College of Business Administration
Department of Management and Organizations
University of Iowa
Iowa City, IA 52242

FINAL REPORT

Establishing Construct Validity for Integrity Tests

by

Deniz S. Ones

DTIC
ELECTE
MAY 0 8 1995
S D
G

June 1993

Prepared for:

Defense Personnel Security Research Center

Under the

Office of Naval Research
N00014-92-J-4041

Approved for Public Distribution: Distribution Unlimited

Review of this material does not imply Department of Defense indorsement of factual accuracy or opinion.

19950505 173

DTIC QUALITY INSPECTED 5

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	Unlimited
4 PERFORMING ORGANIZATION REPORT NUMBER(S) 92-1	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION University of Iowa; 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Defense Personnel Security Research and Education Center
6c ADDRESS (City, State, and ZIP Code) Iowa City, Iowa 52242	7b ADDRESS (City, State, and ZIP Code) 990 Pacific St., Suite 455-E Monterey, CA 93940
8a NAME OF FUNDING/SPONSORING ORGANIZATION Defense Personnel Security Research & Education Center; 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-92-J-4041
8c ADDRESS (City, State, and ZIP Code) 99 Pacific St., Suite 455-E Monterey, CA 93940	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO; PROJECT NO 36549; TASK NO; WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)
Establishing Construct Validity for Integrity Tests

12 PERSONAL AUTHOR(S)
Ones, D.S.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM 9/01/92 TO 6/30/93	June 1993	197

16 SUPPLEMENTARY NOTATION
This report was prepared under an Office of Naval Research Contract

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Personnel selection; personnel integrity & reliability; personnel security; meta-analysis; selection validity; conscientiousness; construct validity

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This dissertation presents extensive evidence related to convergent and discriminant validity of integrity tests in general. The results from both the primary data analyses, and the meta-analyses display significant convergent validity for all integrity tests. The presence of a general factor across integrity tests is confirmed.

Of the Big Five dimensions of personality, conscientiousness was

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Howard Timm, Ph.D.	22b TELEPHONE (Include Area Code) 408-646-2448; 22c OFFICE SYMBOL PERSEREC

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603 — UNCLASSIFIED

found to have the most overlap with integrity tests. However, a more accurate identification of counterproductive individuals is possible by considering scores on agreeableness and emotional stability. It appears that the higher predictive validity of integrity tests arise from the fact that they tap constructs broader than just narrow and construct limited conscientiousness as measured by mainstream personality inventories. With integrity tests, the increased breadth of predictor construct coverage translates itself into better prediction in personnel selection.

Both overt and personality-based integrity test scores seem to be correlated with age . This indicates that younger individuals have the potential to be more counterproductive employees, possibly due to youthful indiscretions and experimentation. Integrity and gender are also correlated. On average, females score higher than males. Overt integrity test scores may be correlated with race (with minorities scoring higher). However, the same results do not apply to personality-based integrity tests. It is clear that using integrity tests in personnel selection, either overt or personality-based ones, will not cause adverse impact.

Integrity test scores and ability test scores are uncorrelated. Using integrity tests in conjunction with measures of ability can lead to substantial incremental validity for all job complexity levels. Employers seeking to maximize work force output should use both integrity tests and measures of general mental ability in making hiring decisions. This combination has the potential for simultaneously reducing adverse impact and enhancing validity and utility, in comparison to selecting on ability alone.

Final Technical Report

Establishing Construct Validity for Integrity Tests

Principal investigator: Deniz S. Ones

Thesis Advisor: Frank L. Schmidt

College of Business Administration

Department of Management and Organizations

University of Iowa

Iowa City, IA 52242

Accesion For	
NTIS CRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

Program Area: Clearance Processes

Personnel Security Research Dissertation Award Program

Institution: University of Iowa

THE CONSTRUCT VALIDITY OF INTEGRITY TESTS

by

Deniz S. Ones

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Business Administration
in the Graduate College of
The University of Iowa

June, 1993

Thesis supervisor: Professor Frank L. Schmidt

Copyright by

Deniz Ones

1993

All Rights Reserved

ACKNOWLEDGMENTS

This dissertation was supported in part by the Office of Naval Research (Contract No: N00014-92-J-4041 to Deniz Ones). I would like to thank Howard Timm and Ralph Carney of the Defense Personnel Security Research and Education Center for their encouragement and support of this research. The content of this dissertation does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

I acknowledge with thanks the contributions to this research of numerous test publishers, authors, and colleagues. I also acknowledge with gratitude the following test publishers and researchers for sharing their research and donating copies of tests to be used for this study: Murray Barrick, Jerry Borofsky, Paul Brooks, Harrison Gough, Jim Harris, Robert Hogan, Robin Inwald, Jack Jones, Michael Mount, George Pajaanen, and Fred Rafilson. I would also like to express our gratitude to the staff of the University of Iowa Examination and Evaluation Services, Xiaoping Wang, Marcia Simmering, Karen Chuin, and Phyllis Yen for their help in scoring, typing, and data management.

My thanks go to my commitee members Duane Thompson, Michael Mount, Murray Barrick, and John Hunter for the extra time they spent helping me with my dissertation and the valuable advice. I would also like to express my heartfelt thanks to my colleague and officemate Vish for the support he provided during all phases of my dissertation.

Above all, I thank my parents for the support they have given me during my education, and my advisor, major professor and mentor Frank Schmidt for his expertise and assistance at every step of the way.

There is one person without whose understanding and consideration, this dissertation would not have been possible. Ates, thank you for everything.

ABSTRACT

This dissertation presents extensive evidence related to convergent and discriminant validity of integrity tests in general. The results from both the primary data analyses, and the meta-analyses display significant convergent validity for all integrity tests. The presence of a general factor across integrity tests is confirmed.

Of the Big Five dimensions of personality, conscientiousness was found to have the most overlap with integrity tests. However, a more accurate identification of counterproductive individuals is possible by considering scores on agreeableness and emotional stability. It appears that the higher predictive validity of integrity tests arise from the fact that they tap constructs broader than just narrow and construct limited conscientiousness as measured by mainstream personality inventories. With integrity tests, the increased breadth of predictor construct coverage translates itself into better prediction in personnel selection.

Both overt and personality-based integrity test scores seem to be correlated with age . This indicates that younger individuals have the potential to be more counterproductive employees, possibly due to youthful indiscretions and experimentation. Integrity and gender are also correlated. On average, females score higher than males. Overt integrity test scores may be correlated with race (with minorities scoring higher). However, the same results do not apply to personality-based integrity tests. It is clear that using

integrity tests in personnel selection, either overt or personality-based ones, will not cause adverse impact.

Integrity test scores and ability test scores are uncorrelated. Using integrity tests in conjunction with measures of ability can lead to substantial incremental validity for all job complexity levels. Employers seeking to maximize work force output should use both integrity tests and measures of general mental ability in making hiring decisions. This combination has the potential for simultaneously reducing adverse impact and enhancing validity and utility, in comparison to selecting on ability alone.

TABLE OF CONTENTS

Page

LIST OF TABLES xi

LIST OF FIGURES xiv

PREFACE xv

CHAPTER

I. INTRODUCTION 1

II. LITERATURE REVIEW 6

Criterion-Related Validities of Integrity Tests 8

Criterion-Related Validities for Predicting Job Performance 8

Criterion-Related Validities for Predicting Counterproductive Behaviors 9

Construct Validity Evidence for Integrity Tests 11

Construct Validity Evidence from Integrity Test Reliabilities 11

Construct Validity Evidence from Factor Analytic Studies 13

Construct Validity Evidence from Contrasted Group Studies 16

Construct Validity Evidence from Correlations with Admissions of Counterproductivity 17

The Big Five Taxonomy of Personality and the Construct of Integrity 19

CHAPTER Page

III. METHODS 31

Modes of Data Collection 32

Primary Data Collection 32

Secondary Data Collection 33

Hypotheses and Procedures Used to Test Them 35

Questions Pertaining to the Relations among Integrity Tests, Big Five Dimensions of Personality and their Implications 35

1. Do overt integrity tests correlate highly with each other? 35

2. Do personality-based integrity tests correlate highly with each other? 36

3. Do both overt and personality-based integrity tests measure the same underlying construct? 38

4. Do integrity tests do correlate with measures of agreeableness, conscientiousness, extroversion, openness to experience, and emotional stability? 39

5. Do integrity tests derive their criterion-related validity from a broadly defined construct of Conscientiousness? 42

6. What levels of operational validity can be obtained by combining measures of the Big Five dimensions of personality (as measured by traditional personality instruments) and integrity (as measured by integrity tests) for personnel selection? 43

Questions Pertaining to Relationships between Integrity Tests and Demographic Variables, and their Implications for Adverse Impact 44

CHAPTER Page

1. Do integrity tests correlate with age? 44

2. Do integrity tests correlate with gender? 44

3. Do integrity tests correlate with race? 45

4. What are the precise implications for adverse impact of using integrity tests ? 45

Questions Pertaining to the Relationship between Integrity and Ability, and its Implications for Incremental Validity 46

1. Do integrity tests correlate with tests of ability? 46

2. What is the precise incremental validity of integrity tests over measures of ability for predicting overall job performance? 47

IV. ANALYSES AND RESULTS 48

Intercorrelations Among Integrity Tests and the Big Five Dimensions of Personality and their Implications 48

Results from the Primary Data 53

Correlations among overt tests (Primary data only) 53

Correlations among personality based tests (Primary data only) 53

Correlations between overt and personality based tests (Primary data only) 54

The relationships between integrity tests and the Big Five dimensions of personality (Primary data only) 57

Results of the Meta-Analyses 60

Correlations between overt and personality based tests (Meta-analytic results) 61

CHAPTER Page

The relationships between integrity tests and the Big Five dimensions of personality (Meta-analytic results) 63

Implications of the meta analytic results for the criterion related validities of integrity tests 73

Implications of the meta-analytic results for combining integrity tests and Big Five personality inventories for personnel selection 77

Integrity-Age, Integrity-Gender, and Integrity-Race Relationships and Implications for Adverse Impact 79

The Relationship between Integrity Test Scores and Age 79

The Integrity-Gender Relationship and Implications for Adverse Impact 80

The Integrity-Race Relationships and Implications for Adverse Impact 81

Overall Summary of Findings Pertaining to Demographic Variables 85

Integrity-Ability Relationship and Implications for Incremental Validity 87

V. DISCUSSION 90

TABLES 101

APPENDIX A. Questionnaire Used in Additional Data Collection 136

APPENDIX B. Instructions for Subjects 144

APPENDIX C. Debriefing Form 146

APPENDIX D. Sample Sizes in Primary Data 147

Page

APPENDIX E. Coding Sheet for Integrity Test Correlations with Other Variables, and Integrity-Personality Intercorrelations for Meta-Analysis....149

APPENDIX F. List of Journal Articles Coded....151

APPENDIX G. Reports and Books Coded....152

APPENDIX H. Classification of Various Personality Scales into the Big Five Dimensions of Personality....153

APPENDIX I. Sources of Data for the Meta-Analyses....177

REFERENCES....188

LIST OF TABLES

Table Page

1. The Big Five Dimensions of Personality from Various Researchers....101

2. The High and Low Ends of Big Five Dimensions of Personality....105

3. Meta-Analyzed Intercorrelations between the Big Five Dimensions of Personality....107

4. Instruments Used in Primary Data Collection....108

5. Integrity Tests Contributing Correlations to the Meta-Analyses....109

6. Hypothesized Intercorrelation Matrix of Integrity and the Big Five Dimensions of Personality....111

7. Personality Inventories Contributing Data to the Analyses Reported....112

8. Descriptive Information on Statistical Artifact Distributions Used to Correct Correlations....114

9. Intercorrelations between Overt Integrity Tests (Obtained from the Primary Data)....115

10. Intercorrelations between Personality-Based Integrity Tests (Obtained from the Primary Data)....116

11. Observed and True Score Intercorrelations between Personality-Based Integrity Tests (Obtained from the Primary data)....117

12. The Correlations between a Linear Composite of Each Dimension of the Big Five and a Linear Composite of Integrity Test Scores (Using the Primary Data)....118

Table Page

13. The Correlations between a Linear Composite of Each Dimension of the Big Five and a Linear Composite of Overt Integrity Test Scores (Using the Primary Data).....119

14. The Correlations between a Linear Composite of Each Dimension of the Big Five and a Linear Composite of Personality-Based Integrity Test Scores (Using the Primary Data).....120

15. Meta-Analyzed Intercorrelations between Overt and Personality-Based Integrity Tests.....121

16. Meta-Analysis of Integrity Test Correlations with the Big Five Dimensions of Personality.....122

17. Meta-Analysis of Overt Integrity Test Correlations with the Big Five Dimensions of Personality.....123

18. Meta-Analysis of Personality-Based Integrity Test Correlations with the Big Five Dimensions of Personality.....124

19. True Score Correlations between Composites of Integrity Test Scores and Composites of the Big Five Dimensions of Personality.....125

20. The True Score Correlation Matrix between the Big Five Dimensions of Personality, Integrity, and Job Performance.....126

21. First-Order Partial Correlations between the Big Five Dimensions of Personality, Integrity, and Job Performance (Conscientiousness Partialed out).....127

22. Meta-Analysis of the Relationship between Integrity and Age.....128

23. Meta-Analysis of the Relationship between Gender and Integrity.....129

24. Meta-Analyses of the Relationships between Race and Integrity.....130

25. List of Cognitive Ability Tests Contributing Data to the Meta-Analyses.....131

26. Descriptive Information on Test Reliabilities Used to Correct Correlations.....132

Table | Page

27. Meta-Analysis of Integrity Test Correlations with General Mental Ability 133

28. Effect of Combining Integrity Tests with Measures of Ability in Predicting Ratings of Overall Job Performance 134

29. Sample Sizes in Primary Data 148

30. Classification of Various Personality Scales into The Big Five Dimensions Of Personality 154

LIST OF FIGURES

Figure Page

1. Conscientiousness and overt integrity tests 36

2. Conscientiousness and personality-based integrity tests 37

3. Conscientiousness and integrity tests 39

4. Conscientiousness and integrity tests: Alternate framework with group factors 40

PREFACE

The research undertaken in this dissertation represents a portion of my ongoing programmatic research about integrity tests. In 1990, when I first started my research on integrity tests, most industrial psychologists were skeptical of them. First, with my colleague C. Viswesvaran and major professor Frank Schmidt, we examined the criterion-related validity of integrity tests. We showed that integrity tests have substantial validity for predicting both job performance and counterproductive behaviors on the job (Ones, Viswesvaran, & Schmidt, in press). We also conducted other research on the criterion-related validity of integrity tests for predicting absenteeism, drug and alcohol abuse, violence on the job, turnover, and accidents. This dissertation is my examination of the construct validity of integrity tests.

I am not and have never been associated with any integrity test publisher, nor is the research presented in this dissertation sponsored by integrity test publishers. My research is about integrity tests in general; I do not endorse any specific integrity test over others for use in personnel selection.

On a methodological note, the database created for this dissertation, with more than 8,000 correlation coefficients, is probably the largest meta-analytic database in existence. As such, the results of the research presented are robust.

CHAPTER I

INTRODUCTION

Employers have long been concerned with honesty, integrity, dependability and trustworthiness of their employees. A recent large scale survey of employers in Michigan (Michigan Department of Education, 1989) found that of 86 employee qualities ranked for importance in entry level employment by over 3000 employers, seven of the top eight qualities were related to integrity, trustworthiness, conscientiousness. The other quality in the top eight (ranked 5th) referred to general mental ability. Clearly employers seem to value integrity in their employees.

Employers are also concerned with a variety of disruptive behaviors on the job. Such disruptive behaviors include theft, absenteeism, drug and alcohol abuse, tardiness, malingering, disciplinary problems and other behaviors that are counterproductive in the work environment. All of these counterproductive behaviors are costly to employers. For example, theft, only one form of disruptive behavior on the job, is estimated to cost about $40 billion per year (American Management Association, 1977), ten times the cost of robberies, burglaries, and street crime combined (Bacas, 1987). To control theft and other disruptive behaviors on the job, employers have resorted to two main methods: (a) surveillance and closer supervision, and (2) personnel selection. The first approach focuses on deterring theft and other counterproductive behaviors of employees by increasing the likelihood of

detection. In contrast, the personnel selection approach aims to limit the proclivity for dishonesty among employees by selecting on the basis of trustworthiness.

In 1948, the first paper and pencil psychological test to assess the integrity of potential employees was developed (Gough, 1948). The instrument was the Personnel Reaction Blank. It was a derivative of what was then called the Delinquency scale of the California Psychological Inventory. Later on the Delinquency scale was renamed the Socialization scale. In 1952, a second type of test aiming to assess honesty of job applicants was developed. This test, the Reid Report, was a compilation of questions that seemed to distinguish honest and dishonest individuals during polygraph examinations Since then several other instruments have been developed and used to select applicants on the basis of integrity. A complete and interesting treatise of the history of integrity tests can be found in Ash (1989) and Woolley (1991). Collectively, the paper and pencil tests that were specifically developed to assess the dependability, integrity, and honesty of job applicants and employees are referred to as "integrity tests" (Sackett & Harris, 1984; Sackett, Burris, & Callahan, 1989). These tests are the focus of this dissertation.

There are three important points to remember about integrity tests. First, they are paper and pencil tests, excluding other methods of assessing honesty such as the polygraph (a physiological method), background investigations, interviews, and reference checks. Second, these tests have been developed for use with applicants and employees (a normal population). Hence instruments such as the MMPI cannot be classified as integrity tests even though some organizations claim to use them for screening out

delinquent applicants. Third, most integrity tests have been initially designed to predict a variety of counterproductive behaviors, even though they may have branched out to predict other criteria such as supervisory ratings of overall job performance.

There is relatively little information about companies that use paper and pencil integrity tests. According to Sackett and Harris (1985) as many as 5,000 companies may use pre-employment integrity tests, assessing about 5,000,000 applicants yearly. A variety of surveys of companies indicate that anywhere between 7 to 20% of all companies in the US could be testing for integrity at least for some jobs (for various estimates see Blocklyn, 1988; American Society for Personnel Administration, 1988; Bureau of National Affairs, Inc., and O'Bannon, Goldinger, and Appleby, 1989). Even by most conservative estimates, millions of people in the US either have been or are being tested using integrity tests. There are at least 43 integrity tests in current use. My personal observation is that of these tests, about a quarter seem to be small operations without much market share and overall 16-19 tests seem to serve the majority of the demand for integrity tests. However, this demand can be expected to increase because in 1988 the Federal Polygraph Act effectively banned the use of the polygraph in employment settings for all but the federal employees.

Employers' desire for trustworthy and conscientious employees has spawned a multimillion dollar industry of integrity testing (see O'Bannon et al., 1989 for prices of various integrity tests three years ago). Employers' concern regarding counterproductive behaviors at work coupled with the recent passage of the Employee Polygraph Protection Act (1988) seems to

indicate that paper and pencil integrity tests will be used more broadly than they are today.

Over the last fifteen years, scientific interest in integrity testing has increased substantially. The publication of a series of literature reviews attests to the interest in this area and its dynamic nature (Guastello & Rieke, 1991; Sackett, Burris, & Callahan, 1989; Sackett & Decker, 1979; Sackett & Harris, 1984). Recently Sackett et al. (1989) and O'Bannon, Goldinger, and Appleby (1989) have provided extensive qualitative reviews and critical observations regarding integrity testing. In addition to these reviews, the US Congressional Office of Technology Assessment (OTA) (1990) and the American Psychological Association (APA) (Goldberg, Grenier, Guion, Sechrest, & Wing, 1991) have each released "papers" on integrity tests. The OTA paper (1990) was in part prompted by the Congress' regulation of the polygraph. It was very critical of integrity tests and research conducted on them. The OTA recommendations were based on the results of only a few "technically competent" studies, ignoring most of the voluminous literature on integrity tests. Compared to the OTA paper (1990), the APA report (Goldberg et al., 1991) was better thought out, more thorough, and insightful. Overall, the APA report (Goldberg et al., 1991) was more scientific than the OTA paper (1990) in its review of integrity tests. It provided a generally favorable conclusion regarding the use of paper and pencil integrity tests in personnel selection.

Most recently, Ones, Viswesvaran, and Schmidt (in press) using psychometric meta-analysis showed that integrity tests have useful, generalizable validities with various job related criteria, including supervisory ratings of job performance, and theft and broad

counterproductive behaviors on the job. Now that we know both overt and personality-based measures of integrity have validity across situations, jobs and settings lending them utility in personnel selection, the questions of paramount importance revolve around the construct validity of integrity tests.

In fact, the American Psychological Association's report on questionnaires used in the prediction of trustworthiness (Goldberg et al., 1991; trustworthiness was their term for integrity) specifically states " . . . we strongly recommend that investigators devote increased research attention to construct-oriented issues" (p. 18). The primary purpose of this dissertation is to establish lines of construct validity evidence for major integrity tests. So far only superficial attention has been directed to construct issues across integrity tests. This proposed research will attempt to offer precise answers to the question "what do integrity tests measure?"

CHAPTER II

LITERATURE REVIEW

Sackett et al. (1989) classify honesty tests into two categories: "Overt integrity tests" and "Personality-based tests." Overt integrity tests (also known as clear purpose tests) are designed to directly assess attitudes regarding dishonest behaviors. Some overt tests specifically ask about past illegal and dishonest activities as well; although for several admissions are not a part of the instrument, but instead are used as the criterion. Overt integrity tests include the London House Personnel Selection Inventory (PSI) (London House, Inc., 1975), Employee Attitude Inventory (EAI) (London House, Inc., 1982), Stanton Survey (Klump, 1964), Reid Report (Reid Psychological Systems, 1951), Phase II Profile (Lousig-Nont, 1987), Milby Profile (Miller & Bradley, 1975), and Trustworthiness Attitude Survey (Cormack & Strand, 1970). According to Sackett et al. (1989), " . . . the underpinnings of all these tests are very similar . . . " (p. 493). Hence, high correlations may be predicted among all these overt integrity measures. One of the purposes of this dissertation is to test this hypothesis.

On the other hand, personality-based measures (also referred to as disguised purpose tests) aim to predict a broad range of counterproductive behaviors at work (e.g., violence on the job, absenteeism, tardiness, drug abuse, in addition to theft) via personality traits, such as reliability, conscientiousness, adjustment, trustworthiness, and sociability. In other

words, these measures have not been developed solely to predict theft or theft-related behaviors. Examples of personality-based measures that have been used in integrity testing include the Personal Outlook Inventory (Science Research Associates, 1983), the Personnel Reaction Blank (Gough, 1954), Employment Inventory of Personnel Decisions Inc.(Paajanen, 1985), and the Hogan's Reliability Scale (Hogan, 1981). Different test publishers claim that their integrity tests measure different constructs, including responsibility, long term job commitment, consistency, proneness to violence, moral reasoning, hostility, work ethics, dependability, depression, and energy level (O'Bannon et al., 1989). The similarity of integrity measures raises the question of whether they all measure primarily a single general construct. The proposed dissertation will test the hypothesis that integrity tests measure a single general construct. Detailed descriptions of all the above tests can be found in the 10th Measurement Yearbook (Conoley & Kramer, 1989) and in the extensive reviews of this literature (O'Bannon et al., 1989; Sackett et al., 1989; Sackett & Harris, 1984).

Another purpose of this dissertation is to establish the nomological network for integrity tests. Specifically, the construct(s) integrity tests tap into will be explored.

In the following literature review sections, I will first review the criterion-related validities of integrity tests, second I will focus on lines of construct validity evidence for integrity tests. The construct validity evidence reviewed will include reliabilities of integrity tests, factor analytic study results, contrasted group studies, and correlations with admissions of counterproductivity. Third, I will review the research on the Big Five

dimensions of personality and present some theories about the place of integrity in that framework.

Criterion-Related Validities of Integrity Tests

Ones et al. (in press) conducted a comprehensive meta-analysis based on 665 validity coefficients across 576,460 data points to investigate whether integrity test validities are generalizable and to estimate differences in validity due to potential moderating influences. Results indicate that integrity test validities are substantial for predicting job performance and counterproductive behaviors on the job such as theft, disciplinary problems, and absenteeism. The estimated mean operational predictive validity of integrity tests for predicting supervisory ratings of job performance is .41. For the criterion of counterproductive behaviors, Ones et al (in press) results indicate that use of concurrent validation study designs may overestimate the predictive criterion-related validity applicable in selection situations. Results from predictive validity studies conducted on applicants and using external criterion measures (i.e., excluding self-reports) indicate that integrity tests predict the broad criterion of organizationally disruptive behaviors better than they predict the narrower criterion of employee theft alone. However, an important conclusion of the Ones et al. (in press) research is that despite the influence of moderators, integrity test validities are positive across situations and settings.

Criterion-Related Validities for Predicting Job Performance

In selection settings, the best estimate of integrity test validities for predicting job performance would be based on (a) predictive studies (b) conducted on samples of applicants. To obtain such an estimate of the mean

validity of integrity tests for selection, Ones et al. (in press) meta-analyzed predictive validities calculated on applicant samples. There were 23 such validities for predicting supervisory ratings of job performance. Across 7,550 people, the best estimate of the mean true validity was .41. The standard deviation of true validity was 0, and the percent variance accounted for was 100%. These findings imply that the average validity that integrity tests may be expected to have in selection settings for supervisory ratings of overall job performance is .41, and that this value is constant across settings. The meta-analysis results presented by Ones et al (in press) also show that overt and personality-based tests produce fairly similar operational validities when the criterion of interest is supervisory ratings of overall job performance.

Criterion-Related Validities for Predicting Counterproductive Behaviors

Ones et al (in press) indicate that validities of integrity tests for predicting counterproductive behaviors on the job appear to be fairly substantial. However, several methodological moderators were identified for this type of criterion: type of test (overt vs. personality based), criterion measurement method (admissions vs. external), criterion breadth (theft vs. broad counterproductivity), validation strategy (predictive vs. concurrent), and validation sample (applicants vs. employees). When the effects of these methodological moderators are controlled , the standard deviations of true validity for integrity tests were found to be no larger than those of ability tests in predicting job performance (e.g., Pearlman, Schmidt, & Hunter, 1980; Schmidt, Hunter, Pearlman, & Shane, 1979), thus indicating similar levels of generalizability and a similar lack of support for purely situational moderators. Some exceptions to this conclusion were concurrent studies of overt tests conducted on employees using externally measured broad

counterproductivity criteria and concurrent studies of overt tests conducted on applicants using both admissions of theft criteria and broad counterproductive behaviors criteria.

For the criterion of counterproductive behaviors, admissions produced much higher correlations than external criteria, and concurrent studies often seemed to overestimate predictive validity (Ones et al., in press). The utility of a selection test depends on its predictive validity; although concurrent validities are relevant to questions of construct validity, the major purpose of concurrent validity in selection research is to estimate predictive validity. Thus, the Ones et al. (in press) finding that in this research domain, at least for overt tests, concurrent validity estimates overestimate predictive validity is potentially important.

In selection research, the best estimate of operational selection validities of integrity tests for predicting theft would be based on predictive studies conducted on applicants. In addition, many would argue for reliance on external criteria in preference to admissions criteria. Considering externally measured theft as the criterion in predictive studies, Ones et al. (in press) found that the mean operational validity of overt integrity tests is estimated at .13. For personality-based tests, no validity estimates for the prediction of theft alone were available. Considering externally measured broad counterproductive behaviors as the criterion in predictive studies conducted on applicants, Ones et al. (in press) found that the mean operational validity of both types of integrity tests is positive across situations and is substantial (.39 for overt tests, and .29 for personality-based tests). Theft appears to be less predictable than broad counterproductive behaviors, although this comparison was made only for overt integrity tests.

