DETECTING INTENTIONAL RESPONSE DISTORTION ON MEASURES OF THE FIVE-FACTOR MODEL OF PERSONALITY: AN APPLICATION OF DIFFERENTIAL PERSON FUNCTIONING

A dissertation presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Charles A. Scherbaum Jr.

August 2003

This dissertation entitled

DETECTING INTENTIONAL RESPONSE DISTORTION ON MEASURES OF THE FIVE-FACTOR MODEL OF PERSONALITY: AN APPLICATION OF DIFFERENTIAL PERSON FUNCTIONING

BY

CHARLES A. SCHERBAUM, JR.

has been approved for

the Department of Psychology

and the College of Arts and Sciences by

Jeffrey B. Vancouver

Associate Professor of Psychology

Leslie A. Flemming

Dean, College of Arts and Sciences

SCHERBAUM, CHARLES A. Ph.D. August 2003. Psychology

Detecting intentional response distortion on measures of the five-factor model of personality: An application of differential person functioning (336 pp.)

Director of Dissertation: Jeffrey B. Vancouver

Despite the increasing use of personality inventories as a basis for personnel decisions, many users of these inventories are seriously concerned about the susceptibility to response distortion (i.e., faking). The basis for the concern is not just that response distortion is possible, but that it will cause deleterious effects to the validity and utility of personality inventories. Because of these concerns, several techniques have been developed to detect individuals who are responding dishonestly on personality inventories. Increasingly, these techniques are based on item response theory (Zickar, 2000).

This paper extends and tests an IRT-based technique called differential person functioning (DPF; Johanson & Alsmadi, 2002) to the detection of response distortion on measures of the five-factor model of personality. DPF is a technique that can be used to identify if the response for a given individual are different for different groups of items.

Two experimental studies were conducted to examine the accuracy and consistency of DPF with other response distortion detection techniques. In the first study, participants completed a personality inventory twice under different types of response instructions (e.g., respond honestly or dishonestly) to determine which of the items on the International Personality Item Pool could be distorted. In the second study, participants completed a personality inventory under one of three different types of response instructions to determine the accuracy of the DPF techniques in identifying individuals who are distorting their responses. The accuracy of the DPF techniques was compared to the accuracy of two response distortion scales and an IRT based person-fit index.

The results of the studies indicate that those who are instructed to respond dishonestly do respond to some items on the personality inventory differentially and that these individuals were detected by the response distortion detection methods, albeit at low levels of accuracy. It was also found that the differential person functioning techniques detected individuals who were responding dishonestly and did so at levels of accuracy comparable to other response distortion detection techniques. The discussion of these studies focuses on the implications of these findings and the differential person functioning techniques for personality assessment and personnel decisions.

Approved: Jeffrey B. Vancouver

Associate Professor of Psychology

Acknowledgments

There are several individuals I want to thank for their role in the successful completion of this dissertation. First, I would like to thank my dissertation director, advisor, and mentor, Jeff Vancouver, for his guidance, patience, and insight throughout this project and the past five years. I would also like to thank the other committee members, Paula Popovich, Keith Markman and Claudia Gonzalez-Vallejo, for their time and thoughtful comments. I would like to thank George Johanson for his collaboration and advice throughout all the stages of this project. Lastly, I would like to thank Karen Scherbaum for the understanding, support, and comments that she provided throughout the project.

Table of Contents

Page

Abstract 4

Acknowledgments 6

List of Tables 11

List of Figures 15

Introduction 18

Personality 21

The Use of Personality Measures in Personnel Decisions 24

Response Distortion 26

The Impact of Response Distortion on the Measurement of Personality in Employment Contexts 28

Frequency and magnitude of response distortion 29

Impact of response distortion on personality factor structures 30

Impact of response distortion on validity and personnel decisions 31

Summary 33

Assessing Response Distortion on Measure of Personality 34

Item Response Theory 36

Applying IRT to Personality Measurement 41

Utilizing IRT to Detect Response Distortion on Personality Measures 46

Differential item functioning 47

Person-fit measures 51

Limitations of Current IRT Approaches for Detecting Response Distortion 54

Differential Person Functioning 56

Present Studies 58

Study 1 59

Method 60

Participants 60

Design 61

Manipulations 61

Response instructions 61

Condition order 62

Measures 62

International Personality Item Pool 62

Demographic characteristics 64

Procedure 64

Results 65

Descriptive statistics 65

Manipulation checks 65

Tests for order effects 66

Tests for sex effects 67

Principle components analysis 69

Parameter estimation and model fit 74

Differential item functioning analysis 76

Study 1 Discussion 87

Study 2 89

Method 91

Participants 91
Design and Manipulations 92
Measures 93
International Personality Item Pool 93
Balanced Inventory of Desirable Responding 93
Marlowe-Crowne Social Desirability Scale 94
Demographic characteristics 95
Procedure 95
Results 96
Descriptive statistics 96
Manipulation check 96
Tests for sex effects 99
Principle components analysis 101
Parameter estimation and model fit 106
Differential person functioning 107
Classification decisions using the BIDR 109
Classification decisions using the person-fit index 110
Comparison of the classification consistency of DPF, the person-fit index, and the BIDR 112
Classification accuracy using DPF 116
Classification accuracy using the BIDR 118
Classification accuracy using the person-fit index 119
Comparison of the classification accuracy of DPF, the person-fit index, and the BIDR 120
Study 2 Discussion 121

General Discussion 126

Implications for Personality Assessment and Response Distortion Detection...... 126

Implications for Personnel Decisions 131

Potential Limitations and Future Directions 133

Conclusions 136

References 138

Appendix A: International Personality Item Pool 156

Appendix B: Experimental Protocol for Session 1 of Study 1 168

Appendix C: Experimental Protocol for Session 2 of Study 1 170

Appendix D: Balanced Inventory of Desirable Responding 172

Appendix E: Marlowe-Crowne Social Desirability Scale 173

Appendix F: Experimental Protocol for Study 2 175

List of Tables

Table *Page*

1. Internal Consistency Reliability for the Personality Factors and Sub-Facets on the IPIP from both Administrations of the Personality Inventory and Both Response Instruction Conditions 178
2. Means and Variances in the Honest and Desirable Response Conditions for Each Item using the Polytomously Scored Data 180
3. Means and Variances in the Honest and Desirable Response Instruction Conditions for Each Factor and Sub-Factor using the Polytomously Scored Data 189
4. Descriptive Statistics and Correlations between Factor Scores on Each Personality Factor 191
5. Eigenvalues from the Principle Components Analysis of the Observed Data from the Honest Response and Desirable Instruction Conditions 192
6. Eigenvalues from the Principle Components Analysis of the Observed Sub-Factor Scores from the Honest Response and Desirable Response Instruction Conditions, and the eigenvalues from the Parallel Analysis 193
7. Factor Loadings of the IPIP Factor Scores for the Honest Response and Desirable Response Instruction Conditions 194
8. Fit Statistics for the 2-PL Model in the Honest Response Instruction Condition 197
9. Fit Statistics for the 2-PL Model in the Desirable Response Instruction Condition 198
10. Fit Statistics for Samejima's Graded Response Model in the Honest Instruction Response Condition 199

11. Fit Statistics for Samejima's Graded Response Model in the Desirable Instruction Response Condition 200
12. Results of the Differential Functioning Item Analysis using the Mantel-Haenszel Procedure on the Dichotomously Scored Items 201
13. Results of the Differential Functioning Item Analysis using the Mantel Procedure on the Polytomously Scored Items 210
14. Linking Coefficients from the Mean-Sigma Method for the Dichotomously and Polytomously Scored Data 219
15. Results of the Differential Functioning Item Analysis using the DFIT Procedure on the Dichotomously Scored Items 220
16. Results of the Differential Functioning Item Analysis using the DFIT Procedure on the Polytomously Scored Items 229
17. Summary of the Decisions about Differential Item Functioning from the Observed Score and Latent Parameter Analyses 238
18. Internal Consistency Reliability for the Personality Factors and Sub-Facets on the IPIP for each Response Instruction Condition 247
19. Means and Variances in the Honest, Desirable, and No Response Instructions Conditions for Each Item using the Polytomously Scored Data 249
20. Means and Variances in the Honest, Desirable, and No Response Instruction Conditions for Each Factor and Sub-Factor using the Polytomously Scored Data . 259
21. Descriptive Statistics and Correlations for the Personality Factor Scores in the Honest, Desirable and No Response Instructions Conditions 261

22. Eigenvalues from the Principle Components Analysis of the Observed Data from the Honest Response, Desirable Response, and No Response Instructions Conditions 263

23. Eigenvalues from the Principle Components Analysis of the Observed Sub-Factor Scores from the Honest Response, and Desirable Response, and No Response Instructions Conditions, and the eigenvalues from the Parallel Analyses 264

24. Factor Loadings of the IPIP Factor Scores for the Honest Response and Desirable Response Instruction Conditions 265

25. Fit Statistics for the 2-PL Model in the Honest Response Condition...................... 269

26. Fit Statistics for the 2-PL Model in the Desirable Response Condition.................. 270

27. Fit Statistics for the 2-PL Model in the No Response Instructions Condition......... 271

28. Fit Statistics for Samejima's Graded Response Model in the Honest Response Condition.. 272

29. Fit Statistics for Samejima's Graded Response Model in the Desirable Response Condition.. 273

30. Fit Statistics for Samejima's Graded Response Model in the No Response Instructions Condition.. 274

31. Results of the Differential Peron Functioning Analysis using the Mantel-Haenszel Procedure on the Dichotomously Scored Items 275

32. Results of the Differential Person Functioning Analysis using the Mantel Procedure on the Polytomously Scored Items 286

33. Person-Fit Statistic, l_z, for Each Participant on Each Personality Factor 297

34. Results of the Classification Consistency Analysis for the DPF, the Person-Fit Index, and the BIDR across all Response Instruction Conditions 307

35. Results of the Classification Consistency Analysis for the DPF, the Person-Fit Index, and the BIDR in the Honest Response Instruction Condition................................ 308

36. Results of the Classification Consistency Analysis for the DPF, the Person-Fit Index, and the BIDR in the Desirable Response Instruction Condition............................ 309

37. Results of the Classification Consistency Analysis for the DPF, the Person-Fit Index, and the BIDR in the No Response Instructions Condition 310

38. Summary of the Decision Accuracy of Each Response Distortion Detection Technique and the Associated True Positive, True Negative, False Positive, and False Negative Rates.. 311

List of Figures

Figure *Page*

1. Item Characteristic Curve 312
2. Option Characteristic Curve 313
3. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Conscientiousness 314
4. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Extroversion 315
5. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Neuroticism 316
6. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Openness 317
7. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Agreeableness 318
8. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Conscientiousness 319
9. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Extroversion 320
10. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Neuroticism 321
11. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Openness 322

12. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Agreeableness 323

13. Distribution of the Standard Errors of the Item Difficulty Parameter in the Desirable Response Instructions Condition 324

14. Distribution of the Standard Errors of the Item Difficulty Parameter in the Honest Response Instructions Condition 325

15. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Conscientiousness 326

16. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Extroversion 327

17. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Neuroticism 328

18. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Openness 329

19. Modified Parallel Analysis Results for the Honest Response Instructions Condition for Agreeableness 330

20. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Conscientiousness 331

21. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Extroversion 332

22. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Neuroticism 333

23. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Openness 334

24. Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Agreeableness 335

25. Decision Theoretic Classification Table 336

Introduction

Industrial and organizational psychologists have long searched for non-cognitive measures that can adequately predict successful job performance (e.g., Bingham, 1922). Unfortunately, the various non-cognitive predictors (e.g., graphology) proposed over the years have been found to be wanting in terms of validity (Schmidt & Hunter, 1998). This is particularly problematic because cognitive predictors (e.g., cognitive ability, IQ), although largely valid, can exhibit adverse impact on certain demographic groups (e.g., females, African-Americans, etc.) that are protected from discrimination by U.S. employment laws (Guion, 1998; Hough & Oswald, 2000). Recently, a number of empirical and meta-analytic studies have established that some measures of five factor model of normal adult personality are a viable alternative that adequately predict job performance for a number of jobs and that demonstrate no demographic group differences (e.g., blacks vs. whites; Barrick & Mount, 1991; Salgado, 1997; Tett, Jackson, & Rothstein, 1991). In spite of the advantage of personality measures as predictors of job performance, organizations have been somewhat reluctant to heavily rely on measures of personality for personnel selection decisions.

To some degree, the use of personality measures in employee selection has been stymied because of managers and personnel specialists' concerns that these measures are susceptible to response distortion and it will negatively impact personnel decisions (Hough & Schneider, 1996; Ones & Viswesvaran, 1998; Zickar & Drasgow, 1996; Whyte, 1956). That is, applicants may distort their responses (i.e., fake) so that they portray themselves in a socially and organizationally desirable manner. The basis for the concern is not just that response distortion is possible, but that it will cause deleterious effects to the validity and utility of personality measures. Even more distressing is that response distortion may negatively influence who actually gets hired. For example, in situations where top-down selection procedures are used, individuals with artificially high scores on a personality inventory may be selected over individuals who actually possess high levels of the trait measured by the personality inventory (Hough, 1998; Rosse, Stecher, Miller, & Levin, 1998).

Because of these concerns, several techniques have been developed for managing response distortion and its potential effects on personnel decisions (Mount & Barrick, 1995). These techniques attempt to detect the individuals who are responding dishonestly and then utilize this information in the decision making process (Hough, 1998). However, the results of studies using these techniques have not lead to definitive conclusions or high levels of detection accuracy. Several authors have argued that the ambiguity concerning the detection of response distortion is, in part, a result of the methods used, as well as, an incorrect level of analysis (Zickar, 2001; Zickar & Drasgow, 1996). For example, much of the research examining response distortion has relied on total score differences between groups (e.g., applicants vs. incumbents). However, it is individuals that distort responses on particular items, not groups distorting total scores. These authors have advocated the use of item response theory (IRT) to examine response distortion on personality measures (Drasgow & Hulin, 1990; Zickar, 2000; Zickar & Drasgow, 1996). To date, the application of IRT to personality measurement and response distortion has been relatively small, but is quickly growing (Zickar, 2001).

The research on response distortion utilizing IRT has taken one of two approaches. The first is the differential item functioning (DIF) approach. DIF can determine which items are being distorted across groups of individuals. If there are differences between groups of individuals (e.g., applicants vs. incumbents) in their response patterns on a given item, there is evidence that response distortion is occurring in one of the groups. The second IRT method is the person-fit approach. Person-fit measures can determine the degree of consistency between the pattern of responses across the entire personality inventory and the estimate of the latent trait for an individual. This approach is a model-fit approach. If the person-fit measure for an individual indicates a high degree of inconsistency, it is taken as evidence of response distortion.

Despite the promise of both approaches, their success has been moderate. In part, this is due to limitations in the information each approach provides. The DIF approach provides information about items and groups, but very little about individuals. In an employment context, decision makers need information about each individual, not groups of individuals. The person-fit approach provides information about each individual, but provides little information about the items. Thus, in an employment context, decision makers know that a given individual has a response pattern that is inconsistent with their estimate of the trait from the whole inventory, but they do not know if this is a result of the individual distorting their responses on the items that can be distorted or from something else.

From the IRT perspective, one would ideally want a technique that combines information about the properties of each individual and the properties of each of the items. The combination of this information would allow for the determination of who is distorting their responses and how they are doing it. Recently, in the area of academic assessment, an IRT-based approach has been offered that combines both of these sources of information. This approach is called differential person functioning (DPF; Johanson & Alsmadi, 2002). DPF is a technique that can be used to identify if the responses for a given individual are different for different groups of items. For example, one could examine an individual's pattern of responses over the items that are verifiable versus those that are not verifiable.

