS2	4.40	.44	.44*	.37*	.58*	1.00						
ECO1	3.49	.82	.24*	.36*	.15	.07	1.00					
ECO2	3.61	.80	.32*	.41*	.12	.18	.67*	1.00				
FAR1	1.43	.32	-.06	-.04	.07	-.01	.02	.03	1.00			
FAR2	1.35	.30	.02	-.02	.12	.05	.04	.03	.36*	1.00		
HON1	4.30	.62	.19	.21*	.43*	.37*	.11	.12	.07	.04	1.00	
HON2	4.50	.50	.36*	.32	.35*	.38*	.18	.23*	.07	.02	.55*	1.00

Note. CFO = Concern for Others; CNS = Conscientiousness; ECO = Emotional Control; FAR = Fairness; HON = Honesty

\* Value is statistically significant (p<.001).

Table 16

*Standardized Parameter Estimates for the Partialled and Non-Partialled Four Factor Models*

Indicator	Non-Partialled Estimates		Partialled Estimates	
	Factor Loading	Error Variance	Factor Loading	Error Variance
CFO1	.78*	.39*	.75*	.44*
CFO2	.82*	.32*	.75*	.44*
CNS1	.76*	.43*	.70*	.50*
CNS2	.82*	.32*	.82*	.32*
ECO1	.79*	.37*	.70*	.50*
ECO2	.92*	.15*	.95*	.10
HON1	.70*	.51*	.65*	.58*
HON2	.87*	.25*	.85*	.29*
CFO	.94*	.12	.88*	.23
CNS	.77*	.41	.75*	.44*
ECO	.64*	.59	.46*	.79*
HON	.72*	.48	.65*	.57*

*Note.* CFO = Concern for Others; CNS = Conscientiousness; ECO = Emotional Control; HON = Honesty

\* Value is statistically significant ( $p < .05$ ).

non-partialled matrix. Again, the ordering of the variances was similar for the partialled and non-partialled models. Honesty indicator one and Conscientiousness indicator one produced the highest measurement error variances in both models. The lowest measurement error variances in both models were for the Emotional Control indicator two and Honesty indicator two.

Just as in the non-partialled model, the partialled model indicated that all of the first-order factors were significantly correlated. However, the partial correlations were lesser in value than those obtained for the non-partialled model (See Table 17). Correlations ranged from .30 for Emotional Control-Honesty to .66 for Concern for Others-Conscientiousness.

*Second order model estimates.* In the partialled model, the four Integrity dimensions (Concern for Others, Conscientiousness, Emotional Control, and Honesty) continued to maintain significant loadings on the Integrity construct (See Table 16). When compared with the non-partialled model, the partialled loadings maintained the same ordering, but were lesser in value.

The error variances for the first-order factors in the partialled model were as follows: .23 for Concern for Others, .44 for Conscientiousness, .79 for Emotional Control, and .57 for Honesty (See Table 16). These variances maintained the same ordering, but were higher in value than those obtained for the non-partialled model.

*Conclusions.* Fit statistics indicated that the partialled four-factor Integrity model provided a reasonable fit to the sample data. Model estimates for the partialled model were similar, though not identical to those obtained in the non-partialled model. Noticeably different between the two models were the reduced factor loading values and

Table 17

*Factor Correlations for Four-Factor Model with Social Desirability Partialled Out*

Factor	CFO	CNS	ECO	HON	INT
CFO	1.00				
CNS	.66*	1.00			
ECO	.41*	.35*	1.00		
HON	.57*	.49*	.30*	1.00	
INT	.88*	.75*	.65	.46	1.00

*Note.* CFO = Concern for Others; CNS = Conscientiousness; ECO = Emotional Control, HON = Honesty;

INT = Integrity

\* Value is statistically significant ( $p < .05$ ).

the increased error variance values. These value changes suggested that social desirability might act as error variance in the four-factor model, masking true relationships among the Integrity dimensions.

In total, the results suggested that social desirability influenced the factor structure of the four-factor Integrity model. This was reflected in the factor loadings and error variances for the second order factor Integrity, fit statistics for the overall model, and correlations among the first-order factors. However, social desirability variance did not destroy the Integrity structure. Fit statistics on the partialled model still indicated a reasonably well fitting model and significant relationships among the first-order factors and Integrity were maintained. As a whole, the results provided partial support for the social desirability hypothesis.



## CONCLUSIONS

The results of this study support the contention that Integrity acts as a second-order factor with multiple first-order dimensions (Green, 1999). Four of the hypothesized first-order dimensions were confirmed in the study: Concern for Others, Conscientiousness, Emotional Control, and Honesty. The inadequacy of the Fairness measures made it impossible to test the relationship of this fifth dimension to the Integrity construct. A discriminant validity analysis failed to support the Integrity definition by indicating that Anticipated Tenure was significantly related to the Integrity construct. Also investigated was the influence of social desirability. Results indicated that social desirability influenced, but did not destroy the factor structure of the Integrity construct.

One of this study's limitations was the inadequate measurement properties of the Fairness indicators. The Fairness items were developed by Meglino et al. (1989) as one dimension of an ipsative, forced-choice work values scale. Past research supported the contention that non-ipsative measurement was possible when the scale was used to measure a single value – which is how the items were used in the present investigation (Meglino, 2000). However, ipsative items are by nature within-subjects measures. Thus, it is possible that the Fairness construct lost meaning and Integrity when used for parametric, between-subject comparisons.