In sum, Ones et al. (in press) found that integrity tests predict overall job performance with moderate and generalizable validity. They also found that integrity tests predict counterproductive behaviors such as theft, absenteeism, tardiness, and disciplinary problems, but validity estimates seem to be affected by several simultaneously operating methodological moderators. All in all, Ones et al. (in press) conclude that the validity of integrity tests is positive and in useful ranges for both overall job performance criteria and counterproductive behaviors criteria.

Construct Validity Evidence for Integrity Tests

Construct Validity Evidence from Integrity Test Reliabilities

In examining the construct of integrity, important insights can be gained from integrity test reliabilities. The amount of measurement error places a limit on the relationships that a measure can have with other measures. Reliabilities of integrity tests as a category have been examined by Ones et al. (1992, in press). Ones et al. (1992) obtained a total of 124 integrity test reliability values from the published literature and the test publishers. Of the 124, sixty-eight were alpha coefficients (55%) and forty seven were test-retest reliabilities.

Test-retest reliabilities address whether integrity test scores are stable over time. Test-retest reliabilities for integrity tests were reported for over periods of time ranging from 1 to 1,825 days (mean = 111.4 days; sd = 379.7 days). The mean of the test-retest reliabilities was .85 (sd = .10). Seventy-two percent of test-retest reliabilities were for overt integrity tests (mean time period between test administrations = 26.5 days, sd = 33.8 days; mean test-retest reliability =.87, sd = .09). Twenty-eight percent of test retest reliabilities

were for personality-based integrity tests (mean time period between test administrations = 331.6 days, sd = 697. 7 days; mean test-retest reliability =.68, sd = .12).

Across all integrity tests, the mean of the coefficient alphas was .81 (sd = .10). 81% of coefficient alphas were reported for overt integrity tests (mean = .82, sd = .09). The mean of the coefficient alphas reported for personality-based tests was .77 (sd = .13). There were 9 reliabilities reported without stating the type of reliability.

The overall mean of the integrity test reliabilities was .81 and the standard deviation was .11. The mean of the square roots of predictor reliabilities was .90 with a standard deviation of .06. The value of .90 for the mean of the square roots of the integrity test reliabilities is an important aspect of construct validity. Using a fundamental principle of classical measurement theory, the maximum true score correlation that integrity tests can have with any other variable cannot exceed .90. Ones et al (in press) report that there were 97 reliabilities for overt tests. The mean of the overt test reliability artifact distribution was .83 and the standard deviation was .09. The mean of the square roots of overt test reliabilities was .91 with a standard deviation of .05. There were 27 reliabilities reported for personality-based tests. The mean of the personality-based test reliability artifact distribution was .72 and the standard deviation was .13. The mean of the square roots of the reliabilities was .85 with a standard deviation of .08.

A question of importance for any variable to be used in personnel selection is related to the stability of a job applicant's performance at two points in time. Specifically, the question is whether the relative rank ordering of candidates change with different administrations of the test.

While it is reasonable to expect a candidate's score to change over a long period of time, any instability in the rank ordering of a set of candidates between two administrations close in time indicates a lack of test-retest reliability (i.e., the traits are stable in the short time intervals between the two administrations). To determine whether integrity is a stable characteristic test-retest reliabilities using parallel forms are appropriate. However, no such reliabilities have been reported for integrity tests. Hence, Ones et al. (1992) meta-analytically cumulated forty seven test-retest reliabilities (N = 4,280) of paper and pencil pre-employment integrity tests to assess the temporal stability of integrity. The mean sample size weighted reliability coefficient was .85. All the observed variance in test-retest reliabilities was attributable to sampling error. Ones et al. (1992) concluded that the results indicated that the construct measured by integrity tests is highly stable.

Construct Validity Evidence from Factor Analytic Studies

Factor analytic investigations have been conducted on a number of integrity tests. In all of these studies, the investigators have aimed to confirm a multiple factorial model of integrity. In other words, factor analysts of integrity tests have never looked for a general factor. This is a major shortcoming.

More factor analytic investigations have been conducted on overt integrity tests than on personality-based integrity tests. Cunningham and Ash (1988) investigated the dimensionality of the Reid Report using principal components analysis using two large samples (N's of 1,281 and 3,071). They found that a solution of four interpretable factors fit the data best (the four factors were labeled self punitiveness, punitiveness toward others, self projection, projection toward others). Jones and Terris (1984) examined the

factor structure of the PSI and found six factors (these were labeled theft temptation and rumination, theft rationalization, projection of theft in others, theft punitiveness, inter-thief loyalty, personal theft admissions). Harris and Sackett (1987) also investigated the factor structure of the PSI Honesty scale (N=849 job applicants) and found four interpretable factors (they labeled these factors temptation and thoughts about dishonest behaviors, actual and expected dishonest activities, norms about the dishonest behaviors of others, impulse control and behavioral tendencies). Martelli (undated) conducted a principal components analysis of the Phase II Profile and found three factors. Hay (1981) and Harris (1987) investigated the factor structure of the Stanton Survey and found seven interpretable factors (these were labeled general theft, opportunism, employee theft, leniency, employee discounting, perceived pervasiveness of dishonesty, and association with dishonest individuals). However, both the attitudes and admissions part of the Stanton Survey were used, a decision that probably clouds the comparison of Stanton Survey factor structure with other overt tests.

As was indicated earlier, the major shortcoming of these factor analytic studies is that no general factorial solution was investigated. In fact, the multiple factors these researchers claim to have found are highly correlated, indicating a problem of overfactoring. This might also be intuitively evident from the labels different researchers used to describe the multiple dimensions (for example, in one study, general theft and employee theft were claimed to be separate dimensions). The results of different factor analytic studies reflect interpretations of various researchers, yet there seems to be a degree of overlap in the construct(s) integrity test tap into. The assertion that overt

integrity tests appear to be multidimensional does not preclude the establishment of a general factor. This interpretation is strengthened by a finding in many of the previously reviewed factor analytic studies. A first factor accounted for a large proportion of the variance when compared to subsequent factors. This fact coupled with high intercorrelations among factors clearly points to the presence of a general factor. Harris and Sackett (1987) explicitly stated that a general factor accounted for most of the variance in their data and further conducted Item Response Theory (IRT) analyses using the one parameter Rasch model. The results suggested that the PSI Dishonesty scale taps into "an underlying construct which may be called dishonesty" (p. 134).

Relatively few studies have investigated the factor structure of personality-based integrity tests. Paajanen (1987) factor analyzed the PDI Employment Inventory. The PDI Employment Inventory has three scales: Performance, Tenure, and Frankness. Of these three scales, only Performance scale is considered to be a personality-based integrity test (even though the observed correlations between the Performance scale and the Tenure scale range between .45-.65). In Paajanen's factor analysis of the PDI Employment Inventory (all three scales combined), a five factor solution provided the best fit to the data. These factors were labeled irresponsibility, sensation seeking, unstable upbringing, frankness and conforming motivation. Similar to the results for overt integrity tests, positive correlations were reported among the dimensions and a large proportion of the variance was accounted for by the first factor "irresponsibility," strengthening an argument for a general factor.

This dissertation will not examine the factor structure of individual integrity tests. Such studies are necessary and useful for refining lines of

construct validity evidence for single instruments, but they are less essential when the focus is on investigating construct validity across measures. In addition, the proprietary nature of scoring keys for most integrity tests makes it impossible to factor analyze them. Positive and often fairly respectable correlations among group factors detected in factor analytic studies appears to be evidence of a general factor and further justifies the approach taken in this dissertation.

Construct Validity Evidence from Contrasted Group Studies

Contrasted group studies aim to illustrate that mean test score differences exist between groups believed to differ in integrity. In these studies people who are decidedly dishonest are contrasted with those who show no signs of lack of integrity. The underlying theory is that if the integrity test is a good measure of the construct of integrity, large differences should be found between the two groups. It is important to note that some integrity tests have used the contrasted groups design for criterion keying of responses in the instrument development stage.

Many contrasted group studies have been reported for integrity tests (for a detailed review see Sackett & Harris, 1984; and Sackett et al., 1989). For example Gough (1972) reported that various student and employee samples (N = 1,626 males, 1,408 females) score one and one and a half standard deviations (females and males, respectively) higher than juvenile delinquent and adult prison inmate samples (505 males, and 114 females). Paajanen (1987) found that juvenile offenders score (N=37) 1.4 standard deviations below the general applicant population mean on the PDI Employment Inventory and .8 standard deviations below the mean of employees fired for theft alone. Similar results are reported by Collins and Schmidt (in press)

comparing the PDI Employment Inventory scores of white collar criminals and a Midwestern sample of white collar workers; by Borofsky, Friedman and Maddocks (1987) comparing the Employee Reliability Inventory scores of convicted thieves and a small sample of nuclear weapons employees; by Terris and Jones (1982) comparing PSI scores of 116 felons imprisoned in county jails with retail store job applicants; by Ash (1974) comparing the Reid Report scores of 187 incarcerated felons and job applicants. A meta-analysis (Ones, Viswesvaran, and Schmidt, unpublished manuscript) is underway examining the results from contrasted group studies.

All in all, these studies suggest that integrity tests seem to successfully discriminate between known criminals and non-criminals, and prisoners and non-incarcerated individuals. A comparison of delinquent or criminal groups with the normal population is an extreme contrast and although it may provide construct validity evidence it does not address the criterion-related validities of integrity tests (Goldberg et al, 1991).

Construct Validity Evidence from Correlations with Admissions of Counterproductivity

The results for overt integrity tests in Ones et al. (in press) indicate that no matter what the content of the criterion measure (theft or broadly defined disruptive behaviors), self-reported criteria tend to result in higher estimates of validities than external criteria. Many may judge that correlations with self-report criteria are not acceptable as estimates of the operational validity of integrity tests; however, it is not entirely clear that external measures of counterproductive behaviors are more valid than admissions of such behaviors. Many thefts and other counterproductive behaviors may go undetected, limiting the validity of external measures. In addition, there is

considerable evidence from research on juvenile delinquency that the correlation between admissions and actual behavior is substantial (about .50; Viswesvaran, Ones, & Schmidt, 1992). On the other hand, when admissions are used as criteria, the difference between reliability and validity becomes potentially tenuous. That is, when tests which include some questions that ask for admissions are validated against admissions, the predictor-criterion correlations indicate in part only that admissions predict other admissions collected at about the same time. In any event, correlations with admissions criteria can be taken as evidence of construct validity (Goldberg et al., 1991).

The meta-analyses of overt test correlations with admissions criteria indicate that correlations are higher for employees than for applicants (Ones et al., in press). For self-reports of theft, the estimated true mean correlation is .54 for the N = 3,217 employee sample and .42 for the N = 68,613 applicant sample. In both cases the standard deviation of true validity is large enough to indicate additional moderators may be operating. However, the positive lower credibility values mean that a positive correlation can be expected between honesty test scores and admissions of theft in concurrent studies for both employee and applicant samples regardless of the setting and situation.

Ones et al. (in press) report that when the admissions criteria include other disruptive behaviors such as tardiness, violence on the job, absenteeism, drug abuse, and alcohol abuse in addition to only theft, mean correlations of overt tests increase to .99 for employee samples (N = 27,887) and .46 for applicant samples (N = 90,527). In both these cases, self-report criteria were collected concurrently with the predictor data. Ones et al. (in press) indicate that the pattern of mean correlations for both theft and broad counterproductive criteria suggest that employees are more willing to admit

negative behaviors than are applicants hoping to obtain a job. Taken together, the results for self-report criteria support the construct validity of overt integrity tests (Ones et al., in press).

The Big Five Taxonomy of Personality and the Construct of Integrity

A variety of individual differences have been of interest to psychologists, among them personality variables. A descriptive model or a taxonomy is vital, if personality psychology is to advance as a science. However, over the years personality has been conceptualized from various taxonomic viewpoints. In fact, there have been so many personality concepts, and so many scales used to measure them that, the characterization of " . . . Babel of concepts and scales . . ." seems appropriate for the personality domain (John, 1990). It is true that " . . . the world does not come to us in neat little packages" (Gould, 1981, p. 158). But without a generally accepted taxonomy, personality research cannot properly communicate empirical findings, let alone systematically cumulate them.

In the field of personality structure, one needs to make the distinction between the development of factor analytic models for representing latent variables underlying individual differences in personality structure and the development of taxonomies of trait descriptive terms of how personality characteristics are encoded in language (Wiggins & Pincus, 1992). The starting place for a common descriptive taxonomy of personality has been with the latter model: the natural language of personality description. The rationale behind this approach is that most important individual differences in human behaviors are encoded in languages. This is the fundamental lexical hypothesis (Goldberg, 1990). Fortunately, the last ten years has seen the

convergence for the structure of personality from these two distinct streams of research.

It is now clear that one cannot take taxonomies of personality which have numerous factors seriously. For example, Cattell's (1948) system with 16 factors and 8 second order factors does not seem to have empirical support. Impressive evidence now exists supporting Big Five dimensions of personality. For example: Banks (1948) offered a less complex solution to the intercorrelations of Cattell (1948); Fiske (1949) found the confirmation of the Big Five dimensions of personality using Cattell's bipolar scales. The Tupes and Christal (1961) factor analyses of Cattell's bipolar scales in an attempt to predict officer effectiveness in the Air Force yielded five factors: Surgency, Agreeableness, Dependability, Emotional Stability, and Culture.

While the Tupes & Christal (1961) study was buried in an obscure Air Force technical report, Cattell's (1947) and Eysenck's (1970) widely accepted structures of personality dominated the literature (recently, Tupes and Christal's technical report was reprinted in the Journal of Personality). In the late 1960's and 1970's, a strong attack was launched upon the field of trait research by Mischel (1968) and others. The claim of these researchers can be summed up as : "Personality is in the heads of researchers." The influence of such behaviorism resulted in the erroneous conclusion that situations, and not traits, dominate human behavior. However, this point seems to have been rebutted by, among many others, Funder and Ozer (1983). The current evidence suggests that traits are quite potent in explaining behavior than previously acknowledged.

During the last decade, a large body of literature has accumulated which provides compelling evidence for the robustness of the five factor

model of personality: across different theoretical frameworks (Goldberg, 1981); using different instruments (e.g. Conley, 1985; Costa & McCrae, 1988; Lorr & Youniss, 1973; McCrae, 1989; McCrae & Costa, 1985;1987); in different cultures (e.g. Bond, Nakazato, & Shiraishi, 1975; Borkenau & Ostendorf, 1990; Noller, Law, & Comrey, 1987, Yang & Bond, 1990); using ratings obtained from different sources (e.g., Digman & Innouye, 1986; Digman & Takemoto-Chock, 1981, Fiske, 1949; McCrae & Costa, 1987; Norman, 1963; Norman & Goldberg, 1966; Watson, 1989); and with a variety of samples. (See McCrae & John, 1992; Digman (1990), John, (1990), Goldberg (1990) and Wiggins & Pincus (1992) for more detailed discussions.) Based on the results of several studies using comprehensive sets of trait terms , using multiple replications, and different factor analytic techniques, Goldberg (1990) summarized the state of the research in the following way: "It now seems reasonable to conclude that analyses of any reasonably large sample of English trait adjectives in either self or peer descriptions will elicit a variant of the Big Five Structure, and therefore virtually all such terms can be represented within this model" (p.1223). Digman and Takemoto-Chock (1981) summarized their conclusions as: "Regardless of whether teachers rate children, officer candidates rate one another, college students rate one another, or clinical staff members rate graduate trainees, the results are pretty much the same" (pp. 164-165).

Table 1 indicates how the different models of personality may be reconciled and even mapped onto the Big Five.

It is widely agreed that the first dimension is Extroversion/ Introversion , although it has also been called Surgency. Traits associated with it include sociability, talkativeness, and assertiveness (on the high end) and being retiring, silent, reserved and cautious on the low end). The second

factor is most frequently called Emotional Stability, or Neuroticism, and is usually defined from the negative pole. Common traits associated with this factor include anxiety, depression, anger, embarrassment, emotion, worry, fearfulness, instability, and insecurity. The third dimension has generally been interpreted as Agreeableness or Likability. Traits associated with this dimension include courteousness, flexibility, trust, good-naturedness, cooperativeness, forgiveness, empathy, soft-heartedness, and tolerance. The fourth dimension has most frequently been called Conscientiousness, although it has also been called Conformity or Dependability. Because of its relationship to a variety of educational achievement measures and its association with volition, it has also been called "Will to Achieve." Traits associated with this dimension reflect both dependability— carefulness, thoroughness, responsibility, organization, efficiency, planfulness, and volition- hard work, achievement- orientation, and perseverance. The last dimension has been interpreted frequently as Openness to Experience , Intellect, or Culture. Traits commonly associated with this dimension include imagination, curiosity, originality, broadmindedness, intelligence, and artistic sensitivity. Each of the Big Five dimensions are regarded as continua and Table 2 describes the positive and negative poles of each.

Among others, Costa & McCrae (1985,1987) demonstrated the presence of the Big Five in Eysenck Personal Inventory, Jackson Personal Research Form, Myers-Briggs Type Indicator, California Q-Set. Only the MMPI was identified as having four dimensions rather than five with Conscientiousness absent. Even though, there seems to be good agreement among personality psychologists on how many dimensions are necessary to account for the entire ocean of personality traits, there is considerable

disagreement on each dimension's meaning. Of the Big Five, the easiest identified and agreed upon dimensions seem to be emotional stability or neuroticism, and extroversion/ introversion. Openness to experience is controversial because some claim it is correlated with intelligence. The empirical evidence seems to support this claim. For example, Hogan and Hogan (1993) report moderate positive observed correlations between the Hogan Personality Inventory's Intellectance scale and various tests of the ASVAB. These observed correlations average about .30. As for Agreeableness and Conscientiousness, Zuckerman, Kuhlman, & Kamac (1988) claim that those two factors can be appropriately subsumed under Eysenck's Psychoticism (Eysenck, 1947). (Eysenck (1947) has a three factor model of personality. These three factors are neuroticism, extroversion, and psychoticism.)

The correlations among the Big Five Dimensions of personality is an important piece of evidence for convergent and discriminant validity. In a recent study, we meta-analytically examined the correlations among the Big Five Dimensions of Personality across major personality inventories (Ones, Viswesvaran, Schmidt, unpublished manuscript). The results of that study are summarized in Table 3.

In the meta-analysis, the observed correlations were corrected for sampling error and unreliability in the personality measures. Table 3 presents the true score intercorrelations. The highest correlations in the table are in the within category diagonal. That is, scales from different inventories tapping the same dimension of the Big Five correlate fairly high. On the other hand, correlations between the Big Five Dimensions are fairly low. In other words, the intercorrelation matrix from 5,703 meta-analyzed

correlations shows adequate convergent and discriminant validity. Describing the personality construct domain in terms of the Big Five dimensions appears appropriate.

Despite these criticisms and suggested modifications, the Big Five dimensions of personality seem to be robust and the evidence supporting it quite voluminous (Goldberg, 1993). One question of interest is where in the Big Five framework do integrity tests fall? Before exploring possible answers to this question, I will first direct attention to the theoretical underpinnings of the construct of integrity.

Different integrity test publishers claim that their tests measure different constructs. However, the hypothesis that I would advance is that there is an underlying construct that runs through all integrity tests. In fact, Hogan and Hogan (1989) have named that underlying construct "organizational delinquency." In Hogan and Hogan's (1989) theoretical perspective, counterproductive acts such as "theft, substance abuse, lying, insubordination, vandalism, sabotage, absenteeism and assaultive actions are elements of a larger syndrome" (p. 273) that can all be grouped under the large umbrella of organizational delinquency.

Along similar lines, Gough (1948, 1954, 1960) proposed that individuals are normally distributed along a continuum of socialization. The high extreme of this construct designates individuals who are extremely obedient, rule compliant and conscientious. The low extreme of this construct designates individuals who are rule breaking and hostile to the rules of the society. This construct first defined as "Wayward Impulse" in 1948 and later as "Socialization" in 1960 by Gough seems to be relevant for the construct of integrity. In fact, Gough (1954) has based an integrity test, the Personnel

Reaction Blank, on the items included in the Socialization scale of the California Psychological Inventory (CPI). The Socialization scale of the CPI was initially called Delinquency (Gough & Petersen, 1952). There have been several studies investigating the structure of the Socialization scale (e.g. Rosen, 1977). It seems that the broad construct behind the scale is comprised of four hierarchically ordered sub factors. The first factor is hostility toward rules and authority, the second factor is thrill seeking impulsiveness, the third factor is social insensitivity, and the fourth factor is alienation.

The theoretical basis for the construct of organizational delinquency, socialization or integrity can be found in Gough's 1948 article entitled "A sociological theory of psychopathy." The theory seems rely on the concept of individual differences in role-taking. The root idea of the theory is that individuals possess varying degrees of self control and varying degrees of sensitivity to reactions of others. People who are delinquents or "psychopaths" fail to adequately anticipate the reactions of others in social situations. This is because even though a person lacking integrity knows society's rules and expectations, he/she is in some way insensitive to them. Insensitivity to expectations and rules of society seems to result from an egotistical inability to perceive the effects of one's behavior on other individuals. One important criticism of this theory is that it does not elaborate on *why* untrustworthy individuals are deficient in sensitivity to rules. However, it seems to be clear that those individuals lacking integrity can also be termed untrustworthy, undependable, irresponsible, rule -breaking, impulsive and ,at the very extreme, psychopathic. Recent studies examining the heritability of personality using the California Psychological Inventory may shed light on the genetic source of individual differences in

integrity (see Arvey & Bouchard, unpublished manuscript for a detailed exposition of heritability of personality variables).

Where would the personality trait "honesty" (as measured by pre-employment integrity tests) fall under the Big Five taxonomy? The similarity of integrity measures indicate that they may all primarily measure a general construct. Different test publishers claim that their integrity tests tap into constructs which include responsibility, long term job commitment, consistency, proneness to violence, moral reasoning, hostility, work ethics, dependability, depression, and energy level (as reported in O'Bannon et al., 1989). Given the descriptions of these constructs and the theory at the root of most integrity tests, it can be hypothesized that this general construct is broadly defined "conscientiousness," one of the Big Five dimensions . Conscientiousness reflects characteristics such as dependability, carefulness, and responsibility. In integrity testing literature, this construct seems to have been viewed from its negative pole (e.g., irresponsibility, carelessness, violation of rules). Inspection of items on several integrity tests supports this notion. The following are some examples of such items from integrity tests: "It is okay to get around the law if you don't actually break it," "People think I am irresponsible," "Sometimes I enjoy going against the rules." Therefore, high correlations among the integrity tests, and between integrity tests and other measures of Conscientiousness would be anticipated.

Establishing construct validity is an involved task which can draw upon factor analytic results, contrasted group studies, correlations with admissions of disruptive behaviors, and correlations with polygraph results, in addition to establishing correlations with other psychological constructs. One focus in this dissertation will be on examining the empirical evidence

on which of the Big Five dimensions of personality constructs integrity tests tap into. Evidence related to convergent and discriminant validity will be developed in order to understand better what is involved in employee honesty.

More than one hundred studies correlating integrity tests with other temperament measures have already been reported. These studies will not be individually reviewed here. (See the results section of this dissertation for a meta-analytic review of this literature.) Below I briefly mention some relevant primary studies that have investigated various integrity-personality relationships.

For example, the Hogan Reliability scale and the Personnel Reaction Blank, both personality-based integrity tests, have been correlated with the CPI (Hogan & Hogan, 1989; Gough, 1954). The highest observed correlations were obtained with the Socialization scale, in high .70's for the female sample and .80's for the male sample.

Nolan (1991) investigated the correlations between an overt integrity test and the Big Five, and found that the highest correlations, unadjusted for unreliability in measures and range restriction , were for Agreeableness and Conscientiousness, r's of .35 and .25, respectively. Collins and Schmidt (in press) studied integrity and personality of white collar criminals contrasted with non criminal white collar workers. The pattern of score differences they found indicated that white collar criminals are more irresponsible, and they disregard rules and social norms. Collins and Schmidt (in press) termed the construct they found "Social Conscientiousness," linking it to the Conscientiousness dimension of the Big Five.

Woolley & Hakstian (in press) examined the construct validity of personality-based and overt measures of integrity using the Costa & McCrae (1985) measure of the Big Five. They concluded that: (1) A single factor unifies three major integrity tests (Personnel Decisions Inc.'s Employment Inventory, Personnel Reaction Blank, and Hogan's Reliability scale), and (2) The construct measured by integrity tests appears as to be closer to Agreeableness rather than Conscientiousness.

The empirical evidence for the construct validity of integrity tests in terms of the Big Five leads to Conscientiousness and to a lesser degree Agreeableness. Before any generalizable robust conclusions can be made about which constructs integrity tests measure, further empirical evidence is necessary from correlations with broader range of personality inventories and the broad Big Five dimensions. It is exactly this type of evidence that this dissertation will provide by pooling results from many construct validity studies to arrive at robust and precise convergent and divergent validity evidence for integrity tests.

Theoretically, a case can be made for a broad concept of Conscientiousness/Dependability. It seems plausible that one can demonstrate that integrity tests tap the low end of Conscientiousness while regular personality inventories contain items measuring the positive end of the same construct. In that case, measures of Conscientiousness from the personality inventories and integrity test scores might be non-linearly related. However, for this to happen, the opposite poles of the construct that conscientiousness scales and integrity tests tap into need to be very extreme. Otherwise, the differential item content from the same construct domain in conscientiousness scales and integrity tests will cause the two be correlated

lower than they really are. In any event, the possibility of a nonlinear relationship between these two measures can be checked by examining the bivariate scatter plot of the two.

So far only superficial attention has been directed to construct issues across integrity tests. The construct validity evidence presented for single integrity tests in terms of the Big Five dimensions of personality (Hogan & Hogan, 1989, Gough, 1972, Nolan, 1991, Collins & Schmidt, 1992, Woolley & Hakstian, in press), reflect only the tip of the iceberg.

An understanding of the constructs measured by integrity tests may be gained by systematically exploring the relationships between integrity tests and other measures in the ability and personality domains. The specific questions to be answered in this dissertation are as follows (a description of how each one of these questions will be answered can be found after the general description of the methods of data collection):

1. Do overt integrity tests correlate highly with each other?

2. Do personality-based integrity tests correlate highly with each other?

3. Do both overt and personality-based integrity tests measure the same underlying construct?

4. Do integrity tests correlate with measures of agreeableness, conscientiousness, extroversion, openness to experience, and emotional stability?

5. Do integrity tests derive their criterion related validity from a broadly defined construct of Conscientiousness?

6. Do integrity tests correlate with tests of ability?

It is also important to establish other relationships or lack thereof in exploring the nomological network of integrity (Cronbach & Meehl , 1959). Some specific questions to be answered are:

1. Do integrity tests correlate with age?
2. Do integrity tests correlate with race?
3. Do integrity tests correlate with gender?

Based on the nomological network investigation outlined above, there are some practical questions that will be answered in this dissertation. These practical questions are:

1. What is the precise incremental validity of integrity tests over measures of ability for predicting overall job performance?
2. What are the precise implications of using integrity tests for adverse impact?
3. What levels of operational validity can be obtained by combining measures of the Big Five dimensions of personality (as measured by traditional personality instruments) and integrity (as measured by integrity tests) for personnel selection?

This dissertation is unique in many ways. It is aimed at developing a theory of conscientiousness and integrity. Yet, at the same time it has practical implications for personnel selection (especially for incremental validity). So far only superficial attention has been directed to construct issues across integrity tests. This dissertation will be the first of its kind to investigate the construct validity of more than four integrity tests at once. This is also the first systematic attempt to compile a fairly broad nomological net for the construct of integrity .