The DPF approach has yet to be utilized to detect response distortion on measures of personality. Thus, the contribution of this paper is five fold. First, it extends the DPF approach to detecting response distortion on measures of personality. Second, this paper extends DPF to situations where the response options are polytomous (i.e., non-binary). Third, this paper examines the classification accuracy of current DPF techniques in identifying individuals who are responding dishonestly on personality inventories. Fourth, it compares the DPF techniques to more commonly used methods of detecting response distortion. Fifth, it assesses the degree to which response distortion can occur on a widely used measure of the five-factor model of personality (i.e., International Personality Item Pool; Goldberg, 1999). In the subsequent sections, the five-factor model of personality, the use of personality in personnel decisions, response distortion on measures of personality, the historical and current approaches to detecting response distortion, the differential person functioning approach, and the methods employed in the paper are discussed.

Personality

Personality has been an enduring concept in psychology (Hogan, 1990). At the heart of the interest in personality has been the belief that personality can be used to understand, explain, and predict behavior (Funder, 2001; Hogan & Roberts, 2001). The belief that personality is a useful means to understand and predict behavior has been met with skepticism (Mischel, 1968). However, the major criticisms of personality (e.g., the situation always determines behavior to a larger extent than personality) as a useful concept in psychology have been addressed (Funder, 2001) and personality is currently viewed as an important determinant of human behavior (Hogan & Roberts, 2001).

There are many different paradigms of personality (e.g., psychodynamic, behaviorist). One of the most enduring is the trait approach to personality. This approach to personality is rooted in the early work of Gordon Allport (Allport, 1937). In this paradigm, personality can be defined as an individual's unique interpersonal characteristics (MacKinnon, 1944). That is, personality is a set of enduring traits or dispositions that describe an individual. The question that many personality psychologists have debated over the yeas is the number of traits in that set. There are several taxonomies of traits that have been offered (e.g., Catell, 1965; Eysenck, 1970; Norman, 1963). However, many sub-fields of psychology have coalesced around a taxonomic structure consisting of five personality factors (i.e., traits; Digman, 1990; Hogan, 1990), but agreement with this taxonomy is not universal (Hough, 1992; 1997; John, 1989). This organization of personality is referred to as the big five or the five-factor model of personality.

The five-factor model was demonstrated empirically by Tupes and Christal (1961) and popularized by Norman (1963). However, the conceptual roots of the five-factor model are much older (e.g., McDougall, 1932). The five factors are conscientiousness, extroversion, neuroticism, agreeableness, and openness to new experiences (Costa &

McCrae, 1992; Digman, 1990; Norman, 1963). Each of these factors is described below using the descriptions provided by Costa and McCrae (1992).

Conscientiousness is a factor capturing competency or character. Individuals high in conscientiousness tend to follow rules, pay attention to detail, are self-disciplined, and hard working. Extroversion is primarily a social dimension. Individuals high in extroversion tend to be talkative, energetic, assertive, and active. Agreeableness is also primarily a social dimension. Individuals high in agreeableness are typically sympathetic, helpful, friendly, and straightforward. Neuroticism is an individual’s level of emotional stability. Individuals high in neuroticism tend to be less able to control their emotions and impulses, experience negative affect, and have difficulty handling stress. Openness is associated with curiosity, imagination, and preference for variety. Individuals high in openness tend to engage in novel stimuli and situations, demonstrate intellectual curiosity, and display aesthetic sensitivity.

A considerable body of evidence has developed demonstrating the robustness of the five-factor model of personality. The five factors emerge in different cultures (Bond, Nakazato, & Shiraishi, 1975), with different instruments (Costa & McCrae, 1988), with ratings from different sources (Norman, 1963), and in many different samples (Digman, 1990). Also, the five personality factors are not practically correlated with cognitive ability (i.e., intelligence; McCrae & Costa, 1987). However, it is interesting to note that there are small differences between men and women on some of these personality characteristics (Costa & McCrae, 1992). Thus, potential sex differences and their impact on the repose distortion detection techniques will be examined.

The Use of Personality Measures in Personnel Decisions

An enduring concern of I/O psychologists is the prediction of future states of organizationally relevant variables (e.g., performance, turnover) for current and potential organizational members, and then to make decisions based on these predictions (Campbell, 1990; Guion, 1998). As such, a wide array of individual difference variables have been identified that can be used as basis of these decisions. Personality has been one of these variables for quite some time (e.g., Bingham, 1922). However, mirroring the trends in the field of psychology as a whole, the view of personality as useful concept for understanding and predicting behavior has waxed and waned amongst personnel psychologists (Hogan & Roberts, 2001).

For many years, the consensus among personnel psychologists was that personality was a poor predictor of job performance and other organizationally relevant criteria (e.g., Ghiselli, 1973; Schmitt, Gooding, Noe, & Kirsch, 1984). Over the last 13 years, the consensus has come full circle (Hogan, 1990). The transformation of opinion is a result of the convergence on the five-factor model of personality and the results of meta-analyses demonstrating that measures of these five factors can adequately predict job performance.

A meta-analysis by Barrick and Mount (1991) found that the meta-analytic correlation with subjective measures of job performance (e.g., supervisor ratings) was 0.26 for consciousness, 0.14 for extroversion, 0.09 for neuroticism, 0.09 for agreeableness, and 0.04 for openness to new experience. For objective measures (e.g., production data), the correlations were in the same direction but smaller in magnitude. Additionally, the magnitude of the relationships varied across occupations. For example, the meta-analytic correlation between extroversion and job performance in sales occupations, across performance criteria, was 0.18. Barrick and Mount (1991) concluded that measures of the five-factor model of personality are valid predictors of performance.

A second meta-analysis published by Tett, Jackson and Rothstein (1991) arrived at conclusions similar to Barrick and Mount (1991). However, Tett , et al. (1991) reported more positive results in terms of the magnitude of the relationships between the five personality factors and job performance. More recently, two different meta-analyses by Salgado (1997; 1998) found that the magnitude of the predictive relationships between the five-factor model and job performance in European samples to be similar to the magnitude of the predictive relationships in U.S. and Canadian samples.

The results of these meta-analyses have been the primary basis for the rapid increase in the use of measures of five-factor model of personality to predict job performance. Although the magnitudes of meta-analytic relationships many seem small to many readers, one must consider that the magnitude of the relationship between job performance and one of the best predictors (i.e., cognitive ability) is only 0.51 for jobs where there is no previous work experience (Schmitt & Hunter, 1998). Further, because personality measures are practically unrelated to cognitive ability, they allow for increased predictive validity when used in conjunction with measures of cognitive ability. The meta-analytic multiple correlation between cognitive ability, conscientiousness and job performance is 0.60. This one of the highest predictive validities achieved by personnel psychologists.

Additionally, the use of personality can increase the fairness of organizational selection procedures (Sackett, Schmitt, Ellingson, & Kabin, 2001). This is because measures of the five-factor model of personality do not share many of the limitations associated with cognitive predictors. For instance, cognitive ability tests demonstrate significant difference in test scores between racial groups (e.g., blacks and whites). The use of predictors that demonstrate racial group differences can lead to legal challenges under a number of state and federal laws (e.g., Civil Right Act of 1964 and 1991). Personality measures, on the other hand, do not demonstrate racial group differences (Mount & Barrick, 1995). Thus, personality measures are one of the few predictors of job performance that possess a degree of validity and fairness (Sackett, et al., 2001).

In addition to job performance, measures of the five-factor model of personality have been found to predict a number of other organizationally relevant criteria, such as team processes (Barrick, Stewart, Neubert, & Mount, 1998), leadership (Judge, Bono, Ilies, & Gerhardt, 2002), motivation (Kanfer & Ackerman, 2000), and job satisfaction (Judge, Heller, & Mount, 2002). Although the use of personality as a predictor of organizationally relevant criteria has clear benefits to organizations, many potential users are greatly concerned that individuals can distort their responses on measures of personality (Zicker, 2001; Mount & Barrick, 1995).

Response Distortion

The potential of individuals to distort (i.e., fake) their response on self-report measures of personality has been a long-standing concern for psychologists and other users of personality inventories (Bass, 1957; Dunnette, McCartney, Carlson, & Kirchner, 1962; Ellis, 1946; Hough & Schneider, 1996; Zickar, 2001). In fact, this concern first appeared shortly after the development of the structured personality inventory (e.g., Kelley, Miles, & Terman, 1936). Formally defined, response distortion on personality inventories refers to a pattern of responding by an individual that inaccurately reflects the degree to which the individual possesses the personality characteristics assessed by the inventory (Stricker, 1963; Zerbe & Paulhus, 1987). In the literature on response distortion, several different terms are used to describe this phenomenon. This includes socially desirable responding, social desirability, dissimilitude, faking, and desirable response set. In this paper, the term response distortion will be used throughout.

Interestingly, there is not one type of response distortion, but two types (Paulhus, 1984; 1991; Zerbe & Paulhus, 1987). Paulhus and his colleagues have argued that response distortion includes both a self-deception and an impression management component. Self-deception is an unconscious tendency to see oneself favorably. It is exhibited as positively biased self-reports of personality. However, the individual actually believes the self-description to be true and can therefore be considered part of individual's personality (Hogan, 1991; Paulhus, 1991). Thus, this is an unintentional form of response distortion. Impression management, on the other hand, is an intentional form of response distortion. In this case, individuals are consciously and deliberately responding in a fashion that will create an overly favorable impression (Cronbach, 1946). Individuals engaging in impression management will exaggerate their positive qualities, understate or deny their negative qualities, or report possessing qualities that they actually do not possess (Leary & Kowalski, 1990). Individuals responding in this manner are aware that they are distorting their responses.

Although there have been a few studies that attempt to examine both types of response distortion (e.g., Barrick & Mount, 1996), the majority of the literature has restricted its focus to impression management (Mount & Barrick, 1995). This is not surprising given that the level of self-deception does not significantly differ between different response groups (i.e., applicants vs. incumbents; Paulhus, Bruce, & Trapnell, 1995), it is actually positively related to job performance (Barrick, & Mount, 1996), as well as, key personality variables (e.g., conscientiousness, Mount & Barrick, 1995). Thus, self-deception has not been considered a major problem for personnel decisions. Impression management, however, does show large differences between groups (e.g., applicants vs. incumbents) and it has not been shown to be related to job performance (Barrick & Mount, 1996, Viswesvaran & Ones, 1998). Because of these findings, impression management is considered a problem for personnel decisions. Thus, the focus of this paper is restricted to deliberate forms of response distortion (i.e., impression management).

The Impact of Response Distortion on the Measurement of Personality in Employment Contexts

Managers and personnel specialists have long cited the occurrence of intentional response distortion and the difficulties of managing it as reasons to avoid using personality measures as a basis for personnel decisions (Mount & Barrick, 1995; Zicker & Drasgow, 1996). The empirical work, to date, indicates that concerns about the occurrence of response distortion may be warranted, but there is little consensus about the extent of the problem. In particular, there is little agreement about the frequency and magnitude of response distortion, the impact of response distortion on the factor structure of personality inventories, and the effects of distortion on validity and personnel decisions (Hogan, 1991; Hough & Oswald, 2000; Hough & Schneider, 1996; Mount & Barrick, 1995; Schmitt & Chan, 1998). Each of these issues is considered below.

Frequency and magnitude of response distortion. One unambiguous finding of the research on response distortion is that individuals can deliberately distort their responses on standardized measures of personality (Ellis, 1946; Guion, 1998; Hogan, 1991; Hogan, Hogan, & Roberts, 1996; Hough & Schneider, 1996; Mount & Barrick, 1995; Viswesvaran & Ones, 1999; Waters, 1965). A number of studies have shown that individuals in research and employment settings can distort their responses when asked or are motivated to do so (e.g., Bass, 1957; Borislow, 1958; Dunnette, et al., 1962; Hough, et al., 1990; McFarland & Ryan, 2000).

What is less clear is how often individuals actually distort their responses (i.e., the prevalence rate) and to what degree (Mount & Barrick, 1995; Schmitt & Chan, 1998). The conclusions about the prevalence rate from studies examining response distortion in employment contexts are conflicting (Mount & Barrick, 1995). There are several authors who argue that the prevalence is unquestionably low (e.g., Costa, 1996; Hogan, 1991; Hogan, et al., 1996; Hough, 1998; Hough & Schneider, 1996; McCrae & Costa, 1983) and there are others that argue the evidence suggests the prevalence is not always low (Mount & Barrick, 1995; Schmitt & Chan, 1998). Surprisingly, several of the same studies are used as support for both positions (e.g., Dunnette, et al., 1962; Michaelis & Eysenck, 1971). A number of studies (e.g., Bass, 1957; Kirchner, Dunnette, & Mousley, 1960; Rosse, et al., 1998) have found that applicants have significantly higher scores on measures of personality than incumbents. Other researchers, however, have found little to no differences (e.g., Hough, 1998; Hough, et al., 1990; Orpen, 1971). Given the difficulty of determining the actual prevalence rate, the issue will not be resolved in the near future. Resolving this issue, however, may not be important because we know that individuals can and do respond dishonestly. Further, organizational users of personality measures are concerned about even a low frequency of dishonest responding (Rosse, et al., 1998).

There is also little agreement about the magnitude of response distortion. That is, how much do individuals raise their scores on personality inventories? In response distortion simulation studies (i.e., experimental studies), where individuals are asked to respond in a certain manner (e.g., fake good or respond honestly), the magnitude of response distortion is appreciable. Regardless of the study design (i.e., within or between subjects), responses on personality inventories can be elevated by at least half a standard deviation (Viswesvaran & Ones, 1999).

In employment contexts the results about the magnitude are less clear. There are several studies that have found or not found significant differences in the magnitude of the responses from applicants and incumbents. Again, the resolution of this issue may be of minor importance (Bartlett & Doorley, 1967). We know that individuals do distort their responses and can do so substantially (Mount & Barrick, 1995; Viswesvaran & Ones, 1999). Therefore, the interest is in determining if a given applicant has responded dishonestly, not if there are mean differences in the responses of applicants and incumbents.

Impact of response distortion on personality factor structures. A second issue with response distortion on measures of personality is whether dishonest responding destroys the factor structure underlying the personality inventory. More simply, does response distortion have a detrimental impact on construct validity? Again, the evidence is mixed (Ellingson, Smith, & Sackett, 2001). Several researchers have found that response distortion erodes the factor structure of personality inventories in both employment (Montag & Comrey, 1990; Schmit & Ryan, 1993) and experimental (Ellingson, Sackett, & Hough, 1999; Michaelis & Eysenck, 1971) contexts. For example, Schmit & Ryan (1993) found that different factor structures of the underlying personality traits emerge when using data collected from applicant and non-applicant samples. Other researchers have found the opposite to be the case in employment (Montag & Levin, 1994) and experimental (Paulhus, et al., 1995) contexts. More recently, Ellingson, et al. (2001) found that the factor structure was unaffected by response distortion in four large organizational data sets that included multiple measures of personality.

Evaluating the evidence presented on the stability or instability of the factor structure is difficult for several reasons. The most problematic is that response distortion was identified in these studies using highly fallible methods (e.g., response distortion scales; Ellingson, et al., 2001) or it was just assumed to exist because of differences in the samples used (e.g., applicants vs. incumbents vs. students; Schmit & Ryan, 1993). Further, for the factor structure to be dissolved there would need to be either a high prevalence rate of response distortion or lower prevalence rates with larger magnitudes of distortion. The occurrence of either is debated. At this point, the evidence indicates that response distortion has a small impact on the factor structure. This is advantageous for the present study because the IRT techniques described below require unidimensionality for each personality factor studied.