Past research on the influence of social desirability on the factor structure of Integrity tests has yielded conflicting results. Some studies have reported considerable influence while others have reported none at all. For example, one study indicated that *partialing social desirability from correlations among the Big Five personality factors did not attenuate the factors' convergent or discriminant validity coefficients* (Ones and

Viswesvaran, 1998). However, Ellingson et al. (2001) reported that when participants were asked to respond in a socially desirable manner, a previously multi-dimensional personality measure was reduced to a single factor. The results of the present investigation fall in the middle of the two extremes. That is, while social desirability influenced the factor structure of the Integrity construct, the Integrity factors and relationships could not be completely explained by the social desirability variance.

These findings can be interpreted in light of the theory offered by Ellingson et al. (2001). These researchers assert that the way in which social desirability is operationalized moderates its relationship to Integrity. Specifically, the researchers assert that while Social Desirability can account for Integrity variance in experimental settings (i.e., when participants are asked to fake good), it fails to do so in applied organizational contexts. The way Social Desirability was operationalized in the present study falls between the true experimental design and the real-world applied setting. Participants were not instructed to respond in a contrived manner, nor were the test results used for organizational decisions. As a result, the modest influence of social desirability on the factor structure of Integrity is in line with the moderator theory presented by Ellingson et al. (2001).

The present investigation was a step towards identifying the true meaning of Integrity. Previous research has focused almost exclusively on the predictive validity of integrity tests. This study represented an important foray into the conceptual and semantic realm of Integrity. Future research should focus on continued testing and refinement of the hypothesized Integrity model. First, a new measure of Fairness is needed to assess this construct's relationship to Integrity. Second, the finding that

Anticipated Tenure was also related to Integrity suggests that the five hypothesized dimensions do not account for all of the true variance in Integrity. Research investigating the possibility of additional Integrity dimensions is warranted.

Integrity is critically important in the organizational setting. It has demonstrated significant predictive validity with multiple measures and manifestations of job performance. Further research into the definition and dimensionality of the construct will enhance our understanding of integrity, integrity's relationship to other performance indicators, and the similarities and differences evident among integrity instruments.

## REFERENCES

- Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*.  
Monterey: Brooks/ Cole.
- Andrews, B. P., & Lilienfeld, S. O. (1993). *Is a commonly used honesty test biased against highly religious individuals?* Unpublished manuscript.
- Ash, P. (1971). Screening employment applicants for attitudes toward theft. *Journal of Applied Psychology, 55*, 161-164.
- Ash, P. (1974). Convicted felons' attitudes towards theft. *Criminal Behavior and Justice, 1*, 1-8.
- Ash, P. (1991). *The Construct of Employee Theft Proneness*. Park Ridge, IL: SRA/ London House.
- Ballard, R. (1992). Short forms of the Marlowe-Crowne Social Desirability Scale. *Psychological Reports, 71*, 1155-1160.
- 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.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: John Wiley and Sons.

Bond, M. H., Nakazato, H., & Shiraishi, D. (1975). Universality and distinctiveness in dimensions of Japanese person perception. *Journal of Cross-Cultural Psychology, 6*, 346-357.

Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences, 11*(5), 515-524.

Borman, W. C., White, L. A., Pulakos, E. D., & Oppler, S. H. (1991). Models of supervisor job performance ratings. *Journal of Applied Psychology, 76*, 863-872.

Brand, C. R. (1984). Personality dimensions: an overview of modern trait psychology. In J. Nicholson, & H. Beloff (Eds.), *Psychology Survey* (pp. 309-337). Leicester, UK: British Psychological Society.

Byrne, B. M. (1998). Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. New Jersey: Lawrence Erlbaum Associates.

Camara, W. J., & Schneider, D. L. (1995). Questions of construct breadth and openness of research in integrity testing. *American Psychologist, 50*, 459-460.

Camman, C., Fichman, M., Jenkins, D., & Klesh, J. (1979). *The Michigan Organizational Assessment Questionnaire*. Unpublished manuscript, University of Michigan, Ann Arbor.

Cascio, W. F. (1991). *Applied psychology in personnel management* (4th ed.). New Jersey: Prentice Hall.

- Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, *12*, 197-220.
- Cattell, R. B., & Digman, J. M. (1962). A theory of the structure of perturbations in observer ratings and questionnaire data in personality research. *Behavioral Science*, *9*, 341-358.
- Chen, X. P., Hui, C., & Segó, D. J. (1998). The role of organizational citizenship behavior in turnover: Conceptualization and preliminary tests of key hypotheses. *Journal of Applied Psychology*, *83*, 922-931.
- Chidester, T. R., Helmreich, R. L., Gregorich, S. E., & Geis, C. E. (1991). Pilot personality and crew coordination: Implications for training and selection. *International Journal of Aviation Psychology*, *1*, 25-44.
- Cochran, T. K. P. (1991). *Preemployment integrity testing by selected retailers as a means of controlling employee theft*. Unpublished doctoral dissertation, Texas Women's College.
- Colarelli, S. M. & Dean, R. A. (1987). Comparative effects of personal and situational influences on job outcomes of new professionals. *Journal of Applied Psychology*, *72*, 558-566.
- Collins, J. D., & Schmidt, F. L. (1993). Personality, integrity, and white collar crime: A construct validity study. *Personnel Psychology*, *46*, 295-311.
- Costa, P. T., Jr., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five factor model. *Journal of Personality and Social Psychology*, *55*, 258-265.

Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd ed.). New York: Harper & Brothers.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Cunningham, M. R., & Ash, P. (1988). The structure of honesty: Factor analysis of the Reid Report. *Journal of Business and Psychology, 3*, 54-66.