CHAPTER III

METHODS

Establishing the network of relationships between integrity tests, personality measures, ability tests and demographic variables such as race, gender and age, involves constructing a matrix of intercorrelations among these variables. In particular, correlations between various integrity tests, and the Big Five dimensions of personality are necessary. Also, intercorrelations among and between the following variables need to be obtained: overt integrity tests (e.g. Personnel Selection Inventory of London House, Stanton Survey, Reid Report), personality-based integrity tests (e.g. Hogan Personality Inventory's Reliability Scale, Personnel Decisions Inc.'s Employment Inventory, Personnel Reaction Blank), and the Big Five dimensions of personality (Conscientiousness, Agreeableness, Neuroticism, Extroversion, Openness to Experience). In this research, various modes of data collection and analysis were employed.

Here, I describe in detail the various modes of data collection used to compose the databases that were used to answer the questions about the construct validity of integrity tests. Data analysis details pertaining to different questions will be provided in the "Analyses and Results" chapter of the dissertation.

Modes of Data Collection

Primary Data Collection

The primary data collection for this study took place between September 1991 and January 1993. Overall there were thirty sessions of data collection. The samples comprised of students from a large Midwestern University (receiving extra credit for participation) and job applicants for skilled manufacturing jobs at Midwestern plants of a Fortune 500 company. Samples were administered various combinations of integrity tests and personality inventories. The primary data collection used the following instruments: London House Personnel Selection Inventory, PDI Inc.'s Employment Inventory, Stanton Survey, Reid Report, Hogan Personality Inventory (including the Reliability Scale), Personnel Reaction Blank, Inwald Personality Inventory, Personal Characteristics Inventory, and Goldberg's Adjective Checklist. Table 4 lists the instruments used and their brief descriptions.

Additional data was collected on the following variables: gender, race, and ratings of the acceptability/fairness of the above instruments. The questionnaire that were used for this additional data collection is provided in Appendix A. Data on the acceptability/fairness of various integrity tests was collected for a separate study and will not be used to answer any of the twelve questions posed in this dissertation. However, it is necessary to include this questionnaire in an Appendix even though no results are being reported, because demographic data on the samples was obtained using questions from this survey. The general instructions that were given to subjects before they took various combinations of the instruments and the debriefing statement provided at the end of the session are provided in Appendices B and C,

respectively. The overall sample size in the primary data collection was 1,365 . A total of 300-500 individuals took any given instrument. Appendix D provides the sample sizes for all the instruments used in the primary data collection. Forty three percent of the student sample were females, 40% were males (with the rest not indicating gender). The racial composition of the sample was as follows: 68% percent were whites, 17% blacks, 7% Hispanics, 6% Asian and 2% other. 82% of the individuals were originally from the Midwest. Of the student sample 26% were freshman, 9% sophomores, 49% juniors, 17% seniors, and .7% graduate students.

Secondary Data Collection

Many studies correlating integrity tests with other measures have been reported. Even though overt and personality-based tests have been correlated with different inventories, the Big Five dimensions of personality provide an appropriate framework for determining what construct(s) integrity tests tap into. All existing correlations were collected between and among: overt integrity tests, personality-based integrity tests, Big Five dimensions of personality (conscientiousness, emotional stability, agreeableness, extroversion, and openness for experience), measures of ability, age, gender, and race. Published and unpublished studies were read and coded. Appendices E, F, and G contain the study coding sheets used. They were developed using the Office of Personnel Management's (OPM) validity generalization database coding sheet. OPM's guidelines were followed in coding the correlations from different studies.

All existing correlations between and among overt integrity tests, personality-based integrity tests, and measures of the Big Five dimensions of personality (conscientiousness, emotional stability, agreeableness,

extroversion, and openness for experience) were used. All unpublished and published technical reports reporting validities, reliabilities, or range restriction information were obtained from integrity test publishers and authors. Some integrity test authors and test publishers responded to my request for construct validity information on their test by sending me computer printouts that had not been written up as technical reports. These were included in the database. Published and unpublished studies were read and coded. The results of the primary data collection described earlier were also included in this database. The list of integrity tests contributing data to this dissertation is provided in Table 5.

The list of studies included in the database for meta-analyses is provided in Appendix I. Some technical reports omitted important and valuable pieces of information necessary to include study results in meta-analyses. I strongly encourage researchers and test publishers to report in the technical reports and test manuals the reliability of the measures used, sample sizes used, validity coefficients, and the specific criterion for which the validity of the test is being reported . Further there should be clear description of the validation strategy used in the study being described (whether the study was predictive or concurrent), the samples used (students, applicants, or employees), and the specific jobs studied. Demographic information on the samples (e.g., age range, sex, race) should also be reported. Failure to report the necessary information results in unnecessary and lengthy correspondence between researchers and test publishers. In most cases I was able to obtain additional information via personal communications with test publishers.

There were more than 8,000 correlation coefficients independently coded by two researchers (Myself and Chockalingam Viswesvaran). Across all coded correlation coefficients, there was 94.7% agreement. Most of the disagreements between the coders resulted from vague reporting of information in technical reports and other unpublished sources. To resolve each disagreement, the test publishers were contacted to inquire about the item of disagreement. In all of the disagreements, the new data obtained from the test publisher resolved the disagreement.

Hypotheses and Procedures Used to Test Them

Questions Pertaining to the Relations among Integrity Tests, Big Five Dimensions of Personality and their Implications

1. Do overt integrity tests correlate highly with each other?

Overt integrity tests were all originally designed to predict theft and theft related activities. These tests assess attitudes toward theft and other illegal activities. Some overt integrity tests include admissions items. The underpinnings of all integrity tests are similar; therefore, all of them can be hypothesized to correlate highly with each other. To test this hypothesis the intercorrelations between overt tests were obtained using data from the primary data collection. Once the correlations between overt tests were established, secondary confirmatory factor analysis was used to show that overt integrity tests intercorrelations arise due to the presence of a single factor. The measurement model consisted of the Reid Report, the Stanton Survey, and the Personnel Selection Inventory of London House. Figure 1 shows the hypothesized measurement model for overt tests.

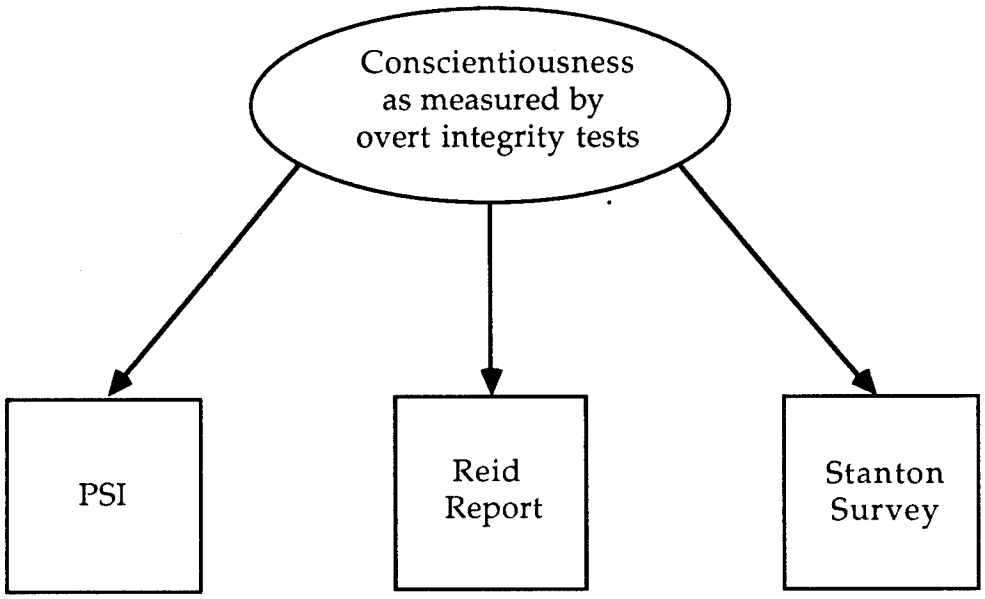

Figure 1. Conscientiousness and overt integrity tests.

Also, meta-analytic procedures were used to cumulate all correlations reported among overt tests (statistical artifact distributions used to correct the correlations are described in the analyses and results section). The data came from studies reporting such correlations and from the primary data collection effort described earlier. Overall, for this meta-analysis there were 56 correlations across 7,424 individuals.

2. Do personality-based integrity tests correlate highly with each other?

Personality-based integrity tests were all originally designed to predict broadly defined counterproductive behaviors and not theft. These tests are all derived from various personality inventories . The underpinnings of all personality-based integrity tests are similar. The common thread that spans all personality-based integrity tests seems to be trustworthiness, integrity, and conscientiousness. Therefore, all of them can be hypothesized to correlate

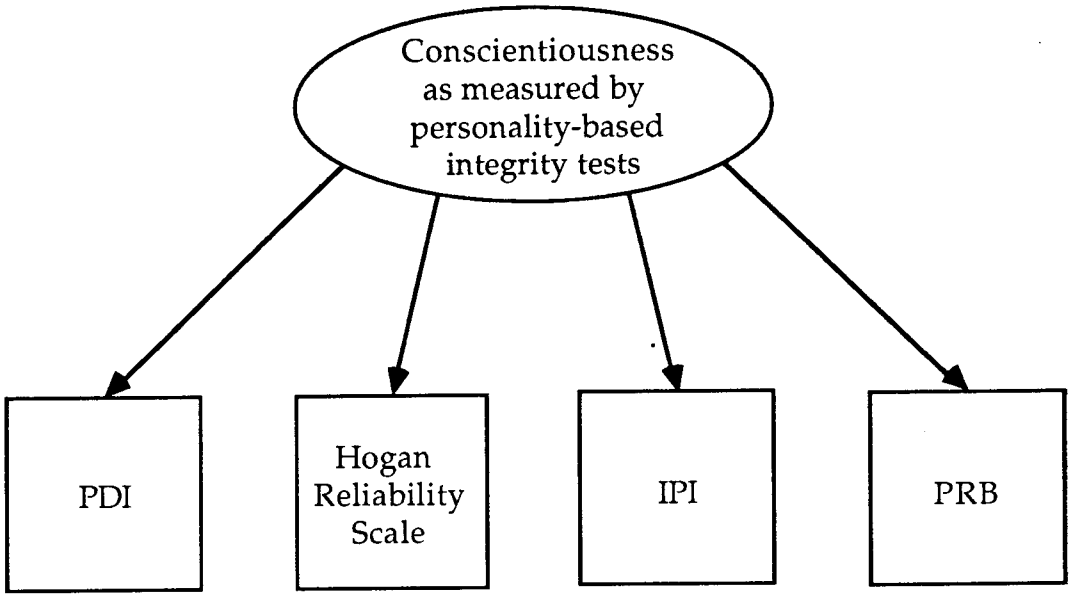

Figure 2. Conscientiousness and personality-based integrity tests.

highly with each other. To test this hypothesis the intercorrelations between personality-based tests used in the primary data collection were obtained. Once the correlations between personality-based tests were computed, confirmatory factor analysis was used to show that correlations between personality-based integrity tests result from the presence of a single factor. The measurement model will consist of the PDI Employment Inventory, the Personnel Reaction Blank, the Inwald Personality Inventory, and the Hogan Personality Inventory's Reliability scale. Figure 2 shows the hypothesized measurement model for personality-based tests.

Additional data was obtained from studies reporting correlations between various integrity tests. In this case, data from other studies and from the primary data collection effort described earlier were combined to obtain a pool of correlations to meta-analyze. There were 37 correlations available,

across 7,062 individuals. Psychometric meta-analysis was used to cumulate the correlations (the artifact distributions used to correct the correlations are reported in the analyses and results part of this dissertation).

3. Do both overt and personality-based integrity tests measure the same underlying construct?

This question can be answered by showing that overt and personality-based tests correlate highly. The correlations between overt and personality-based integrity tests were computed using data from the primary data of this dissertation. Confirmatory factor analysis was used on the intercorrelation matrix of integrity tests to calculate the factor loadings of each integrity test and two categories of tests on a general factor. The possibility of group factors by integrity measurement method (overt vs. personality-based) was also explored. Two alternate factor structures to be tested are depicted in Figures 3 and 4. Figure 3 depicts the case where there is a single general factor being tapped into across integrity tests and the lack of any group factors. In contrast, Figure 4 depicts the case where in addition to a general factor there are two group factors based on the type of integrity test.

Also, correlations between these overt and personality-based test types were obtained from published and unpublished studies and the primary data described earlier to conduct a meta-analysis (statistical artifact distributions used to correct the correlations are described in the analyses and results part of the dissertation). There were 117 correlations for this meta-analysis across 15,978 individuals.

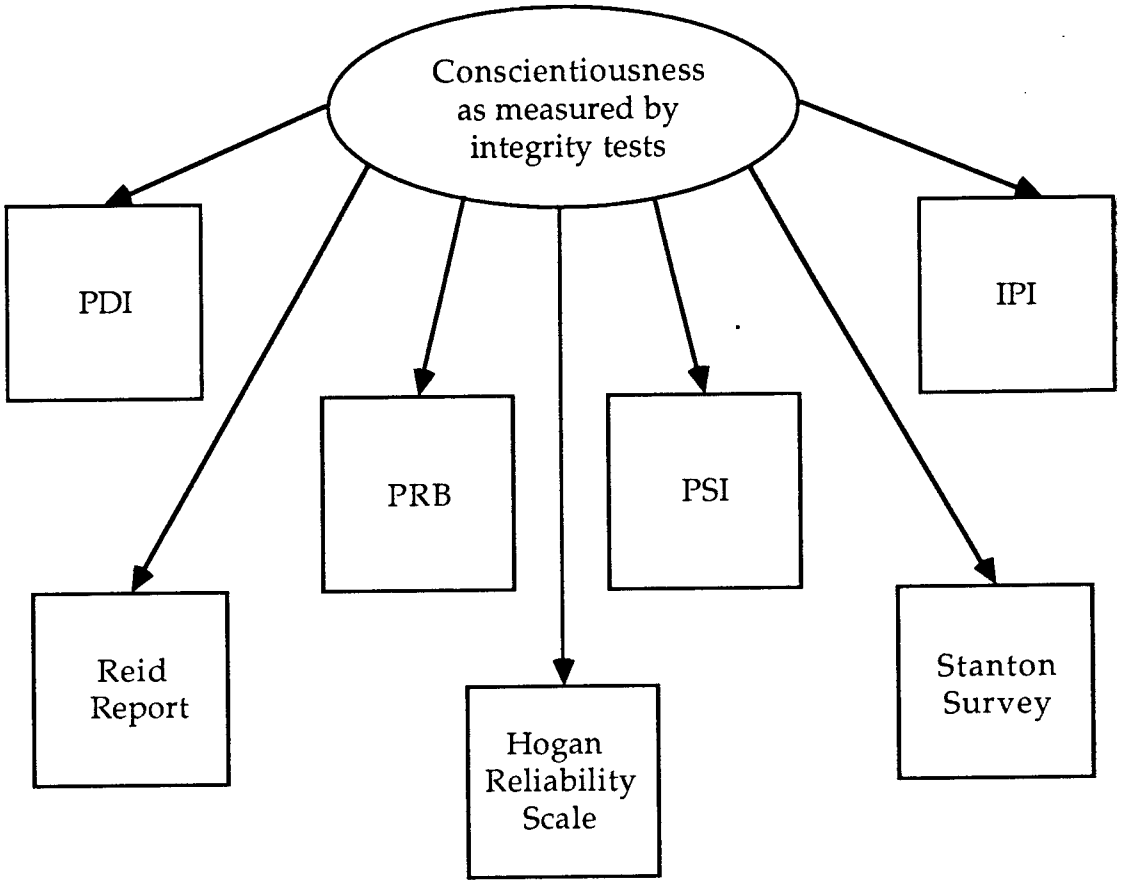

Figure 3. Conscientiousness and integrity tests.

4. Do integrity tests do correlate with measures of agreeableness, conscientiousness, extroversion, openness to experience, and emotional stability?

Questions 1, 2, and 3 of this dissertation are aimed at determining whether (1) overt integrity tests measure a single construct, (2) personality-based integrity tests measure a single construct, and (3) overt and personality-base tests are highly correlated. The next question is what overall construct integrity tests measure. To systematically examine the dimension of

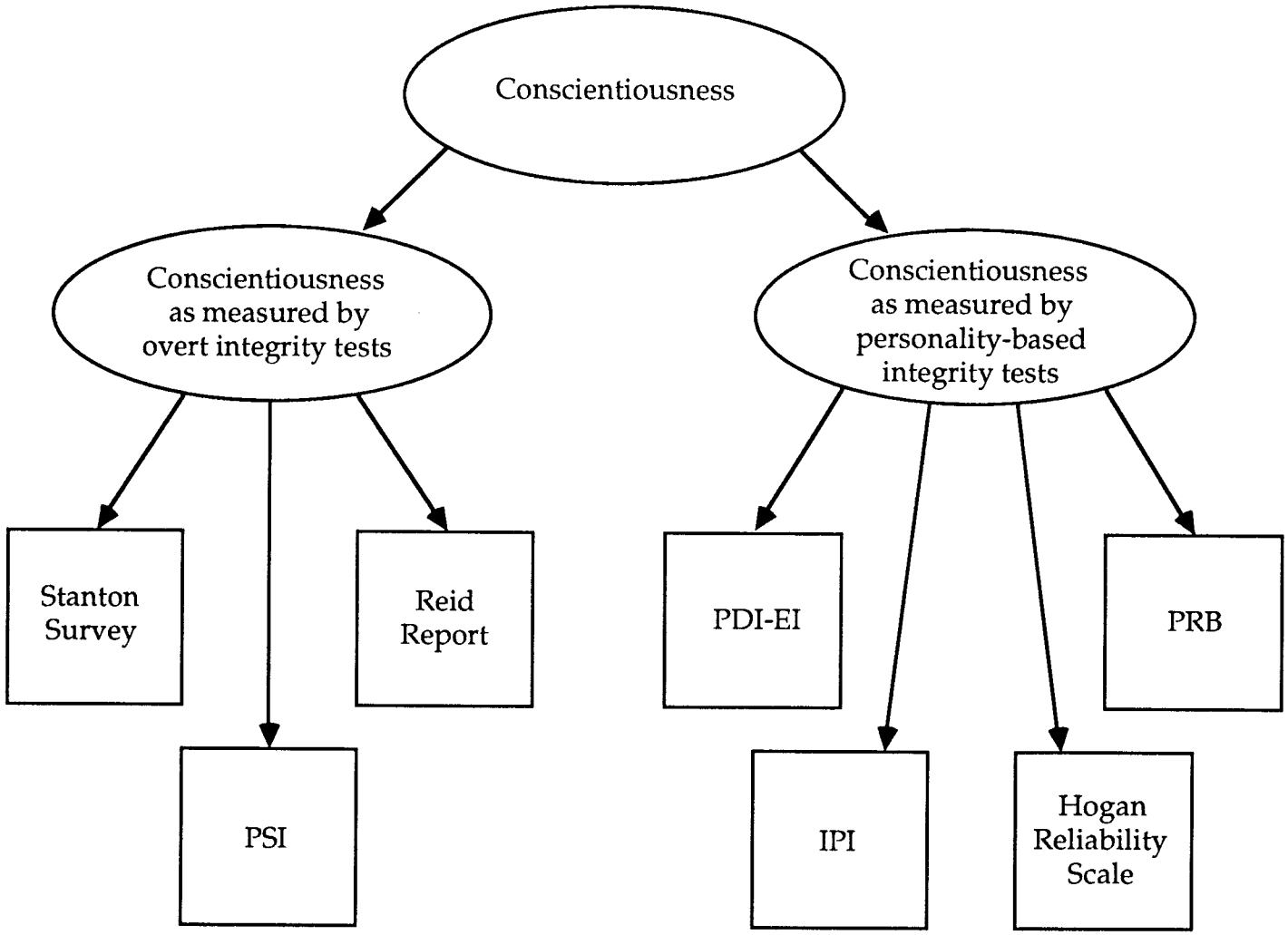

Figure 4. Conscientiousness and integrity tests: Alternate framework with group factors.

personality integrity tests tap into, the Big Five dimensions of personality were used. Correlations between integrity tests and the Big Five Personality dimensions were computed using the primary data. Correlations between integrity tests and the Big Five Personality dimensions were also meta-analytically examined. Overall, there were 1,506 correlations reported between integrity tests and personality inventories measuring the Big Five.

Of the Big Five dimensions of personality, conscientiousness can be hypothesized to be the most consistent correlate of integrity tests. Given the construct validity studies so far (e.g. Woolley, 1991), agreeableness is also anticipated to correlate with integrity tests. Inwald (1988) and others have hypothesized that integrity test scores may reflect sensation seeking and risk taking tendencies, both aspects of extroversion. If this is true, some correlations may be expected between integrity test scores and extroversion. No or very low correlations are expected between emotional stability, openness to experience, and integrity. Table 6 depicts the hypothesized intercorrelation matrix of integrity and the Big Five dimensions of personality. The hypothesized relationships or lack thereof are indicated. Once the correlations between integrity tests and each of the Big Five dimensions are obtained , these correlations can be compared to those found between conscientiousness and the other Big Five dimensions. If the hypothesis that integrity tests measure a broadly defined construct of conscientiousness is correct, these two sets of correlations should be about the same.

After correlations between integrity tests and measures of Big Five are examined, a regression equation can be used to predict integrity from the Big Five dimensions of personality. For previously discussed reasons, in this

regression equation the largest weight is expected for conscientiousness; the second largest weight is expected for agreeableness and the third largest weight is expected for extroversion. Such a weighted combination gives the weighted combination that best predicts and describes integrity.

5. Do integrity tests derive their criterion-related validity from a broadly defined construct of Conscientiousness?

Stated alternately , this question aims to find an answer to the question whether integrity test validities for predicting supervisory ratings of job performance can be explained by a broadly defined construct of conscientiousness. The meta-analytically obtained intercorrelation matrix of integrity tests and personality measures can be used to compute the partial correlation between integrity and job performance, if the hypothesis that integrity tests measure conscientiousness is correct, this partial correlation should be negligibly small. Furthermore, if integrity tests and conscientiousness measure the same construct, the multiple correlation of integrity and conscientiousness with job performance should not be higher than the integrity job performance correlation. If integrity tests and conscientiousness scales measure the same trait but sample from different aspects of the domain, the sum of the two should be a better measure of the overall construct they measure.

6. What levels of operational validity can be obtained by combining measures of the Big Five dimensions of personality (as measured by traditional personality instruments) and integrity (as measured by integrity tests) for personnel selection?

This practical question can be answered if the correlations among integrity test and the Big Five dimensions of personality (as measured by personality inventories) are known. These correlations were meta-analytically obtained using data from the primary and secondary data collections described earlier. Then the standard formula for a multiple correlation was used to calculate the multiple correlation for supervisory ratings of job performance, when the predictors are tests of integrity tests and Big Five scales from personality inventories. If integrity tests mainly measure conscientiousness, adding conscientiousness as measured by personality inventories to a battery that already contains integrity tests should not improve prediction. The results of this analysis will differ depending on whether the correlations are corrected for measurement error (see Schmidt, Hunter, & Caplan, 1981, for details). In personnel selection, operational validity is appropriately calculated without correction for measurement error in the predictors. In answering the question of what levels of operational validity can be obtained by combining measures of the Big Five and integrity for personnel selection, I will compute operational validity without correction for measurement error in predictors. This is because I am interested in the answer to a practical application question in personnel selection in this particular case.

Questions Pertaining to Relationships between Integrity Tests and Demographic Variables, and their Implications for Adverse Impact

1. Do integrity tests correlate with age?

To answer this question, correlations between integrity test scores and age were meta-analytically cumulated. Evidence from individual studies conducted by integrity test publishers and previous dissertations (e.g. Lasson 1992; Schlessinger, 1993; Woolley, 1991) suggest that age and integrity test scores are slightly correlated (Goldberg et al., 1991). That is, younger individuals score somewhat lower. This is possibly due to experimentation and youthful indiscretions that characterize late teenage years and early adulthood.

2. Do integrity tests correlate with gender?

Review of studies of adverse impact of integrity tests suggest that scores on integrity tests are somewhat correlated with gender (Goldberg et al., 1991). Females on average score slightly higher than males on integrity tests. To precisely estimate the exact magnitude of the relationship between gender and integrity test scores, all published and unpublished studies (including the primary data analyses of this dissertation) reporting correlations between gender and integrity tests were read and coded. Meta-Analysis was used to cumulate these correlations. Then the cumulated correlation were expressed in terms of mean differences between males and females in standard deviation units.

An interesting possibility is interaction between gender and age. That is, it is plausible that the differences in integrity between males and females may diminish by age. Alternately stated, the correlation between sex and

integrity test scores could be lower at older ages. Testing this interaction would require the correlations between integrity test scores and gender for various age groups. No such correlations have been reported in the published or the unpublished literature on integrity tests; therefore, unfortunately this dissertation can only focus on the bivariate relationship between integrity test scores and gender.

3. Do integrity tests correlate with race?

Review of studies of adverse impact of integrity tests suggest that scores on integrity tests do not cause any adverse impact for minority groups. To examine how similarly minorities and whites score on integrity tests, all published and unpublished studies (including the primary data analyses of this dissertation) reporting correlations between race and integrity were read and coded. Meta-Analysis was used to precisely estimate the correlation between race and integrity test scores. This correlation was then used to answer questions about adverse impact of integrity tests in general.

4. What are the precise implications for adverse impact of using integrity tests ?

Hunter and Hunter (1984) indicate that it may be possible to identify other predictors that will add to the validity of general mental ability and at the same time reduce adverse impact. Integrity test publishers have devoted considerable research to examining the question of adverse impact. Many studies have found no adverse impact for minorities (e.g., Arnold, 1989; Bagus, 1988; Cherrington, 1989; Moretti & Terris, 1983; Strand & Strand, 1986; Terris & Jones; 1982). Sackett et al. (1989, p. 499) concluded " . . . minority

groups are not adversely affected by either overt integrity tests or personality oriented measures."

From the ability testing and personnel selection literature, we know that blacks average about one standard deviation below whites on tests of general mental ability. This difference of one standard deviation between blacks and whites on general mental ability tests can also be expressed as a correlation between ability and race, r = .45. This correlation is obtained using the standard formula for converting effect sizes (1 sd here) to correlations (.45). The correlations between ability and integrity scores and race and integrity scores were estimated. The correlation between race and an optimally weighted (using regression weights) composite of ability and integrity was then computed. This correlation was then be converted to an effect size , an indicator of the mean difference between blacks and whites on an optimally weighted composite of ability and integrity test scores.

Questions Pertaining to the Relationship between Integrity and Ability, and its Implications for Incremental Validity

1. Do integrity tests correlate with tests of ability?

Evidence from studies so far indicate that integrity tests do not correlate with measures of ability (Ones et al., in press). This is important because ability tests have the highest operational validities of all personnel selection devices and if integrity tests are to have incremental validity above and beyond measures of ability, integrity and ability should not correlate highly. To test the hypothesis that measures of integrity and ability do not correlate with each other, correlations between ability and integrity tests were cumulated using meta-analysis. Correlations from primary data collected

were also be included among the correlations meta-analyzed. (In the primary data collection the Wonderlic Personnel Test is used to measure ability.)

2. What is the precise incremental validity of integrity tests over measures of ability for predicting overall job performance?