Impact of response distortion on validity and personnel decisions. The most contentious issue in the literature on response distortion has been the impact of dishonest responding on the criterion-related validity of personality inventories in predicting job performance (Mount & Barrick, 1995). Several authors have presented evidence that response distortion attenuates criterion-related validities (Dunnette, et al., 1962; Kluger, Reilly, & Russell, 1991; Zickar & Drasgow, 1996). Intuitively, this makes sense. Inflated scores on the predictor are not necessarily accompanied by a corresponding inflation on the criterion. Thus, the criterion-related validity would suffer.

Despite the intuitive appeal of the changing validity premise, the majority of the empirical and meta-analytic evidence suggests that this is not the case (Costa, 1996; Hough, 1997; 1998; Hough & Oswald, 2000; Hough & Schneider, 1996; Mount & Barrick, 1995; Schmitt & Chan, 1998). Several empirical studies have found that predictive validities are not attenuated by response distortion (e.g., Christiansen, Goffin, Johnston, & Rothstein, 1994; Barrick & Mount, 1996; Hough, et al., 1990). That is, the correlations between personality measures and job performance do not deteriorate in the presence of response distortion. Further, meta-analytic studies have found that response distortion has no meaningful impact on the predictive validity of personality measures (Ones & Viswesvaran, 1998; Ones, Viswesvaran, & Reiss, 1996).

However, as is elaborated below, the techniques used in the studies included in these meta-analyses to determine if individual were distorting their responses have questionable construct validity (Kroger & Turnbull, 1975; Paulhus, 1991; Stark, Chernyshenko, Chan, Lee, & Drasgow, 2001; Zickar & Robie, 1999). Specifically, a major criticism of the construct validity is that individuals' responses on these scales can be distorted (Kroger & Trunbull, 1975; Stark, et al., 2001). That is, individuals can fake the faking scale. Therefore, it is difficult to ascertain what is being measured: dishonest responding or poorly performed dishonest responding. Because these measures are the basis of the meta-analytic work, it is not clear whether response distortion does not impact criterion-related validity or that poorly performed response distortion does not impact criterion-related validity.

Additionally, there is evidence suggesting that response distortion can change the rank ordering of individuals in top-down selection, especially at the upper and lower ends of the applicant pool (e.g., Ellingson, et al., 1999; Hough, 1998; Rosse et al., 1998). Moreover, the correlation coefficient, for particular ranges of a bivariate distribution, may not be sensitive to changes in rank order (Drasgow & Kang, 1984; Zickar, Rosse, Levin & Hulin, 1997). Thus, from a decision theoretical framework, response distortion may impact which individuals are identified as true positives, true negatives, false positives, and false negatives even though it is not impacting the overall shape of the bivariate distribution (i.e., criterion-related validity). This situation can have a substantial impact on the usefulness of personality as a predictor of job performance.

Summary. Although there are several points of debate concerning response distortion on measures of personality, it is clear that it does happen and it can have an impact on the usefulness of personality as a predictor of future job performance. Thus, intentional response distortion continues to be a problem, in the eyes of organizational decision makers, associated with using personality measures as a basis for personnel decisions (Barrick & Mount, 1996). Clearly, this is an undesirable state of affairs given the advantages of using personality in personnel decisions (Hogan, 1991; Hough & Oswald, 2000; Mount & Barrick, 1995).

As noted above, the primary purpose of this paper is to present and test a response distortion detection technique (i.e., differential person functioning) that organizational decision makers and personnel specialists can employ to manage the potential problems that response distortion presents in personnel decisions by identifying the individuals who are responding dishonestly. However, there are a variety of other techniques that have also been developed to detect response distortion on measures of personality. The most notable are those based on special response distortion scales and those based on item response theory. Because this study will employ both of these methods in conjunction with differential person functioning, they are reviewed below.

Assessing Response Distortion on Measures of Personality

The most frequently used techniques to detect response distortion are scales that purportedly identify the individuals who are responding dishonestly. Interest in developing these scales is almost as old as the development of standardized self-report measures of personality (e.g., Hathaway & McKinley, 1943). Historically, researchers were interested in developing scales that could be embedded in larger inventories of abnormal personality (e.g., Minnesota Multiphasic Personality Inventory; MMPI). For instance, the MMPI has three different sub-scales (e.g., the F, L, and K scales) that were designed to detect dishonest responding. These scales contain items with response options that are unlikely to be true, very rare, or impossible (e.g., “I have never told a lie”, “I have never used profanity”). If an individual is responding honestly, they would endorse none or very few of the items on these scales. On the other hand, individuals who endorse unlikely responses on a large number of these items are identified as responding untruthfully.

These early developments in the measurement of abnormal personality were quickly adopted for use with normal adult personality (e.g., Crowne & Marlowe, 1960; Edwards, 1957). The scales to detect response distortion on measures of normal personality go by a variety of names, including social desirability scales, lie scales, detection scales, validity scales, impression management scales, or unlikely virtues scales. For the purpose of consistency, these scales will be referred to as response distortion scales in this paper.

In contrast to the response distortion scales used in the measurement of abnormal personality, the scales used for normal adult personality are not necessarily specific to or embedded within a given personality inventory. Therefore, they could be used with any measure of personality. Further, the items on these scales have a slightly more subtle flavor than those on abnormal personality inventories. For example, the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960) includes items such as, "I am sometimes irritated by people who ask favors of me" and "On occasion I have doubted my ability to succeed in life". As with the sub-scales for the MMPI, the scale user looks for individuals who score highly on these items.

The use of these scales became widespread very quickly (e.g., Edwards, 1957) and they are still included in many commercial personality inventories (e.g., 16PF, Hogan Personality Inventory). However, their use and effectiveness has been controversial. Some authors claim that these that scales are effective in identifying individuals who are responding dishonestly (e.g., Hogan, et al., 1996; Hough, 1998; McCrae & Costa, 1983; Ones & Viswesvaran, 1999). In fact, several experimental studies have found these scales to be useful (Hough, et al., 1990; McCrae & Costa, 1983). In a meta-analysis of response distortion, Viswesvaran and Ones (1998) found that the effect sizes of response distortion are largest on response distortion scales. They concluded that these scales are highly sensitive to response distortion and capture it very well.

However, other authors have argued that the usefulness and success of these scales in detecting response distortion is questionable (Kroger & Turnbull, 1975; Paulhus, 1991; Robie, Zickar, & Schmit, 2001; Zickar & Robie, 1999). Although there may be data supporting their usefulness, the data and scales are limited in fundamental ways. Most notable is that individuals' responses on these scales can be distorted (Kroger & Trunbull, 1975; Stark, Chernyshenko, Chan, Lee, & Drasgow, 2001). That is, individuals can fake the faking scale. Thus, it is difficult to ascertain who is being identified: the dishonest responders or those who are not very good at responding dishonestly. Further, respondents believe that it is easy to distort one's response to many of the items on these scales (e.g., Dwight & Alliger, 1997). Thus, the construct validity of the scales is debated (Stark, et al., 2001).

Recently, alternative techniques to detect response distortion have been presented (e.g., reaction time, item response theory, physiological measures). In terms of accuracy and practicality, the most promising of these detection techniques is item response theory (Zickar, 2001). Item response theory is a model-based (i.e., mathematical) approach to psychological and cognitive measurement (Hulin & Drasgow, 1990; Lord, 1980). Several authors have argued that item response theory has the potential to provide insight into response processes, especially when used with experimental manipulations (e.g., Thissen & Steinberg, 1988). Below, item response theory and its use with personality measurement are discussed.

Item Response Theory

Item response theory (IRT) is a psychometric approach to understanding the relationships between individual characteristics (e.g., ability, traits), item characteristics (e.g., difficulty), and the response patterns of individuals (Drasgow & Hulin, 1990; Lord, 1980). Specifically, IRT relates the characteristics of individuals and the properties of items on a test or inventory to the probability of a particular response to an item. This relationship is typically presented as an item characteristic curve (ICC). ICCs graph the relationship between the latent trait (labeled θ in IRT), underlying the responses on the instrument, and the probability of particular response to an item (labeled $P(\theta)$ in IRT). The ICCs in most IRT application take a 'S' shaped curve, such as the logistic or normal ogive (Crocker & Algina, 1986).

Figure 1 presents a generic representation of an ICC. In the figure, the latent trait, θ, is along the x-axis and the probability of a particular response, $P(\theta)$, is along the y-axis. The values of θ are normally distributed and expressed as z-scores. Thus, an individual with $\theta = 1$, has a value on the latent trait that is one standard deviation above the sample mean. For example, consider the case where IRT is being used to model the ICC between cognitive ability and the probability of correctly responding to an item on an aptitude test. Using Figure 1, we see that as the latent trait increases (i.e., moves to the right), the probability of correctly answering the item increases. For any level of θ, we can determine the $P(\theta)$ by finding the height of the curve at that particular value of θ.

The IRT model selected determines the exact shape of the ICC. These models differ based on the number of item parameters estimated and the scoring of the sample data (i.e., dichotomous or polytomous) used to estimate the parameters. For ease of exposition, we will first consider the models used with dichotomously scored data, and then discuss IRT models for polytomously scored data in the subsequent section. Dichotomously scored data usually assigns the value of '1' to the correct response and a value of ‘0’ to incorrect responses. For example on a four option multiple-choice question, one of the options is scored as ‘1’ and the remaining options are scored as ‘0’. However, with personality and attitudinal data, the data are scored such that a ‘1’ represents a higher level of the trait or attitude and a ‘0’ represents a lower level of the trait or attitude. When dichotomous data is used, the models are referred to as logistic models.

There are three possible parameters of the items that can be estimated in the logistic models. The first parameter is an item difficulty parameter. Logistic models that include only a difficulty parameter are call 1-PL or Rasch models. Mathematically, these model are expressed as,

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \qquad (1)$$

where $P_i(\theta)$ is the probability of an examinee with θ selecting the option(s) scored as ‘1’ on item i, b_i is the item difficulty parameter, and e is the mathematical constant 2.718.

In this model, the probability of selecting the option(s) scored as ‘1’ is a function of the level of the latent trait and the difficulty of the item. Specifically, the item difficulty parameter is the point on the θ scale where $P(\theta)$ is .50. Larger values of b_i indicate that the item is more difficult and smaller values of b_i indicate that the item is less difficult. Therefore, as b_i increases, the level of θ necessary for a 50% chance of selecting the option(s) scored as ‘1’ also increases. The difficulty parameter and estimate of the latent trait are on the same scale (i.e., z-scores). The value of b_i typically ranges between –2.0 and 2.0 (Hambleton, Swaminanthan, & Rogers, 1991). The value of b_i determines the location (i.e., left to right) of the ICC on x-axis.

The second parameter is an item discrimination parameter. This parameter represents how well an item discriminates between individuals with different level of θ. These models are called 2-PL models. Mathematically, the 2-PL model is expressed as,

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (2)$$

where $P_i(\theta)$ is the probability of an examinee with θ selecting the option(s) scored as '1' on item i, b_i is the item difficulty parameter, a_i is the item discrimination parameter, D is a scaling constant equal to 1.7 that makes the logistic function more similar to the normal ogive function, and e is the mathematical constant 2.718.

In this model, the probability of selecting the option(s) scored as '1' is a function of the level of the latent trait, the difficulty of the item, the ability of the item to discriminate between individuals with different levels of the latent trait. Larger values of a_i indicate that the item is better able to discriminate between individuals with different levels of θ and lower values indicate the item is less able to discriminate between individuals with different levels of θ. The value of a_i can theoretically range between $-\infty$ and $+\infty$ (Hambleton, et al., 1991). However, in practice, items with negative values of a_i are not used and rarely is the value of a_i greater than 2.0. The value of a_i determines the slope of the ICC.

It is highly unlikely that the P(θ) is zero for individuals with very low levels of the latent trait because of guessing behavior. Thus, the third parameter is a pseudo-guessing parameter. This parameter represents the impact of guessing behavior on the ICC. These models are called 3-PL models. Mathematically, the 3-PL model is expressed as,

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \tag{3}$$

where $P_i(\theta)$ is the probability of an examinee with θ selecting the option(s) scored as '1' on item *i*, b_i is the item difficulty parameter, a_i is the item discrimination parameter, D is a scaling constant equal to 1.7, c_i is the pseudo-guessing parameter, and e is the mathematical constant 2.718.

In this model, the probability of selecting the option(s) scored as '1' is a function of the level of the latent trait, the difficulty of the item, the ability of the item to discriminate between individuals with different levels of the latent trait, and the amount of guessing on the item. Larger values of c_i indicate a larger proportion of individuals with very low levels θ are selecting the option(s) scored as '1' due to chance. The value of c_i tends to be very small (i.e., $c_i < .25$). The value of c_i determines the non-zero lower bound for the ICC.

When utilizing IRT, the a, b and c parameters, depending on the model, must be estimated for each item from the sample of data. The most commonly used estimation procedures in IRT application are the maximum likelihood procedures (Hambleton, et al., 1991). Because IRT is a model-based approach, the model must fit the data for the resulting estimates to be of use.

The use of IRT models requires that two primary assumptions be made. The first is that the construct underlying the responses to the item is unidimensional. The second is local independence. That is, after controlling for the latent trait, there is no relationship between the item responses. If the unidimensional assumption is met, the local independence assumption is also met (Hambleton, et al., 1991). The degree to which the results of an IRT analysis are valid rests on the degree to which these assumptions are met.

Applying IRT to Personality Measurement

IRT has been widely used with the measurement of cognitive abilities and academic achievement. These applications have been well developed and established guidelines are available for a number of issues concerning the use of IRT in these situations (see Embretson & Reise, 2000, Lord, 1980, or Hambleton, et al., 1991 for reviews). The translation of work on the measurement of cognitive abilities and academic achievement to the measurement of personality is fairly straightforward in several aspects. For other aspects, the translation requires a number of additional considerations (Drasgow, 1982; Smith, 2001) and requires a brief elaboration before discussing the application of IRT to response distortion.

First, the choice of an IRT model can become increasingly complex with personality data. With cognitive abilities, items are typically scored dichotomously with one correct option and one or more incorrect options. In these instances, the 1, 2, or 3 parameter logistic models can be chosen. With personality data, the scoring of the options depends on the particular operationalization of personality. Thus, the IRT model chosen will depend on how personality is conceived and measured. Some measures of personality, such as Goldberg's (1992) Adjective Checklist, are scored dichotomously (e.g., describes me/does not describe me). For these measures, the 2-PL or 3-PL IRT models are the most appropriate (Zickar, 2001). Although the application of the 2-PL or 3-PL model is relatively clear-cut, the interpretation of the parameters, especially the difficulty parameter, is not well understood (Zickar & Ury, 2002).

Other operationalizations of personality use polytomously scored items. For items on these inventories, there are a number of response options for each item and no one response option is necessarily ‘correct’. The polytomous measures of personality differ, however, in whether or not the response options are ordered. With measures with ordered response options, the value of the latent trait is assumed to be higher for individuals who choose the higher numbered response options. Inventories based on the five-factor model of personality; such as the NEO-PI-R (Costa & McCrae, 1992) or IPIP (Goldberg, 1999) are examples of these types of instruments.

When the response options are ordered, graded response IRT models may be chosen. Graded response models start with the assumption that the value of the latent trait for an individual who chooses the first response option is smaller than the value of the latent trait of an individual who chooses the second response option. For example, an individual who selects the response option labeled ‘1’ on a item assessing extroversion is less extroverted than an individual who selected the response option labeled ‘2’. Because the level of analysis is the response option, these IRT models produce option characteristic curves (OCC) instead of ICCs.