Cunningham, M. R., Wong, D. T., & Barbee, A. P. (1994). Self-presentation dynamics on overt integrity tests: Experimental studies of the Reid Report. *Journal of Applied Psychology, 79*, 643-658.

Digman, J. M. (1988). *Classical theories of trait organization and the Big Five Factors of personality*. Presented at annual meeting of the American Psychological Association, Atlanta, GA.

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.

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

Driskell, J. E., Hogan, J., Salas, E., & Hoskin, B. (1994). Cognitive and personality predictors of training performance. *Military Psychology, 6*, 31-46.

Ellingson, J. E., Smith, D. B., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology, 86*, 122-133.

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

Frost, A. G., & Rafilson, F. M. (1989). Overt integrity tests versus personality-based measures of delinquency: An empirical comparison. *Journal of Business and Psychology, 3*, 269-278.

George, J. M., & Jones, G. R. (1996). The experience of work and turnover intentions: Interactive effects of value attainment, job satisfaction, and positive mood. *Journal of Applied Psychology, 81*, 318-325.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26*, 461-477.

Ghiselli, E. E., & Barthol, R. P. (1953). The validity of personality inventories in the selection of employees. *The Journal of Applied Psychology, 37(1)*, 18-20.

Goldberg, L. R. (1981). Language and individual differences: the search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (pp. 141-165). Beverly Hills, CA: Sage.

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

Goldberg, L. R., Grenier, J. R., Guion, R. M., Sechrest, L. B., & Wing, H. (1991). *Questionnaires used in the prediction of trustworthiness in pre-employment selection*



*decisions: An APA Task Force report*, Feb. 1991. Washington, DC: American Psychological Association, Inc.

Gough, H. G. (1972). *Manual for the Personnel Reaction Blank*. Palo Alto, CA: Consulting Psychologists Press.

Green, A. P. (1999). [Critical incident interviews: How high and low integrity is demonstrated at work]. Unpublished raw data.

Guion, R. M. (1965). *Personnel Testing*. New York: McGraw Hill.

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

Harris, M. M., & Sackett, P. R. (1987). A factor analysis and item response theory analysis of an employee honesty test. *Journal of Business and Psychology*, 2, 122-135.

Harris, W. G. (1987). A components analysis of a preemployment integrity measure: A replicated study. *Psychological Reports*, 60, 1051-1055.

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

Hogan, J., & Brinkmeyer, K. (1997). Bridging the gap between overt and personality-based integrity tests. *Personnel Psychology*, 50, 587-599.

Hogan, R., Curphy, G. J., & Hogan, H. (1994). What we know about leadership: Effectiveness and leadership. *American Psychologist*, 49, 493-504.

Hogan, J., & Hogan, R. (1989). How to measure employee reliability. *Journal of Applied Psychology*, 74, 273-279.

- Hogan, J., Hogan, R., & Busch, C. M. (1984). How to measure service orientation. *Journal of Applied Psychology, 69*, 167-173.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist, 51*, 469-477.
- Hogan, J. & Ones, D. S. (1997). Conscientiousness and Integrity at work. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 849-870). San Diego: Academic Press.
- Hollinger, R. C. (1988). Working under the influence (WUI): Correlates of employees' use of alcohol and other drugs. *Journal of Applied Behavioral Science, 24*, 439-454.
- Hunter, J. E. (1980). *Validity generalization for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB)*. Washington, DC: U.S. Department of Labor, U.S. Employment Service.
- Hunter, J. E. & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.
- International Personality Item Pool (2001). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (<http://ipip.ori.org/>). Internet Web Site.
- John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 261-271). New York: Springer-Verlag.
- Jones, J. W. (1980). Attitudinal correlates of employees' deviance: Theft, alcohol use, and nonprescribed drug use. *Psychological Reports, 47*, 71-77.

Jones, J. W., Joy, D. S., & Martin, S. L. (1990). A multidimensional approach for selecting child care workers. *Psychological Reports, 67*, 542-553.

Jones, J. W., Joy, D. S., Werner, S. H., Orban, J. A. (1991). Criterion-related validity of a preemployment integrity inventory: A large scale between groups comparison. *Perceptual and Motor Skills, 72*, 131-136.

Jones, J. W., & Terris, W. (1983). Predicting employee theft in home improvement centers. *Psychological Reports, 52*, 187-201.

Joreskog, K. G., & Sorbom, D. (1989). *LISREL 7: A guide to the program and applications* (2<sup>nd</sup> ed.). Chicago, IL: SPSS.

Joreskog, K. G. & Sorbom, D. (2001). LISREL (Version 8.52) [Computer Software]. Lincolnwood, IL: Scientific Software International.

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

Kochkin, S. (1987). Personality correlates of a measure of honesty. *Journal of Business and Psychology, 1*, 236-247.

Landy, F. J. (1986). Stamp collecting versus science: Validation as hypothesis testing. *American Psychologist, 41*, 1183-1192.

Lasson, E. D. (1992). *Preemployment honesty testing: Construct validity issues and a test of the person-situation question*. Unpublished doctoral dissertation, Wayne State University.

Lilienfeld, S. O. (1994). Response to comment from D. W. Arnold. *Skeptical Inquirer, 18*, 320-321.

Lilienfeld, S. O., Andrews, B. P., & Stone-Romero, E. F. (1994). The relations between a self-report honesty test and personality measures in prison and college samples. *Journal of Research in Personality, 28*, 154-169.

Logan, T. G., Koettel, R. C., & Moore, R. W. (1986). Personality correlates of a test of honesty. *Psychological Reports, 59*, 1015-1018.