The correlations between measures of integrity and measures of ability were meta-analyzed in order to obtain more precise estimates of the magnitude of the incremental validity of integrity tests. An unanswered question in Ones et al (in press) is the size of the precise increment in validity from adding integrity tests to general mental ability tests in predicting overall job performance in personnel selection. Available studies suggest that the correlations between integrity measures and ability measures are low and negligible. For example, Jones and Terris (1983) found that the correlations between an overt integrity test and a measure of general mental ability were -.02 for the theft admissions subscore and -.03 for the theft attitudes subscore; Gough (1972) reported that a vocabulary test correlated -.05 with the Personnel Reaction Blank; Werner, Jones, and Steffy (1989) reported that integrity test scores are unrelated to educational level (an arguable proxy for ability); and Hogan and Hogan (1989) reported correlations of .07 and -.09 between the Hogan Reliability Scale (a personality-based integrity test) and the quantitative and verbal portions of the Armed Services Vocational Aptitude Battery (ASVAB), respectively. Meta-analytic cumulation of correlations between ability and integrity measures can be expected to show that the correlation between the two is zero or negligibly small. Then the expected incremental validity of integrity tests in predicting supervisory ratings of overall job performance can be calculated.

CHAPTER IV
ANALYSES AND RESULTS

Intercorrelations Among Integrity Tests and the Big Five Dimensions of Personality and their Implications

The first four questions to be answered in this dissertation are: (1) Do overt integrity tests correlate highly with each other? (2) Do personality-based integrity tests correlate highly with each other? (3) What is the correlation between overt and personality based integrity tests? (4) Do integrity tests correlate with measures of agreeableness, conscientiousness, extroversion, openness to experience, and emotional stability?

Establishing the network of relationships between integrity tests, and temperament measures involves constructing a matrix of intercorrelations among various integrity tests, and the Big Five dimensions of personality. In particular, intercorrelations among and between the following variables were obtained: overt integrity tests (e.g., PSI, Stanton Survey, Reid Report), personality-based integrity tests (e.g., Hogan Reliability Scale, PDI Employment Inventory, Personnel Reaction Blank), and the Big Five dimensions of personality (Conscientiousness, Agreeableness, Neuroticism, Extroversion, Openness to Experience). Data came from the primary data collection described earlier and from studies reporting correlations of integrity tests with other measures.

Psychometric meta-analysis (Hunter & Schmidt, 1990) was conducted using secondary data to establish a matrix of precisely estimated intercorrelations among the following constructs: overt integrity and personality-based integrity tests, and measures of conscientiousness, agreeableness, neuroticism, extroversion, and openness to experience. Even though overt and personality-based tests have been correlated with different inventories, the Big Five dimensions of personality provide an appropriate framework for determining what construct(s) integrity tests tap into. The list of all personality inventories scales of which have been correlated with integrity tests is provided in Table 7.

Various scales of personality inventories were classified into the Big Five dimensions of personality following the scheme established by Barrick and Mount (1991) and Hough et al. (1990). Appendix H provides the summary of which personality scales were assigned to which of the Big Five dimensions.

A total of 1,716 correlation coefficients comprised the portion of the database used to determine the intercorrelations between integrity tests and integrity tests and the Big Five dimensions of personality. The total sample size across 1,716 correlations was 367,970.

Several sets of artifact distributions were compiled: 3 distributions for the reliability of the integrity tests, one distribution for the reliability of personality scales, and one distribution for range restriction to correct for range restriction due to selection on integrity. (It was not necessary to correct for range restriction in personality variables because there was no evidence of direct range restriction on personality variables; the sample standard deviations reported for various scales were not any more different than the

referent group standard deviations than would be expected by chance.) Descriptive information on the artifact distributions are provided in Table 8.

A total of 124 integrity test reliability values were obtained from a database created by Ones et al. (in press). Of the 124, 68 were alpha coefficients (55%) and 47 were test-retest reliabilities over periods of time ranging from 1 to 1,825 days (mean = 111.4 days; sd = 379.7 days). The mean of the coefficient alphas was .81 (sd = .10) and the mean of the test-retest reliabilities was .85 (sd = .10). There were 9 reliabilities reported without stating the type of reliability. The ideal estimate of reliability for purposes of this meta-analysis is coefficient alpha or the equivalent. However, test-retest reliability estimates over relatively short time periods provide reasonably close approximations to alpha coefficients. Further, in this case the means of the two reliability types were similar. The overall mean of the integrity test reliability artifact distribution was .81 and the standard deviation was .11. The mean of the square roots of integrity test reliabilities was .90 with a standard deviation of .06. Two other predictor reliability distributions were constructed: one for overt integrity tests and another for personality-based integrity tests. There were 97 reliabilities reported for overt tests. The mean of the overt test reliability artifact distribution was .83 and the standard deviation was .09. The mean of the square roots of overt test reliabilities was .91 with a standard deviation of .05. There were 27 reliabilities reported for personality-based tests. The mean of the personality-based test reliability artifact distribution was .72 and the standard deviation was .13. The mean of the square roots of the reliabilities was .85 with a standard deviation of .08. Each one of these integrity test reliability distributions were used in analyses in corresponding meta-analytic categories. That is, when correlations of overt

tests were being cumulated, the reliability distribution for overt tests was used. But when correlations of personality-based tests were being meta-analyzed, the reliability distribution for personality-based tests was used. Finally, when the analyses involved both overt and personality-based tests (i.e., all integrity tests), the overall integrity test reliability distribution was used.

Reliability estimates for the personality scales were taken from the studies that contributed to the database for this meta-analysis and Barrick and Mount (1991). Initially, a separate reliability distribution was created for each of the Big Five dimensions of personality; however, the means and standard deviations of these five reliability distributions were not any different than would be expected due to sampling error. Hence, all the reliabilities of personality scales were combined to form one robust overall distribution. The overall distribution of personality scale reliabilities had 372 values. The Barrick and Mount (1991) study provided 156 of these reliabilities. The studies that contributed to the database for this meta-analysis provided 216 reliabilities. Of the 372 personality scale reliabilities, 203 were alpha coefficients (55%) and 13 were test-retest reliabilities (4%). There were 156 reliabilities reported without stating the type of reliability (from Barrick and Mount, 1991). The overall mean of the personality scale reliability artifact distribution was .76 and the standard deviation was .11. The mean of the square roots of personality scale reliabilities was .87 with a standard deviation of .07.

Because integrity tests are used to screen applicants, the correlations calculated using employee samples may be affected by restriction in range. The range restriction ratio is calculated as the ratio of study to reference group

standard deviations (s/S). The distribution of integrity test range restriction values was the same one used by Ones et al. (in press) in their meta-analysis of integrity test validities. The mean ratio of the restricted sample's standard deviation to the unrestricted sample's standard deviation used was .81 for integrity tests . This range restriction value indicates that there is considerably less range restriction in integrity domain than is the case for cognitive ability (Alexander, Carson, Alliger, & Cronshaw, 1989). Thus, range restriction corrections were much smaller in present research than in meta-analyses in the abilities domain.

The hypotheses were tested using the Hunter-Schmidt (1990, p. 185) psychometric meta-analytic procedure. In this study, psychometric meta-analysis was used to calculate true score correlations among overt and personality-based integrity tests and measures of the Big Five. The artifact distributions described above were used to correct biases in the observed validities caused by statistical artifacts. The artifacts operating across studies include sampling error, unreliability in the measures, and range restriction. Meta-analysis provides the most accurate obtainable estimate of the mean true score correlation between variables. In this study, the interactive meta-analysis procedure was used (Hunter & Schmidt, 1990a, p.165; Schmidt, Hunter, & Gast-Rosenberg, 1980). One reason why meta-analytically obtained true score corrections between constructs will be used in this research rather than observed correlations is that, uncorrected correlations cannot estimate relations between constructs (Schmidt, 1993). Observed correlations reveal relations between imperfect measures of constructs. That is why for exploring the questions posed in this research true score correlations were used.

Results from the Primary Data

Correlations among overt tests (Primary data only)

To answer the first question of whether overt integrity tests correlate highly with each other, the intercorrelations among overt integrity tests were computed using the primary data. The results are presented in Table 9.

The observed correlations among overt integrity tests are high. Furthermore, when these correlations are corrected for unreliability in the overt tests, the true score correlations average to .85. These results seem to indicate that there is a high degree of overlap between overt integrity tests.

Confirmatory factor analysis was used to confirm that overt integrity tests intercorrelations arise due to the presence of a single factor. The measurement model consisted of the Reid Report, the Stanton Survey, and the Personnel Selection Inventory of London House as manifest variables (the measurement model for this was presented in Figure 1). The presence of a general factor was confirmed. The factor loadings of the of the individual tests on this general factor were as follows: .82 for the Reid Report, 1.00 for the Stanton Survey, and .87 for the Personnel Selection Inventory. The residual matrix of intercorrelations consisted entirely of zeros. Therefore, it can be concluded that these three overt integrity tests measure the same general factor.

Correlations among personality based tests (Primary data only)

To answer the second question of whether personality-based integrity tests correlate highly with each other, the intercorrelations among personality-based integrity tests were computed using the primary data. The results are presented in Table 10.

The observed correlations among personality-based integrity tests are fairly high. When these correlations are corrected for unreliability in the personality-based tests, the average true score intercorrelation among personality-based tests is .75. These results seem to indicate that there is a high degree of overlap between personality-based integrity tests; however, this overlap is somewhat less than the one found among overt integrity tests.

Confirmatory factor analysis was used to confirm that personality-based integrity tests intercorrelations arise due to the presence of a single factor. The measurement model consisted of the PDI- Employment Inventory, IPI Risk factor, the Personnel Reaction Blank, and the Hogan Personality Inventory's Reliability scale as manifest variables (the measurement model for this was presented in Figure 2). Confirmatory factor analysis indicated that these four personality-based integrity tests tap into a general construct. The factor loadings of the of the individual tests on this general factor were as follows: .75 for the PDI- Employment Inventory, .90 for the IPI Risk factor, .87 for the Personnel Reaction Blank, and .86 for the Hogan Personality Inventory's Reliability scale. The largest residual correlation was .06. It seems reasonable to conclude that the four personality-based integrity tests used here measure the same general factor.

Correlations between overt and personality based tests (Primary data only)

To investigate the relationship between overt and personality based integrity tests, I computed the intercorrelations between the seven integrity tests that were used in the primary study of this research. The intercorrelations between three overt (PSI, Reid Report, and the Stanton Survey) and four personality-based integrity tests (PRB, PDI-EI Performance scale, IPI critical scale, Hogan Reliability scale) are provided in Table 11.

In general, the correlations presented in Table 11 are lower than those presented in both Tables 9 and 11. Overt tests seem to correlate higher among themselves than they do with personality-based tests. Personality-based tests seem to correlate higher among themselves than they do with overt tests.

A composite of the three overt tests (PSI, Reid Report, and the Stanton Survey) and a composite of the four personality-based integrity measures (PRB, PDI-EI Performance scale, IPI critical scale, Hogan Reliability scale) was formed. The rationale behind forming the composites was to define "integrity' as what is shared across overt tests in one case and what is shared across personality-based tests in the other. In other words, in forming composites factors specific to each test were treated as error. Also, the reliabilities for the three overt tests and for the four personality-based tests were computed. This was done using the formula for reliability of a composite (standardized coefficient alpha, which assigns specific factor variance to measurement error). The reliability of the overt tests' composite is .93. The reliability of the personality-based tests' composite is .89.

The observed correlation between the composite of overt integrity tests and the composite of the personality-based integrity tests is .55. The true score correlation between the composite of overt integrity tests and the composite of the personality-based integrity tests is .61. The true score correlation of .61 strengthens the inference of a large common factor across overt and personality-based integrity tests.

Confirmatory factor analysis was used on the intercorrelation matrix of the seven integrity tests used in the primary data collection of this dissertation to calculate the factor loadings of each integrity test on a general factor. This was done to test whether integrity tests intercorrelations arise due

to the presence of a single factor. The measurement model consisted of three overt tests and four personality-based integrity tests; the Reid Report, the Stanton Survey, and the Personnel Selection Inventory of London House, PDI- Employment Inventory, IPI Risk factor, the Personnel Reaction Blank, and the Hogan Personality Inventory's Reliability scale (the measurement model for this was presented in Figure 3). The factor loadings of the of the individual tests on a general factor were as follows: .69 for the Reid Report, .86 for the Stanton Survey, .87 for the Personnel Selection Inventory, .66 for the PDI- Employment Inventory, .83 for the IPI Risk factor, .77 for the Personnel Reaction Blank, and .82 for the Hogan Personality Inventory's Reliability scale. An inspection of the residual matrix of intercorrelations indicated that the fit of the model could be improved by the respecification of the model in terms of two correlated factors being tapped into by integrity tests. The model was respecified such that overt integrity tests loaded on one factor and personality-based tests loaded on another factor (these two factors were allowed to correlate with each other). The factor loadings for the overt tests on the first factor were as follows: .80 for the Reid Report, 1.00 for the Stanton Survey, .94 for the Personnel Selection Inventory. The factor loadings of the personality-based tests on the second factor were as follows: .72 for the PDI-Employment Inventory, .91 for the IPI Risk factor, .85 for the Personnel Reaction Blank, and .89 for the Hogan Personality Inventory's Reliability scale. The correlation between the two factors (one factor defined by overt tests and the other defined by personality-based integrity tests) is .66. It appears that there is evidence for a general factor being tapped into across integrity tests but in addition to a general factor, there are two group factors based on the type of integrity test.

The relationships between integrity tests and the Big Five dimensions of personality (Primary data only)

Using the primary data the relationships between integrity tests and the Big Five dimensions of personality were investigated. As was indicated in the methods section of this dissertation, to check whether integrity test scores are nonlinearly related to scores on conscientiousness scales, the bivariate plots between these two variables were examined. There was no evidence of nonlinearity. Correlations between integrity tests and each of the Big Five dimensions of personality were computed.

A linear composite of the seven integrity tests used in this research (PSI, Reid Report, Stanton Survey, Personnel Reaction Blank, Hogan Reliability scale, IPI, and PDI Employment Inventory) was formed. Also, composites for the Big Five dimensions of personality using scale scores from three personality inventories tapping the Big Five (Personal Characteristics Inventory, Hogan Personality Inventory, and Goldberg's Adjective Checklist) were formed. In other words, corresponding Big Five scales from the three personality inventories were combined into composites. Then, for each one of the Big Five dimensions, the correlation between the composite of integrity tests and the composite of the corresponding scales of the three personality inventories was computed. The rationale behind forming the composites was to define "integrity" as what is shared across integrity tests, overt and personality-based, and to define each dimension of the Big Five as what is shared across corresponding scales from different personality inventories. The correlations between the integrity composite and each of the Big Five composites are presented in Table 12.

Conscientiousness is the strongest correlate of integrity tests (ρ=.91), overt and personality-based tests combined. This confirmed our earlier hypothesis that integrity tests primarily tap into conscientiousness. It seems that individuals who are dependable, responsible and who follow rules are also high on integrity tests. However, Agreeableness and Emotional Stability are also correlated with integrity tests scores (true score correlations of .61 and .31, respectively). Individuals who are agreeable , cooperative and not hostile tend to be high on integrity tests. Individuals who are emotionally stable and not neurotic also tend to score high on integrity tests. The linear composite of extroversion correlates .31 (true score correlation) with linear composite of overt and personality-based integrity test scores, indicating that there is a somewhat weak positive relationship between extroversion and integrity. Finally, Openness to Experience correlates .08 with integrity, indicating a minimal overlap between the two constructs.

Next, the relationships between overt integrity tests and the Big Five dimensions of personality were examined. A linear composite of the three overt integrity tests used in this research (PSI, Reid Report, and Stanton Survey) was formed. Also, composites for the Big Five dimensions of personality using scale scores from three personality inventories tapping the Big Five (Personal Characteristics Inventory, Hogan Personality Inventory, and Goldberg's Adjective Checklist) were formed. In other words, corresponding Big Five scales from the three personality inventories were combined into composites. The reason behind forming the composites was to define "integrity" as what is shared across overt tests and to define each dimension of the Big Five as what is shared across corresponding scales from personality inventories. Then, for each one of the Big Five dimensions, the

correlation between the composite of integrity tests and the composite of the corresponding scales of the three personality inventories was computed. The results are presented in Table 13.

The composite for conscientiousness correlates .68 with the composite for overt integrity tests (ρ=.68). Of the Big Five, the composite of conscientiousness correlates highest with the composite of overt integrity tests. However, composites of Agreeableness and Emotional Stability are also correlated with the composite of overt integrity tests scores (true score correlations of .47 and .31, respectively). The linear composite of extroversion correlates .19 (true score correlation) with the linear composite of overt integrity test scores, indicating that there is a somewhat weak positive relationship between extroversion and integrity. Finally, the composite for Openness to Experience correlates .08 with the composite of overt integrity test scores, indicating a minimal overlap between the two constructs. It is important to note that this pattern of correlations is identical to that presented for integrity tests, overt and personality-based combined, even though the magnitudes are somewhat different.

Next the pattern of correlations between the Big Five dimensions of personality and "integrity" as defined by personality-based tests was examined. A linear composite of the four personality-based integrity tests used in this research (PRB, PDI-EI Performance scale, IPI critical scale, Hogan Reliability scale) was formed. Also, composites for the Big Five dimensions of personality using scale scores from three personality inventories tapping the Big Five (Personal Characteristics Inventory, Hogan Personality Inventory, and Goldberg's Adjective Checklist) were formed. Then, for each one of the Big Five dimensions, the correlation between the composite of

integrity tests and the composite of the corresponding scales of the three personality inventories was computed. The results are presented in Table 14.

The composite for conscientiousness correlates .86 with the composite for personality-based integrity tests (ρ=.86). However, composites of Agreeableness and Emotional Stability are also correlated with the composite of personality-based integrity tests scores (true score correlations of .58 and .51, respectively). The linear composite of extroversion correlates .37 (true score correlation) with the linear composite of personality-based integrity test scores, indicating that there is a somewhat weak positive relationship between extroversion and integrity. Finally, the composite for Openness to Experience correlates .03 with the composite of personality-based integrity test scores, indicating virtually no overlap between the two constructs. It is important to note that this pattern of correlations is identical to that presented for integrity tests, overt and personality-based combined, as well as those presented for overt tests.

In all the analyses of the primary data, for integrity tests, overt tests alone, and personality-based tests alone, conscientiousness emerged as the highest correlate of integrity tests. It was followed by agreeableness and emotional stability, in that order. Extroversion had a weak positive relationship to integrity tests. Finally, the relationship between openness to experience and integrity test scores was negligibly small. This same pattern of correlations for both overt and personality-based tests strengthens the case for these two types of tests tapping into the same overall construct.

Results of the Meta-Analyses

One important question is whether the results presented above from the primary data hold up in other settings. Do overt integrity tests correlate

highly with each other? Do personality-based integrity tests correlate highly with each other? What is the correlation between overt and personality based integrity tests? The aim of these sets of analyses was to answer these questions with precise estimates using meta-analysis.

Correlations between overt and personality based tests (Meta-analytic results)

The intercorrelations between overt and personality-based integrity tests were meta-analytically obtained. The results are presented in Table 15.

Overt integrity tests were all originally designed to predict theft and theft related activities. These tests assess attitudes toward theft and other illegal activities. Some overt integrity tests include admissions items. The underpinnings of all integrity tests are similar; therefore, all of them can be hypothesized to correlate highly with each other. Indeed this hypothesis received some support in the meta-analysis. Across 56 correlation coefficients and 7,424 data points, the true score correlation between various overt tests was .45. The meta-analysis presented here includes three correlations from the primary data (those presented in Table 9). This mean true score correlation is much lower than the true score correlations presented in Table 9 for primary data. One reason for this discrepancy is that the correlations included in the meta-analysis came from a diverse set of overt integrity tests. The three overt tests used in the primary data collection (PSI, Reid Report, and Stanton Survey) are possibly more similar to each other than are most overt integrity tests. All three pose similar questions to job applicants about their attitudes toward workplace theft. There are other overt tests with somewhat different item format and content. For example, the True Test presents integrity-related vignettes to test takers, and requires them to choose the reaction they would have in a real world setting. Other overt tests

include questions about other workplace counterproductive behaviors in addition to theft alone.

Nevertheless, the meta-analytic results indicate that overt tests on the average correlate .45 with each other. To a certain extent, we can conclude that overt tests seem to share a general common core construct.

Personality-based integrity tests were all originally designed to predict broadly defined counterproductive behaviors and not theft. These tests are all derived from various personality inventories. The underpinnings of all personality-based integrity tests are similar. The common thread that spans all personality-based integrity tests seems to be trustworthiness, integrity, and conscientiousness. Therefore, all of them can be hypothesized to correlate highly with each other. To test this hypothesis the true score correlation between personality-based tests was obtained via meta-analysis. There were 37 correlations across 7,062 data points that were used in this meta-analysis. The true score correlation between personality-based integrity tests is .70. Compared to the true score correlation between overt integrity tests, this correlation is fairly high, indicating that personality-based tests share a common construct to a greater degree. This fairly high correlation between various personality-based tests is not all that surprising because most personality-based integrity tests on the market were patterned after the Personnel Reaction Blank (Gough, 1948).

To precisely estimate the relationship between overt and personality-based tests, the correlations between these two types of tests were meta-analytically cumulated. Across 117 correlation coefficients involving 15,978 individuals, the true score correlation between overt and personality-based tests is .39. When I formed a linear composite of overt tests and a linear

composite of personality-based tests, the correlation between these two composites was .64. These results indicate that even though there is a large common core to both overt and personality-based integrity tests, the two types of tests are not construct-wise identical (i.e. perfectly correlated). Ones et al. (in press) found that the validities of overt and personality-based tests are similar in predicting supervisory ratings of overall job performance, coupled with the finding here that there is a large degree of overlap between overt and personality-based tests, the hypothesis that both types of integrity tests derive their validity from the general factor they share is strengthened.

The relationships between integrity tests and the Big Five dimensions of personality (Meta-analytic results)

The next question of importance for the nomological net of integrity tests is establishing what overall construct or constructs integrity tests measure. To systematically examine the dimension of personality integrity tests tap into, the correlations between integrity tests and the Big Five Personality dimensions were meta-analytically examined. Tables 16, 17, and 18 summarize the results for integrity tests (overt and personality-based tests combined), overt tests and personality-based tests.

An inspection of Tables 16, 17, and 18 indicates that magnitudes of the correlations for overt and personality-based tests with the Big Five dimensions of personality are not identical. However, for both the highest three correlations are with conscientiousness, agreeableness, and emotional stability, and in that rank order. Also, for both, the relations with extroversion and openness to experience are negligible. This pattern of results is identical to the one found using data form the primary data. It is interesting to note that the correlations between personality-based tests and

all the Big Five dimensions are slightly larger than those between overt tests and the Big Five dimensions. One possible reason for this is the fact that personality-based integrity tests have their roots in personality inventories. Even though there are size differences in true score correlations of overt and personality-based tests with the Big Five, the differences are small enough that the correlations of integrity tests (overt and personality-based tests combined) with the Big Five are meaningful. Furthermore, since the pattern of correlations with the Big Five are identical regardless of integrity test type, the overall results will be presented and interpreted.

Confirming expectations, integrity tests are correlated with the Conscientiousness dimension of personality. The cumulation of 423 integrity-conscientiousness correlations across 91,360 data points resulted in a true score correlation of .42. However, the true score correlation of integrity tests with Agreeableness was almost as strong (.40). The third highest correlation was found between integrity test scores and Emotional Stability (.33). The correlations with extroversion and openness to experience were low (-.08 and .12, respectively).

The fact that both overt and personality-based tests display the highest correlations with conscientiousness, followed by agreeableness, followed by emotional stability indicates that these three constructs permeate all types of integrity tests in varying degrees. Even though overt tests on the surface tap into "attitudes," it is clear from the results presented in Table 17 that they tap into personality constructs, and with a pattern similar to personality-based integrity tests. It is interesting to note that , previously I indicated that overt and personality-based integrity tests correlate at about .39, this is the exact same correlation found between overt tests and conscientiousness. For

personality-based tests, the correlation with conscientiousness is somewhat higher (.45).

It appears that integrity tests identify those individuals who will be conscientious, responsible and dependable on the job (i. e., conscientious). They also identify individuals who will be cooperative, agreeable and not hostile in work settings. Correlations with emotional stability suggest that people who are high on integrity also tend to be emotionally stable (i. e., not neurotic). There seems to be little association between being intellectual, open to experiences, and being cultured and integrity. Also, evidence suggests that there is a weak negative relationship between integrity and extroversion. That is , extroverts tend to be slightly more dishonest. However, the relationship between extroversion and integrity is different for overt and personality-based integrity tests. There seems to be no relation between integrity and extroversion. Individuals who score high on personality-based integrity tests tend to be more introverted.

Measures of conscientiousness from various personality inventories correlate .47 among themselves (from Table 3); various integrity tests correlate .46 among themselves (the sample size weighted average of all true score correlations presented in Table 15; this value is also the true score correlation obtained, when a psychometric meta-analysis was conducted cumulating all correlations between integrity tests). The meta-analytically obtained true score correlation between integrity tests and measures of conscientiousness is .42. Thus, there appears to be minimal discriminant validity for conscientiousness and integrity.

The correlations between integrity tests and each of the Big Five dimensions can be compared to those found between conscientiousness and

the other Big Five dimensions. If the hypothesis that integrity tests measure a broadly defined construct of conscientiousness is correct, these two sets of correlations will be about the same. Emotional stability scales correlate .26 with conscientiousness scales and .33 with integrity tests. Scales of Extroversion correlate -.08 with conscientiousness scales, but .00 with integrity tests. Openness to experience scales correlate -.06 with scales of conscientiousness and .12 with integrity tests. Finally, measures of agreeableness correlate .27 with measures of conscientiousness, but .40 with measures of integrity. The pattern of correlations between conscientiousness and the other Big Five dimensions is not identical to the pattern of correlations between integrity and the same four dimensions of the Big Five. This indicates that integrity tests are tapping into other constructs than just conscientiousness.

Agreeableness measures correlate .40 with integrity tests. Measures of agreeableness from various personality inventories correlate .53 among themselves. (Recall that various integrity tests correlate .46 among themselves.) There seems to be much evidence for convergent validity between measures of agreeableness and integrity.

Emotional Stability measures correlate .33 with integrity tests. Measures of Emotional Stability from various personality inventories correlate .63 among themselves. (Recall that various integrity tests correlate .46 among themselves.) The discriminant validity of integrity tests and emotional stability scales is discernible, but weak.

Measures of extroversion from various personality inventories correlate .41 among themselves; as was stated earlier, various integrity tests correlate .46 among themselves. The meta-analytically obtained true score

correlation between integrity tests and measures of extroversion is -.08. There is excellent convergent and discriminant validity evidence for measures of extroversion and integrity. Clearly, integrity tests and extroversion scales tap into different constructs.

Openness to Experience measures correlate .12 with integrity tests. Measures of Openness to Experience from various personality inventories correlate .43 among themselves. (Various integrity tests correlate .46 among themselves.) The discriminant validity of integrity tests and Openness to Experience scales is strong. Similar to the results for extroversion, integrity tests and Openness to Experience scales tap into different constructs.

Finally, using the meta-analytic results, the relationships between linear composites of integrity tests and linear composites for each of the Big Five dimensions of personality were investigated. The rationale behind forming the composites was to define "integrity' as what is shared across integrity tests on one hand and to define each dimension of the Big Five as what is shared across corresponding scales of various personality inventories on the other. In other words, in forming composites factors specific to each integrity test and personality scale were treated as error. Composite reliabilities for integrity tests (overt and personality-based), overt tests, personality-based tests and each of the Big Five dimensions was formed using the formula for the reliability of a composite (standardized coefficient alpha, assigning scale or scale specific variance to measurement error). The reliability of the integrity test composite is .98. The reliability of the overt test composite is .98. The reliability of the personality-based test composite is .94. The reliability of the personality-based test composite is .94. The reliabilities of composites of each of the Big Five Dimensions are as follows: .99 for

conscientiousness and agreeableness, .98 for openness to experience, .95 for extroversion, and .99 for emotional stability.