In OCCs, the latent ability is related to the probability of selecting a particular option. Each response option will have a unique OCC. A generic set of OCCs is presented in Figure 2. In the figure, the latent trait, θ, is along the x-axis and the probability of selecting a particular response option, $P(\theta)$, is along the y-axis. The values of θ are normally distributed and expressed as z-scores. Thus, an individual with $\theta = 2$, has a value on the latent trait that is two standard deviation above the sample mean.

For example, consider the case where a graded response model is being used to model the OCCs between extroversion and the probability of selecting any of the 3 options on a personality inventory. Using Figure 2, we see that for option 1, the probability of selecting that option decreases as the latent trait increases (i.e., moves to the right). Thus, only individuals with low levels of the trait tend to select this option. For option 2, the probability of selecting that option at first increases then decreases as the latent trait increases (i.e., moves to the right). Individuals with moderate levels of the trait tend to select option 2. For option 3, the probability of selecting that option at increases as the latent trait increases (i.e., moves to the right). Therefore, individuals with high levels of the trait tend to select option 3. For any level of θ, we can determine the P(θ) for a given option by finding the height of the curve at that particular value of θ.

In graded response models (GRM), an option difficulty parameter and an option discrimination parameter are estimated from the data. In GRMs, the option difficulty parameter is called the threshold parameter. The option difficulty parameter and estimate of the latent trait are on the same scale (i.e., z-scores). Thus, like the item difficultly parameter of logistic models, the option difficulty parameter is the point on the θ scale where P(θ) is .50. That is, it is the location on the scale of the latent trait where there is a 50% chance that the option will be selected. Larger values indicate that only individual with high levels of the trait are likely to select the option. Smaller values indicate that the option is likely to be selected by individuals with lower levels of the trait.

This discrimination parameter represents how well an option discriminates between individuals with different level of θ. Larger values indicate that the option is better able to discriminate between individuals with different levels of θ and smaller values indicate the option is less able to discriminate between individuals with different levels of θ. In homogeneous GRMs, the discrimination parameters are constant across options within an item, while the difficulty parameters vary across the options.

Thus, an item can have certain options that are more difficult for individuals with lower levels of the latent trait, but the options will be equally discriminating between individuals with different level of the latent trait. It is important to note that non-homogenous GRM does not require that the option discrimination parameters be constant. GRMs also make the assumptions of unidimensionality and local independence.

The most widely used graded response model in personality assessment is Samejima's (1969) graded response model (Zicker, 2001). In Samejima's (1969) GRM, the number of response options is labeled m_i. Difficulty and discrimination parameters are estimated for $m_i - 1$ boundary response functions. Boundary response functions are the cumulative probability of selecting a response option equal to or higher than the current response option. Therefore, on a personality inventory with 5 response options, 4 boundary response functions will be estimated. Mathematically, the boundary response function is expressed as,

$$P^*_{ik}(\theta) = \frac{e^{Da_i(\theta - b_{k_i})}}{1 + e^{Da_i(\theta - b_{k_i})}} \tag{4}$$

where $P^*_{ik}(\theta)$ is the probability of an examinee with θ responding to option i on item k, b_{ik} is the option difficulty parameter, a_i is the option discrimination parameter, D is a scaling constant equal to 1.7, and e is the mathematical constant 2.718. Thus, the probability of selecting the option i is a function of the level of the latent trait, the difficulty of the option, the ability of the option to discriminate between individuals with different levels of the latent trait. However, the discrimination is constant across options.

From the boundary response functions, the option characteristic curves (OCCs) can be estimated. As noted above, an OCC describes the probability of selecting a given response option as a function of the latent trait. The probability of selecting a particular option, $P_{ik}(u_i = k|\theta)$, is determined by the boundary response functions. For example, on a three option item $P_{ik}(u_i = k|\theta)$ is as follows,

$$P_{ik} = \begin{cases} 1 - P^*_{i1}(u_i = k|\theta) & \text{if k =1,} \\ P^*_{i(k-1)}(u_i = k|\theta) & \text{if k = m, and} \\ P^*_{i(k-1)}(u_i = k-1|\theta) - P^*_{ik}(u_i = k|\theta) & \text{for all else,} \end{cases} \quad (5)$$

where the probability of selecting an option is a function of the conditional probability of responding above the threshold parameter for option *k*-1 minus the conditional probability of responding above the threshold parameter for option *k*. For the first option, $P_{ik}(u_i = k|\theta)$ is $1 - P^*_{i1}(u_i = k|\theta)$. For the third option, $P_{ik}(u_i = k|\theta)$ is $P^*_{i(k-1)}(u_i = k|\theta)$. Lastly, $P_{ik}(u_i = k|\theta)$ is $P^*_{i(k-1)}(u_i = k-1|\theta) - P^*_{ik}(u_i = k|\theta)$ for the second option. A generic OCC is presented in Figure 2.

There are no definitive guidelines, as to which IRT model (e.g., logistic or graded response) is best when using personality data (Zickar, 2001). It is interesting to note that some studies have found that the 2-PL models fit personality data better than the graded response models (e.g., Chernyshenko, Stark, Chan, Drasgow, & Williams, 2001). However, the choice of the model should ultimately be guided by what is appropriate given conceptual concerns and potential uses of the results (Zicker, 2001).

The second issue with using IRT with personality measures is that personality inventories are likely to violate the IRT assumption of unidimensionality to some degree.

Most personality inventories measure several dimensions of personality, contain several subscales, and are not completely orthogonal. The research discussed below has been able to fit unidimensional models to this type of data with some success, indicating that the models are somewhat robust to violations of unidimensionality (e.g., Trippe & Harvey, 3003; Reckase, 1979; Zickar, 2001).

The above discussion is by no means intended to be a complete treatment of the possible IRT models that can be applied to personality data or the issues involved with fitting IRT models to personality data. A more complete review of some of these issues can be found in Zickar (2001) or Chernyshenko, et al. (2001).

Utilizing IRT to Detect Response Distortion on Personality Measures

The IRT approach was offered as an alternative to response distortion scales that can meet the needs for accuracy and practicality (Drasgow & Guertler, 1987). There have been two conceptual perspectives and subsequent approaches to response distortion. The first conceptual perspective of response distortion is called the changing items paradigm. This perspective assumes that the latent trait does not change, but the perceptions of items differ between individuals who are responding dishonestly and those who are responding honestly. It is hypothesized that the changes in perceptions are likely related to expectations of consequences for endorsing particular options (Zickar, 2000).

Moreover, some items may be easier to distort than others. Therefore, the changes in perceptions may be related to transparency, content, or difficulty of the item. From this perspective, detecting options or items that are being dishonestly answered is analogous to detecting differential item functioning.

The second conceptual perspective on response distortion is called the changing person paradigm. This perspective assumes that on some items individuals respond as if their value of theta was increased by a certain amount. That is, the response pattern across items for some individuals is inconsistent with their estimated level of the latent trait as measured by the whole test (Drasgow & Hulin, 1990). Assuming that the model fits the data, a family of methods called person-fit or appropriateness measures can be applied to identify the specific individuals who are responding dishonestly. The differential item functioning and person-fit approaches are described below.

Differential item functioning. When using a test or inventory with different groups of respondents (e.g., applicants and incumbents), a situation that can arise is that the items on the test or inventory function differentially for each of the groups. That is, different groups of respondents (e.g., applicants and incumbents) may have different ICCs. Therefore, individuals from different groups with same level of the latent trait will have different probabilities of selecting particular options of the item (Camilli & Sheppard, 1994). This situation is called differential item functioning (DIF). For example, we could find that an item on a personality inventory assessing agreeableness may function differently for individuals with the exact same level of the trait, but who are instructed to respond honestly versus dishonestly.

There are several ways to determine if DIF is present (see Camilli & Sheppard, 1994 or Cohen, Kim & Baker, 1993, for reviews). To utilize any of the DIF techniques, one first needs 2 well-defined groups. In the response distortion literature, applicant and incumbent samples, or samples that are instructed to respond honestly or dishonestly are typically used. It is important to note that DIF analyses examine differences on an item across groups of respondents.

There are two general classes of methods to identify DIF. There are the techniques based on the observed scores and the techniques based on the estimates of the latent parameters (Millisap & Everson, 1993; Potenza & Dorans, 1995). Both classes of techniques assume that the trait underlying the responses is unidimensional. They differ in what is used as the estimate of the trait. Observed score methods use a total score on the items assessing the trait and latent parameter methods use the estimate of the latent trait that is derived from the IRT analysis.

The observed score methods use the actual responses of the individuals in the different groups to detect DIF. These methods are considered an approximation to the IRT procedures (Crocker & Algina, 1986). The general logic behind most of these methods is that there should be no relationship between group membership and the response to an item after controlling for the trait. There are a variety of specific statistical techniques (e.g., logistic regression, discriminant function analysis) that can be used to test if a relationship exists between group membership and an item response after controlling for the trait.

One of the most popular is the Mantel-Haenszel chi-square procedure (Camilli & Sheppard, 1994). The Mantel-Haenszel procedure (Mantel & Haenszel, 1959) and its extensions to the generalized Mantel-Haenszel procedure (Somes, 1986; Zwick, Donoghue, & Grima 1993) and the Mantel (1963) procedure examine the relationship between two variables in a 2 x K frequency table while stratifying on (i.e., controlling for) the level of a third variable. In these analyses, K represents the number of response options.

For example, we could examine the association between employment status (i.e., applicant vs. incumbent) and response (i.e., the option indicating a high level of the trait vs. the option indicating a low level of the trait) for a particular item from a personality inventory, while controlling for different levels of the particular personality trait. This relationship is assessed through an odds ratio. For the example above, we are testing if the odds ratio of correctly responding is different for applicants and incumbents with the same level of the trait. The odds ratio over all of the levels of the stratification variable is a measure of the effect size of the DIF.

Several studies have used the observed score methods, such as the Mantel-Haenszel procedures, to examine DIF on measures of personality. A study by Stark, et al. (2001) examined DIF for applicant and non-applicant samples on 4 scales from the 16PF personality inventory (Conn & Reike, 1994) using the Mantel-Haenszel procedure. Stark, et al. (2001) found that a large number of items demonstrated DIF according to this method. Stark, et al. also examined DIF on a response distortion scale that was included with the personality inventory. Interestingly, they found that three fourths of the items on the response distortion scale demonstrated DIF when comparing the applicant and non-applicant samples. These authors conclude that the observed score methods are a viable option for detecting response distortion, and, in fact, were able to detect DIF on the measure that was supposed to detect response distortion.

The latent parameter methods use the IRT item and ability parameters from a focal (e.g., applicants) and a reference (e.g., incumbents) group to detect DIF. When using these methods, item and ability parameters are separately estimated for the two groups and then the parameters are placed on the same metric. This can be done by equating the parameter estimates from the focal group to the scale of the referent group

(Drasgow & Hulin, 1990). Using the latent parameter methods, DIF can be identified in several ways (Cohen, Kim, & Wollack, 1998; Millisap & Everson, 1993; Potenza & Dorans, 1995). The general logic behind most of these methods is that there should be no significant differences in the item parameters, and hence, the ICC or OCC between the focal and reference groups if the item is not functioning differentially after controlling for the latent trait.

The item parameters for the focal and reference groups can be compared using chi-square tests (e.g., Lord's [1980] χ^2). If the parameters are not different for the focal and reference group, the chi-square will not be significant. Also, area and distance measures can be used to compare the expected scores (i.e., $\Sigma P(\theta)$*test length) for examinees at the same level of the latent trait from the focal and reference group (e.g., Raju, van der Linden, & Fleer, 1995). If the expected scores are the same at each level of the latent trait, the ICC or OCC should not be significantly different. There are several statistics that can be used to assess significant differences in area or distance, but typically a chi-square test is used. A commonly used distance measure is Raju, et al.'s (1995) differential functioning of items and tests (DFIT). DFIT is based on the expected score method described above. In the context of response distortion, significant differences on the latent parameter techniques are taken to mean that one of the groups is distorting their responses.

There have been several studies that have examined the ability of the latent parameter methods to identify which items and options are being distorted on personality inventories. For example, a study by Zickar and Robie (1999) examined DIF between groups of U.S. Army enlistees who were induced to fake good, who were given coaching on how to fake good, and those who were induced to respond honestly on three scales from the ABLE personality inventory (White, Nord, Mael, & Young, 1993). Samejima's (1969) graded response model (GRM) was fit to the personality data used in the study. Zickar and Robie (1999) found that a very small number of items were identified as differentially functioning. There was some consistency in which items demonstrated DIF when comparing each of the two faking groups to the group of honest respondents. Further, for the items that demonstrated DIF, the second difficulty estimate was lower in the groups of faking respondents. That is, the respondents in the honest group had a lower probability of selecting the second option (i.e., the higher theta option). In a similar study that compared applicant and non-applicant samples using a proprietary personality measure of personality, Robie, et al (2001) also found that very few items demonstrated DIF.

Person-fit measures. The general purpose of the person-fit techniques are to identify individuals whose response patterns are inconsistent with their estimated level of the latent trait, as measured by the whole test (Drasgow & Hulin, 1990). Assuming that the model fits the data, a family of methods called person-fit or appropriateness measures can be applied to identify the specific individuals whose response patterns are aberrant.

Although there are a number of person-fit measures (see Birenbaum, 1985; Drasgow, Levine, & McLaughlin, 1991; Levine & Drasgow, 1988 for reviews), the standardized likelihood function, l_z is generally considered the most accurate index, and is the index that is most often used. Mathematically, the l_z index is expressed as,

$$l_z = \frac{l_o - E(l_o)}{\sqrt{var(l_o)}} \qquad (6)$$

where l_o is the likelihood function as defined by Levine and Rubin (1979), $E(l_o)$ is the expected value of l_o, and $var(l_o)$ is the variance of l_o. Each can be computed using the following formulas,

$$l_o = \sum_{i=1}^{n} \{u_i \ln P_i(\hat{\theta}) + (1 - u_i)\ln[1 - P_i(\hat{\theta})]\} \tag{7}$$

$$E(l_o) = \sum_{i=1}^{n} \{[P_i(\hat{\theta}) \ln P_i(\hat{\theta})] + [1 - P_i(\hat{\theta})] \ln[1 - P_i(\hat{\theta})]\} \tag{8}$$

$$Var(l_o) = \left\{\sum_{i=1}^{n} P_i(\hat{\theta})[1 - P_i(\hat{\theta})]\{\ln P_i(\hat{\theta})/[1 - P_i(\hat{\theta})]\}\right\}^2 \tag{9}$$

where n is the number of items on the inventory, u_i is the response (0 or 1) of the individual to item *i*, $P_i(\hat{\theta})$ is the probability of the response on item *i* given the theta estimate. The response patterns of individuals with l_z values that are less than –2.0 are considered inconsistent with their theta estimate. Thus, it is deemed that these individuals are responding dishonestly.

The proponents of the l_z index have argued that l_z has approximately a standard normal distribution across all levels of theta (Drasgow, Levine, & Williams, 1985). Thus, it can be interpreted like a z-score. However, evidence from other studies indicates that this may not always be the case (Reise, 1995). One consideration with the use of l_z is that large sample sizes are needed. Another is that the accuracy of l_z decreases as the length of the test decreases. The test length can be problematic because personality inventories typically measure a number of personality factors, and thus each factor will have a limited number of items.