Lorr, M., & Youniss, R. P. (1974). An inventory of interpersonal style. *Journal of Personality Assessment, 37*, 165-173.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness of fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.

Marsh, H. W., & Hovecar, D. (1988). A new, more powerful approach to multitrait-multimethod analyses: Application of second-order confirmatory factor analysis. *Journal of Applied Psychology, 73*, 107-117.

McCrae, R. R. (1989). Why I advocate the five-factor model: Joint factor analyses of the NEO-PI with other instruments. In D. M. Buss, & N. Cantor (Eds.), *Personality Psychology* (pp. 237-245). New York: Springer-Verlag.

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

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

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175-215.

McDaniel, M. A., & Jones, J. W. (1988). Predicting employee theft: A quantitative review of a standardized measure of honesty. *Journal of Business and Psychology, 2*, 327-345.

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85*, 812-821.

Meglino, B. M. (2000). *Use of the Comparative Emphasis Scale*. Retrieved December 15, 2000 from University of South Carolina, Darla Moore School of Business Web site: <http://129.252.51.50/faculty/vita/comparuse.htm>.

Meglino, B. M., Ravlin, E. C., & Adkins, C. L. (1989). A work values approach to corporate culture: A field test of the value congruence process and its relationship to individual outcomes. *Journal of Applied Psychology, 74*, 424-432.

Meglino, B. M., Ravlin, E. C., & Adkins, C. L. (1992). The measurement of work value congruence: A field study comparison. *Journal of Management, 18*, 33-43.

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.

Murphy, K. R. (1990). *Performance appraisal: An organizational perspective*. Boston: Allyn & Bacon.

Murphy, K. R. (1993). *Honesty in the workplace*. Pacific Grove, CA: Brooks/Cole.

Murphy, K. R., & Lee, S. L. (1994). Personality variables related to integrity test scores: The role of conscientiousness. *Journal of Business and Psychology*, 8, 413-424.

Nolan, K. A. (1991). *The relationship between socially desirable responding and personnel integrity testing: The unique effect of self-deception and impression management*. Unpublished master's thesis, Georgia State University, Atlanta, GA.

Noller, P., Law, H., & Comrey, A. L. (1987). Cattell, Comrey, and Eysenck personality factors compared: More evidence for the five robust factors? *Journal of Personality and Social Psychology*, 53, 775-782.

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

O'Bannon, R. M., Goldinger, L., & Appleby, G. S. (1989). *Honesty and integrity testing: A practical guide*. Atlanta, GA: Applied Information Resources.

Ones, D. S. (1993). *The construct validity of integrity tests*. Unpublished doctoral dissertation, University of Iowa, Iowa City.

Ones, D. S. & Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology*, 83, 35-42.

Ones, D. S. & Viswesvaran, C. (2001). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. *International Journal of Selection and Assessment*, 9, 31-39.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660-679.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*(4), 679-703.

Paaianen, G. E., Hansen, T. L., & McLellan, R. A. (1993). *PDI Employment Inventory and PDI Customer Service Inventory Manual*. Minneapolis, MN: Personnel Decisions, Inc.

Personnel Decisions, Inc. (1985). *Development and validation of the PDI Employment Inventory*. Minneapolis, MN: Author.

Personnel Selection Inventory (1977). Parkridge, IL: London House: Author

Ravlin, E. C., & Meglino, B. M. (1987). Issues in work values measurement. In W. C. Frederick (Ed.) *Research in Corporate Social Performance and Policies, Vol. 9* (p. 153-183).

The Reid Report (1984). Chicago, IL: Reid Psychological Systems.

Reilly, R. R., & Chao, G. R. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology, 35*, 1-62.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38*, 119-125.

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second order factor analysis. *Multivariate Behavioral Research, 23*, 51-67.

Rosse, J. G., Miller, H. E., & Barnes, L. K. (1991). Combining personality and cognitive ability predictors for hiring service-oriented employees. *Journal of Business and Psychology, 5*, 431-445.

Sackett, P. R. (2002). The structure of counterproductive work behaviors: Dimensionality and relationships with facets of job performance. *International Journal of Selection and Assessment, 10*, 5-11.

Sackett, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology, 42*, 491-529.

Sackett, P. R., & Harris, M. M. (1984). Honesty testing for personnel selection: A review and critique. *Personnel Psychology, 37*, 221-245.

Sackett, P. R., & Wanek, J. E. (1996). New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel Psychology, 49*, 787-829.

Schmidt, F. L., Viswesvaran, C., & Ones, D. S. (1997). Validity of integrity tests for predicting drug and alcohol abuse: A meta-analysis. *National Institute on Drug Abuse Research Monograph Series, 170*, 69 – 95.

Skinner, H. A. (1982). The drug abuse screening test. *Addictive Behaviors, 7*, 363-371.

Terris, W., & Jones, J. (1982). Psychological factors related to employees' theft in the convenience store industry. *Psychological Reports, 51*, 1219-1238.

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



Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality, 60*, 225-251.

U.S. Congress, Office of Technology Assessment (1990). *The use of integrity tests for pre-employment screening* (OTA-SET-442). Washington, DC: U.S. Government Printing Office.

Viswesvaran, C. (2002). Absenteeism and measures of job performance: a meta-analysis. *International Journal of Selection and Assessment, 10*, 12-17.

Wanek, J. E. (1995). *The construct of integrity: Item-level factor analysis of the dimensions underlying honesty testing and Big-Five measures of personality*.

Unpublished doctoral dissertation, University of Minnesota.