The true score correlations between integrity composites and composites for each of the Big Five dimensions are presented in Table 19.

Again, conscientiousness emerges as having the most overlap with integrity tests (true score correlation of .87 between the composites of integrity and conscientiousness). This confirmed the generalizability of the earlier finding from the primary data that integrity tests primarily tap into conscientiousness. The linear composite of Agreeableness is highly correlated with the linear composite of integrity tests scores (true score correlation of .78). The linear composite of Emotional Stability is also correlated with the composite of integrity tests scores (ρ = .59), but to a lesser degree than conscientiousness and integrity, and agreeableness and integrity. The linear composite of extroversion correlates -.18 (true score correlation) with integrity test scores, indicating that there is a weak negative relationship between extroversion and integrity. Finally, the composite of Openness to Experience scales correlates .26 with the composite of integrity test scores, indicating a small overlap between the two constructs.

An inspection of Table 19 indicates that magnitudes of the correlations for overt and personality-based test composites with the composites of the Big Five dimensions of personality are not identical. However, for both types of tests the highest three correlations are with composites of conscientiousness, agreeableness, and emotional stability, and in that rank order. Also, for both, the relations with extroversion and openness to experience are negligible. Based on these patterns of correlations, it appears that the overall construct tapped into by both overt and personality-based tests is identical.

For overt integrity tests, the results from composite correlations indicate that the common factor shared across overt tests correlates .81 with the common construct across conscientiousness scales of major personality inventories. The corresponding values for agreeableness and emotional stability are .66 and .50, respectively. Overt integrity tests and conscientiousness scales seem to share a common core construct. However, agreeableness and emotional stability scales also seem to share a common theme with overt tests, although not to the same extent as conscientiousness scales.

For personality-based integrity tests, the results from composite correlations indicate that the common factor shared across personality-based integrity tests correlates .75 with the common construct across conscientiousness scales of major personality inventories. The corresponding composite correlations for agreeableness and emotional stability are .69 and .54, respectively. Personality-based integrity tests and conscientiousness scales seem to share a common core construct. However, in this case agreeableness and emotional stability scales also seem to share a common theme with personality-based tests.

Next, confirmatory factor analysis was used to confirm that overt integrity tests, personality-based integrity tests, and scales of conscientiousness from mainstream personality inventories tap into the same general overall construct. The meta-analyzed correlations were used in this confirmatory factor analysis. The correlation between overt and personality-based tests came from Table 15 (ρ = .39). The correlation between overt integrity tests and conscientiousness scales came from Table 17 (ρ = .39). The correlation between personality-based tests and conscientiousness scales came from Table

18 (ρ = .45). The factor loadings of the overt tests, personality-based integrity tests, and conscientiousness scales on this general factor are .87, 80, and .98 , respectively. These high factor loadings coupled with the fact that the residual intercorrelation matrix for this confirmatory factor analysis consisted entirely of zeros, lead to the conclusion that there is a strong general factor that explains the intercorrelations between overt tests, personality-based tests, and conscientiousness.

Given that both overt and personality-based integrity tests seem to share a common theme that is strongly related to conscientiousness and to a lesser degree to agreeableness, it may be useful to examine the correlation between the sum of scores on integrity tests and the composite of scores on conscientiousness plus agreeableness. This correlation is .95. This extremely high correlation may be taken as proof that the construct shared across integrity tests is virtually identical to the construct shared across combinations of conscientiousness and agreeableness scales of personality inventories.

A regression equation can be used to predict integrity from the Big Five dimensions of personality. The initial hypothesis in this dissertation was that in this regression equation the largest weight would be expected for conscientiousness; the second largest weight for agreeableness and the third largest weight for extroversion. Based on the primary study results and the meta-analytic results, this initial hypothesis must be modified. Based on zero-order correlations, in the regression equation to predict integrity using the Big Five dimensions of personality, the largest weight is to be expected for conscientiousness; the second largest weight for agreeableness and the third largest weight for emotional stability. This modified combination gives the

weighted combination that best predicts and describes integrity. The regression equation for predicting integrity from the Big Five (C= Conscientiousness; A= Agreeableness; ES = Emotional Stability; EX = Extroversion; OE = Openness to Experience) is as follows:

$$\text{Integrity} = .30\,C + .29\,A + .20\,ES - .19\,EX + .11\,OE$$

The multiple correlation for the above equation is .58. In predicting integrity using dimensions from the Big Five, conscientiousness is the best predictor. However, one question of interest is whether improvement in prediction of integrity can be obtained by adding agreeableness, emotional stability and other Big Five dimensions to conscientiousness. This question was answered by computing multiple correlations. The multiple correlation between integrity and a linear combination of measures of conscientiousness and agreeableness is .52. When emotional stability is added to the combination of conscientiousness and agreeableness, the multiple correlation increases to .54. However, the addition of extroversion and openness to experience to the linear combination of the other Big Five dimensions increases the multiple correlation to .58.

This portion of the dissertation was aimed to answer four questions: (1) Do overt integrity tests correlate highly with each other? (2) Do personality-based integrity tests correlate highly with each other? (3) Do both overt and personality-based integrity tests measure the same underlying construct? (4) Do integrity tests do correlate with measures of agreeableness, conscientiousness, extroversion, openness to experience, and emotional stability?

Based on the comprehensive analyses presented the answers to these questions are as follows:

Overt integrity tests correlate .45 with each other. There is evidence of a general construct that underlies all overt integrity tests. Confirmatory factor analysis indicated that overt integrity tests intercorrelations arise due to the presence of a single factor.

Personality-based integrity tests correlate .70 with each other. Theoretically, the underpinnings of all personality-based integrity tests are similar. Empirically, the presence of a general single construct that spans personality-based integrity tests is confirmed.

Overt and personality-based tests correlate .39. The linear composites of overt integrity tests and personality-based integrity tests correlate .64. This result indicates that the construct shared by overt tests correlates highly with the construct shared by personality-based integrity tests. The second order factor analytic results confirm this conclusion.

True score correlations between integrity tests and the Big Five dimensions of personality indicate that integrity tests have strong empirical foundations in the personality domain. In all the analyses, Conscientiousness was found to be the most consistent correlate of both types of integrity tests. Further, analyses confirmed that the construct of integrity largely overlaps with the construct of conscientiousness. However, agreeableness and emotional stability are also related to integrity. An accurate identification of individuals who would score high on integrity tests is possible by examining scores on three of the Big Five Scales: conscientiousness, agreeableness, and emotional stability, in that order.

Implications of the meta-analytic results for the criterion related validities of integrity tests

One crucial question that still has not been addressed is: Do integrity tests derive their criterion-related validity from a broadly defined construct of Conscientiousness?

Criterion-related validities of integrity tests indicate that job applicants who score poorly on integrity tests also turn out to be poor employees (Ones et al., in press). Why do integrity tests predict job performance? In particular, the question here is whether integrity test validities for predicting supervisory ratings of job performance can be explained by a broadly defined construct of conscientiousness. It is impossible to conclusively determine the answer to this question because both conscientiousness scales and integrity tests might be "narrowly" defined measures of conscientiousness. Barrick and Mount (1991) found that the personality dimension with the highest validity for job performance is conscientiousness and this validity is generalizable across situations.

If the hypothesis that integrity tests derive their validity from a broadly defined construct of conscientiousness is true, then controlling for this broad construct of conscientiousness should reduce the validity of integrity tests obtained by Ones et al. (in press) to negligible. However, as it was indicated earlier in this dissertation, conscientiousness is defined very narrowly in traditional scales from personality inventories, in which case partialing conscientiousness from the integrity-job performance relationship may have minimal impact on the size of the validity of integrity tests for predicting overall job performance.

To test the hypothesis that the conscientiousness dimension of the Big Five fully or partially explains the correlation between integrity test scores and supervisory ratings of overall job performance, the true score correlations between integrity, conscientiousness, and job performance are necessary.

The complete matrix of zero order correlations and the first order partial correlations between the Big Five dimensions of personality, integrity and job performance are provided in Tables 20 and 21. Table 20 was obtained by bringing together results from Barrick and Mount (1991), Ones et al. (in press), and this dissertation. It is important to note that in the correlations presented in Table 20 were obtained by using narrow measures of both conscientiousness (i. e., as measured by traditional personality inventory scales) and integrity (i. e., as measured by integrity tests).

The true score correlation between integrity and job performance was obtained from Ones et al. (in press) (true score correlation of .46). The true score correlation between narrow conscientiousness and job performance was obtained from Barrick and Mount (1991) (true score correlation of .23). Finally, the true score correlation between integrity and conscientiousness was obtained from this dissertation (true score correlation of .42). Based on these three true score correlations, the partial correlation between integrity and job performance, after conscientiousness has been partialed out, was computed. In other words, the relationship between integrity and job performance controlling for conscientiousness was obtained. This partial correlation is .41. Partialing out conscientiousness from the integrity-job performance relationship, reduces the true score correlation from .46 to .41. This is a reduction of 12% in the true score correlation. This results suggests

that conscientiousness as measured by narrow personality inventory scales only partially explains the validity of integrity tests for job performance.

In this and the next paragraph I briefly mention some other interesting findings that emerged from the first order true score partial correlations. The correlation between agreeableness and job performance that Barrick and Mount found (.06) is reduced down to .00 when conscientiousness is partialed out. The correlation between Emotional Stability and job performance that Barrick and Mount found (.07) diminishes to .01 when conscientiousness is partialed out.

The first order partial correlations also indicate that the true score correlation between integrity and extroversion that was found in this dissertation is purely due to conscientiousness. That is, the small correlation for extroversion and integrity that was meta-analytically obtained in this dissertation is reduced to .00 when conscientiousness is removed from the relationship. However, when conscientiousness is partialed out from the Integrity-Agreeableness relationship and the Integrity-Emotional Stability relationship, the partial correlations are not zero. The partial correlation between integrity and agreeableness is .33 (when conscientiousness is partialed out). The partial correlation between integrity and emotional stability is .25 (when conscientiousness is partialed out).

As was stated earlier, a construct of narrowly defined conscientiousness only partially explains the relationship between integrity test scores and job performance. An obvious alternate question to ask is whether integrity tests explain the conscientiousness-job performance relationship found by Barrick and Mount (1991). If conscientiousness as measured by integrity tests is the construct producing the conscientiousness-job performance relationship,

once integrity test scores are partialed from the relationship, the correlation should be close to zero. To determine if this is true or not, the partial correlation between conscientiousness and job performance , partialing out integrity test scores, was computed. This true score partial correlation is .05. Conscientiousness as measured by integrity tests seems to be the construct that produces the conscientiousness-job performance relationship that Barrick and Mount (1991) found. That is, measures of conscientiousness from mainstream personality inventories tap into a broader construct measured by integrity tests , lending them the criterion related validity for job performance. This finding is in line with the thinking that integrity tests tap into a construct of conscientiousness much broader than what traditional personality inventories tap into. Stated alternately, most personality inventories may be deficient in the their measurement of the all the facets of the total domain of conscientiousness. Given that integrity tests predict job performance with a true score validity of .46 and conscientiousness scales predict job performance with a true score validity of .23, it would be prudent to expand the item coverage of conscientiousness scales to include the type of items that make up most integrity tests.

The results presented above indicate that adding conscientiousness as measured by personality inventories to a battery that already contains integrity tests does not improve prediction. This conclusion was confirmed by calculating the multiple correlation of integrity tests and conscientiousness scales for job performance. This multiple correlation is .41, representing no improvement over the operational validity of integrity tests with the same criterion.

Implications of meta-analytic results for combining integrity tests and Big Five personality inventories for personnel selection

What levels of operational validity can be obtained by combining measures of all the Big Five dimensions of personality (as measured by traditional personality instruments) and integrity tests in personnel selection?

This practical question was answered using the correlations among integrity, the Big Five dimensions of personality, and job performance. The results of this analysis will differ depending on whether the correlations are corrected for measurement error (see Schmidt, Hunter, & Caplan, 1981, for details). In personnel selection, operational validity is appropriately calculated without correction for measurement error in the predictors. The meta-analytically obtained correlations form Barrick and Mount (1991), Ones et al. (in press), and this dissertation were used to compute the multiple correlation for job performance when the predictors are the integrity and the Big Five dimensions of personality. This operational multiple correlation was .46, representing a 12% increase in the predictive validity of integrity tests alone. The regression equation for predicting job performance using integrity (I) and the Big Five (C= Conscientiousness; A= Agreeableness; ES = Emotional Stability; EX = Extroversion; OE = Openness to Experience) is presented below:

$$\text{Job Performance} = .45\,I + .09\,C - .11\,A - .06\,ES + .15\,EX - .07\,OE$$

The interesting thing in the above equation is that two of the Big Five dimensions emerge as suppressors: Agreeableness and Emotional Stability. Agreeableness correlates moderately with the other predictors (see Tables 3 and 12 of this dissertation); it has a very low correlation with the criterion of job performance (.05 from Barrick and Mount, 1991); and has a negative

regression weight. Emotional Stability correlates moderately with the other predictors in the equation (see Tables 3 and 12 of this dissertation); it has a fairly low correlation with the criterion of job performance (.06 from Barrick and Mount, 1991); and has a negative regression weight. Hence, it appears that the increase of 12% in predictive validity, when all the Big Five and integrity are used to predict job performance compared to when only integrity and conscientiousness are used, is at least partially due to the suppression effects from agreeableness and emotional stability.

What levels of operational validity can be obtained by combining of all the Big Five scales from mainstream personality inventories in personnel selection?

The meta-analytically obtained correlations form Barrick and Mount (1991), and this dissertation were used to compute the multiple correlation for job performance when the predictors are the Big Five dimensions of personality. This operational multiple correlation was .46, representing a 12% increase in the predictive validity of integrity tests alone. The regression equation for predicting job performance using the Big Five (C= Conscientiousness; A= Agreeableness; ES = Emotional Stability; EX = Extroversion; OE = Openness to Experience) is presented below:

$$\text{Job Performance} = .20\,C - .00\,A + .01\,ES + .09\,EX - .03\,OE$$

The operational multiple correlation for the above equation is .22. In other words, combining Big Five scales from mainstream personality inventories for predicting job performance produces a lower multiple correlation (.22) than integrity tests combined with similar Big Five scales (.46). By adding an integrity test to a personnel selection battery that already contains Big Five scales increases predictive validity by 109%.

As was indicated earlier, in addition to conscientiousness, agreeableness and emotional stability are also highly related to integrity. With integrity tests, the increased breadth of predictor construct coverage translates itself into better prediction in personnel selection.

Integrity-Age, Integrity-Gender, and Integrity-Race Relationships and Implications for Adverse Impact

The Relationship between Integrity Test Scores and Age

Do integrity tests correlate with age? To answer this question, correlations between integrity test scores and age were obtained and meta-analytically cumulated (correlations of integrity test scores and age obtained using same age individuals or individuals within the same age group, for example only high school students, were excluded). All published and unpublished studies reporting correlations between age and integrity tests were obtained and coded. Where correlations were not reported other statistics such as Cohen's effect size, t-and F values were converted to correlations. In cases where either the integrity test scores or age were dichotomized, corrections for dichotomization were applied. The appropriate correlation to look at in this instance is the continuously expressed variable of integrity and the continuously expressed variable of age.

Psychometric meta-analysis was used to cumulate the correlations. The results are presented in Table 22.

The results indicate that older individuals tend to score higher than younger individuals on integrity tests. The observed correlation between the continuous variable of integrity and age is .16. This correlation is .17 for

overt tests and .15 for personality-based tests. That is, the integrity-age relationship is similar for both overt and personality-based integrity tests.

These results, that younger individuals score somewhat lower on integrity tests, may be possibly due to experimentation and youthful indiscretions that characterize late teenage years and early adulthood. Overall, from a practical point of view using integrity tests in personnel selection would result in increased hiring rates for older individuals.

The Integrity - Gender Relationship and Implications for Adverse Impact

Review of studies of adverse impact of integrity tests suggested that scores on integrity tests are somewhat correlated with gender (Goldberg et al., 1991). Females on average score slightly higher than males on integrity tests. To precisely estimate the exact magnitude of the relationship between gender and integrity test scores a meta-analysis was conducted. All published and unpublished studies reporting correlations between gender and integrity tests were obtained and coded. Where correlations were not reported other statistics such as Cohen's effect size, t-and F values were converted to correlations. In some instances, based on the integrity test means and standard deviations reported for males and females, I computed d values, which were converted to correlations. Psychometric meta-analysis was used to cumulate these correlations. After I obtained the mean correlation between integrity test scores and gender, I expressed it in terms of mean differences between males and females in standard deviation units (d values). The results are presented in Table 23.

Females tend to score .65 standard deviation units higher than males on integrity tests. This value is .82 standard deviation units for overt tests and .39 standard deviation units for personality-based tests. These results

indicate that using integrity tests in personnel selection would result in increased hiring rates for females.

The Integrity-Race Relationships and Implications for Adverse Impact

Adverse impact studies of integrity tests suggest that using integrity tests does not result in any adverse impact for any minority groups. In other words, on average, minorities do not score lower than whites on integrity tests. To precisely estimate the relationship between race and integrity, a series of meta-analyses were conducted. All published and unpublished studies reporting correlations between race (white vs. black, or white vs. Hispanics, or white vs. Asians) and integrity were obtained and coded. Where correlations were not reported other statistics such as Cohen's effect size, t-and F values were converted to correlations. In very few of the studies the percentages of individuals in various racial groups comprising the sample were reported. Therefore, the correlations coded for this set of meta-analyses were not corrected for splits on the race variable. I have recently contacted all the test publishers that provided data for the race-integrity investigations, inquiring about the proportions of different racial groups in the data that they reported. Because none of the race-integrity correlations were corrected for the splits in the race variable, these results may be distorted and should be considered preliminary only. When the proportion of minorities in each sample contributing data to the meta-analysis is known, then all integrity-race correlations need to be converted to a 50-50 split. This is necessary for the d values to be correct. (In examining the integrity test-race relationships, the main focus is on d values.) In any event, since the integrity-race correlations are not corrected for the race composition splits, the results reported below are not final (in fact may be badly distorted) and should be viewed with caution.

Our database for the race-integrity meta-analyses had 96 correlations, across 712,214 individuals. Psychometric meta-analysis was used to cumulate these correlations. After the mean correlations between race and integrity were obtained, these were expressed in terms of mean differences between whites and blacks in standard deviation units (a d value). The results are presented in Table 24.

The results are quite different for overt and personality-based tests. On overt integrity tests: Blacks tend to score .61 standard deviation units higher than whites. Hispanics seem to score .43 standard deviation units higher than whites. Asians tend to score .22 standard deviation units higher than whites. One important fact to note is that in the latter two of these three analyses for overt tests, the 90% credibility interval included zero (when Hispanics and Asians constitute the minority group). This was not true for the case where blacks constituted the sample. In any event, the magnitudes of the differences between races on integrity that were calculated here may not be applicable to applicant populations in general.

In the construction phase of some overt tests items that favored whites were intentionally removed. This resulted in tests that were made up of items either showing no race differences or favoring minority groups. This was true for tests constructed in 1970's and for tests that were revised during the same time period. The rationale behind such an item selection strategy was to ensure that the tests did not create adverse impact for minorities. This type of item selection strategy was carried out only for some overt tests and not others. Some integrity tests (such as the True Test developed separate scoring keys for whites and minorities). It appears that no such item selection

strategies to eliminate items favoring whites was employed in the construction of personality-based integrity tests.

The results for personality-based integrity tests are quite different than those found for overt tests. Blacks tend to score .16 standard deviation units lower than whites (observed d value of -.16). Hispanics tend to score -.06 standard deviation units lower than whites (observed d value of -.06). Asians tend to score .06 standard deviation units higher than whites. One important fact to note is that in all the analyses for personality-based integrity and race relationships, the 90% credibility interval included zero. Hence, it seems reasonable to conclude that racial differences are negligible on personality based integrity tests.

However, until the proportions of minorities in each of the samples that contributed to these meta-analyses are available, the results for race differences on overt and personality-based tests should not be compared. In any event, one finding that emerges from the race-integrity relationship meta-analyses is that minority groups are not adversely affected by the use of either overt or personality-based integrity use in personnel selection. These results also indicate that using integrity tests in personnel selection could potentially result in increased hiring rates for minorities.

From the ability testing and personnel selection literatures, we know that blacks average about one standard deviation below whites on tests of general mental ability. This difference of one standard deviation between blacks and whites on general mental ability tests can also be expressed as a correlation between ability and race, r = .45. This correlation is obtained using the standard formula for converting effect sizes (1 sd here) to correlations (.45). The evidence indicates that ability and integrity scores are correlated .02

(see the next section of this dissertation for this finding). Race (blacks vs. whites) and overt integrity test scores are correlated -.29 (from Table 24, the negative sign of the correlation is to indicate that the blacks score higher than whites on overt integrity tests). The correlation between race and an optimally weighted (using regression weights) composite of ability and overt integrity tests can then be computed. This correlation is .11 which converts to an effect size of .22. In other words, the mean difference between blacks and whites on an optimally weighted composite of ability and overt integrity test scores and race is .22 standard deviations. Thus, when ability and overt integrity test scores are optimally weighted, the black-white difference in standard deviation units is much reduced in comparison to ability tests used alone. This reduction can be expected to translate into a greater reduction in adverse impact (reduction in adverse impact depends on the selection ratio as well). By way of example, suppose all those above the white mean were selected (i.e., a selection ratio of .50 for whites). In this case, assuming normality of the scores, the percentage of blacks selected based solely on ability, without an integrity test, would be 15.9%. However, if an overt integrity and an ability test were used together, with scores optimally weighted in a regression equation, the percentage of minorities selected would increase to 41.3%. This is an increase in the hiring rate of minorities of 160%. This increase is effected with no reduction in mean job performance of selected applicants.

Similar calculations were made for personality-based integrity tests. Overt and personality-based integrity tests show differential relationships with race. Race (blacks vs. whites) and personality-based integrity test scores are correlated .08 (from Table 24, the positive sign of the correlation is to

indicate that the blacks score lower than whites on personality-based integrity tests). The correlation between race and an optimally weighted (using regression weights) composite of ability and personality-based integrity tests was computed. This correlation is .37 which converts to an effect size of .80. In other words, the mean difference between blacks and whites on an optimally weighted composite of ability and personality-based integrity test scores and race is .80 standard deviations. Thus, when ability and personality-based integrity test scores are optimally weighted, the black-white difference in standard deviation units is somewhat reduced in comparison to ability tests used alone. This reduction can be expected to translate into a greater reduction in adverse impact (reduction in adverse impact depends on the selection ratio as well). By way of example, suppose all those above the white mean were selected (i.e., a selection ratio of .50 for whites). In this case, assuming normality of the scores, the percentage of blacks selected based solely on ability, without a personality-based integrity test, would be 15.9%. However, if a personality-based integrity and an ability test were used together, with scores optimally weighted in a regression equation, the percentage of minorities selected would increase to 21.2%. This is an increase in the hiring rate of minorities of 33.6%. Even though race and personality-based integrity test scores correlate marginally (.08), important reductions in adverse impact can be realized by combining ability test scores with personality-based integrity test scores in personnel selection.

Overall Summary of Findings Pertaining to Demographic Variables

The results presented in this section of the dissertation answered four questions: Do integrity tests correlate with age? Do integrity tests correlate

with gender? Do integrity tests correlate with race? What are the adverse impact implications of using integrity tests in personnel selection?

Integrity test scores are somewhat correlated with age (true score correlation of .21). Overt and personality-based tests do not differentially correlate with age. Meta-analytic results presented indicate that scores on integrity tests are correlated with gender. Females on average score .82 standard deviations higher than males on overt integrity tests. Females on average score .39 standard deviations higher than males on personality-based integrity tests. One question is whether there is an interaction between gender and age. That is, there is the possibility that the differences in integrity between males and females may diminish by age. Alteranately stated, the correlation between sex and integrity test scores could be lower at older ages. Unfortunately, testing this interaction would require the correlations between integrity test scores and gender for various age groups. No such correlations have been reported in the published or the unpublished literature on integrity tests (although one integrity test publisher has agreed to share a primary database that can address this question for future research).

Hunter and Hunter (1984) indicate that it may be possible to identify other predictors that will add to the validity of general mental ability and at the same time reduce adverse impact. Integrity test publishers have devoted considerable research to examining the question of adverse impact. The results from the meta-analyses of race-integrity relationships indicate that minority groups are not adversely affected by either overt integrity tests or personality oriented measures. Employers can use integrity tests in conjunction with measures of ability and can much reduce the adverse impact of tests of ability alone.

Integrity-Ability Relationship and Implications for Incremental Validity

Do integrity tests correlate with tests of ability? What is the precise incremental validity of integrity tests over measures of ability for predicting overall job performance? Available studies suggest that the correlations between integrity measures and ability measures are low and negligible. For example, Jones and Terris (1983) found that the correlations between an overt integrity test and a measure of general mental ability were -.02 for the theft admissions subscore and -.03 for the theft attitudes subscore; Gough (1972) reported that a vocabulary test correlated -.05 with the Personnel Reaction Blank; Werner, Jones, and Steffy (1989) reported that integrity test scores are unrelated to educational level (an arguable proxy for ability); and Hogan and Hogan (1989) reported correlations of .07 and -.09 between the Hogan Reliability Scale (a personality-based integrity test) and the quantitative and verbal portions of the Armed Services Vocational Aptitude Battery (ASVAB), respectively. To precisely estimate the relationship between integrity and ability test scores, I conducted a meta-analysis. A database of all available integrity-ability correlations was formed.

Overall, there were 106 correlations. Twenty nine of these correlations were for overt tests and 77 were for personality-based tests. The majority of these correlations were obtained from various test publishers. In fact, one test publisher responded to my request of integrity-ability correlations by sending raw data from eight separate applicant samples, which I used to compute the necessary correlations. The list of ability tests contributing data to this meta-analysis is presented in Table 25.

In this database of integrity-general mental ability, correlations from the primary study described in the methods section were also included. These

correlations were between the Wonderlic Personnel test and the following integrity tests: Personnel Selection Inventory of London House, Stanton Survey, Reid Report, Personnel Reaction Blank, Inwald Personality Inventory, Hogan Reliability scale, and the PDI- Employment Inventory.

The Hunter-Schmidt (1990, p. 185) psychometric meta-analytic procedures were used to obtain the true score correlation between integrity and ability. The statistical artifacts biasing the observed correlations were corrected for. The reliability distributions used in the corrections are provided in Table 26. The results of the meta-analysis are presented in Table 27.

The true score correlation between integrity tests and tests of ability is .02 (N = 23,306, K = 106). This is a negligible relationship. Furthermore, overt and personality-based tests do not seem to be differentially correlated with integrity tests (true score correlations of -.01 and .03, respectively). General mental ability does not correlate with integrity, indicating no overlap between the construct domains covered by the two types of tests. These results also point out the fact that new information can be gained by using integrity tests in the selection process where applicants might already be tested for aptitude.

That ability and integrity are uncorrelated is important because ability tests have the highest operational validities of all personnel selection device. A key unanswered question is the size of the increment in validity from adding integrity tests to general mental ability tests in predicting overall job performance in personnel selection. Now that we know that integrity and ability correlate only .02, the expected incremental validity of integrity tests in predicting supervisory ratings of overall job performance can be calculated. Table 28 presents the predicted incremental validity of integrity tests in

predicting supervisory ratings of overall job performance for each of the five job complexity levels used by Hunter (1980).