Several studies have examined the ability of the person-fit measures to detect response distortion. For example, Zickar and Drasgow (1996) presented a study in which they compared the ability of the l_z index and a response distortion scale to detect response distortion on the ABLE personality inventory (White, et al., 1993) in two samples of U.S. Army enlistees who were told to either answer honestly or fake good. In this study, Zickar and Drasgow (1996) varied the number of items that could be distorted and the false positive rate to determine the effects on the l_z index and the response distortion scale. Using a 2-PL model, they found that at low false positive rates the l_z index was better at detecting response distortion. As the false positive rate increased, the response distort scale became more accurate than the l_z index in detecting response distortion. The ratio of items that can be distorted to those that are distortion resistant also impacted the detection rates. When the number of items resistant to distortion was large, response distortion was difficult to detect. Based on these findings, Zickar and Drasgow concluded that for both sets of techniques the correct identification of individuals who were distorting their responses was too low to justify their use in selection contexts.

Robie, et al. (2001) conducted a study where applicant and incumbent samples were used to examine the ability of the l_z index to detect response distortion. Similar to Zickar and Drasgow (1996), Robie, et al. (2001) were able to identify a very small number of individuals as faking in both the applicant and incumbent samples (4.5% vs. 4.4%, respectively). However, their conclusions about the usefulness of person-fit measures were less pessimistic than that of Zickar and Drasgow (1996). Ferrando and Chico (2001) performed a study similar to Robie, et al. (2001) that manipulated the response instructions (i.e., respond honestly or dishonestly) with students in Spain. They found that the ability of the l_z index to detect response distortion was lower than the detection rate of a response distortion scale.

Despite the low detection rates reported in the Zickar and Drasgow (1996), Ferrando and Chico (2001), and the Robie, et al. (2001) studies, Schmitt, et al. (1999) found that the l_z index is correlated (i.e., $r = 0.26$) with test taking motivation. The importance of this finding is that test taking motivation has been found to be one of the reasons why response distortion occurs (Schmit & Ryan, 1992). Therefore, the l_z index may be useful, but the evidence is not overwhelming (Brown & Harvey, 2003).

Limitations of Current IRT Approaches for Detecting Response Distortion

Despite the potential advantages of the IRT approaches for detecting response distortion on personality measures, the empirical work, to date, is ambiguous in terms of the degree to which the potential advantages can be realized. As discussed above, the studies using these techniques have been able to identify respondents that are distorting their responses and items that are being distorted, but the percentage of items and individuals identified is much less than what one would expect it to be. I argue that this is, in part, due to some potential limitations of these approaches. Two of these limitations are discussed below.

First, the usable information that each approach provides, from personnel decision-makers' perspective, is limited. The DIF approach provides information about items and groups, but very little about individuals. In an employment context, decision makers need information about each individual, not groups of individuals. The person-fit approach provides information about each individual, but little information about the items. Thus, in employment context, decision makers know that a given individual has a response pattern that is inconsistent with their estimate of the trait from the whole inventory, but they do not know if this is a result of the individual distorting their responses on the items that can be distorted or from something else. Person-fit only indicates that, given the data, the estimated parameters are inappropriate The inconsistent response pattern may be a result of careless responding, incorrectly marking the answer sheet, incorrect data entry or something more substantial, such as response distortion. Although some possibilities may be ruled out (e.g., incorrect data entry), others cannot. Thus, the reason for a response pattern that is inconsistent with the estimate of the latent trait may never be determined. It can only be assumed. Also, these approaches have a number of operational requirements that are rarely met in organizational settings.

Second, researchers have been unable to link item content to aberrant responding (Zickar & Ury, 2002). That is, at this point, we are unable to predict which items or types of items are likely to be distorted. This is a result of some practical limitations facing researchers using IRT methods to detect response distortion. For example, item content is often confidential in employment contexts. In the instances where the item content is available (e.g., Stark, et al., 2000), there is no apparent relationship between the content features of items (e.g., item subtly, transparency) and the occurrence DIF. Although a specific relationship has not been found, it is apparent that content does play a role in which items are responded to dishonestly (Zickar & Drasgow, 1996). Clearly, more research is needed that attempts to link the content of the item to the functioning of the item.

From a personnel decision-maker's perspective, one would ideally want a technique that combines information on the properties of each individual and the properties of each of the items. The combination of this information would allow for the determination of who is distorting their responses and how they are doing it. Also, an approach that links the item content to the response patterns would be highly desirable. Recently, in the area of academic assessment, an approach has been offered that combines these sources of information and links item content to response patterns. This approach is called differential person functioning (Johanson & Alsmadi, 2002).

Differential Person Functioning

Differential person functioning (DPF; Johanson & Alsmadi, 2002) is a technique developed in the domain of cognitive and academic assessment. The purpose of this technique is to determine if there are different response patterns between groups of items for a given individual. That is, does an individual respond differently to different types of items (e.g., verifiable vs. non-verifiable items)? This approach combines information about individuals with information about the items and links the item content to the response patterns. Thus, this approach offers several advantages over the DIF and person-fit approaches for examining response distortion.

Conceptually, DPF is similar to DIF. Both methods attempt to determine if differential responding is present. With DIF, the focus is on how item responses differ over groups of individuals. That is, are there differences between applicants and incumbents in their responses to an item? With DPF, the focus on is how an individual's responses differ over groups of items. That is, are there differences in a given individual's responses over the items that can and cannot be distorted? In essence, DPF transposes the typical person-item matrix that is used in DIF to an item-person matrix. Instead of examining an ICC, a person characteristic curve is now examined. For example, one could examine how an individual's responses vary over items at the beginning or end of a personality inventory or items that can be distorted and cannot be distorted.

The analytical methods used to detect DIF can also be used to detect DPF. Particularly useful are the Mantel-Haenszel chi-square procedures described above (Alsmadi, & Alnabhan, 2000; Johanson & Alsmadi, 2002). However, in this case, the analyses are used to examine the association between the type of item and the response for a given individual after controlling for some property of the item (e.g., item difficulty). Thus, significant values on these tests indicate that an individual is responding differentially over the groups of items. These tests are performed for each individual. Further, the resulting effect sizes will be larger for DPF methods than for DIF methods. It is important to note that this technique will not work in situations where individuals are distorting all of the items to the same degree. However, because this situation is highly unlikely in most employment settings (McFarland & Ryan, 2000; Mount & Barrick, 1995), this limitation is a minor concern.

DPF is conceptually similar to the person-fit approach. Both attempt to identify the individuals for which the ability estimates produced by the IRT analysis may not be appropriate given the pattern of responding. However, the DPF approach provides information about groups of items, not the whole test, and because it is not a model fitting approach, it has less stringent restrictions. Moreover, DPF results are more interpretable than the results of the person-fit indices. Person-fit only indicates that, given the data, the estimated parameters are inappropriate. It does not provide information about why this is the case. As noted above, it may simply be careless responding or something more substantial, such as response distortion. The reason for the misfit is typically assumed. Because the DPF technique links properties of the individual with properties of the items (e.g., verifiability), the reasons for the aberrant response pattern are more transparent.

To date, only a few studies have utilized the DPF approach. Johanson and Alsmadi (2002) used DPF to examine differential responding of sixth grade students to items from different content domains within the mathematics section on the California Achievement Test. Johanson and Osborn (2000) used DPF to examine differential responding of individuals to positively and negatively keyed survey items. Matthews-López, Larkin, and Johanson (2002) used DPF to examine differential responding of raters to item presented during the first and second half of a standard-setting session for a teacher licensure exam. The DPF approach has yet to be applied to problem of response distortion on measures of personality. Nor, has the DPF techniques been used with polytomous item scoring.

Present Studies

The present studies are an attempt to extend the differential person functioning technique (DPF) to the detection of response distortion on measures of personality. To that end, two experimental studies were conducted where individuals completed personality inventories under different response instructions (e.g., respond honestly or dishonestly). The purpose of the first study was to determine which items on an inventory measuring the five-factor model of personality could be distorted using DIF techniques. Using the results from these analyses, items were identified as being capable of being distorted or not capable of being distorted. Identifying this property of the items is necessary to apply the DPF technique. The purpose of the second study was to utilize the DPF technique to determine its accuracy at identifying individuals who are distorting their responses on measures of personality. Further, the decision consistency and accuracy between the DPF technique and more traditional techniques (i.e., two response distortion scales and the person-fit index, l_z) was examined. Both studies are described in the subsequent sections.

Study 1

In the first study, the aim was to determine which items could be distorted on an inventory that measures the five-factor model of personality. In the first study, a repeated measures design was utilized. That is, the participants in the study completed the personality inventory twice. In one administration, the participants were asked to complete the inventory honestly. In the second administration, the participants were asked to complete the instrument in a manner that will make them look very desirable to a potential employer (i.e., fake good). Half of the participants received the honest response instructions first and the other half received the desirable response instructions first. The responses from the two administrations were compared using DIF techniques to determine which items on the inventory were subject to differential responding across the two experimental groups.

Both observed score and latent parameter methods were used to detect DIF. For the latent parameter methods, both graded response and logistic IRT models were used. Specifically, Samejima's (1969) graded response model and the 2-PL model were used. Once the item and ability parameters were estimated and equated, Raju, et al.'s (1995) DFIT technique was used. The observed score method utilized was the traditional Mantel-Haenszel and Mantel procedures. Based on the results of both the observed score and latent parameter analyses, the items were classified as can be distorted or cannot be distorted.

It is important to note that the use of repeated measures designs to examine response distortion has been criticized (e.g., Schwab, 1971; Skrzypek & Wiggins, 1966). These authors have argued that these designs fail to control for several threats to internal validity. However, many of these criticisms are related to methodological flaws present in the earlier studies (e.g., lack of counterbalancing the conditions), which have been eliminated in more recent studies (e.g., McFarland & Ryan, 2000). Further, response distortion studies utilizing within-subjects designs produce more accurate estimates of response distortion (Viswesvaran & Ones, 1999), and, if done properly, avoid potential instruction-by-subject interactions associated with between-subjects designs (Ones & Viswesvaran, 1998). Given the need in this study for high levels of precision and statistical power for group comparisons, the potential benefits of the within-subjects design outweighed the potential costs.

Method

Participants

During the first administration, 212 undergraduate students enrolled in introductory psychology courses volunteered to participate in exchange for course credit. Participants were recruited through the psychology subject pool. Of the 212 original participants, only 170 participants returned for the second administration of the personality. Additionally, the data from 16 participants were excluded from the analysis. Seven of these sixteen participants were excluded because of incorrectly marked identification numbers, which prevented matching the data from both administrations. The remaining 9 participants were excluded because they indicated that they received the same response instruction manipulation during both administrations. Thus, the final sample size for Study 1 was 154 participants in each response instruction condition.

Approximately two thirds of the final sample was female (67.8%) and one third were male (31.5%). The average age of the participants was 19.20 years with a standard deviation of 2.31 years. The overwhelming majority of the sample had held a job at some point in their life (98.1%).

Design

A 2 x 2 mixed design was utilized in this study. The response instruction manipulation was the within-subjects factor and the order of the instructions was the between-subjects factor.

Manipulations

Response instructions. The response instructions given to the participants were manipulated. Participants were instructed to either respond honestly or to respond in a manner that made them look very desirable to a potential employer. Each session was assigned to one of the response manipulation conditions. The order of the manipulations was counter-balanced in the first session. In the second session, the participants received the opposite set of instructions. The manipulation in each session was given to all the participants at once. The experimenter read the response instructions to the participants. The specific instructions for the honest condition were as follows,

> "Please complete this personality inventory as honestly as you can. The results will be completely anonymous and will be used for research purposes only. It is *very* important that you respond to this survey by describing yourself as you really are and not as you want to be or as you want others to see you."

The specific instructions for the desirable condition were as follows,

> "Please complete the personality inventory as if you were applying for a job you

> really want. To increase your chances of being hired, you should respond in ways that will make you look good to the organization. But, do not respond in a way that will look like you were obviously faking your responses."

The wording of this manipulation was designed so that it would induce a more realistic and sophisticated type of response distortion, instead of maximal response distortion (Paulhus, et al., 1995).

Condition order. The order of the initial instructions was manipulated between subjects. Each experimental session was assigned to the condition of honest instructions presented first or desirable instructions presented first. All the participants in the session were assigned to the condition. A random number generator was used to determine the order of the conditions. Each experimental session was assigned to the condition. Of the 155 participants, 84 received the honest response instructions first and 71 received the desirable response instructions first.

Measures

International Personality Item Pool. The International Personality Item Pool (IPIP; Goldberg, 1999) is a public domain measure of the five-factor model of normal adult personality. It was designed to be similar to a commercial instrument that measures the five-factor model of adult personality, the NEO-PI-R (Costa & McCrae, 1992). Specifically, the IPIP measures an individual's level of conscientiousness, extroversion, neurotism, openness, and agreeableness.

The specific IPIP inventory used was the NEO 30 Factor. This inventory contains 5 scales with one for each of the five personality factors. Each scale is comprised of six sub-scales measuring the sub-facets of each of the five factors. Each sub-scale contains

10 items. Thus, each factor is composed of 60 items and the inventory as a whole contains 300 items. This particular configuration of the scales and the sub-scales mimics the configuration of the NEO-PI-R (Costa & McCrae, 1992).

There is evidence that the scores on the IPIP demonstrate satisfactory psychometric properties (Goldberg, 1999; Trippe & Harvey, 2003). The average internal consistency across the subscales is high ($\alpha = .80$). In fact, scores on the IPIP demonstrate higher levels of reliability than scores on the NEO-PI-R. Also, there is evidence for the construct validity of the IPIP. The mean correlation between the sub-scales of the IPIP and the NEO-PI-R is .73. When corrected for unreliability, the mean correlation between the IPIP and the NEO-PI-R becomes .94. Additionally, there is evidence that only five factors are present in the inventory and the items load on the correct factor (Goldberg, 1999). Thus, the underlying factor of the IPIP is similar to the NEO-PI.

The instructions provided with the inventory ask respondents to indicate how well each item describes them. The items are rated on a 5-point scale with the anchors of “very inaccurate” at 1 and “very accurate” at 5. The items are contained in Appendix A. Several items on the inventory are reverse coded. A score for each of the five personality factors was computed by summing scores for the items that comprise that particular factor. Additionally, a score was computed for the sub-facets of each personality factor.

The scores on the IPIP demonstrated satisfactory reliability (i.e., α) in both administrations. In the first administration, the reliabilities at the factor level ranged from .95 to .86. In the second, the reliabilities ranged from .96 to .83. A complete list of the reliabilities for each factor and sub-facet for both administrations of the personality inventory and for each response instruction condition are presented in Table 1.

In IRT analyses requiring polytomously scored data, the responses to the IPIP, as provided by the participant on the original 5-point scale, were used. For the analyses requiring dichotomously scored data, the responses were recoded into a binary format. To dichotomize the data, the response options indicating high levels of the trait were coded as ‘1’ and the remaining options were coded as ‘0’. Specifically, the response options of 1, 2, and 3 were recoded as 0, and the response options 4 and 5 were recoded as 1. This recoding scheme is consistent with previous research using measures of the five-factor model of personality to examine response distortion (e.g., Reisse & Waller, 1993; Schmitt, Chan, Sacco, McFarland, & Jennings, 1999; Stark, et al., 2000; Zickar & Drasgow, 1996).

Demographic characteristics. The participants were asked their sex, age, and if they have ever held a job.

Procedure

Participants volunteered via sign-up sheets on the psychology subject pool board. The sessions were run in classrooms in Porter Hall with a maximum of 30 participants in each session. When the participants arrived, they were informed that this was a 2-session study in which they would need to return for a second session about 2 weeks later. If they were willing to return for a second session, they were asked to read and sign the consent form.