Williams, M., & Sanchez, J. I. (1997). Customer service-oriented behavior: Person and situational antecedents. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Missouri, April.

Woolley, R. M., & Hakstian, A. R. (1992). An examination of the construct validity of personality-based and overt measures of integrity. *Educational and Psychological Measurement, 52*, 475-489.

Yoon, K., Schmidt, F., & Ilies, R. (2002). Cross-cultural construct validity of the five-factor model of personality among Korean employees. *Journal of Cross-Cultural Psychology, 33*, 217-235.

Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items. *Journal of Applied Psychology, 84*, 551-563.

## APPENDIXES

## Appendix A

## CRITICAL INCIDENT INTERVIEW

## Participant Background Information

INSTRUCTIONS: Please complete the following background items. Thank-you for your participation!

1. Are you currently employed? \_\_\_\_\_
2. If you answered YES to Question 1, how long have you worked for this company?  
\_\_\_\_\_
3. What is your current position in this company? \_\_\_\_\_
4. What type of work do you perform in this company? \_\_\_\_\_
5. How long have you worked in this line of work? \_\_\_\_\_
6. How many hours do you usually work in a week? \_\_\_\_\_
7. What is your ETHNICITY/ RACE? \_\_\_\_\_
8. Are you MALE or FEMALE? \_\_\_\_\_
9. What is your AGE? \_\_\_\_\_

Interview

INSTRUCTIONS TO PARTICIPANT: Think about the job you are currently performing. I'd like you to take a moment and think of an instance when you, someone you know, or someone you heard about demonstrated high or low integrity. Just let me know when you've thought of something and I'd like to ask you questions about it. (Wait for person to think of incident.)

**If participant requests a definition:** Actually, we're most interested in what you think about when you hear the word integrity; not what the dictionary or someone else thinks. We consider you the expert.

**If participant requests a second time:** The dictionary defines integrity as morally correct as evidenced in character and actions; strict regard for what is right.

Check if definition is given: \_\_\_\_\_

## PART I

---

A. Example Number: 1 2 3 4 5

B. Did this incident occur in your current job? YES NO

C. IF NO:

Did this incident occur in a previous job of yours? YES NO

Is this an incident that you heard about from someone else? YES NO

What type of work was being performed? \_\_\_\_\_

## PART II

---

1. What were the circumstances leading up to this incident?

- Can you be more specific?
- Who was involved?
- Was this something you witnessed or heard about?
- What happened exactly?

2. What did the employee do that makes you think he has high or low integrity?

- Can you be more specific?
- What was this employee's position?

3. What were the consequences of the employee's behavior in this incident?

- Can you be more specific?
- What else happened?
- Is that all? Did anything happen later?
- Can you still see, feel, the effects?

4. On a scale from one to five, with one being a poor example and 5 being a perfect example, how good of an example do you think this is of High/Low integrity?

1

2

3

4

5

POOR EXAMPLE

PERFECT EXAMPLE

5. Can you think of another time when you, someone you know, or someone you heard about demonstrated high or low integrity?

YES

NO

## Appendix B

## CONCERN FOR OTHERS ITEM SELECTION SURVEY

**THANK-YOU FOR JOINING IN THIS STUDY!**  
**YOUR TIME AND EXPERTISE ARE GREATLY APPRECIATED.**

**INSTRUCTIONS:** *The personality construct CONCERN FOR OTHERS is defined below. Please review the definition, along with the adjectives used to describe the positive and negative pole of this construct.*

*On the following page, you will find a list of 20 to 30 items. Please rate the relevance of each item to the CONCERN FOR OTHERS construct, as it has been defined. In other words, we are asking you to tell us how well you believe each item assesses an individual's CONCERN FOR OTHERS.*

*All ratings are completed on a 5-pt. scale ranging from (1) barely relevant to (5) extremely relevant. A rating of '1' indicates that an individual's response to this item would tell us very little about this person's concern for others. A rating of '5' indicates that an individual's response to the item would tell us a great deal about the person's concern for others. Please check the box NOT RELATED if the item is completely unrelated to the CONCERN FOR OTHERS construct. You may use this page as a reference while you rate the individual items.*

*It should be noted that if the items on the following pages are used in future personality assessments, individual items would require different response formats (e.g., objective, forced choice, or Likert-type ratings). The response format that will be used in future personality assessments is indicated by the two capital letters in parentheses following each item. This information may or may not assist you in rating the relevance of individual items.*

*(LT) = Likert-type      (OB) = Objective      (FC) = Forced Choice*

### CONCERN FOR OTHERS:

Acting in a manner that promotes positive interactions and cooperation, and demonstrates a concern for others, especially as such behaviors contribute to work goal accomplishment.

(+) positive/we can do attitude cooperative polite empathetic	(-) negative/ pessimistic attitude uncooperative rude not interested/ affected by others
---	--

ITEM	RELEVANCE-RATING					NOT RELEVANT
	1 Slightly Relevant	2	3 Moderately Relevant	4	5 Extremely Relevant	
<b>SAMPLE:</b>						
<b>Ski often (LT)</b>	1	2	3	4	5	√
Am on good terms w/ nearly everyone (LT)	1	2	3	4	5	
Have no time for others (LT)	1	2	3	4	5	
Approach others in a positive manner (LT)	1	2	3	4	5	
Don't put much faith in the opinions of others (LT)	1	2	3	4	5	
Like to be of service to others (LT)	1	2	3	4	5	
Am able to cooperate with others (LT)	1	2	3	4	5	
Lay down the law to others (LT)	1	2	3	4	5	
Can't be bothered with others' needs (LT)	1	2	3	4	5	
Acknowledge others' accomplishments (LT)	1	2	3	4	5	
Respect the opinion of others (LT)	1	2	3	4	5	
Look down on others (LT)	1	2	3	4	5	
Working with others is usually more trouble than it's worth (LT)	1	2	3	4	5	