In Table 28, the second column of multiple correlations shows the combined validity of integrity and general mental ability test scores. For example, for medium complexity jobs (complexity level 3), the multiple correlation is .65. This is an increase in validity of 27% compared to ability alone, and an increase in validity of 59% compared to integrity alone. The third column of multiple correlations in Table 24 reports the combined validity of general mental ability, psychomotor ability, and integrity. The correlations between general mental ability and psychomotor ability necessary to calculate the multiple correlations were obtained from Hunter (1980); they are approximately .30 in each of the job complexity levels. The multiple correlation for predicting overall job performance is .63 for the lowest complexity jobs (level 5), .66 for medium complexity jobs (level 3) and .70 for highest complexity jobs (level 1). These results indicate that using integrity tests in conjunction with measures of ability can lead to substantial incremental validity for all job complexity levels.

CHAPTER V
DISCUSSION

This study expanded the nomological net for integrity tests. The APA report on integrity tests stated "We strongly recommend that investigators devote increased research attention to construct-oriented issues" (Goldberg et al., 1991, p.18). This dissertation presented extensive evidence related to convergent and discriminant validity of integrity tests in general.

The results from both the primary data analyses, and the meta-analyses display significant convergent validity for all integrity tests. Average true score intercorrelations among overt tests is .45. The average true score intercorrelation among personality-measures is higher at .70. A linear composite of overt tests and a linear composite of personality-based tests correlate .64. The presence of a general factor across integrity tests was confirmed. Even though, there has been a general tendency in I/O psychology to treat integrity tests as separate from personality inventories, the results here clearly point out that integrity tests are best regarded as an extension of personality testing. This is true even for overt tests.

As Goldberg at al. (1991) pointed out some integrity test publishers acknowledge that their instrument measures integrity, whereas others deny this. It appears that whatever the claims of integrity test publishers, integrity tests investigated in this research tap into a shared common construct.

Our results confirm Hogan and Hogan's (1989) finding that the best indicator of integrity is conscientiousness. Conscientiousness has the most overlap with integrity tests (true score correlation of .87 between the composites of integrity and conscientiousness). However, a much more accurate identification of counterproductive individuals is possible by considering scores on agreeableness and emotional stability. Integrity tests identify individuals who are dependable, trustworthy, cooperative and stable. Low scorers on integrity tests seem to be irresponsible, impulsive, hostile, and not well adjusted.

Taken together, the results of this research also seem to indicate that the construct that integrity tests measure cannot be conceptualized in terms of a single dimension from the Big Five framework. In fact, integrity test scores seem to tap into the constructs measured by conscientiousness, agreeableness, and emotional stability in that order. The recent Ones et al. (in press) meta-analytic study has shown that integrity tests are valid predictors of job performance in a variety of settings (operational validity of .41). On the other hand, Barrick and Mount (1991) found that the dimension of Big Five that has generalizable validity across settings is conscientiousness , and it has a true score correlation of .23 with job performance. However, conscientiousness scales that Barrick and Mount (1991) meta-analyzed are poor (narrow and therefore construct limited) measures of the broad construct of conscientiousness. It seems reasonable that the higher predictive validity of integrity tests arise from the fact that they tap constructs broader than just narrow conscientiousness as measured by mainstream personality inventories. Furthermore, in addition to conscientiousness, agreeableness and emotional stability are also highly related to integrity. With integrity

tests, the increased breadth of predictor construct coverage translates itself into better prediction in personnel selection. When integrity tests are used in a selection battery along with personality inventories tapping into the Big Five, Agreeableness and Emotional Stability act as suppressor variables. The multiple correlation for predicting job performance using Big Five scales and integrity tests is .46 (12% higher than the operational validity of integrity tests used alone).

Integrity test scores seem to be correlated with age (a corrected correlation of .21). This indicates that younger individuals have the potential to be more counterproductive employees, possibly due to youthful indiscretions and experimentation. Using integrity tests in personnel selection may result in an older group of applicants being selected. Integrity and gender are also correlated. On average, females score .65 standard deviations above males. This would translate itself into increased selection rates for females. One interesting hypothesis is age and gender interaction with integrity test scores. That is as individuals grow older, the male and female differences on integrity tests may be reduced. However, the testing of this hypothesis will have to wait for further data to be available.

Overt integrity test scores may be correlated with race. In fact, minorities seem to score higher than whites. However, the same results do not hold up for personality-based integrity tests. One problem in these analyses may be the proportions of minorities represented in each of the samples contributing data to the race-integrity meta-analyses. In any event, it is clear that using integrity tests in personnel selection, either overt or personality-based ones, will not cause adverse impact.

The results presented in this dissertation confirm that integrity test scores and ability test scores are uncorrelated. Using integrity tests in conjunction with measures of ability can lead to substantial incremental validity for all job complexity levels. In other words, in personnel selection situations, using ability and integrity tests together will provide employers with increased information to base hiring decisions on. The predictive validity of combining integrity and ability tests is high, pointing to substantial utility gains when these two types of tests are used together.

Even though the use of integrity tests alone can be expected to produce increased hiring of minorities, it can be expected to result in a loss in utility of at least 37% in comparison to use of ability and integrity tests in combination. Using a composite of ability and integrity tests in selection can be expected to result in improved utility of at least 58% compared to integrity alone. Hence, the implication of this research is that employers seeking to maximize work force output should use both integrity tests and measures of general mental ability in making hiring decisions. This combination has the potential for simultaneously reducing adverse impact and enhancing validity and utility, in comparison to selecting on ability alone. In addition to these increases, use of integrity tests can be expected to result in utility gains from reductions in counterproductive behaviors.

The next question that needs to be answered is : What levels of operational validity can be obtained by combining measures of ability, Big Five dimensions of personality (as measured by traditional personality instruments) and integrity (as measured by integrity tests) for personnel selection?

This practical question can be answered if the correlations among measures of cognitive ability, integrity , the Big Five (as measured by personality inventories) and job performance are known. The results of this dissertation provide the meta-analytically obtained correlations among the Big Five measures and integrity tests. Other meta-analyses provide the validities of ability tests, integrity tests, and Big Five scales for predicting job performance. The only unavailable set of relationships is those between the Big Five dimensions of personality and ability. There have been a number of studies that report ability-personality relationships. It appears that Openness to Experience is correlated with general mental ability (See Hogan and Hogan, 1993 for some moderate correlations). Furthermore, other personality dimensions appear to be differentially related to crystallized and fluid intelligence (Goff and Ackerman, 1992). However, conscientiousness does not seem to be one of them. We now have research underway meta-analytically investigating the relationships between the Big Five dimensions of personality and ability.

Critics of integrity tests have suggested that honesty tests do not in fact measure honesty. For prediction purposes, it does not matter what "integrity tests" measure. It has been shown that they have substantial predictive validity (Ones et al., in press). However, construct validity is as important as (maybe more important than) criterion-related validity. Construct validity is important from a theoretical and a practical perspective.

What is so special about integrity tests? Why should they be used in personnel selection? Ones et al. (in press) offered the most important answer to these questions: Because they work. What this research has done is prove that the fact that they work is doubly important because we have not

reinvented the wheel. Integrity tests tap into a construct not tapped into by the best predictor of job performance, ability.

Despite massive evidence for the predictive validity (Ones et al., in press) and for construct related validity of integrity tests, many still have emotional reactions to the use of integrity tests in personnel selection. Critics have expressed concern that if integrity tests are used in personnel selection a large portion of job applicants may be denied jobs. All selection instruments are intended to ensure that the most promising people from the applicant pool are hired. Hence, depending on the selection ratio, all selection instruments have the potential for excluding a large fraction of job applicants. That is the point in selection. The use of any valid selection predictor will result in a lower false positive (and false negative) rate than its non use. Integrity tests have higher validity than many other available selection instruments and (more importantly) can add incremental validity over and above other procedures.

It has even been suggested that screening by integrity tests would mean that millions of people would be excluded from the work force. This is a straw man argument. If integrity tests were widely used in the US economy, they would help allocate more honest people to jobs requiring high levels of security and less honest people to jobs where security is not as crucial (or is counterbalanced by close supervision). Therefore, using integrity tests would not cause any unemployment, but the efficient allocation of human resources to jobs. The number of jobs that must be filled would remain unchanged, and so would be the number of people obtaining jobs. The key point that is missing from this argument is that in an economy with low unemployment it is not possible for most people not to get a job.

Is it a concern that some individuals may be false-positives on integrity tests? Any selection device, short of having perfect validity, will result in some false positives and false negatives. False positives occur on ability tests as well. It has been claimed that false-positives on some aptitude test might at least get a lower-level job on which they can demonstrate their ability and win promotion but there is less chance for appeal form a false verdict on an honesty test. However, it is not unreasonable to believe that false-positives on an integrity test can get lower level jobs where they can demonstrate their honesty and get promoted. Furthermore, integrity tests are almost never used alone but in conjunction with other selection devices such as ability tests and application blanks.

There have been calls for closer monitoring of integrity tests. In arguments against integrity tests, there is concern that the innocent should not be presumed guilty until proven so. Yet, there is sometimes the presumption of worthlessness for integrity tests despite the massive evidence presented to the contrary both in this dissertation and Ones et al. (in press). Does failing an integrity test have more severe implications than failing an aptitude test? Failing a general mental ability and being told that one has low intelligence is more damaging to the ego than to be told that one has undesirable attitudes toward theft of counterproductive job behaviors. One's ability scores are probably even more stable from test to test, employer to employer, etc. Furthermore, regardless of the selection device used by a company, applicants are not usually given the precise reason why they are not hired. Further, it is unreasonable to suspect that any company would share honesty (or ability) test score information with any outsiders. Finally, there are requirements of confidentiality for all psychological test scores (including

integrity tests). Integrity tests should be held exactly to the same standards as all other psychological tests, and failing an integrity test is no worse than failing an aptitude test.

As the APA Task force report on integrity tests also points out, integrity tests should be evaluated on the same basis and using the same criteria as other psychological tests (Goldberg et al., 1991). There seems to be an application of different standards to integrity tests and ability tests and more mainstream personality tests. Most of the same arguments against integrity tests can be made against any psychological test used in personnel selection. In evaluating the merits of various test used in employment settings, the most important factors involve reliability and validity.

Integrity tests have fairly high validities when compared to the other predictors used in personnel selection. Integrity tests have a mean predictive validity of .41 for supervisory ratings of job performance. Mean operational validities of other widely used personnel selection devices for the same criterion are as follows: Ability/Aptitude ρ = .53 (Hunter, 1980), job knowledge ρ = .48 (Schmidt, Hunter, Outerbridge,1986), biographical data ρ = .39 (Rothstein, Schmidt, Erwin, Owens, Sparks, 1990), assessment center ratings ρ = .37 (Gaugler, Rosenthal, Thornton, & Bentsen, 1987), unstructured interviews ρ = .40 (McDaniel, Whetzel, Schmidt, Maurer, & Hunter, in press), reference checks ρ = .26 (Hunter & Hunter,1984), amount of job experience ρ = .18 (Schmidt, Hunter, & Outerbridge, 1986); and education ρ = .10 (Hunter & Hunter,1984) (see Schmidt, Ones, & Hunter (1992) for a current summary of meta-analyzed validities of many frequently used predictors). Given a validity of .41 for integrity tests, using integrity tests in personnel selection

would improve prediction and decrease both the false positive and false negative rates.

It has been argued that screening procedures such as integrity tests discourage good management and surveillance at the workplace. Good management techniques are important in organizations. But given better workers (brighter, more honest and hard-working workers), good management can produce even better results. It is likely that close surveillance at work can produce aversion, frustration, and low morale in workers—the syndrome of "why don't my superiors trust me? Why do they watch me all the time?" So it is apparent that using integrity tests before hiring is a good investment that can reduce unnecessary surveillance costs later on (and the added psychological costs).

One important question about integrity tests is whether applicants can be taught to appear honest on integrity tests. Can test takers alter their scores on integrity tests? When instructed some people can alter their scores on integrity tests. However, there is evidence from the personality domain that faking by applicants is infrequent (Hough et al., 1990). Furthermore, response distortion, if it exists at all, does not destroy the criterion-related validities of integrity tests (Ones et al., in press). It seems that individuals cannot be instructed to pass integrity tests without revealing to them the exact nature of the questions used in the inventories.

One major criticism of integrity testing relates to the unwillingness of some integrity test publishers to allow independent scrutiny of their tests by independent researchers. It must be pointed out that most integrity test publishers approached for data or tests to be used in this dissertation were very cooperative and forthright. Most test publishers are willing to allow the

use of their tests by interested researchers as long as the security of the proprietary scoring keys is not compromised. The magnitude of the secondary construct validity data accumulated for this dissertation attests to the fact that integrity test publishers do conduct good research studies on their tests and are willing to share them with interested independent researchers.

The uniqueness of this dissertation arises from the fact that it is the most comprehensive comparative construct validity study conducted on integrity tests. No other study has examined relationships between integrity tests to the extent that this study has. This is also the most comprehensive study examining the Big Five dimensions of personality as they relate to the construct of integrity. Also the precise estimates of the correlations between integrity and ability have not been provided elsewhere.

Ones et al. (in press) observed that it was not simply enough to show that integrity tests have generalizable predictive validities for job performance and counterproductive behaviors on the job; construct validity evidence is critical if the theoretical meaning of the criterion-related validities of integrity tests is to be determined. The results presented in this dissertation indicate that integrity may be the most important noncognitive individual differences variable predicting and explaining job performance as well as work-place counterproductive behaviors. Theoretically much can be gained by including integrity in models of job performance. Practically, integrity tests can provide employers with information on applicants not obtained by ability tests.

When I started my research on the construct validity of integrity tests in 1991, many were expressing disappointment in the small amount of independent research about what integrity tests measure (Goldberg et al., 1991;

Sackett et al., 1989; O'Bannon, Goldinger, & Appleby, 1989; and Woolley & Hakstian, in press). Many industrial psychologists were skeptical of integrity tests used in industry. The results presented in this dissertation based on more than 8,000 correlations have hopefully alleviated the construct validity related concerns and reservations of industrial psychologists about integrity tests.

Table 1

The Big Five Dimensions of Personality from Various Researchers

Researcher(s)	I	II	III	IV	V
Fiske (1949)	confident self-expression	social adaptability, conformity	will to achieve	emotional control	inquiring intellect
Cattell (1957)	exvia	cortertia	superego strength	anxiety	intelligence
Tupes & Christal (1961)	surgency	agreeableness	dependability	emotional stability	culture
Norman (1963)	surgency	agreeableness	conscientiousness	emotional stability	culture
Borgotta (1964)	assertiveness	likability	task interest	emotionality	intelligence
Eysenck (1970)	extroversion	psychoticism		neuroticism	
Guilford (1975)	social activity	paranoid disposition	thinking introversion	emotional stability	

table continues

Table 1—continued

Researcher(s)	I	II	III	IV	V
Digman & Takemoto-Chock (1981)	extroversion	friendly compliance	will to achieve	ego strength anxiety	intellect
Goldberg (1981,1989)	surgency	agreeableness	conscientiousness	emotional stability	intellect
Buss & Plomin (1984)	activity	sociability	impulsivity	emotionality	
McCrae & Costa (1985)	extroversion	agreeableness	conscientiousness	neuroticism	openness to experience
Conley (1985)	social extroversion	agreeableness	impulse control	neuroticism	intellectual interests

table continues

Table 1—continued

Researcher(s)	I	II	III	IV	V
Tellegen (1985)	positive emotionality		constraint	negative emotionality	
Hogan (1986)	sociability & ambition	likability	prudence	adjustment	intellectance
Lorr (1986)	interpersonal involvement	level of socialization	self-control	emotional stability	independent
De Raad et al. (1988)	extroversion	agreeableness vs. cold hearted-ness	conscientiousness	emotional stability	culture
Digman (1989)	extroversion	friendly compliance	will to achieve	neuroticism	intellect

table continues

Table 1—continued

Researcher(s)	I	II	III	IV	V
Botwin & Buss (1989)	extroverted	agreeable-stable	conscientious	dominant-assured	intellectance-culture
Field & Millsap (1989)	extroversion	agreeableness		satisfaction	intellect
Peabody & Goldberg (1989)	power	love	work	affect	intellect

Note. This table is compiled based on John (1990) and Digman (1989).

Table 2

The High and Low Ends of Big Five Dimensions of Personality

Dimension	High End	Low End
Emotional Stability	Calm	Emotional
	Placid	Worrying
	Poised	Easily Upset
	Not neurotic	Neurotic
Extroversion	Sociable	Retiring
	Talkative	Silent
	Assertive	Reserved
	Adventurous	Cautious
Openness to Experience	Intellectual/Cultured	Boorish
	Original	Conventional
	Imaginative	Practical
	Polished	Clumsy, awkward
Agreeableness	Good Natured	Spiteful
	Cooperative	Obstructive
	Not Jealous	Jealous
	Trustful	Suspicious

table continues

Table 2—continued

Dimension	High End	Low End
Conscientiousness	Conscientious	Unscrupulous
	Responsible	Frivolous
	Persevering	Quitting
	Dependable	Undependable

Note. High and low end adjectives are the most consistent definers of the Big Five.

Table 3

Meta-Analyzed Intercorrelations between the Big Five Dimensions of Personality

		ES	EX	OE	A	C
Emotional	ρ	.63				
Stability	K	319				
(ES)	N	148,721				
Extroversion	ρ	.19	.41			
(EX)	K	710	437			
	N	440,440	294,515			
Openness to	ρ	.16	.17	.43		
Experience	K	423	418	127		
(OE)	N	254,937	252,004	47,177		
Agreeableness	ρ	.25	.17	.11	.53	
(A)	K	561	243	236	102	
	N	415,679	135,529	144,205	79,303	
Conscientiousness	ρ	.26	.00	−.06	.27	.47
(C)	K	587	632	338	344	226
	N	490,296	683,001	356,680	162,975	288,512

Note. The values presented in each cell are true score correlations (ρ), number of coefficients in the meta-analysis (K), and the total sample size (N). The diagonal contains the true score correlations between different personality scales tapping the same dimension of the Big Five.

Table 4

Instruments Used in Primary Data Collection

Name of Instrument	Description
Personnel Selection Inventory (PSI)	overt integrity test
Stanton Survey	overt integrity test
Reid Report	overt integrity test
PDI Employment Inventory	personality-based integrity test
Inwald Personality Inventory (IPI)	personality-based integrity test
Hogan's Reliability Scale	personality-based integrity test
Personnel Reaction Blank	personality-based integrity test
Personal Characteristics Inventory (PCI)	measure of Big Five dimensions of personality
Goldberg-Adjective Checklist	Goldberg's Measure of Big Five dimensions of personality
Hogan Personality Inventory	Measure of 7 dimensions of personality

Table 5

Integrity Tests Contributing Correlations to the Meta-Analyses

Test Name
1. Accutrac Evaluation System[a]
2. Applicant Review[a]
3. Employee Attitude Inventory (London House)[a]
4. Employee Reliability Inventory[a]
5. Employment Productivity Index[b]
6. Hogan Personnel Selection Series (Reliability Scale)[b]
7. Inwald Personality Inventory[b]
8. P.E.O.P.L.E. Survey[a]
9. Personnel Decisions Inc. Employment Inventory[b]
10. Personal Outlook Inventory[b]
11. Personnel Reaction Blank[b]
12. Personnel Selection Inventory (London House)[a]
13. Phase II Profile[a]
14. P.O.S. Preemployment Opinion Survey[a]
15. Preemployment Analysis Questionnaire[a]
16. Reid Report and Reid Survey[a]
17. Rely[a]
18. Stanton Survey[a]
19. True Test[a]

table continues

Table 5—continued

Test Name
20. Trustworthiness Attitude Survey; PSC Survey; Drug Attitudes/Alienation Index[a]

Note. The list of publishers and authors of these tests are available in O'Bannon et al. (1989).

[a]Overt integrity test.

[b]Personality-Based integrity test.

Table 6

Hypothesized Intercorrelation Matrix of Integrity and the Big Five Dimensions of Personality

	I	EX	A	C	ES	OE
Integrity (I)	—	a	b	c	d	d
Extroversion (EX)		—	d	d	d	d
Agreeableness (A)			—	b	d	d
Conscientiousness (C)				—	d	d
Emotional Stability (ES)					—	d
Openness to experience (OE)						—

Note. When integrity tests are scored so that a high score indicates being more honest, a represents a low but slightly negative relationship, b represents a low but positive relationship, c represents a high positive relationship, and d represents a positive or a negative negligibly small relationship.

Table 7

Personality Inventories Contributing Data to the Analyses Reported

Personality Inventory Name
ABLE Substantive Scales
Adjective Checklist (Gough, 1952)
Adjective Checklist (Goldberg Big-Five Factor Markers)
Bernreuters Personality Inventory
California Psychological Inventory
CAQ Clinical Analysis Questionnaire
Comrey Personality Scales
Differential Personality Questionnaire
Edwards Personal Preference Schedule
Emotional Maladjustment Questionnaire
Eysenck Personality Questionnaire
Gordon Personal Inventory
Gordon Personal Profile
Guildford-Martin Personnel Inventory
Guildford-Zimmermann Temperament Survey
Hogan Personality Inventory
Interpersonal Adjective Scales
Inventory of Personal Motives
Inwald Personality Inventory
Jackson Personality Inventory
Jackson Personnel Research Form

table continues

Table 7—continued

Personality Inventory Name
Manifest Needs Questionnaire
Minnesota-Multiphasic Personality Inventory
Multi-dimensional Personality Questionnaire
Myers-Briggs
NEO-PI
Omnibus Personality Inventory
Personal Characteristics Inventory
Personality Plus
Personality Research Form
Profiles
Protestant Work Ethic
Self Descriptive Inventory
Sixteen Personality Factor Questionnaire
Special Assessment Battery
Thurstone Temperament Schedule

Table 8

Descriptive Information on Statistical Artifact Distributions Used to Correct Correlations

	No. of values	Mean	Standard deviation	Mean of the square roots of reliabilities	Standard deviation of the square roots of reliabilities
Integrity test reliabilities					
Overall distribution	124	.81	.11	.90	.06
Overt	97	.83	.09	.91	.05
Personality-Based	27	.72	.13	.85	.08
Personality scale reliabilities					
	372	.76	.11	.87	.07
Artifact distribution for range restriction corrections					
U[a] for integrity	79	.81	.19	—	—

[a]U refers to the ratio of the selected group standard deviation to the referent group standard deviation.

Table 9

Intercorrelations Between Overt Integrity Tests (Obtained from the Primary Data)

	PSI Honesty	Reid Report	Stanton Survey
PSI Honesty	—	.59	.79
Reid Report	.71	—	.74
Stanton Survey	.95	.89	—

Note. The correlations above the diagonal are the observed values; the values reported below are the correlations corrected for unreliability in the measures.

Table 10

Intercorrelations between Personality-Based Integrity Tests (Obtained from the Primary Data)

	PRB	PDI-EI	Reliability Scale	IPI RISK
Personnel Reaction Blank (PRB)	—	.40	.57	.57
PDI-Employment Inventory (PDI-EI)	.56	—	.47	.53
Hogan Personality Inventory-Reliability Scale (Reliability)	.79	.65	—	.53
Inwald Personality Inventory (IPI-RISK)	.79	.74	.74	—

Note. The correlations above the diagonal are the observed values; the values reported below are the correlations corrected for unreliability in the measures.

Table 11

Observed and True Score Intercorrelations between Personality-Based Integrity Tests (Obtained from the Primary Data)

	Reid Report	Stanton Survey	PSI
Personnel Reaction Blank (PRB)	.20 (.26)	.37 (.48)	.55 (.71)
PDI-Employment Inventory (PDI-EI)	.28 (.36)	.21 (.37)	.48 (.62)
Hogan Personality Inventory-Reliability Scale (Reliability)	.37 (.48)	.52 (.68)	.44 (.57)
Inwald Personality Inventory (IPI-RISK)	.40 (.52)	.52 (.68)	.39 (.51)

Note. The correlations without parentheses are the observed values; the values in parentheses reported following the observed correlations are corrected for unreliability in the measures.

Table 12

The Correlations between a Linear Composite of Each Dimension of the Big Five and a Linear Composite of Integrity Test Scores (Using the Primary Data)

Personality Dimension	ρ with Integrity	r with integrity
Emotional Stability	.50	.46
Extroversion	.31	.25
Openness to Experience	.08	.06
Agreeableness	.61	.53
Conscientiousness	.91	.85

Note. A composite was formed for each one of the Big Five dimensions by combining the appropriate scales from the Personal Characteristics Inventory, Hogan Personality Inventory and Goldberg's Adjective Checklist. The composite for integrity was formed by combining the seven integrity tests used in the primary data collection. ρ = true score correlation with integrity, r= observed correlation with integrity.

Table 13

The Correlations between a Linear Composite of Each Dimension of the Big Five and a Linear Composite of Overt Integrity Test Scores (Using the Primary Data)

Personality Dimension	ρ with overt tests	r with overt tests
Emotional Stability	.31	.28
Extroversion	.19	.15
Openness to Experience	.08	.06
Agreeableness	.47	.40
Conscientiousness	.68	.58

Note. A composite was formed for each one of the Big Five dimensions by combining the appropriate scales from the Personal Characteristics Inventory, Hogan Personality Inventory and Goldberg's Adjective Checklist. The composite for integrity was formed by combining the three overt integrity tests used in the primary data collection. ρ = true score correlation with overt tests, r= observed correlation with overt tests.

Table 14

The Correlations between a Linear Composite of Each Dimension of the Big Five and a Linear Composite of Personality-Based Integrity Test Scores (Using the Primary Data)

Personality Dimension	ρ with personality-based tests	r with personality-based tests
Emotional Stability	.51	.45
Extroversion	.37	.29
Openness to Experience	.03	.02
Agreeableness	.58	.49
Conscientiousness	.86	.72

Note. A composite was formed for each one of the Big Five dimensions by combining the appropriate scales from the Personal Characteristics Inventory, Hogan Personality Inventory and Goldberg's Adjective Checklist. The composite for integrity was formed by combining the four personality-based integrity tests used in the primary data collection. ρ = true score correlation with personality-based tests, r= observed correlation with personality-based tests.

Table 15

Meta-Analyzed Intercorrelations between Overt and Personality-Based Integrity Tests

		Overt	Personality-Based
Overt	Total N	7,424	15,978
	K	56	117
	r_{mean}	.32	.25
	ρ	.45	.39
Personality Based	Total N	—	7,062
	K	—	37
	r_{mean}	—	.43
	ρ	—	.70

Note. N = Total sample size; K = number of correlations; r_{mean} = mean observed correlation; ρ = true score correlation.

Table 16

Meta-Analysis of Integrity Test Correlations with the Big Five Dimensions of Personality

	ρ	K	N	r_{mean}
Emotional Stability	.33	378	78,651	.22
Extroversion	–.08	233	59,030	–.05
Openness to Experience	.12	200	46,368	.08
Agreeableness	.40	272	62,097	.26
Conscientiousness	.42	423	91,360	.28

Note. ρ = true score correlation; K = number of correlations; N = Total sample size; r_{mean} = mean observed correlation; .

Table 17

Meta-Analysis of Overt Integrity Test Correlations with the Big Five Dimensions of Personality

	ρ	K	N	r_{mean}
Emotional Stability	.28	142	17,265	.18
Extroversion	.03	94	12,836	.02
Openness to Experience	.09	97	13,608	.06
Agreeableness	.34	105	13,885	.23
Conscientiousness	.39	160	22,422	.26

Note. ρ = true score correlation; K = number of correlations; N = Total sample size; r_{mean} = mean observed correlation.

Table 18

Meta-Analysis of Personality-Based Integrity Test Correlations with the Big Five Dimensions of Personality

	ρ	K	N	r_{mean}
Emotional Stability	.37	236	61,386	.23
Extroversion	–.11	139	46,194	–.07
Openness to Experience	.14	103	32,760	.09
Agreeableness	.44	167	48,212	.28
Conscientiousness	.45	263	68,942	.28

Note. ρ = true score correlation; K = number of correlations; N = Total sample size; r_{mean} = mean observed correlation.