Once all the participants completed the consent form, they were provided with the IPIP inventory and an optically scanned answer sheet. The participants were instructed to mark all of their responses on the answer sheet. Next, the participants were asked to mark their sex, year of birth, and whether or not they had ever held a job on the answer sheet.

They were also instructed to make a mark on the response form that indicated the response instruction condition and the session number (i.e., 1st or 2nd). Then, they were asked to record the 7-digit id number on their answer sheets on the back of their experiment card. This identification number was used in both sessions of the study to match the individual's data from each session. The participants were then given the response instruction manipulation assigned to that session. After the instructions were presented, the participants were asked to begin completing the personality inventory.

When the participants completed the inventory for the first time, they were given one credit for participation, and were scheduled to return for the second administration of the personality inventory. All participants were given a reminder card with the date, time, and location of the second session. The second session was similar to the first in most respects. In the second session, the participants were given the opposite set of response instructions from their first session. When they completed the inventory, they were given their second credit and a debriefing form. Each session lasted between 40-50 minutes. The experimental protocols for the first and second session are presented in Appendix B and Appendix C.

Results

Descriptive statistics. Descriptive statistics were computed using the polytomous data. The mean and variance for each item in the honest response and desirable response conditions are presented in Table 2. The means and variances for the sub-factors are presented in Table 3. The descriptive statistics and correlations between the personality factor scores for each experimental session are presented in Table 4.

Manipulation checks. The effectiveness of the response instruction manipulation was examined before performing the DIF analyses. Specifically, mean differences in the scores of the personality factors between the honest response and desirable response conditions were examined. To test for these potential differences, a MANOVA was performed with the response instruction manipulation as the independent variable and the five personality factor scores as the dependent variables. If the manipulation led the participants to respond differently in each condition, there should be significant difference between the experimental groups on the scores for each personality factor. Specifically, the desirable responding condition should have significantly higher scores for the conscientiousness, extroversion, agreeableness, and openness factors, and significantly lower scores on the neuroticism factor than the honest responding condition. This would be a first indication that differential responding was occurring.

The overall multivariate effect was significant, Pillai's Trace = 0.217, $F(5, 304) = 16.847$, $p < .001$, partial $\eta^2 = .217$. The subsequent univariate F-tests for conscientiousness ($F[1, 308] = 61.937$, $p < .001$, partial $\eta^2 = 0.167$), extroversion ($F[1, 308] = 9.342$, $p < .002$, partial $\eta^2 = 0.029$), neuroticism ($F[1, 308] = 54.930$, $p < .001$, partial $\eta^2 = 0.151$), and agreeableness ($F[1, 308] = 12.725$, $p < .001$, partial $\eta^2 = 0.040$) were significant at $p < .05$. The univariate F-test for openness ($F[1, 308] = 3.306$, $p < .07$, partial $\eta^2 = 0.011$) was not. Further, the differences in the means between the response instruction conditions were all in the correct direction (see Table 3 for the means and variances). These results indicate that the response instruction manipulation did lead to differences in the participant's responses between the two administrations of the personality inventory.

Tests for order effects. Next, the impact of the order of the response instruction manipulation on the responses of the participants was examined. If the counter-balancing eliminated this threat to internal validity, no significant effect of the order of the manipulation on the scores for each of the personality factors should be present. A MANOVA was utilized with the order of the instructions entered as the independent variable and the five personality factor scores entered as the dependent variables.

Surprisingly, the overall multivariate effect was significant, Pillai's Trace = 0.054, $F(5, 304) = 3.458$, $p < .005$, partial $\eta^2 = 0.054$. Although significant, the effect accounted for very little variance, as indicated by the value of the partial η^2. Further, the subsequent univariate F-tests for conscientiousness ($M_1 = 226.09$, $M_2 = 224.59$, $F[1, 308] = 0.180$, $p < .672$, partial $\eta^2 = 0.001$), extroversion ($M_1 = 225.48$, $M_2 = 221.24$, $F[1, 308] = 2.372$, $p < .125$, partial $\eta^2 = 0.008$), neuroticism ($M_1 = 151.31$, $M_2 = 144.60$, $F[1, 308] = 2.924$, $p < .088$, partial $\eta^2 = 0.009$), openness ($M_1 = 207.27$, $M_2 = 210.07$, $F[1, 308] = 1.442$, $p < .231$, partial $\eta^2 = 0.005$), and agreeableness ($M_1 = 219.21$, $M_2 = 218.45$, $F[1, 308] = 0.076$, $p < .783$, partial $\eta^2 = 0.000$) were not significant at $p < .05$. Given the effect sizes of the multivariate and univariate tests, and the non-significance of the univariate tests, it appears the order of the experimental instructions did not have a meaningful effect on the responses of the participants.

Tests for sex effects. To determine if the sex of the participants was affecting the responses to the personality items, a MANOVA was performed. Sex was entered as the independent variable and the five personality factor scores as the dependent variables. These analyses were performed within each response instruction condition. If the sex of the participants was not exerting an influence on the responses, the overall multivariate effect and the effect from each of five univariate ANOVAs should not be significant.

However, there is some evidence that men and women do differ in their mean levels on some of the personality factors. Thus, significant results would not be completely surprising.

Indeed, there was a significant multivariate effect in the honest response condition, Pillai's trace = 0.194, $F(5, 142) = 6.82$, $p < .001$, partial $\eta^2 = 0.194$. The subsequent univariate F tests revealed that the men and women were significantly different on extroversion, $M_M = 207.83$, $M_F = 221.28$, $F(1, 146) = 8.61$, $p < .01$, partial $\eta^2 = 0.056$, and agreeableness, $M_M = 201.04$, $M_F = 217.17$, $F(1, 146) = 13.96$, $p < .001$, $\eta^2 = .087$. For both factors the women scored higher than the men. No significant differences were found for the other personality factors (Conscientiousness, $M_M = 207.09$, $M_F = 212.93$; Neuroticism, $M_M = 154.83$, $M_F = 163.96$; Openness, $M_M = 201.96$, $M_F = 207.02$).

Within the desirable response instruction condition, there was also a significant multivariate effect for sex, Pillai's trace = 0.237, $F(5, 142) = 8.84$, $p < .001$, partial $\eta^2 = 0.237$. The subsequent univariate F tests again revealed that the men and women were significantly different on extroversion, $M_M = 218.15$, $M_F = 232.32$, $F(1, 146) = 15.75$, $p < .001$, partial $\eta^2 = 0.097$, and agreeableness, $M_M = 210.89$, $M_F = 229.19$, $F(1, 146) = 25.24$, $p < .001$, partial $\eta^2 = 0.147$. Additionally, there was a significant difference on the conscientiousness factor, $M_M = 225.38$, $M_F = 243.51$, $F(1, 146) = 13.23$, $p < .001$, partial $\eta^2 = 0.083$. For all three factors the women scored higher than the men. No significant differences were found for the other personality factors (Neuroticism, $M_M = 139.77$, $M_F = 132.13$; Openness, $M_M = 208.00$, $M_F = 211.93$).

The results for agreeableness and extroversion in both response instruction conditions are not surprising given previous research demonstrating that women tend to score higher on these factors than men (Costa & McCrae, 1992). The difference between men and women on conscientiousness, on the other hand, was not expected. However, as was demonstrated here, the differences between men and women on measures of the five-factor model of personality tend to be small (Mount & Barrick, 1995). More importantly for this study, previous research has found that sex differences in the ability to respond desirably are very small (i.e., $d < .20$; Ones & Viswesvaran, 1998). Further, in this study, men and women were their own controls and within each sex there were significant multivariate effects for the response instruction manipulation. Moreover, the magnitude of change (i.e., η^2 associated with the multivariate effect) within each sex between the response instruction conditions was not dramatically different (men = .13 vs. women = .27). Thus, the sex differences found here do not complicate the interpretation of the differential item functioning analyses.

Principle components analysis. Before the analyses using IRT can be performed, it is first necessary to determine if the unidimensionality assumption was met. That is, there should be only one factor underlying the responses. However, others have argued that that a less stringent criterion is appropriate (Trippe & Harvey, 2003). That is, there should be only one dominant factor. In this case, the first factor should account for at least 20% of the variance for the item parameters to be stable (Reckase, 1979). To test this assumption, a principle components analysis was performed. If this assumption is met, then only five factors should exist in the data and the items should load on the correct factors.

Table 5 presents the first and second eigenvalue from the PCA using the sub-factor scores and the percent of variance explained. As can be seen in the Table, a dominant factor emerged for each of the personality factors within each response condition. The first component from the PCA analyses for conscientiousness, extroversion, and neuroticism in both response instruction conditions accounted for more than 20% of the variance, as recommended by Reckase (1979). In the honest response condition, agreeableness accounted for slightly less than 20%, but more than 20% in the desirable response condition. In both conditions, openness accounted for less than 20% of the variance. However, each personality factor did have a clear dominant first factor.

To determine the number factors that were present in the data, modified parallel analysis (MPA; Drasgow & Lissak, 1983) was used. MPA is an extension of Humphries and Montanelli's (1975) parallel analysis. In this procedure, the eigenvalues estimated from the observed data are compared to the eigenvalues from a synthetic dataset that satisfies the unidimensionality assumption. The synthetic dataset is created from the item parameters (i.e., item discrimination and difficulty) from the observed data set. If the second eigenvalue from the observed data is greater than the second eigenvalue from the synthetic data, multidimensionality is said to exist. MPA was performed using the SAS routine developed by R. J. Harvey (Trippe & Harvey, 2003).

The comparison of the eigenvalues resulting from a PCA on the observed data and the modified parallel analysis (MPA) for the responses from the desirable response instructions condition are presented in Figures 3 through 7 for conscientiousness, extroversion, neuroticism, openness, and agreeableness, respectively. In each of these figures, the unidimensionality criterion is not met. The second eigenvalue from the observed data is greater than the second eigenvalue from the simulated data. The comparison of the eigenvalues resulting from the PCA and the MPA using the responses from the honest response instructions condition are presented in Figures 8 through 12 for conscientiousness, extroversion, neuroticism, openness, and agreeableness, respectively. Again, the unidimensionality criterion is not met. The second eigenvalue from the observed data is greater than the second eigenvalue from the simulated data. Although, each personality trait possess one clear dominant factor and each accounts for more than 20% of the variance, the results of the MPA indicate that strict unidimensionality is not observed.

When strict unidimensionality is not met, the question becomes “how much can this assumption be violated and still produce stable parameter estimates”? Reckase (1979) has presented evidence suggesting that when a dominant first factor is present, IRT models estimate the first factor. Also, the size of the first factor is related to model fit. As the size of the first factor increases, so does model fit (Reckase, 1979). Further, Parson and colleagues (Drasgow & Parsons, 1983; Parsons & Hulin, 1982) have presented evidence that logistic IRT models are surprisingly robust to substantial violations of the unidimensionality assumption. Similar evidence has been presented for graded response IRT models (e.g., Kirisci, Hsu, & Yu, 2001; Reise & Yu, 1990) for test with over 40 items. Given, these findings and the moderate degree to which this assumption is violated, the parameter estimation is unlikely to be compromised. Moreover, the dimensionality of the data seen in this analysis is comparable to studies with larger sample sizes examining the IPIP with IRT (e.g., Trippe & Harvey, 2003).

Additionally, the principle component analysis allows the stability of the factor structure and loadings to be examined. To examine the stability of the factor structure, a PCA with a promax rotation was performed for each of the response instruction groups.

A promax rotation was used because the five personality traits are not completely orthogonal (Costa & McCrae, 1992). To mimic, the analysis performed by Schmit & Ryan (1993), the analysis was performed using the scores for the sub-factors of each personality trait. If the factor structure is stable, the resulting factor structure should be the same for each group, contain only five factors, and the items should load on the correct factor.

The number of components in the dataset was determined using parallel analysis (Horn, 1965). Parallel analysis is similar to MPA. However, in parallel analysis the item and person parameters are not taken into account when creating the synthetic data set, as is done in MPA. Kaufman and Dunlap's (2000) parallel analysis program was used to perform the parallel analysis. This program computes a random dataset with the same number of items and persons as the observed dataset. It then performs a factor analysis on the random dataset to determine the size of the eigenvalues that would occur due to chance. The observed eigenvalues are compared to the eigenvalues from the random data set. If the observed eigenvalue is larger than the random eigenvalue, it is consider a real factor. If the observed eigenvalue is less than the random eigenvalue, it is not considered to be a real factor. As can be seen in Table 6, only five factors exists in the dataset. The first 5 factors are all larger than the eigenvalue due to chance. The sixth factor is not.

The factor loading matrices were examined to determine if different factor structures emerged in the different response instruction conditions. As can be seen in Table 7, the rotated factor structure matrix for the honest response condition conforms reasonably well to a five-factor model. The majority of the sub-scales load on the correct factor. The exceptions are the fourth sub-scale of extroversion and the third sub-scale of openness. The fourth sub-scale of extroversion is labeled “activity”. The activity items loaded on the factor for conscientiousness. The third sub-scale of openness is label “feelings”. These items loaded on the factor for neuroticism. However, this pattern of loading is not abnormal. In factor loading matrix presented by Costa & McCrae (1992) for the NEO-PI-R, the loadings for the feelings items on neuroticism and the activity items on conscientiousness were both greater than 0.30.

Interestingly, the pattern of factor loadings for the desirable response condition was scattered (see Table 7). Although there were five factors, conscientiousness and neuroticism loaded on the same factor, albeit in different directions. Thus, there were really only four factors in the desirable responding condition. This finding could be interpreted as the opposite of Schmit & Ryan (1993). Instead of an expanded factor structure, a compressed structure emerged.

Additionally, two of the sub-scales from consciousness, extroversion, and openness loaded on wrong factor. For consciousness, sub-factors number two (order) loaded on the same factor as openness and number three (dutifulness) loaded on the same factor as agreeableness. For extroversion, sub-factors number three (assertiveness) and number four (activity) loaded on the same factor as conscientiousness. For openness, sub-factors number four (actions) and five (ideas) also loaded on the same factor as conscientiousness. Again, the pattern of loadings for activity is consistent with evidence presented by Costa & McCrae (1992). Also, the loadings for neuroticism sub-factors 3-6 (assertiveness, ideas, and dutifulness) are in line with the evidence presented by Costa and McCrae (1992).

In summary, it appears there are clear differences in the factor structure between the honest responding and desirable responding conditions. However, the presence of desirable response condition did not completely destroy the five factor structure. Only the neuroticism factor was compromised. For the other factors, the majority of the sub-scales loaded on the correct factor.

Parameter estimation and model fit. The MULTILOG program (version 7; Theissen, 2003) was used to compute the marginal maximum likelihood estimates of the item parameters and the expected a posteriori estimates of the theta parameters in Samejima's (1969) graded response model. The BILOG program (version 2; Mislevy & Bock, 1991) was used to compute the marginal maximum likelihood estimates of the item parameters and the expected a posteriori estimates of the theta parameters in the 2-PL model. For both models, the parameters were estimated for each of the response instruction conditions and each of the five personality factors separately. Thus, ten sets of item and person parameters were estimated for each IRT model. The program defaults were used in all analyses. The results of the IRT analyses for the polytomously and dichotomously scored items for each individual and each item are available upon request.