There's not always time to be kind to others (LT)	1	2	3	4	5	
Am easy to live with (LT)	1	2	3	4	5	
Point out others shortcomings (LT)	1	2	3	4	5	
Inquire about others' well being (LT)	1	2	3	4	5	
Have a sharp tongue (LT)	1	2	3	4	5	
Am polite to strangers (LT)	1	2	3	4	5	
Am quick to judge others (LT)	1	2	3	4	5	
Cut others to pieces (LT)	1	2	3	4	5	
Appreciate people who wait on me (LT)	1	2	3	4	5	
Sympathize with the homeless (LT)	1	2	3	4	5	
Put down others' ideas (LT)	1	2	3	4	5	
Try to outdo others (LT)	1	2	3	4	5	
Hang up the phone on people (LT)	1	2	3	4	5	



## Appendix C

## ITEM SELECTION SURVEY ITEMS

(LT) = Likert-type (OB) = Objective (FC) = Forced Choice

## CONCERN FOR OTHERS

Acting in a manner that promotes cooperation, and demonstrates a concern for others, especially as such behaviors contribute to work goal accomplishment.

- |   |      |
|---|------|
| 1. Am on good terms w/ nearly everyone                          | (LT) |
| 2. Have no time for others                                      | (LT) |
| 3. Approach others in a positive manner                         | (LT) |
| 4. Don't put much faith in the opinions of others               | (LT) |
| 5. Like to be of service to others                              | (LT) |
| 6. Am able to cooperate with others                             | (LT) |
| 7. Lay down the law to others                                   | (LT) |
| 8. Can't be bothered with others' needs                         | (LT) |
| 9. Acknowledge others' accomplishments                          | (LT) |
| 10. Respect the opinion of others                               | (LT) |
| 11. Look down on others   | (LT) |
| 12. Working with others is usually more trouble than it's worth | (LT) |
| 13. There's not always time to be kind to others                | (LT) |
| 14. Am easy to live with  | (LT) |
| 15. Point out others shortcomings                               | (LT) |
| 16. Inquire about others' well being                            | (LT) |
| 17. Have a sharp tongue   | (LT) |
| 18. Am polite to strangers                                      | (LT) |
| 19. Am quick to judge others                                    | (LT) |
| 20. Cut others to pieces  | (LT) |
| 21. Appreciate people who wait on me                            | (LT) |
| 22. Sympathize with the homeless                                | (LT) |
| 23. Put down others' ideas                                      | (LT) |
| 24. Try to outdo others   | (LT) |
| 25. Hang up the phone on people                                 | (LT) |

## CONSCIENTIOUSNESS

Tendency to meet and exceed work expectations

- |                                     |      |
|-------------------------------------|------|
| 1. Get to work at once              | (LT) |
| 2. Try to follow the rules          | (LT) |
| 3. Pay my bills on time             | (LT) |
| 4. Get others to do my duties       | (LT) |
| 5. Do the opposite of what is asked | (LT) |

6. Go straight for the goal (LT)
7. Work hard (LT)
8. Know how to get around the rules (LT)
9. Would never cheat on my taxes (LT)
10. Do more than what's expected of me (LT)
11. Have difficulty starting tasks (LT)
12. Get things done quickly (LT)
13. Do dangerous things (LT)
14. Set high standards for myself and others (LT)
15. Do just enough work to get by (LT)
16. Need a push to get started (LT)
17. Hang around doing nothing (LT)
18. Am always busy (LT)
19. Work on improving myself (LT)
20. Excel in nothing at all (LT)
21. Pay attention to details (LT)
22. Accomplish my work on time (LT)
23. Have difficulty starting tasks (LT)
24. Am often late to work (LT)
25. Am careful to avoid making mistakes (LT)
26. Check over my work (LT)
27. Follow through on my commitments (LT)
28. Excel in what I do (LT)

#### EMOTIONAL CONTROL

Controlling and displaying emotions in a professional and non-destructive manner

1. Keep my cool (LT)
2. Get irritated easily (LT)
3. Use swear words (LT)
4. Act quickly without thinking (LT)
5. Snap at people (LT)
6. Get angry easily (LT)
7. Remain calm under pressure (LT)
8. Handle tough tasks smoothly (LT)
9. Do things I later regret (LT)
10. Get stressed out easily (LT)
11. Make rash decisions (LT)
12. Panic easily (LT)
13. Take offense easily (LT)
14. Lose my temper (LT)
15. Am calm even in tense situations (LT)
16. Don't lose my head (LT)
17. Can stand criticism (LT)
18. Keep my emotions under control (LT)

- 19. Shoot my mouth off (LT)
- 20. Am able to control my emotions (LT)
- 21. Lash out physically when I'm angry (LT)

### FAIRNESS

Acting consistently across people and times.