Table 19

True Score Correlations between Composites of Integrity Test Scores and Composites of the Big Five Dimensions of Personality: Using Meta-Analytic Results

	Integrity tests	Overt tests	Personality-Based tests
Emotional Stability	.69	.50	.54
Extroversion	−.18	.07	−.20
Openness to Experience	.26	.19	.24
Agreeableness	.78	.66	.69
Conscientiousness	.87	.81	.75

Table 20

The True Score Correlation Matrix between the Big Five Dimensions of Personality, Integrity, and Job Performance

	C	A	ES	OE	EX	I	JP
C	1.00	.27	.26	.16	.19	.42	.23
A	.27	1.00	.25	.11	.17	.40	.06
ES	.26	.25	1.00	.16	.19	.33	.07
OE	.16	.11	.16	1.00	.17	.12	–.03
EX	.19	.17	.19	.17	1.00	–.08	.10
I	.42	.40	.33	.12	–.08	1.00	.46
JP	.23	.06	.07	–.03	.10	.46	1.00

Note. The correlations presented are true score correlations based on meta-analyses C= Conscientiousness; A = Agreeableness; ES = Emotional Stability; OE = Openness to Experience; EX = Extroversion; I = Integrity; JP = Job Performance.

Table 21

First-Order Partial Correlations between the Big Five Dimensions of Personality, Integrity, and Job Performance (Conscientiousness Partialed out)

	A	ES	OE	EX	I	JP
A	1.00	.19	.07	.13	.33	.00
ES	.19	1.00	.13	.15	.25	.01
OE	.07	.13	1.00	.15	.04	–.07
EX	.13	.15	.15	1.00	.00	.06
I	.33	.25	.04	.00	1.00	.41
JP	.00	.01	–.07	.06	.41	1.00

Note. A= Agreeableness; ES = Emotional Stability; OE = Openness to Experience; EX = Extroversion; I = Integrity; JP = Job Performance.

Table 22

Meta-Analysis of the Relationship between Integrity and Age

	r_{obs}	ρ	K	N
All integrity tests	.16	.21	36	18,043
Overt tests	.17	.22	27	7,320
Personality-Based tests	.15	.21	.15	10,723

Note. Positive values indicate that older individuals score higher on integrity tests. r_{mean} = mean observed correlation between age and integrity; ρ = mean corrected (for unreliability in integrity) correlation between age and integrity; K = number of correlations; N = Total sample size.

Table 23

Meta-Analysis of the Relationship between Gender and Integrity

	d_{obs}	δ	K	N	r_{mean}
All integrity tests	.65	.93	90	256,991	.31
Overt tests	.82	1.12	70	167,935	.38
Personality-Based tests	.39	.58	20	89,056	.19

Note. Positive values indicate that females score higher on integrity tests. d_{mean} = difference between male/female means in standard score form; δ = d value corrected for unreliability in integrity tests; K = number of correlations; N = Total sample size; r_{mean} = mean observed correlation between gender and integrity.

Table 24

Meta-Analyses of the Relationships between Race and Integrity

	d_{obs}	δ	K	N	r_{mean}
		Whites vs. Blacks			
All integrity tests	.20	.24	42	312,751	.10
Overt tests	.61	.68	37	152,702	.29
Personality-Based tests	–.16	–.18	5	292,387	–.08
		Whites vs. Hispanics			
All integrity tests	.18	.20	39	262,060	.09
Overt tests	.43	.47	35	132,120	.21
Personality-Based tests	–.06	–.04	4	129,940	–.03
		Whites vs. Asians			
All integrity tests	.08	.10	15	137,403	.04
Overt tests	.22	.26	11	15,770	.11
Personality-Based tests	.06	.08	4	121,633	.03

Note. Positive values indicate that minorities score higher on integrity tests. d_{mean} = difference between minority group/white means in standard score form; δ = d value corrected for unreliability in integrity tests; K = number of correlations; N = Total sample size; r_{mean} = mean observed correlation between race and integrity.

Table 25

List of Cognitive Ability Tests Contributing Data to the Meta-Analyses

Test	Scale
ASVAB	Word Knowledge, Arithmetic Reasoning, Math Knowledge, Electronics Information, Mechanical Comprehension, General Science, Various Composites
Basic Skills Tests	Reading Comprehension, Computation, Following Written Directions, Coding, Computations, Visual Scanning, Language Skills, Vocabulary, Problem Solving, Decision Making, Oral Directions, Written Directions, Form Checking, Reasoning, Classifying, Filing Names, Visual Speed & Accuracy, Memory, Filing Numbers
Differential Aptitude Test	Numerical Reasoning, Verbal Reasoning
Flanagan Industrial Tests	Inspection, Precision, Assembly
Watson Glaser Critical Thinking	—

Table 26

Descriptive Information on Test Reliabilities Used to Correct Correlations

	Number of values	Mean	Standard deviation	Mean of the square roots of reliabilities	Standard deviation of the square roots of reliabilities
			Integrity test reliabilities		
Overall distribution	124	.81	.11	.90	.06
Overt	97	.83	.09	.91	.05
Personality-Based	27	.72	.13	.85	.08
			Ability test reliabilities		
	100	.80	.08	.89	.05

Table 27

Meta-Analysis of Integrity Test Correlations with General Mental Ability

	ρ	K	N	r_{mean}
All integrity tests	.02	106	23,306	.01
Overt tests	–.01	29	4,903	–.00
Personality-Based tests	.03	77	18,386	.02

Note. ρ = true score correlation; K = number of correlations; N = Total sample size; r_{mean} = mean observed correlation between gender and integrity.

Table 28

Effect of Combining Integrity Tests with Measures of Ability in Predicting Ratings of Overall Job Performance

Job complexity level [a]	Validity of general mental ability[b] (GMA)	Validity of psychomotor ability[b] (PA)	Validity of integrity[c] (I)	Multiple correlations		
				GMA+PA[d]	GMA+I	GMA+PA+I
Complexity level 1 (Highest)	.58	.21	.41	.58	0.70	0.70
Complexity level 2	.56	.30	.41	.55	0.69	0.69
Complexity level 3	.51	.32	.41	.53	0.65	0.66
Complexity level 4	.40	.43	.41	.50	0.57	0.64
Complexity level 5 (Lowest)	.23	.48	.41	.49	0.47	0.63

Note. The multiple correlations reported in this table are computed by taking both general mental ability and psychomotor ability correlations with integrity as .02 (the correlation between integrity and ability, meta-analytically computed in this study).

table notes continue

Table 28 notes—continued

[a]Job complexity levels are those used by Hunter (1980). [b]Validities are from Hunter (1980). [c]Predictive validity of integrity tests for supervisory ratings of overall job performance calculated using applicants (See Ones et al., in press). [d]From Hunter (1980).

APPENDIX A

QUESTIONNAIRE USED IN ADDITIONAL DATA COLLECTION FROM THE SUBJECTS

BEFORE YOU LEAVE...

We are interested in your reactions to the questionnaires that you completed. Also we would like to know a little more about your demographic characteristics. Use the half form University of Iowa computerized answer sheet to respond to the following questions. If you come to a question about an inventory or survey that you did not fill out, skip the question.

Questions on the PSI-7ST

1. In general , how reasonable is it for organizations to use the PSI-7ST to select applicants for jobs?
    a. Very reasonable
    b. Reasonable
    c. Somewhat reasonable
    d. Unreasonable

2. How "fair" do you think it is to use the PSI-7ST to select employees?

   a. Very fair

   b. Fair

   c. Somewhat fair

   d. Not fair

Questions on the PDI Employment Inventory

3. In general , how reasonable is it for organizations to use the PDI Employment Inventory to select applicants for jobs?

   a. Very reasonable

   b. Reasonable

   c. Somewhat reasonable

   d. Unreasonable

4. How "fair" do you think it is to use the PDI Employment Inventory to select employees?

   a. Very fair

   b. Fair

   c. Somewhat fair

   d. Not fair

Questions on the Hogan Personality Inventory

5. In general , how reasonable is it for organizations to use the Hogan Personality Inventory to select applicants for jobs?

   a. Very reasonable

   b. Reasonable

   c. Somewhat reasonable

   d. Unreasonable

6. How "fair" do you think it is to use the Hogan Personality Inventory to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Other Questions

7. What is your gender?

    a. Male

    b. Female

8. Please go on to question nine.

9. What is your ethnic background?

    a. Caucasian

    b. African American

    c. Hispanic

    d. Oriental

    e. Other

10. Which part of the US are you originally from?

    a. Northeast

    b. Southeast

    c. South

    d. West Coast

    e. Midwest / Middle US

11. What is your academic standing?

    a. Freshman

    b. Sophomore

    c. Junior

    d. Senior

Questions on the IPI (Inwald Personality Inventory)

12. In general , how reasonable is it for organizations to use the IPI (Inwald Personality Inventory) to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

13. How "fair" do you think it is to use the IPI (Inwald Personality Inventory) to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on the PCI

14. In general , how reasonable is it for organizations to use the PCI to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

15. How "fair" do you think it is to use the PCI to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on "How Accurately Can You Describe Yourself"

16. In general , how reasonable is it for organizations to use "How Accurately Can You Describe Yourself" to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

17. How "fair" do you think it is to use "How Accurately Can You Describe Yourself" to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on Employee Reliability Inventory

18. In general , how reasonable is it for organizations to use Employee Reliability Inventory to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

19. How "fair" do you think it is to use Employee Reliability Inventory to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on Reid Report

20. In general , how reasonable is it for organizations to use Reid Report to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

21. How "fair" do you think it is to use Reid Report to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on the Stanton Survey

22. In general , how reasonable is it for organizations to use Stanton Survey to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

23. How "fair" do you think it is to use Stanton Survey to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on the Personnel Reaction Blank

24. In general , how reasonable is it for organizations to use Personnel Reaction Blank to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

25. How "fair" do you think it is to use Personnel Reaction Blank to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

Questions on the Wonderlic Test

26. In general , how reasonable is it for organizations to use the Wonderlic test to select applicants for jobs?

    a. Very reasonable

    b. Reasonable

    c. Somewhat reasonable

    d. Unreasonable

27. How "fair" do you think it is to use the Wonderlic test to select employees?

    a. Very fair

    b. Fair

    c. Somewhat fair

    d. Not fair

APPENDIX B

INSTRUCTIONS FOR SUBJECTS

PLEASE READ THE INSTRUCTIONS BELOW FIRST:

You are being asked to participate in a study assessing the relationship between several pre-employment selection questionnaires and personality characteristics of individuals. You will be asked to complete screening surveys anonymously. All these surveys are currently being used to select employees for various organizations.

Your participation in this study is completely voluntary. Both your instructor and the researchers of this project have judged it to be a valuable experience. You are free to discontinue the experiment anytime, if you feel any type of discomfort or for any other reason.

All the information provided for this experiment will be anonymous using assigned identification numbers. Do not use your name or social security number on any forms that you fill out.

The researcher for this project is Deniz Ones of the department of management and organizations. The data collection for the project will continue through September 1992. Individual results will not be available to you because of the anonymous nature of the responses. However, if you are interested in the research findings, please contact Deniz Ones at 335-0972 .

The names of the people who complete the questionnaire will be provided to your instructor for class purposes. So make sure that you sign the attendance sheet, when you turn in your experimental materials.

Please do not start answering questions before the experimenter explains the questionnaires.

You will be timed on the Wonderlic test. But you will be allowed to work at your own pace for the other questionnaires.

It is important that you are cooperative and complete with your answers.

Thank you for your participation.

APPENDIX C

DEBRIEFING FORM

In this study, we are interested in how some pre-employment screening surveys relate to the personal characteristics of individuals. In particular, we will be studying (1) if the pre-employment screening surveys measure what they say they measure, and (2)which other variables these questionnaires relate to. We are also interested in the reactions of the survey participants to these questionnaires. Your responses will contribute to a large database of many other pre-employment surveys used in selecting employees across many individuals. Ultimately more than 1000 people will be included in the database.

To determine if the surveys of personal characteristics measure what they claim to measure, we will use advanced statistical techniques including correlational analyses, meta-analyses and factoring.

APPENDIX D

SAMPLE SIZES IN PRIMARY DATA

Table 29

Sample Sizes in Primary Data

	PSI	Stanton	Reid	PDI	HPI	PRB	IPI	PCI	Wonder	Gold
PSI	500									
Stanton	112	320								
Reid	125	164	443							
PDI	140	123	124	423						
HPI	229	134	93	167	471					
PRB	97	157	84	113	180	446				

table continues

Table 29—continued

	PSI	Stanton	Reid	PDI	HPI	PRB	IPI	PCI	Wonder	Gold
IPI	302	99	124	222	222	127	487			
PCI	157	121	66	98	105	181	177	487		
Wonder	112	309	162	142	108	208	133	136	366	
Gold	167	112	114	194	147	133	187	194	104	446

Note. PSI = Personnel Selection Inventory; Stanton = Stanton Survey; Reid = Reid Report; PDI = Personnel Decisions, Inc.; Hogan = Hogan Personality Inventory; PRB = Personnel reaction Blank; IPI = Inwald Personality Inventory; PCI = Personality Characteristics Inventory; Wonder = Wonderlic; Gold = Goldberg's Adjective Checklist.

APPENDIX E

CODING SHEET FOR INTEGRITY TEST CORRELATIONS WITH OTHER VARIABLES, AND INTEGRITY-PERSONALITY INTERCORRELATIONS FOR META-ANALYSIS

Study Number: _____

Unique Sample Number: _____

Test or Variable 1: _____

Scale: _____

Test or Variable 2: _____

Scale: _____

Sample Size: _____

Adjusted Sample Size: _____

Correlation: _____

Was it reported as above? _____

Time interval between the administrations of instruments: _____

Ees/Apps./Students/General Population?? _____

Industry: _____

Specific Job: _____

Reliability of Test 1: _____

Type of Reliability: _____

Reliability of Test 2: _____

Type of Reliability: _____

Range Restriction (SD_{res}/SD_{un}) on Test 1: _____

Range Restriction (SD_{res}/SD_{un}) on Test 2: _____

Sample Demographics: _____

Age: _____

Sex: _____

Race: _____

_____

Location: _____

Test 1 Anonymously Completed? _____

Self Rating? _____

Research Purpose/Administrative Purpose? _____

Test 2 Anonymously Completed?: _____

Self Rating? _____

Research Purpose/Administrative Purpose? _____

Published Study? _____

Sponsored Research? _____

APPENDIX F

LIST OF JOURNAL ARTICLES CODED

Author(s): _____

Year: _____

Title: _____

_____

Journal: _____

Volume: _____

Pages: _____

Study Number: _____

_____

APPENDIX G

REPORTS AND BOOKS CODED

Author(s): _____

Year: _____

Title: _____

Pages: _____

Publisher: _____

Other: _____

_____

_____

Study Number: _____

_____

APPENDIX H

CLASSIFICATION OF VARIOUS PERSONALITY SCALES INTO THE BIG FIVE DIMENSIONS OF PERSONALITY

Table follows

Table 30

Classification of Various Personality Scales into The Big Five Dimensions Of Personality

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
ABLE Substantive Scales	Emotional Stability	Dominant, Energy Level	Traditional Values	Cooperative	Conscientious, Internal Control, Non delinquent, Work Orientation

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Adjective Checklist (Gough,1952)	Abasement, Succorance	Affiliation, Dominance, Exhibition	Autonomy, Change, Convention-ality, Endurance, Imagination, Intraception, Liability	Aggression, Deference, Nurturance	Achievement, Order, Self-Control

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
California Psychological Inventory	Self-Control, Well-Being	Dominance, Sociability, Social Presence	Flexibility, Self-Acceptance, Tolerance	Psychological Mindedness	Achievement via Conformance, Achievement via Independence, Responsibility, Socialization, Status

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Bernreuters Personality Inventory	Neurotic Tendencies, Self-Confidence	Dominance, Extroversion, Self-Sufficiency, Sociability			

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Clinical Analysis Questionnaire	Agitated Depression, Anxious Depression, Bored Depression, Hypochondriasis, Paranoia, Psychasthenia, Psychological Inadequacy, Schizophrenia, Suicidal	Low Energy			Psychopathic Deviate

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Comrey Personality Scales	Emotional Stability	Activity, Extroversion		Empathy, Trust	Orderliness, Social Conformity
Differential Personality Questionnaire	Alienation, Harm Avoidance, Stress Reaction, Well-Being	Social Potency	Traditionalism	Affiliation, Aggression, Social Closeness	Achievement, Control

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Edwards Personal Preference Schedule	Abasement, Succorance	Dominance, Exhibition	Autonomy, Change, Deference	Affiliation, Aggression, Intraception, Nurturance	Achievement, Endurance, Order

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Emotional Maladjustment Questionnaire	Depression, Fear & Anxiety, Inferiority Feelings, Organic Reaction, Sexual Disturbances	Withdrawal	Rationalization	Hostility	

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Eysenck Personality Questionnaire	Neuroticism	Extroversion		Psychoticism	
Guildford Martin Personnel Inventory	Inferiority, Nervousness	Ascendance, General Activity			
Goldberg Big Five Factor Markers	Emotional Stability	Surgency	Intellect	Agreeableness	Conscientiousness

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Gordon Personal Inventory			Cautiousness, Original Thinking	Personal Relations	Vigor
Gordon Personal Profile	Emotional Stability	Ascendance, Sociability			Impulsive, Responsibility
Guildford-Martin Personnel Inventory	Objectivity	Ascendance, Dominance, Social Leader	Tolerance	Agreeableness, Cooperative-ness	

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Guildford-Zimmermann Temperament Survey	Emotional Stability, Objectivity	Ascendance, General Activity, Sociability	Thoughtfulness	Friendliness, Personal Relations	Restraint
Hogan Personality Inventory	Adjustment	Sociability	Intellectance, School Success	Likability	Ambition, Prudence

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Interpersonal Adjective Scales		Aloof-Introverted, Assured-Dominant, Gregarious-Extroverted	Unassuming-Ingenious	Arrogant-Calculating, Cold-Hearted, Warm-Agreeable	
Inventory Personal Motives		Power, Recognition	Aesthetic, Scientific	Affiliation, Altruism	Security, Tradition

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Inwald Personality Inventory	Anxiety, Depression, Hyperactivity, Illness concerns, Obsessive Personality, Phobic Personality, Treatment Programs	Lack of Assertiveness, Loner Type	Rigid Type	Antisocial Attitudes, Interpersonal Difficulties, Sexual Concerns, Spouse/Mate Conflicts, Trouble With the Law, Undue Suspiciousness	Absence Abuse, Alcohol Abuse, Driving Violations, Drug Use, Job Difficulties, Substance Abuse

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Jackson Personality Inventory	Anxiety	Energy Level, Social Adroitness, Social Participation	Breadth of interest, Complexity, Innovation, Tolerance, Value Orthodoxy	Interpersonal Affect	Conformity, Organization, Responsibility, Risk-taking

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Jackson Personnel Research Form		Ascendancy	Intellect	Interpersonal	Need for Achievement, Impulse, Work Orientation
Life Experiences Inventory	Personal			Social	Childhood
Mallar Tests		Inhibition		Cooperation	Persistence
Manifest Needs Questionnaire		Dominance	Autonomy	Affiliation	Achievement

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Minnesota Mutiphasic Personality Inventory	Depression, Hypochondriasis, Paranoia, Psychostenia, Schizophrenia	Extroversion		Defensiveness, Hysteria	McAndrew Add, Mania, Psychopathic Deviate

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Multi-Dimensional Personality Questionnaire	Alienation, Harm Avoidance, Negative Emotionality, Positive Emotions, Stress Reaction, Well-Being	Social Potency	Traditionalism	Aggression, Social Closeness	Achievement, Constraint, Hard Work, Impulsiveness

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Myers-Briggs		Introversion-Extroversion	Intuition-Sensing, Perceiving-Judging	Feeling-Thinking	
NEO-PI	Neuroticism	Extroversion	Openness	Agreeableness	Conscientiousness

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Omnibus Personality Inventory	Anxiety Level, Personal Integration	Social Extroversion	Autonomy, Complexity, Estheticism, Practical Outlook, Religious Orientation, Theoretical Orientation, Thinking Introversion		Impulse Expression

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Personality Research Form	Defendance	Dominance, Exhibition, Play	Autonomy, Change, Cognitive Structure, Harm Avoidance, Sentience, Understanding	Abasement, Affiliation, Aggression, Nurturance, Social Recognition	Achievement, Endurance, Impulsivity, Order
Personality Plus	Sensitivity, Tension	Assertiveness, Recognition	Flexibility, Imagination	Social Need	Competitive, Organization

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Profile	Argument, Arrogance, Dependence, Fear of Failure, No Common Sense, Passive-Aggressive, Unstable Relationships	Attention Seeking		Interpersonal Insensitivity	Perfectionism, Trustworthiness
Protestant Work Ethic					Conethic, Proethic

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Self Descriptive Inventory	Maturity		Self-Actualization	Interpersonal Affinity	Achievement, Decisiveness, Initiative

table continues

Table 30—continued

Inventory	Emotional Stability	Extroversion	Openness to Experience	Agreeableness	Conscientiousness
Personality Factor Questionnaire	Anxiety, Apprehensive, Emotional Stability, Tender minded, Tense	Assertive, Bold, Dominant, Enthusiastic, Extroversion, Happy-go-lucky, Outgoing, Venturesome	Experimenting, Imaginative, Independence, Self-sufficient	Suspicious, Warm	Conscientiousness, Controlled, Ego Strength
Special Assessment Battery	Adjustment	Surgency		Likability	Conscientiousness, Prudence, Work Orientation

APPENDIX I

SOURCES OF DATA FOR THE META-ANALYSES

Alexander, R. (March, 1993). Personal communication.

Arnold, D. & Brooks, P. (1988). Personality, Correlates of Honesty in a Public Safety Setting. RPS Research Memorandum Number 14. (Tech. rep.). Chicago: Reid Psychological Systems.

Ash, P. (1986). Fakability of the Reid Report. RPS Research Memorandum Number 5. (Tech. rep.). Chicago: Reid Psychological Systems.

Ash, P. (1991). The Construct of Employee Theft Proneness. SRA/London House.

Ash, P. (1992, March). Prediction of Employee Honesty from Youthful and from Adult Experiences. Paper presented at the meeting for the Academy of Criminal Justice Sciences, Denver, CO.

Barbrisons Management Systems, Inc. (1988). Accutrac Evaluation System for trucking transportation personnel: Concurrent validation and adverse Impact studies for the positions of company driver and owner/operator and the Accutrac measure of violation of company policy and property damage accidents for "----", Inc. (Unpublished tech. rep.). Cincinnati, OH: Author.

Barge, B. N. & Skilling, N. J. (1986). An Analysis of Requirements, Employee Characteristics and Performance in Kelly Assisted Living Assignments. Personality Decisions Research Institute. (Rep. #121). Minneapolis: Personality Decisions Research Institute.

Barrick, M. R. & Mount, M. K. (in press). Autonomy as a moderator of relations between the Big Five Personality Dimensions and Job Performance. Journal of Applied Psychology.

Behrens, G. M. (19??). Construct validity of the PSI honesty scale with selection measures for drug store applicants. (Tech. rep.). Park Ridge, IL: London House.

Berland, A. & Muchinsky, P. (1991). Psychometric Assessment of Honest Testing for Personnel Selection. Unpublished manuscript, Iowa State University, Ames, Iowa.

Berte, Moretti, Julko, & Leonard. (1984). An investigation of a combined withdrawal and counterproductive behavior decision process model. Paper presented at 2nd Annual National Graduate Student Convention in I/O Psychology - OB, East Lansing, MI. (Tech. rep. #13). Park Ridge, IL: London House.

Board of Directors of The Association of Personnel Test Publishing (1990). A critical review of OTA's "The Validity of Integrity Tests in Personnel Selection." Unpublished manuscript.

Borofsky, G. (1993, 30 April). Personal communication.

Borofsky, G. L. (1990). User's Manual for the Employee Reliability Inventory. Boston, MA: Bay State Psychological Associates.

Borofsky, G. L. (1992). Assessing the likelihood of reliable workplace behavior: Further contributions to the validation of the Employee Inventory. Psychological Reports, pp. 563-592.

Brannick, J. (1993, 13 April). Personal communication.

Brooks, P. & Arnold, D. (1988). Report and Survey Norm Modification and Lack of Adverse Impact. (Tech. rep. #17).

Brooks, P. & Arnold, D. (1989). Reid Report Examiner's Manual, 3rd ed. Chicago: Reid Psychological Systems.

Brooks, P. & Arnold, D. (1989). The Reid Report's Lack of Impact. RPS Research Memorandum 22.

Brooks, P. (1990). Predictive Relationship of the Reid Report to Measures of Employee Work Performance in a Retail Setting. RPS Research Memorandum Number 36.

Brooks, P. (1992, 1 March). Personal communication.

Broucek, M. J. (1992, 23 Nov.). Personal communication. Minneapolis: Personnel Decisions, Inc.

Carpenter, B. N. (1986). Reliability and motivational distortion in the Wilkerson Pre-employment Audit. (Unpublished tech. rep.). Houston: Team Building Systems, Inc.

Caspy, T., Reisler, A., & Mendelson, Y. (1987). MMPI and Intelligence Correlates: Assessment of relationships via multivariate analyses. Journal of Clinical Psychology, 43, pp. 415-421.

CHC Forecast Inc. (1989). The Applicant Review; Administrator's Manual. Orem, Utah: Author.

Chrisman, F. P. (1976). Attitude Psychology and Public Opinion. Pennsylvania State University Press.

Clarke, W. V. & Hassler, K. R. (1967). Differentiation of criminals and non-criminals with a self concept measure. Psychological Reports, 20, 623-632.

Collins, J. M., & Schmidt, F. L. (in press). Personality, integrity, and white collar crime: A construct validity study. Personnel Psychology.

Cooke, D. K., Bernardin, J. H., & Salerno, L. J. (1991). The predictive validity of an integrity test for convenience store employees. Paper presented at Academy of Management Meetings, Miami Beach, Florida.

Cortina, J. M., Doherty, M. L., Schmitt, N., Kaufman, G., Smith, R. G. (1992). The "Big Five" personality factors in the IPI and MMPI: Predictors of Police Performance. Personnel Psychology, 45, pp. 119-140.

Costa, P. T., Busch, C. M., Zonderman, A. B., & McCrae, R. R. (1986). Correlations of MMPI Factor Scales with measure of the Five Factor Model of Personality. Journal of Psychological Assessment, 50, pp. 640-650.

Cunningham, M. The effects of honesty test content information on performance: Testing the resistance of the Reid Report faking. Unpublished manuscript, University of Louisville, Louisville, KY.

Cunningham, M., Trucott, M., & Wong, D. (1990). Paper presented at the Annual Convention of the American Psychological Association, Boston, MA.

Cunningham, M. R. (1989). Test-taking motivations and outcomes of a standardized measure of on the job integrity. Journal of Business and Psychology, 4(1), pp. 119-127.

Curphy, G. (1991, 25 April). Personal communication.

Day, D. V. & Siverman, S. B. (1989). Personality and job performance: Evidence of incremental validity. Personnel Psychology, 42, pp. 25-36.

Deardorff, P. A., Finch, A. J., Kendall, P.C., Lira, F., & Inderisano. (1975). Empathy and socialization in repeat offenders, first offenders, and normals. Journal of Counseling Psychology, pp. 453-455.

Durbrow, Brian R. (1985). Accutrac Evaluator System for Hotel/Restaurant Personnel Predictive Validation Study and Minority Adverse Impact. (Unpublished tech. rep.). Cincinnati: Accutrac Evaluation System.