There are numerous criteria that can be used to establish the fit of an IRT model to the data (Hambleton, et al., 1991). Model fit was assessed using the procedures recommended by Drasgow, Levine, Tsien, Williams, and Mead (1995). Specifically, Drasgow, et al.'s chi-square statistics were used. These tests are indices of the difference between the expected frequency of responses for the options (derived from the ICC or OCC) and the observed frequency of responses for the options. Chi-squares are computed for single items (singles), sets of two items (doubles), and sets of three items (triples).

For the singles, the chi-square statistics are computed based on the expected frequency of responses for option *m*. For the doubles, the chi-square statistics are computed from the expected and observed probabilities for selecting particular options from two different items. For the triples, the chi-square statistics are computed from the expected and observed probabilities for selecting particular options from three different items. Large values for the chi-square statistics indicate poor model fit to the data. Drasgow, et al. recommends that χ^2 values greater than 3 be viewed as indicating poor fit. The computation of the doubles and triples allows for the examination of the potential interactions between the items. Moreover, they are more sensitive to misfit than single items. Examining the interactions is especially important when the test or inventory many not demonstrate strict unidimensionality. This analysis was performed using the MODFIT program.

The results of the χ^2 tests of fit for the 2-PL model are presented in Table 8 and Table 9 for the honest and desirable response conditions, respectively. The Tables displays the frequency of the chi-square values for the singles, doubles, and triples. As can be seen in the table, the fit was modest in both conditions. There were several χ^2 values that were above 3 for the singles, doubles and triples. The results of the χ^2 tests for Samejima's graded response model are presented in Table 10 for the honest response condition and in Table 11 for the desirable response condition. As can be seen in the table, the fit for the honest response condition was quite good. The majority of the χ^2 values were below 3 for the singles, doubles and triples. The fit for the desirable response condition on the other hand was modest. There were several χ^2 values above 3 for the singles, doubles and triples. Thus, there is a modest degree of confidence in the stability of the parameters from the 2-PL model. The degree of confidence for the graded response model is higher, but caution in the interpretation of the parameters is warranted.

Additionally, the results of the factor analyses provide some evidence about the fit of the model to the data (Hambleton, et al., 1991). Because only a limited number of techniques were used to establish model fit, and the same sample was used to calibrate the parameters and evaluate DIF, the results should be view as preliminary.

Differential item functioning analyses. To determine which items were functioning differentially across the honest and desirable response instruction conditions, differential item functioning (DIF) analyses were performed. DIF was first examined using the observed score methods described above. Specifically, the Mantel-Haenszel procedure and the Mantel procedure were utilized to identify items that were functioning differentially across the experimental groups.

As noted above, the Mantel-Haenszel and Mantel procedures examine the relationship between two variables in a 2 x K frequency table (K = the number of response options) while controlling for the level of a third variable. The relationship is assessed through an odds ratio and the odds ratio over all of the levels of the stratification variable is a measure of the effect size.

The effect size can be used to identify the degree of DIF. The Educational Testing Service has identified three categories of DIF based on the odds ratio and the significance of the chi-square (Camilli & Sheppard, 1994). For each category, the odds ratio is transformed into the Educational Testing Service's delta scale as follows,

$$D = -2.35[\ln(\Omega_{MH})] \tag{10}$$

where ln is the natural log and Ω_{MH} is the odds ratio from the Mantel-Haenszel procedure. Dorans and Holland (1993) call this estimate of DIF the MH D-DIF. Items that have a non-significant chi-square and an absolute value of D that is less than one are called “A items”. These items are not considered a problem in that they show no DIF. Items that have a significant chi-square and the absolute value of D that is greater than 1 but less than 1.5 are called “B items”. These items show a small to moderate degree of DIF. Items that have a significant chi-square and an absolute value of D that is greater than 1.5 are called “C items”. These items demonstrate a large degree of DIF. For the purposes of classification in this study, items falling into category B or C were identified as functioning differentially.

With the dichotomously scored data, the traditional Mantel-Haenszel procedure (Mantel & Haenszel, 1959) was performed. The Mantel-Haenszel procedure is approximately distributed as a chi-square with one degree of freedom. The null hypothesis is that the common odds ratio is 1.0. Values significantly greater than 1.0 indicate that the group coded ‘1’ has a higher odds of success. Significant values are interpreted as evidence of DIF. For example, if job incumbents were coded ‘0’ and job applicants were coded ‘1’ and a common odds ratio significantly greater than 1.0 was found, the conclusion would be that the odds of applicants successfully answering the item are greater than the odds of the incumbents, even after controlling for the level of the trait (Holland & Thayer, 1988). Thus, the item functions differently for individuals with the same level of the latent trait, but are in different groups. Common odds ratio significantly less than 1.0 indicate that the group coded ‘1’ has a lower odds of success. Using the example above, applicants, relative to incumbents, would have lower odds of responding successfully.

To perform the Mantel-Haenszel test with these data, a 2 x 2 frequency table was used. The data in this analysis were scored dichotomously (i.e., 0, 1) and the groups are coded ‘0’ for the reference group (e.g., honest responders) and ‘1’ for the focal group (e.g., desirable responders). The stratification variable in each analysis was the composite factor score (i.e., sum of the items) for the particular personality factor. For example, only the composite (i.e., total score) for the extroversion items was used in the analysis for extroversion. It is likely that using every observed value for the total factor score for the stratification variable will result in overly small sample sizes within each level of the stratification variable. In this event, the levels of the stratification variable are collapsed (i.e., thickened) to produce larger sample sizes within each level of the stratification variable. In this study, the stratification variable (i.e., total personality factor score) was collapsed to produce samples sizes of at least 5 in each response instruction condition at each level of the stratification variable. This resulted in six categories of personality factor scores. That is, the total score on each personality factor was collapsed into six categories. These categories were used as the stratification variable in both the Mantel-Haenszel and Mantel procedures described below.

The results of the Mantel-Haenszel procedure are presented in Table 12. As can be seen in the Table, 66 of the 300 items (22%) displayed differential functioning using the B item and C item categories of Educational Testing Service’s DIF classification. The odds ratios for 41 of the 66 differentially functioning items indicated that the odds for the desirable response condition of endorsing the option associated with higher levels of the trait were greater than the odds for the honest response condition. It is important to note that for the items measuring neuroticism, the odds ratio is smaller for the desirable response condition because of the scoring of the trait is in the opposite direction of the other four traits. At least five items from each factor were identified as functioning differentially. The greatest number of differentially functioning items came from the openness factor (12 items) and the least came from the conscientiousness factor (5 items).

For the remaining 25 items, the odds of the desirable response condition endorsing the option associated with higher levels of the trait were less than honest response condition. The majority of these differentially functioning items came from the openness factor (9 items) and the least came from the neuroticism factor (0 items).

With the polytomously scored data, the Mantel procedure (Mantel, 1963) was utilized. This procedure is very similar to traditional Mantel-Haenszel procedure. In this procedure, the groups are still coded '0' for the reference group (e.g., honest responders) and '1' for the focal group (e.g., desirable responders). However, the response categories can be larger than two. Thus, it can be used with personality and attitudinal data where there are typically five or more response categories. In this case, a 2 x 5 frequency table was used.

In the Mantel procedure, the response options are explicitly treated as ordered. This procedure assigns an index number to the response option category and then compares the item means for the reference and focal groups who have been matched on the stratification variable. The stratification variable is the total score on the particular personality factor. As noted above, the stratification variable was collapsed to form six categories of personality factor scores.

The Mantel procedure has a chi-square distribution with one degree of freedom. Significant values indicate that individuals in the focal and reference group differ in their performance on an item, even after controlling for the latent trait. In contrast, to the traditional Mantel-Haenszel procedure, there is no easily interpretable measure of the effect size of DIF (Potenza & Dorans, 1995).

However, Zwick, et al. (1993) and others (e.g., Potenza & Dorans, 1995) have argued that the standardized mean difference (SMD) method for polytomous items proposed by Dorans and colleagues (Dorans & Schmitt, 1991; Dorans, Schmitt & Bleistein, 1992) is an appropriate descriptive statistic of the degree of DIF that can supplement the Mantel procedure. This statistic compares the means of the reference and focal groups, controlling for the level of the latent trait. The results of the traditional Mantel-Haenszel procedures and the SMD are often in agreement (Dorans & Holland, 1993).

The formulas for the SMD presented by Potenza and Dorans (1995) were used here. SMD is computed in several steps. In the first step, the expected item scores for the focal and reference groups are computed. The formulas for the computations at the first step are as follows,

$$E_{Fm}(Y|X) = \sum_{k} N_{Fmk} Y_k / N_{Fm} \tag{11}$$

$$E_{Rm}(Y|X) = \sum_{k} N_{Rmk} Y_k / N_{Rm} \tag{12}$$

where E_{Fm} is the expected item score for the focal group, E_{Rm} is the expected item score for the reference group, Y is the response to the item with k possible response options (k=1, 2, 3, 4, 5), X is the stratification variable (i.e., the level of the latent trait) with m possible levels, N_{Fmk} is the number of individuals in the focal groups at level m of the stratification variable that selected option k, N_{Rmk} is the number of individuals in the reference groups at level m of the stratification variable that selected option k, N_{Fm} is the number of individuals in the focal groups at level m of the stratification variable, and N_{Rm} is the number of individuals in the reference groups at level m of the stratification variable.

The second step in computing SMD is to take the difference between the expected items score for the focal and reference groups at each level of the stratification variable. This difference is then weighted by the sample size of the focal group at the particular level of the stratification variable. This product is then divided by the sample size of the focal group. The formulas are as follows,

$$D_m = E_{Fm}(Y|X) - E_{Rm}(Y|X) \tag{12}$$

$$SMD = \sum_m N_{Fm} D_m / N_F \tag{13}$$

where D_m is the difference in the expected item score between the reference and focal group and N_F is the sample size of the focal group.

Values of SMD can range from –1.0 to 1.0. Negative values indicate that the reference group has a higher mean score than the focal group after controlling for the latent trait. Positive values indicate that the focal group has a higher mean score than the reference group after controlling for the latent trait. Millsap and Everson (1993) argue that SMD values between –0.05 to 0.05 are considered as demonstrating insignificant degrees of DIF, values between ±0.05 and ±0.10 are considered as demonstrating moderate degrees of DIF, and values greater than ±0.10 are considered as demonstrating substantial degrees of DIF. Dorans, et al. (1992) have presented evidence that these values work well in actual testing situations.

The results of the Mantel procedure are presented in Table 13. As can be seen in the Table, 76 of the 300 items (25.3%) displayed differential functioning using the criteria presented by Millsap and Everson (1993). The SMD for 47 of the 76 differentially functioning items indicated that after controlling for latent trait, the item means for the desirable response condition were higher than the item means for the honest response condition. It is important to note that for the items measuring neuroticism, the sign of the SMD will be negative for the desirable response condition and positive for honest response condition because neuroticism is scored such that the desirable level of the trait is on the opposite end of the response scale from the other four traits. That is, lower levels of neuroticism are desirable, while higher levels of the other four traits are desirable. At least eight items from each factor were identified as functioning differentially. The greatest number of differentially functioning items was 12 from the openness factor and the least was 8 from the conscientiousness, neuroticism, and agreeableness factors.

For the remaining 29 items, items indicated that after controlling for latent trait, the item means for the honest response condition were higher than the item means for the desirable response condition. The majority of these differentially functioning items came from the openness factor (11 items) and the least came from the neuroticism and conscientiousness factors (1 item).

As can be seen by comparing Tables 12 and 13, there was a great degree of overlap in the items identified by the Mantel-Haenszel and Mantel procedures. In fact, 55 items were identified as functioning differentially by both procedures. Additionally, the direction of the odds ratio and SMD were in agreement for these 55 items.

Next, the latent parameter differential item functioning analyses were performed. The analyses were performed using the item and latent trait estimates produced from the IRT analyses described above. Before performing the analyses, the item and person parameter estimates for each group were put on a common metric.

Lord (1980) demonstrated that under IRT the relationship between two calibrations of the same items from subgroups of a population is linear. Thus, the item parameters (i.e., discrimination and difficulty parameters) from one sample can be linked to the metric of another sample using a linear equation. The transformation equations are as follows,

$$a^*_{j2} = a_{j2} / \alpha \tag{14}$$

$$b^*_{j2} = \alpha b_{j2} + \beta \tag{15}$$

where a is the discrimination parameter for item j, b is the difficulty parameter for item j, α is the slope coefficient, β is the intercept coefficient, 2 indicates that the parameters come from the second of the two samples to be linked, and * represents a transformed value. Correspondingly, the values of theta can be transformed as follows,

$$\hat{\theta}^*_{I2} = \alpha \hat{\theta}_{i2} + \beta \tag{16}$$

where θ is the estimated value of the latent trait for person i.

There are a variety of methods that can be used to determine the slope and intercept term (i.e., the linking constants) of the linear equation. The method used here is the mean-sigma method (Marco, 1977). The mean-sigma method is appropriate when two samples of examinees respond to a common set of items. The means and standard deviations of the common items from both samples are used to determine the slope and intercept coefficients for the linking equation. For the 2-PL IRT models, the coefficients are computed as follows,

$$\alpha = \frac{s_{1c}}{s_{2c}} \tag{17}$$

$$\beta = \bar{b}_{1c} - \alpha \bar{b}_{2c} \tag{18}$$

where s_{1c} is the standard deviation of the difficulty parameters for the first sample on the common items, s_{2c} is the standard deviation of the difficulty parameters for the second sample on the common items, $\bar{b}_{1c}$ is the mean of the difficulty parameters for the first sample on the common items, $\bar{b}_{2c}$ is the mean of the difficulty parameters for the second sample on the common items, and c refers to the common items. These linking coefficients can be used to transform the parameter estimates of the second sample to the metric of the first sample using equations 14, 15 and 16.

Cohen and Kim (1998) extended the mean-sigma method to graded response IRT models. Equations 17 and 18 are used to compute the linking coefficients, however, the means and standard deviations are now taken over all of the item difficulty parameters. It is important to note that this method does not take into account that the items will be estimated with differing levels of accuracy (i.e., standard errors of differing magnitude). Thus, if the item parameters are poorly estimated, the slope and intercept coefficients will be biased. The standard errors of the item parameters were examined to determine the degree to which this was a limitation of the linking method used here.

The derived linking coefficients for each data type are presented in Table 14. These linking coefficients were used as input to the DFIT analysis described below. In general, the standard errors were relatively small. Of course, there were some items that were poorly estimated, but the majority of items were adequately estimated. Consider, for example, Figures 13 and 14, which present the distribution of the standard errors of the item difficulty parameters for all of the items for the 2-PL model. As can be seen in the Figure, the majority of the standard errors are less than 0.50 in both the honest response and the desirable response conditions. The standard errors of the option difficulty parameters for the graded response model demonstrated a similar pattern. Thus, it is unlikely that the quality of the parameter estimates is substantially reducing the quality of the linking coefficients. However, caution is still warranted give the moderate degree of fit demonstrated by the IRT models.

The specific latent parameter DIF technique used was Raju, et al.'s (1995) DFIT. The non-compensatory DIF (NCDIF) index was used in this study. The NCDIF is a measure of DIF that examines the potential differential functioning of an item without including information about the DIF from the other items. This is a parametric procedure that is similar to Lord's (1980) chi-square goodness-of-fit test for logistic models (Flower, Oshima, & Raju, 1999). The NCDIF index for an item is a function of the difference in expected scores, and hence item probabilities, for the focal and reference group (i.e., desirable and honest response conditions) after accounting for the latent trait. Large NCDIF index values are achieved when large differences exist between focal and reference groups in expected scores after controlling for the latent trait. A NCDIF is computed for each item, as well as, an associated chi-square. Items are identified as demonstrating DIF when the chi-square statistic is significant at $p < .01$, and the NCDIF is > .006 for two response options and NCDIF > .096 for five response options. Raju's (1998) DFIT4 program was used to compute NCDIF and the associated test statistics.