- 1. Care about justice (FC)
- 2. Act at the expense of others (LT)
- 3. Listen to others viewpoints (LT)
- 4. Am able to settle disagreements (LT)
- 5. Do not believe you can be fair to others all the time (FC)
- 6. Believe that some people deserve to be treated differently (LT)
- 7. Impartial in dealing with others (FC)
- 8. Consider different viewpoints before taking action (LT)
- 9. Make decisions which are fair to all concerned (FC)
- 10. Impartial in judging disagreements (FC)
- 11. Give everyone an equal opportunity to work (FC)
- 12. Judge people fairly based on their abilities rather than their personalities only (FC)
- 13. Make sure that work assignments are fair to everyone (FC)
- 14. Attempt to bring out a fair solution to disputes (FC)
- 15. Give rewards in the fairest way possible (FC)
- 16. Provide fair treatment for all employees (FC)
- 17. Ensure each employee has an equal chance to get rewards (FC)

### HONESTY

Acting in a truthful manner at all times.

- 1. Return borrowed items (LT)
- 2. Trust others (LT)
- 3. Return extra change when a cashier makes a mistake (LT)
- 4. Do things behind other people's backs (LT)
- 5. Cheat to get ahead (LT)
- 6. Believe that people seldom tell the whole truth (LT)
- 7. Try to fool others (LT)
- 8. Stand behind my actions (FC)
- 9. Tell the truth (FC)
- 10. Break my promises (LT)
- 11. How much in merchandise or goods have you stolen from your employers/ jobs in the last five years? (OB)
- 12. How much money have you stolen from your employers/ jobs in the last five years? (OB)
- 13. Have you taken anything from a store in the last five years without paying for it? (OB)

14. What percentage of employees do you think steal from their companies? (LT)
15. Take actions which represent my true feelings (LT)
16. Speak my mind even my views aren't popular (LT)
17. Express my true opinion when asked (LT)
18. Admit to my mistakes (FC)
19. Truthful in dealing with others (FC)
20. Accept the consequences for my errors (FC)
21. Take credit for others' ideas (FC)
22. Admit responsibility for errors made (FC)
23. Lie to make myself look good (LT)
24. Take a stand for what I believe in (LT)
25. Will not do something I think is wrong (FC)
26. Hold true to my convictions (FC)

## Appendix D

MEAN RELEVANCE RATINGS FOR CONCERN FOR OTHERS ITEMS ON THE  
ISS<sup>a</sup>

Item	Mean	SD
<b>Items Included in Final Survey</b>		
Can't be bothered with others' needs	5.00	0.00
Like to be of service to others	4.50	0.76
Inquire about others' well being	4.13	1.13
Approach others in a positive manner	4.13	0.83
There's not always time to be kind to others	4.00	0.93
Cut others to pieces	3.88	0.99
Sympathize with the homeless	3.75	0.71
Am polite to strangers	3.75	0.89
Am able to cooperate with others	3.75	1.67
Look down on others	3.63	1.41
Hang up the phone on people	3.50	1.20
Acknowledge others' accomplishments	3.50	1.31
<b>Items Not Included in Final Survey</b>		
Am on good terms with nearly everyone	3.38	0.74
Put down others' ideas	3.38	0.92
Don't put much faith in the opinions of others	3.38	0.74

---

Item	Mean	SD
Have a sharp tongue	3.13	0.83
Appreciate people who wait on me	3.00	1.20
Point out others' shortcomings	3.00	0.76
Working with others is usually more trouble than it's worth	3.00	1.31
Lay down the law to others	2.75	1.39
Try to outdo others	2.63	0.92
Am quick to judge others	2.38	1.06
Am easy to live with	2.38	1.19

---

<sup>a</sup> Item Selection Survey

## Appendix E

MEAN RELEVANCE RATINGS FOR CONSCIENTIOUSNESS ITEMS ON THE ISS<sup>a</sup>

Item	Mean	SD
<b>Items Included in Final Survey</b>		
Work hard	4.83	0.41
Follow through on my commitments	4.67	0.82
Accomplish my work on time	4.67	0.82
Set high standards for myself and others	4.50	0.84
Check over my work	4.33	1.21
Pay attention to details	4.17	1.17
Try to follow the rules	4.00	1.10
Work on improving myself	4.00	1.10
Do more than what's expected of me	3.83	1.60
Am often late to work	3.83	1.60
Do just enough work to get by	3.83	1.60
Pay my bills on time	3.83	.98
<b>Items Not Included in Final Survey</b>		
Excel in what I do	3.83	1.47
Need a push to get started	3.83	1.47
Get to work at once	3.83	0.75

Item	Mean	SD
Have difficulty starting tasks	3.50	1.38
Would never cheat on my taxes	3.50	1.22
Go straight for the goal	3.50	1.38
Do the opposite of what is asked	3.50	1.52
Have difficulty starting tasks	3.33	1.21
Get others to do my duties	3.17	1.47
Get things done quickly	3.00	1.26
Am always busy	2.83	1.17
Know how to get around the rules	2.67	1.37
Excel in nothing	1.83	1.83
Do dangerous things	1.33	1.21

<sup>a</sup> Item Selection Survey



## Appendix F

MEAN RELEVANCE RATINGS FOR EMOTIONAL CONTROL  
ITEMS ON THE ISS<sup>a</sup>

Item	Mean	SD
<b>Items Included in Final Survey</b>		
Keep my emotions under control	5.00	0.00
Am calm even in tense situations	4.86	0.38
Keep my cool	4.86	0.38
Get angry easily	4.71	0.49
Get physical when I'm angry	4.57	0.79
Snap at people	4.43	0.53
Get irritated easily	4.43	0.53
Panic easily	4.29	0.49
Get stressed out easily	4.00	1.41
Take offense easily	3.57	1.27
Shoot my mouth off	3.43	1.27
Act quickly without thinking	3.43	1.27
<b>Items Not Included in Final Survey</b>		
Can stand criticism	3.43	.98
Make rash decisions	3.14	.69
Do things I later regret	3.00	1.00