Durbrow, Brian R. Summary of Study Results for "----". (Unpublished confidential tech. rep.). Cincinnati: Accutrac Evaluation System.

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(3), pp. 269-278.

George, J. M. (1991). State or Trait: Effects of Positive Mood on Prosocial Behaviors at Work. Journal of Applied Psychology, 76, pp. 299-307.

Goldberg, L. R. (1991). The Development of Markers for the Big-Five Factor Structure. Unpublished manuscript.

Gough, H. G (1971). The assessment of wayward impulse by means of the Personnel Reaction Blank. Personnel Psychology, 24, pp. 669-677.

Gough, H. G. (1965). Conceptual Analysis of Psychological Test scores and Other Diagnostic Variables. Journal of Abnormal Psychology, 70(4), pp. 294-302.

Gough, H. G. (1965). Cross-cultural validation of a measure of asocial behavior. Psychological Reports, 17, pp. 379-387.

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

Hare, R. D. (1985). Comparison of Procedures for the Assessment of Psychopathy. Journal of Counseling and Clinical Psychology, 53, pp. 7-16.

Harris, J. (1993, 25 March). Personal communication.

Heilbun, A. B. (1982). Cognitive Models of Criminal Violence Based upon Intelligence and Psychopathy Levels. Journal of Consulting and Clinical Psychology, 50, pp. 546-557.

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

Hogan, J. (1993, 16 April). Personal communication.

Hogan, J. (in press). Training Predictors. Journal of Military Psychology.

Hogan, R. & Hogan, J. (1992). Hogan Personality Inventory Manual. Tulsa, OK: Hogan Assessment Systems.

Homer, B. C. & Reed. J. R. (1985). A Survey of Employment-Related Behavioral Traits. Stanton Survey Validation Studies. Charlotte, N. C.: Stanton Corporation.

Hooland, J. (1959). The prediction of college grades from the California Psychological Inventory and the Scholastic Aptitude Test. Journal of Educational Psychology, 50, pp. 135-142.

Hunter, J. E. (1982). Personality, Cognitive Ability and Executive Performance. Unpublished manuscript.

Inwald, R. (1993, 8 March). Personal communication.

Jones, J. W. & Fay, L. (1987). Predicting Child Abuse Potential with the Personnel Selection Inventory for Childcare Workers. (Tech. rep. #54). Park Ridge, IL: London House.

Jones, J. W. & Terris, W. (1982). The Employee Attitude Inventory: A Validity Study on Theft by Current Employees. Paper presented at the 8th Annual Meeting of the Society of Police and Criminal Psychology, Nashville. (Tech. rep. #E1).

Jones, J. W. & Terris, W. (1983). Personality Correlates of Theft and Drug Abuse among Job Applicants. Proceedings of the 3rd International Conference on the 16 PF Test. (Tech. rep. #32), pp. 85-94.

Jones, J. W. & Terris, W. (1983). Predicting Employees; Theft in Home Improvement Centers. Psychological Reports, 52, pp. 187-201.

Jones, J. W. (1979). Employee Deviance. Attitudinal Correlates of Theft and on the Job Alcohol Abuse. Paper presented at the 5th Annual Convention of the Society of Police and Criminal Psychology. (Tech. rep. #5). Park Ridge, IL: London House.

Jones, J. W. (1980). Attitudinal Correlates of Employees' Deviance: Theft, alcohol use, and nonprescribed drug use. Psychological Reports, 47, pp. 71-77.

Jones, J. W. (1981). Dishonesty, Staff Burnout, and Employee Theft. Paper presented at 89th Annual APA Meeting, Los Angeles, CA. (Tech. rep. EQ). Park Ridge, IL: London House.

Jones, J. W. (1981). Dishonesty, Staff Burnout, and Employee Theft. Paper presented at the 89th Annual Conference of APA, Los Angeles, CA.

Jones, J. W. (1983). Dishonesty, Staff Burnout, and Unauthorized Work Break Extensions. (Tech. rep.#11). Park Ridge, IL: London House.

Jones, J. W. (1983). Personality Correlates of Theft and Drug Abuse among Job Applicants. (Tech. rep. 332). Park Ridge, IL: London House.

Jones, J. W. (1991). Basic Psychometric Properties of a Pre-employment Honesty Test: Reliability, Validity, and Fairness. In J. W. Jones (Ed.) Preemployment Honesty Testing, pp. 65-88.

Jones, J. W., & Terris, W. (1981). Predictive validations of a dishonesty test that measures theft proneness. Paper presented at the XVII Interamerican Congress of Psychology, Santo Domingo.

Jones, J. W., Joy, D. S., & Martin, S. L. (1990). A Multidimensional Approach for Selecting Child Care Workers. Psychological Reports, 67, pp. 543-553.

Jones, J. W., Joy, D. S., & Terris, W. (1991) Psychological Correlates of Illicit Drug Use among Job Applicants. In J. W. Jones (Ed.) Preemployment Honesty Testing, pp. 159-170.

Joy, D. S. & Raspenda, K. M. (1990). Effectiveness and Fairness of the Personnel Selection Inventory for a restaurant franchise. (PSI tech. rep. #89). Park Ridge, IL: London House.

Joy, D. S. (1991). Reducing Turnover through Personnel Selection. In J. W. Jones (Ed.) Preemployment Honesty Testing, pp. 195-219.

Joy, D. S., Frost, A. G. & Cook, M. S. Predicting counterproductive behaviors at a major metropolitan Transit Authority using the PSI. (Tech. rep. #63). Park Ridge, IL: London House.

Joy, D. S., Frost, A. G., & Boomhower, D. (1987). Using the Personnel Selection Inventory (PSI-7ST) to Predict Successful Performance of Employees in the Automotive Rental Industry. (Tech. rep. #55). Park Ridge, IL: London House.

Kobs, S. W. & Arvey, R. D. (1972). Distinguishing deviant and non-deviant nurses using the Personnel Reaction Blank. Unpublished manuscript.

Kochkin, S. (1981). Some by-products associated with screening with the Reid Report. (Tech. rep.). Chicago: United Human Resources Department.

Kochkin, S. (1987). Personality Correlates of a Measure of Honesty. Journal of Business Psychology, 1, pp. 236-247.

Krug, S. E. (1978). Further Evidence on 16PF Distortion Scales. Journal of Personality Assessment, 42(5), pp. 513-518.

Kurtines, W. M., Ball, L. R., & Wood, G. H. (1978). Personality characteristics of long term recovered alcoholics: A Comparative Analysis. Journal of Consulting and Clinical Psychology, 46, pp. 971-77.

Lasson, Elliot D. (1992). Preemployment Honesty Testing: construct validity issues and a test of the person-situation question. Unpublished doctoral dissertation, Wayne State University, Detroit.

Lavelli M. A. (1985). Psychological Predictors of Employees' Theft and Counterproductivity. Unpublished master's thesis, De Paul University, Chicago.

Lilenfied, S,, Andrews, B. P., Stone, E. F., & Stone, D. (1991). Examination of the construct validity of the Reid Report in prisoners. Paper presented at 3rd Annual conference of APS, Washington, DC.

LoBello, S. (1993). Fakability of a Commercially Produced Pre-employment Integrity Test. Unpublished manuscript.

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

London House, Inc. (1990). Adverse Impact Analysis Summaries for the Honesty Scale of the London House Personnel Selection Inventory. (Unpublished tech. rep. #85).

Maller, J. B. (1932). General and specific factors in character. Paper presented at the 10th International Congress of Psychology, Copenhagen.

Martelli, T. A. (1988, Aug.). Preemployment screening for honesty: The construct validity, criterion related validity and test retest reliability of a written integrity test. Unpublished doctoral dissertation, Ohio University.

Martin, S. (1992, 25 November). Personal communication.

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

Miner, J. B. & Capps, M. H. (1988). Using the polygraph to develop integrity test measures. (Tech. rep.). Intergram Inc..

Moore, R. W. & Stewart, R. M. (1989). Evaluating Employee Integrity: Moral and methodological problems. Employee Rights and Responsibilities Journal, 2(3), pp. 203-215.

Martin, D. M. & Terris, W. (1983). The Personnel Selection Inventory: A predictive validation study conducted for Bloomingdale's. (Tech. rep.). Park Ridge, IL: London House.

Murphy, K. R. & Lee, Sandra L. L. (1991). Does Conscientiousness Explain the Relationship between Integrity and Job Performance? Unpublished manuscript.

Nawn, J. (1992, 18 Dec.). Personal communication.

Needham, R., Orsolin, R. A., & Astor, S. D. (1986). Aims Ed. 3 Validation Studies. Downer's Grove, IL: Aims Review, Inc.

Nelson, G. J., Adams, D. R., Foster, S. F. & Farley, F. H. (1980). Socialization and empathy as predictors of moral judgment: Test of a model. Counseling and Values, 24, pp. 264-271.

Nolan, K. (1991). The Relationship between Socially Desirable Responding & Personnel Integrity Testing: The unique effect of self perception and impression management. (Master's thesis, Georgia State University).

Paajanen, G. (1991). Construct Validity. Unpublished manuscript.

Paajanen, G. (1993, 4 March). Personal communication.

Paajanen, G. E. (1986). Development and validation of the PDI Employment Inventory. Paper presented at the 94th Annual Convention of the American Psychological Association, Washington, DC.

Paajanen, G. E. (1986). Development and validation of the PDI Employment inventory. Paper presented at the 94th APA, Washington, DC.

Parker, J. P., Wiskoff, M. F., McDaniel, M. A., Zimmerman, R. A., & Sherman, F. (1989, April). The Development of a Prescreening Instrument for Maine Security Guards. Paper presented at the 4th Annual Conference of the Society for Industrial & Organizational Psychology.

PDI Inc. (1992). Employment Inventory Selection Impact Analysis. Unpublished manuscript.

Performance Improvement Corporation. (1988). Extended Validation Study abstract: Personality Plus indicator. (Unpublished tech. rep.).

Personnel Decisions, Inc. (1989). A concurrent validation study of the Employment Inventory for Eckerd Drug Company. (Confidential tech. rep.). Minneapolis: Author.

Personnel Decisions, Inc. (1990). Using the EI doesn't discriminate against older workers. EI Report - Newsletter for users of the PDI Employment Inventory.

Personnel Decisions, Inc. (1990). Validity of the PDI Employment Inventory for Drug Stores. (Tech. rep.). Minneapolis: Author

Personnel Decisions, Inc. (1990, May). Validity of the PDI Employment Inventory for Jack Eckerd Corporation. (Confidential Tech. Rep.). Minneapolis: Author.

Personnel Decisions, Inc. (1992). PDI Employment Inventory Selection Impact Analysis Report. Unpublished manuscript.

Personnel Decisions, Inc. (Deniz Ones) (1992). Construct Validity of PDI Employment Inventory: Qualitative Review of Trends across Studies. Unpublished manuscript.

Porter, D. B., Phil D., & Hamel, S. M. (1991). Personality Correlates of Aircraft Commanders' Perceived Effectiveness. Poster presented at 3rd Annual APS Conference, Washington, DC.

Psychological Surveys Corporation. (1986). Validity, reliability, compliance and efficacy studies of the R. A., A. I., E. S., D. A., and P. A. S. S. Surveys. (Unpublished tech. rep.). Oak Brook, IL: Author.

Rafilson, F. M. & Callans, M. C. (1992). Validity, classification, and response distortion in attitudinally based and criterion-keyed predictors of honest workplace behavior: A preliminary study. Unpublished manuscript.

Rafilson, F. M. & Frost, A. G. (1989). Overt integrity measures versus personality based measures of delinquency: An empirical comparison. Journal of Business & Psychology, 3, pp. 269-277.

Ryan, A.M. & Sackett, P. R. (1987). Pre-employment honesty testing: Fakability, Reactions of test takers and company image. Journal of Business and Psychology, 1(3), pp. 248-256.

Terris, W. & Jones, J. (1980). Attitudinal and personality correlates of theft among supermarket employees. Journal of Security Administration, pp. 1-4.

Terris, W. & Jones, J. W. (1984). Psychological Correlates of Employee Theft in Department Stores. (Tech. rep. #22). Park Ridge, IL: London House.

Terris, W. & Jones, J. W. (1984). Validation of the PSI with Managerial and Hourly Employees in a Fast Food Company. (Tech. rep. #22). Park Ridge, IL: London House.

Terris, W. (1979). Attitudinal Correlates of Theft, Violence, & Drug Use. Paper presented at the 17th Interamerican Congress of Psychology. (Tech. rep. #2). Park Ridge, IL: London House.

Terris, W. (1986). The development and validation of EPI-3. (EPI-3 tech. rep. #01). Park Ridge, IL: London House.

Upmanyu, V. V. & Kaur, K. (1986). Diagnostic utility of word association emotional indicators. Psychological Studies, 32(2), pp.71-78.

Urban, G. D. & McDaniel, M. A. (1990). Factor and Cluster Analysis of the Special Assessment Battery. Educational and Psychological Measurement, 50, pp. 663-671.

Werner, S., Jones, J. W., & Steffy, B. D. (1989). The relationship between intelligence, honesty and theft admission. Educational & Psychological Measurement, 49, pp. 921-927.

Werner, S. H. & Joy, D. S. (1990). The Relationship between PSI-7 Scores & a Measure of Basic Math Skills. (Tech. rep. #85). Park Ridge, IL: London House.

Werner, S. H., Jones, J. W. & Steffy, B. D. (1989). The relationship between intelligence, honesty and theft admissions. Educational and Psychological Measurement, 49, pp. 921-927.

Wiechman, D. J. & Bae, R. P. (1983). The Personnel Selection Inventory: An Empirical Investigation. (Tech. rep. #24). Park Ridge, IL: London House.

Wonderlic & Associates, Inc. (1988). Wonderlic Personnel Test: Performance norms summary. Northfield, IL: Authors.

Wonderlic & Associates, Inc. Personnel Test Manual. (Tech. rep.). Chicago: Author.

Woolley, R. M. & Hakstian, A. R. (in press). An examination of the construct validity of personality-based and overt measures of integrity. Educational & Psychological Measurement.

Woolley, R. M. (1990, March). An examination of the construct and criterion-related validity of overt and personality-oriented predictors or counterproductivity. Unpublished master's thesis, University of British Colombia, Canada.

Wright, P. M., Kacmar, K. M., McMahan, G. C., & Delauw, K. (1992). P=f (M x A): Cognitive ability as a moderator of the relationship between personality and job performance. Paper presented at Academy of Management Meeting, Las Vegas, NV.

Wuebker, L. J. (1987). The safety focus of control scale: A construct validation study. Unpublished manuscript.

Wuebker, L. J. (1988). Relationship of PSI-3S safety scale scores to Rotler's general locus of control measure. Unpublished manuscript.

REFERENCES

Alexander, R. A. , Carson, K. P. , Alliger, G. M., Cronshaw, S. F. (1989). Empirical distributions of range restricted SD_X in validity studies. Journal of Applied Psychology, 74, 253-258.

American Management Association (1977). Crimes against business project: Background and recommendations. New York, NY: AMA.

American Society for Personnel Administration (1988, June). Most employers test new job candidates, ASPA survey shows. Resource, Arlington, VA: Author.

Arnold, D. W. (1989). The Reid Report's lack of adverse impact (Research Memorandum #22). Chicago, IL: Reid Psychological Systems.

Arvey, R. D. & Bouchard, T. J. (in press). Genetics, Twins, and Organizational Behavior. In Staw, B. A. and Cummings, L. L. (Eds.). Research in Organizational Behavior, Vol. 16.

Ash, P., (1989). The construct of employee theft proneness. Rosemont, IL: SRA/London House, Inc.

Bacas, H. (1987). To stop a thief. Nation's Business, 75 (6), 16-23.

Bagus, K. K. (1988). Industry specific affirmative action studies. Stanton Survey Report. Charlotte, NC: Stanton Corporation.

Banks, C. (1948). Primary personality factors in women: A re analysis. British Journal of Statist. Sect., 1, 204-218.

Barrick, M. R. , & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 41, 1-26.

Blocklyn, Paul L. (1988, February). Consensus: Preemployment testing. Personnel, 66 - 68.

Bond, M. H. , Nakazato, H. S., & Shiraishi, D. (1975). Universality and distinctiveness in dimensions of Japanese person perception. Journal of Cross-Cultural Psychology, 6, 346-355.

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

Borofsky, G. L, Friedman, J., & Maddocks, A, Jr. (1986). User's Manual for the Employee Reliability Inventory. (Tech. rep.). Boston: Bay State Psychological Associates.

Bureau of National Affairs, Inc. (1988, May). Recruiting and selection procedures, personnel policy forum survey no. 146. Washington, DC: Author.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors, Psychometrika, 12, 197-220.

Cattell, R. B. (1948). The primary personality factors in women compared with those in men. British Journal of Psychology, 1, 114-130.

Cherrington, D. (1989). Validation of the Applicant Review: Analysis of adverse impact (Research Report). Standfordville, NY: CHC Forecast Inc.

Collins, J. M. , & Schmidt, F. L. (1992). Integrity , Personality, and White Collar Crime. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Montreal.

Collins, J. M., & Schmidt, F. L. (in press). Integrity, personality, and white-collar crime: A construct validity study. Personnel Psychology.

Conley, J. J. 1985. A personality theory of adulthood and aging. In R. Hogan & W. H. Jones (Eds.), Perspectives in Personality (Vol. 1, pp. 81-115).

Conoley, J. C., & Kramer, J. J. (Eds.). (1989). The Tenth Mental Measurements Yearbook. Lincoln, NE: Buros Institute of Mental Measurements.

Cormack, R. W., & Strand, A. L. (1970). Trustworthiness Attitude Survey. Oakbrook, IL: Personnel Systems Corporation.

Costa, P. T., & 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. & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, pp. 281-302.

Cunningham, M. R. & Ash, P. (1988). The structure of honest: Factor analysis of the Reid Report. Journal of Business and Psychology, 3(1), pp. 54-66.

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

Digman, J. M. & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis and comparison of six major studies. Multivariate Behavioral Research, 16, 146-170.

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

Eysenck, D. W. (1970). The structure of human personality. London: Methuen.

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

Funder, D. C. , & Ozer, D. J. (1983). Behavior as a function of the situation. Journal of Personality and Social Psychology, 44, 107-112.

Gaugler, B. B. , Rosenthal, D. B. ,Thornton, G. C. , Bentson, C. (1987). Meta-analysis of assessment center validity. Journal of Applied Psychology, 72, 493-511.

Goff, M. , & Ackerman, P. L. (1992). Personality-Intelligence Relations: Assessment of typical intellectual engagement. Journal of Educational Psychology, 4, 537-552.

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

Goldberg, L. R. (1990). An alternative "Description of Personality": The Big-Five Structure. Journal of Personality and Social Psychology, 59. 1216-1219.

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 A.P.A. task force report. Washington DC

Goldberg, L. R. (1993). The Structure of Phenotypic Personality Traits. American Psychologist, 48(1), pp.26-34.

Gough, H. G. (1948). A Sociological theory of psychopathy. American Journal of Sociology, 53, 359-366.

Gough, H. G. (1954). Personnel Reaction Blank. Palo Alto, CA : Consulting Psychologists Press.

Gough, H. G. (1954). Systematic validation of a test for delinquency. American Psychologist, 9, 381.

Gough, H. G. (1960). Theory and Measurement of Socialization. Journal of Consulting Psychology, 24, 23-30.

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

Gough, H. G., & Petersen, D. R. (1952). The identification and measurement of dispositional factors in crime and delinquency. Journal of Consulting Psychology, 16, 207-212.

Gould, S. J. (1981). The mismeasure of man. New York: Norton.

Guastello, S. J., & Rieke, M. L. (1991). A review and critique of honesty test research. Behavioral Sciences and the Law, 9, 501-523.

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

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

Hay, D. W. (1981). An exploration of the dimensionality of a personnel selection instrument designed to measure honesty. Unpublished master's thesis, Illinois Institute of Technology, Chicago.

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

Hogan, R. & Hogan, J. (1992). Hogan Personality Inventory Manual. Tulsa, OK: Hogan Assessment Systems.

Hogan, R. (1981) Hogan Personality Inventory Manual. Tulsa, OK: Hogan Assessment Systems.

Hough, L. M. ,Eaton, N. K. , Dunnette, M. D. , Kamp, J. D. , McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. Journal of Applied Psychology [Monograph], 75, 581-595.

Hunter, J. E. , & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, CA: Sage.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-95.

Inwald, R. E. (1988). Five-Year Follow-Up Study of Departmental Terminations as Predicted by 16 Pre-employment Psychological Indicators. Journal of Applied Psychology, 73, pp. 703-710.

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

Jones, J. W., & Terris, W. (1983). Personality correlates of theft and drug abuse among job applicants. Proceedings of the Third International Conference on the 16PF Test, pp. 85-94.

Jones, J. W. & Terris, W. (1984). Validity and factor structure of a dishonesty test that measures attitudes toward theft. (Tech. rep. #20). Park Ridge, IL: London House.

Klump, C. S. (1964). Stanton Survey. Charlotte, NC : Stanton Corporation.

Lasson, E. D. (1992). Preemployment Honesty Testing: Construct Validity Issues and a Test of the Person-Situation Question. Doctoral Dissertation, Wayne State University, Detroit, MI

London House Inc. (1975). Personnel Selection Inventory. Park Ridge, IL.: Author.

London House Inc. (1982). Employee Attitude Inventory. Park Ridge, IL. : Author.

Lorr, M. & Youniss, R. P. (1973). An inventory of interpersonal style. Journal of Personality Assessment, 37, 165-173.

Lousig-Nont, G. M. (1987). Phase II Profile. Las Vegas, NV: Author.

Martelli, T. A. (1988, Aug.). Preemployment screening for honesty: The construct validity, criterion related validity and test retest reliability of a written integrity test. Unpublished doctoral dissertation, Ohio University.

McCrae, R. R. (1989). Why I advocate the five factor model: Joint factor analyses of the NEO-PI with other instruments. In Buss DM, Cantor N (Eds.), Personality Psychology: Recent Trends and emerging directions. New York: Springer-Verlag.

McCrae, R. R. , & John, O. P. (1992). An introduction to the Five Factor model and its applications. Journal of Personality, 60, 175-216.

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

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

McDaniel, M. A. , Whetzel, D. L. , Schmidt, F. L. , Maurer, S. ,Hunter, J. E. (in press). The validity of employment interviews: A review and meta-analysis. Journal of Applied Psychology.

Michigan Department of Education. (1989). The Michigan Employability Survey.

Miller, J. F., & Bradley, P. (1975). Milby Profile. Minneapolis, MN. : Milby Systems Inc. Lousig-Nont, G. M. (1987). Phase II Profile. Las Vegas, NV. : Author.

Mischel, W. (1968). Personality and Assessment. New York: Wiley.

Moretti, D. M., & Terris, W. (1983). The Personnel Selection Inventory: A predictive validation study. Park Ridge, IL: London House Inc.

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

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, 574-582.

Norman, W. T. , & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. Journal of Personality and Social Psychology, 4, 681-691.

O'Bannon, R. M., Goldinger L. A., & Appleby, G. S. (1989). Honesty and Integrity Testing. Atlanta, GA: Applied Information Resources.

Ones, D. S. , Viswesvaran, C. , Schmidt, F. L. Construct validity evidence for integrity tests: Evidence from contrasted group studies. Unpublished manuscript.

Ones, D. S. , Viswesvaran, C. , Schmidt, F. L. Convergent and divergent validity evidence for the Big Five dimensions of personality: A meta-analysis across major personality inventories. Unpublished manuscript.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1992).

Ones, D. S., Viswesvaran, C., Schmidt, F. L. (in press). Meta-Analysis of Integrity Test Validities: Findings and Implications for Personnel Selection and Theories of Job Performance. Journal of Applied Psychology [Monograph].

Paajanen, G. (1985). PDI Employment Inventory. Minneapolis, MN. : Personnel Decisions, Inc.

Paajanen, G. E. (1985). Development and Validation of the PDI Employment Inventory, Minneapolis, MN: Personnel decisions Inc.

Paajanen, G. E. (1987). The prediction of counterproductive behavior by individual and organizational variables. Unpublished doctoral dissertation, University of Minnesota.

Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. Journal of Applied Psychology, 65, 373-406.

Reid Psychological Systems. (1951). Reid Report. Chicago, IL: Author.

Rosen, A., (1977). On the dimensionality of the California Psychological Inventory Socialization scale. Journal of Consulting and Clinical Psychology, 45, 583-591.

Rothstein, H. R. , Schmidt, F. L. , Erwin, F. W. , Owens, W. A. , Sparks, C. P. (1990). Biographical data in employment selection: Can validities be made generalizable? Journal of Applied Psychology, 75, 322-327.

Sackett, P. R., & Decker, P. J. (1979). Detection of deception in the employment context: A review and critical analysis. Personnel Psychology, 32, 487-506.

Sackett, P. R., & Harris, M. M. (1984). Honesty testing for personnel selection: A review and critique. Personnel Psychology, 37, 221-246.

Sackett, P. R., & Harris, M. M. (1985). Honesty testing for personnel selection: A review and critique. In H. J. Bernardin & D. A. Bownas (Eds.), Personality assessment in organizations. New York: Praeger.

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

Schlessinger, P. (April 1992). Personal communication.

Schmidt, F. L. , Hunter, J. E. , and Gast-Rosenberg, I. (1980). Validity generalization results for computer programmers. Journal of Applied Psychology, 65, 643-661.

Schmidt, F. L. , Ones, D. S. , & Hunter, J. E. (1992). Personnel Selection. Annual Review of Psychology, 43, 627-670.

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N. (1986). The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. Journal of Applied Psychology, 71, 432-439.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Shane, G. S. (1979). Further tests of the Schmidt-Hunter Bayesian validity generalization procedure. Personnel Psychology, 32, 257-281.

Schmidt, F. L., Hunter, J. E., & Caplan, J. R. (1981). Validity generalization results for two occupations in the petroleum industry. Journal of Applied Psychology, 66, pp.261-273.

Science Research Associates. (1983). Personnel Outlook Inventory. Park Ridge IL: Author.

Strand, A. L., & Strand, M. L. (1986). Validity, reliability, compliance, and efficacy studies of the T. A., A. I., E. S., D. A., and P. A. S. S. Surveys. Oak Brook, IL: Psychological Surveys Corporation.

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

Tupes, E. C. , & Christal, R. C. (1961). Recurrent personality factors based on trait ratings (Tech. Rep. No. ASD-TR-61-97). Lackland Air Force Base, TX: US Air Force.

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

Viswesvaran, C. , Ones, D. S. , & Schmidt, F. L. (1992, May). Appropriateness of self-reported criteria for integrity test validation: Evidence from beyond. Paper presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, Montreal, Canada.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self report. Journal of Personality & Social Psychology, 57, 120-128.

Werner, S. H., Jones, J. W., & Steffy, B. D. (1989). The relationship between intelligence, honesty, and theft admissions. Educational and Psychological Measurement, 49, 921-927.

Wiggins, J. S., , & Pincus, A. L. (1992). Personality: Structure and Assessment. Annual Review of Psychology, 43, 473-504.

Woolley, R. (1991). An examination of the construct and criterion-related validity of overt and personality oriented predictors of counterproductivity. Unpublished masters thesis, University of British Columbia, Vancouver, Canada.

Woolley, R. M., & Hakstian, A. R. (in press). An examination of the construct validity of personality-based and overt measures of integrity. Educational and Psychological Measurement.

Yang, K. , & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: the Chinese case. Journal of Personality and Social Psychology, 58, 1087-1095.

Zuckerman, M., Kuhlman, D. M., Camac, C. (1988). What lies beyond E and N factor? Factor analyses of scales believed to measure basic dimensions of personality. Journal of Personality and Social Psychology, 54, 96-107.