Item	Mean	SD
Handle tough tasks smoothly	2.86	1.35
Use swear words	2.86	1.77

<sup>a</sup>Item Selection Survey

## Appendix G

MEAN RELEVANCE RATINGS FOR FAIRNESS ITEMS ON THE ISS<sup>a</sup>

Item	Mean	SD
<b>Items Included in Final Survey</b>		
Impartial in dealing with others <sup>b</sup>	4.88	0.35
Impartial in judging disagreements <sup>b</sup>	4.63	0.52
Give everyone an equal opportunity to participate <sup>b</sup>	4.38	0.74
Judge people based on their abilities rather than their personalities <sup>b</sup>	3.75	1.04
Give rewards based on performance <sup>b</sup>	3.75	0.71
Consider different viewpoints before taking action <sup>b</sup>	3.50	1.77
<b>Items Not Included in Final Survey</b>		
Believe that some people deserve to be treated Differently	4.63	0.52
Care about justice	4.37	1.06
Take into account many perspectives when making Decisions	3.50	1.77
Do not believe you can be fair to others all the time	3.50	0.76

---

Item	Mean	SD
Act at the expense of others	3.38	1.06
Listen to others' view points	3.12	1.55
Can bring about a win/win solution to disputes	3.00	1.51
Am able to settle disagreements	2.25	0.89

---

<sup>a</sup> Item Selection Survey

<sup>b</sup> Items from the Comparative Emphasis Scale

## Appendix H

MEAN RELEVANCE RATINGS FOR HONESTY ITEMS ON THE ISS<sup>a</sup>

Item	Mean	SD
<b>Items Included in Final Survey</b>		
Tell the truth	5.00	0.00
Lie to make myself look good	4.80	0.45
Admit responsibility for errors made	4.80	0.45
Truthful in dealing with others	4.80	0.45
Cheat to get ahead	4.80	0.45
Do things behind other people's backs	4.80	0.45
Have you taken anything from a store in the last five years without paying for it?	4.60	0.89
How much money have you stolen in the last five years?	4.60	0.89
Return extra change when a cashier makes a mistake	4.60	0.55
Take credit for others' ideas	4.40	0.55
Express my true opinion when asked	4.40	0.55
Stand behind my actions	4.20	0.84
Break my promises	4.00	1.22

Item	Mean	SD
<b>Items Not Included in Final Survey</b>		
How much in merchandise or goods have you stolen in the last five years?	4.40	0.89
Hold true to my convictions	3.80	1.10
Take actions which represent my true feelings	3.80	0.84
Try to fool others	3.60	1.67
Speak my mind even if my views aren't popular	3.40	0.89
Will not do something I think is wrong	3.20	0.84
Accept the consequences of my errors	3.20	2.05
Return borrowed items	3.20	1.10
What percentage of employees do you think steal from their companies?	2.80	1.92
Believe that people seldom tell the whole truth	2.40	1.82
Take a stand for what I believe in	2.40	1.48
Trust others	1.40	1.14

<sup>a</sup> Item Selection Survey

## Appendix I

PROCEDURE FOR CREATING SUB-SCALES FOR MAXIMUM LIKELIHOOD  
ESTIMATION IN STRUCTURAL EQUATION MODELING

Subscales are critical in maximum likelihood estimation because they create multiple indicators for each latent variable. Multiple indicators allow LISREL to estimate the measurement error variance of the sample variance covariance matrix (Bollen, 1989). The procedure for creating subscales is outlined below.

1. Submit all items for a given scale to a principal components factor analysis.
2. To create two subscales, use the factor loadings and assign the following items to the first subscale:

- Item with the highest loading
- Item with the lowest loading
- Item with the fourth highest loading
- Item with the fourth lowest loading

3. Examine the factor loadings and assign the following items to the second subscale:

- Item with the second highest loading
- Item with the second lowest loading
- Item with the third highest loading
- Item with the third lowest loading

4. Assign any remaining items randomly to either the first or second subscale, alternating between the two scales.

## VITA

ARLENE P. GREEN

Department of Psychology  
 Old Dominion University  
 Norfolk, VA 23509

## EDUCATION

***BA, Psychology***

James Madison University, May 1994

***MS, Psychology***

Old Dominion University, December 1996

***Ph.D., Industrial/ Organizational Psychology***

Old Dominion University, August 2003

## TEACHING EXPERIENCE

***Instructor***

Introduction to Psychology, Old Dominion University, Fall 1997

***Lab Instructor***

Experimental Research, Old Dominion University, Summer 1996, Summer 1997

## PUBLICATION

***Kelley, M. L., Smith, T. S., Green, A. P., Berndt, A. E., & Rogers, M. C. (1998).***  
 Importance of fathers' parenting to African-American toddlers' social and  
 cognitive development. *Infant Behavior and Development*, 21, 733-744.

## PRESENTATIONS

***Bernas, K. H., Pace (Green), A., & McIntyre, R. M. (Spring 1998)***

Another Look at the Prevalence of Data Sharing  
 Conference for the Society of Industrial and Organizational Psychologists

***Pace (Green), A. & Bartle, S. A. (Spring 1997)***

Understanding the Importance of Levels of Analysis  
 Graduate Student Conference of I/O and O/B Psychologists

## AWARDS AND HONORS

***Commonwealth Fellowship***

State Council for the Higher Education of Virginia  
 Fall 1994 – Spring 1995, Fall 1995 – Spring 1996

***Outstanding Minority Student Doctoral Assistantship***

Old Dominion University, Norfolk, VA  
 Fall 1996 – Spring 1997