These analyses were performed separately for each of the five personality factors.

The results of the DFIT analysis for the dichotomous items are presented in Table 15. As can be seen in the table, 72 of the 300 items (24%) were identified as differentially functioning using the criteria of NCDIF > 0.006 and a significant χ^2 at $p < .01$. The majority of the differentially functioning items assessed either extroversion or openness. Very few differentially functioning items were found for neuroticism.

The results of the DFIT analysis for the polytomous items are presented in Table 16. As can be seen in the table, 66 of the 300 items (22%) were identified as differentially functioning using the criteria of NCDIF > 0.096 and a significant χ^2 at $p < .01$. The majority of the differentially functioning items assessed extroversion and very few items assessing conscientiousness were found to be functioning differentially.

As can be seen by comparing Tables 15 and 16, there was a considerable degree of overlap in the items identified by the DFIT procedure with the dichotomously and polytomously scored data. In fact, there were 34 items that were identified as functioning differentially by both procedures. However, it is interesting to note that less agreement was found among the latent parameter procedures than among the observed score procedures. In part, this is a result of the increasingly conservative standard for the NCDIF with the polytomous data. There were several items that were slightly below this standard that did provide evidence of DIF using the dichotomously scored data. This may also be a result of the increased complexity in the steps that are necessary (e.g., linking procedure, larger standard errors, moderate model fit) to perform the latent parameter analyses.

A summary of the four differential item function analyses is presented in Table

17. Items that demonstrated evidence of differential functioning on at least three of the four procedures were classified as items that can be distorted. This resulted in 52 of the 300 items (17.33%) being classified as capable of being distorted. However, it is important to note that the majority of these 52 items demonstrated DIF on the observed score methods and the DFIT with the dichotomous data. Very few of these items were classified as functioning differentially as a result of the polytomous DFIT technique. The remaining items were classified as not capable of being distorted. These classification decisions are the basis for the item groupings for the differential person functioning analyses in Study 2.

Study 1 Discussion

The purpose of Study 1 was to identify items on the IPIP that could be classified as capable or not capable of being distorted using differential item functioning (DIF) analyses. This classification is the basis of the differential person functioning analysis that is utilized in Study 2. Using two observed score and two latent parameter methods of identifying DIF, 52 of the 300 items (17.33%) were identifying as functioning differentially between the two response instruction conditions.

Although this may be interpreted as a particularly low detection rate, it is not dramatically less than some of the values published in previous experimental research. For example, in a study using Army recruits, Zickar and Robie (1999) found 29% of the items functioned differentially between an honest response and a desirable response instructions condition. Moreover, the response instruction manipulation used in this study was more subtle than what is typically used in research on intentional response distortion (Brown & Harvey, 2003). Thus, a lower percentage of items identified as functioning differentially was not surprising. Also, this identification rate is across at least three different detection techniques. The results of any one of the DIF detection techniques were very much in line with the findings of previous studies. By using the results of four DIF detection methods, the detection rate is a conservative estimate of the number of items that are functioning differentially.

In the majority of the DIF analyses, at least one item from each personality factor was identified as differentially functioning. Yet, a surprising result of the DIF analyses was the small number items from the conscientiousness factor that were identified as functioning differentially. Many personnel specialists (e.g., Behling, 1998; Schmidt & Hunter, 1998) see conscientiousness as the most important personality trait for predicting job performance and other work behaviors. This factor also is most in line with common folk beliefs about desirable employee characteristics (e.g., being diligent, hardworking, and detailed). Thus, it was expected that a larger number of these items would be identified as differentially functioning.

However, it may be the case that these items are the most transparent in terms of their importance in a job application situation. For example, conscientious item present stems such as, “I shirk my duties.” Under the subtle response instruction manipulation, participants may have felt their dishonest responding would have been too apparent by consistently endorsing the upper end of the response scale on these items. Although the DIF results for conscientiousness were less than expected, the number of items from each personality factor that were identified as functioning differentially is not important for the analyses in Study 2 which draw on these results. The differential person functioning analyses are performed across the personality factors. As long as each personality factor is represented, this is not a potential limitation of the results of the analysis based on these item classifications.

A potential concern with the results of this study is that they are based on a sample size that is smaller than the recommend level for an IRT study (e.g., Zickar, 2000). Although the sample size is less than the recommend level, it is within the range of the sample sizes of studies using IRT to examine response distortion on measures of personality (e.g., Zickar, & Ury, 2002) and the results are line with larger sample size studies using IRT to examine the IPIP (e.g., Trippe & Harvey, 2003) and other related measures of the five-factor model of personality (e.g., Chernyshenko, et al., 2001). Further, the tests of the IRT assumptions were not grossly violated. However, the fit of the model in the 2-PL condition was modest. The fit of the graded response model was better, but not ideal. Because four different methods were used to make classification decisions, and there was a substantial amount of agreement between the methods, the results of this study provide reasonable item classification decisions for use in Study 2.

Study 2

In the second study, the purpose was to determine if the DPF techniques can be used to identify the individuals who are distorting their responses on the personality inventory and what is the accuracy at identifying these individuals. In this study, a between-subjects design was utilized. That is, the participants completed the IPIP in only one of the experimental conditions. In this study, a third experimental condition was added. This additional experimental condition is described below.

Despite the advantages of within-subjects (i.e., repeated measures) designs that were noted above, a between-subjects design was utilized in this study for several reasons. First, the purpose of this study was to determine the classification accuracy for a given individual using the DPF technique and not compare group differences in mean item scores. Therefore, the potential problems associated with between-subjects designs (e.g., group non-equivalence or subject-by-instructions interactions) were less of a concern. Second, a between-subjects design was more efficient in this case. Using a within-subjects design with three response instruction manipulations that are completely counter-balanced would result in 6 experimental cells. Given the need for large sample sizes in each cell (i.e., $n > 100$), the use of within-subjects designs becomes impractical. Lastly, the between-subjects design will provide a more conservative test of the classification accuracy of the DPF technique because the effect sizes of response distortion tend to be smaller for these types of designs (Ones & Viswesvaran, 1998). Thus, the between-subjects design was selected.

The usefulness of the DPF techniques for identifying individuals who were distorting their responses was assessed in several ways. Specifically, the classification consistency between the DPF and other response distortion detection techniques was assessed, the classification accuracy of the DPF technique was examined, and the classification accuracy of the DPF was compared to the classification accuracy of other response distortion detection techniques. First, the DPF technique was performed for each individual. The results were used to make a decision as whether the participant was distorting or not distorting their responses. A decision about the participant's responses was also made using the other three response distortion methods (the BIDR scale, the Marlowe-Crowne scale, and the l_z index). Next, the consistency of these decisions was examined. That is, does each detection method lead to a similar decision? Third, the accuracy of the decisions made for each participant using the DPF method was examined using the known response instruction manipulation for the participants. The question being asked is: Are the decisions made using the DPF method correct? Finally, the accuracy of the decisions made using all four of the distortion detection methods were compared. The question asked here is: Are the decisions made using one of the detection methods more accurate than the others? If the DPF offers advantages over the other methods, as is argued in this paper, it should result in more accurate decisions.

Method

Participants

Participants in this study were 393 undergraduate students enrolled in introductory psychology courses. Participants volunteered to participate in exchange for course credit Participants were recruited through the psychology subject pool. Eighteen of the 393 participants were excluded from the analyses. These 18 participants failed to complete the personality inventory (i.e., IPIP) in the allotted time. Thus the sample size was 375. There were 129 participants in honest response instructions condition, 125 participants in desirable response instructions condition, and 121 participants in no response instructions condition.

Because of the time limits of the study (i.e., 55 minutes) there were several participants that completed the IPIP but did not complete the one or both of the other two scales. Of the 375 participants, 9 did not complete either of the other two scale and 78 only completed one of the two scales. The remaining 288 participants completed the IPIP and the other two scales in the allotted time period. It is important to note that all 375 participants were utilized in the subsequent analyses.

Seventy-three percent of the final sample was female and 27 percent were male. The average age of the participants was 19.37 years with a standard deviation of 2.04 years. The overwhelming majority of the sample had held a job at some point in their life (98.7%).

Design and Manipulations

The response instruction manipulation was a between-subjects factor with three levels. Participants were instructed to either respond honestly, respond desirably, or were given no specific instructions about the direction of responding. The assignment of the conditions to the experimental sessions was randomly determined using the same procedure as in Study 1. The experimenter read the instructions to the participants. The specific instructions for the honest condition were as follows,

> "Please complete this personality inventory as honestly as you can. The results will be completely anonymous and will be used for research purposes only. It is *very* important that you respond to this survey by describing yourself as you really are and not as you want to be or as you want others to see you."

The specific instructions for the desirable condition were as follows,

> "Please complete the personality inventory as if you were applying for a job you really want. To increase your chances of being hired, you should respond in way that will make you look good to the organization. But, do not respond in a way that will look like you were obviously faking your responses."

The instructions given to the no response instructions group were as follows,

> "Please complete the personality inventory as if you were applying for a job."

Measures

International Personality Item Pool. The International Personality Item Pool (IPIP; Goldberg, 1999) was also used in the second study. Before conducting the analyses, the reverse-scored items on the inventory were recoded. A score for each of the five factors was computed by summing scores for the items that comprise the factor. Additionally, a score was computed for each of the sub-facets. The scores on the IPIP demonstrated satisfactory internal consistency (i.e., α). The reliability across the three experimental groups at the factor level ranged from .87 to .95. The complete list of reliabilities for each factor and sub-facet on the IPIP for each experimental condition are presented in Table 18.

For the IRT analyses requiring polytomously scored data, the responses to the IPIP, as provided by the participant on the original 5-point scale, were used. For the analyses requiring dichotomously scored data, the responses were recoded into a binary format. To dichotomize the data, the response options indicating high levels of the trait were coded as ‘1’ and the remaining options were coded as ‘0’. Specifically, the response options of 1, 2, and 3 were recoded as 0, and the response options 4 and 5 were recoded as 1.

Balanced Inventory of Desirable Responding. The Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984) assesses the two forms of desirable responding identified by Paulhus and colleagues (Paulhus, 1984; 1991; Zerbe & Paulhus, 1987). The first form is self-deception, which is called self-deceptive positivity on this scale. The second form is impression management.

The items on the inventory present propositions to the respondent that he or she rates as either describing them or not describing them. The inventory contains 40 items. There are 20 items for each form of desirable responding. Approximately half of the items for each type of desirable responding are reverse scored. The items are rated on a 7-point scale with "not true" and "very true" as the anchors at 1 and 7, respectively. Items on the BIDR are scored a 1 if the respondent endorses the rating option of 6 or 7. All other response options are scored as 0. The responses are summed to create a total score for each respondent. Thus, the total scores for each scale can range between 0 and 20. Because this study is focused on deliberate response distortion, only the impression management items were used. The items are presented in Appendix D. If a participant was missing any data on this measure, a total score was not computed. Three response distortion groups were created using the top 5%, 10%, and 15% of the scores on the scale.

Paulhus (1991) has reported that the internal consistency of the scores on the BIDR range from .75 to .86. A test-retest reliability of .65 has also been reported by Paulhus. Further, the scores on the impression management scales of the BIDR correlate highly with other measures of response distortion (Paulhus, 1991). In this study, the scores on the BIDR demonstrated satisfactory internal consistency (i.e., α). The reliability across the three experimental groups was .82. The reliability of the scores from the respondents in the honest response instructions condition was .93. The reliability of the scores from the respondents in the desirable response instructions condition was .77. The reliability of the scores from the respondents in the no response instructions condition was .78.

Marlowe-Crowne Social Desirability Scale. The Marlowe-Crowne Social

Desirability Scale (MCSD; Crowne & Marlowe, 1960) presents items concerning desirable, but uncommon behaviors (e.g., always admitting mistakes) and undesirable, but common behaviors (e.g., gossiping). The MCSD contains 33 true or false items. Approximately half of the items are reverse scored. The responses are summed to create a total score for each respondent. Thus, scores on the MCSD can range from 0 to 33. The items are presented in Appendix E. If a participant was missing any data on the measure, a total score was not computed. Three response distortion groups were created using the top 5%, 10%, and 15% of the scores on the scale.

The estimates of internal consistency reported in previous research range from .73 to .88 and the estimates of test-retest reliability range from .84 to .88 (Paulhus, 1991). The MCSD has demonstrated moderately high correlations with other response distortion scales (Crowne & Marlowe, 1964; Crowne & Marlowe, 1960). In this study, the scores on the MCSD demonstrated modest internal consistency (i.e., KR-20). The reliability across the three experimental groups was .643. The reliability of the scores from the respondents in the honest response instructions condition was .704. The reliability of the scores from the respondents in the desirable response instructions condition was .596. The reliability of the scores from the respondents in the no response instructions condition was .588.

Demographic characteristics. The participants were asked their sex, age, and if they have ever held a job.

Procedure

Participants volunteered via sign-up sheets on the psychology subject pool board. Study 2 was run in classrooms in Porter Hall with a maximum of 30 participants in each session. When the participants arrived, they were asked to read and sign the consent form. Once all the participants completed the consent form, they were provided with the personality inventory, the response distortion measures, and an optically scanned answer sheet. The participants were instructed to mark all of their responses on the answer sheet. Next, the participants were asked to mark their sex, age, and whether or not they had ever held a job on the answer sheet. They were also asked to make a mark on the response sheet that indicated the response instruction condition assigned to that session. Then, the participants were then given the response instruction manipulation assigned to that particular session. After the instructions were presented, the participants were asked to begin completing the personality inventory. The order of the measures was constant. The participants first completed the IPIP, the BIDR second, and the MCSD last. When the participants completed the inventory, they were given one credit for their participation, and debriefed. Each session lasted between 40-50 minutes. The experimental protocol is presented in Appendix F.

Results

Descriptive statistics. The means and variances for each item within each condition are presented in Table 19. The means and variances for each factor and sub-factor are presented in Table 20. The descriptive statistics and correlations between the personality factors and the response distortion scales are presented in Table 21 for each response instruction condition.

Manipulation check. The effectiveness of the response instruction manipulation was examined. If the manipulation led the participants to respond differently in each condition, there should be significant difference between the response instruction conditions on the scores for each personality factor. To test for these potential differences, a MANOVA was performed with the instruction manipulation as the independent variable and the five personality factor scores as the dependent variables.

The overall multivariate effect was significant, Pillai's Trace = 0.158, $F(10, 738) = 6.32$, $p < .001$, partial $\eta^2 = .079$. The subsequent univariate F-tests for conscientiousness ($F[2, 372] = 18.95$, $p < .001$, partial $\eta^2 = 0.092$), extroversion ($F[2, 372] = 3.85$, $p < .03$, partial $\eta^2 = 0.020$), neuroticism ($F[2, 372] = 29.07$, $p < .001$, partial $\eta^2 = 0.135$), and agreeableness ($F[2, 372] = 6.95$, $p < .001$, partial $\eta^2 = 0.036$) were significant at $p < .05$. The univariate F-test for openness ($F[2, 372] = 1.87$, $p < .160$, partial $\eta^2 = 0.01$) was significant at $p < .16$. Further, the differences in the means between the response instruction conditions were all in the correct direction (see Table 20 for means and variances). The desirable response instruction condition had the highest means, followed by the no instructions condition, and the lowest means were found in the honest response condition. Additionally, a planned contrast between the honest and desirable response instruction conditions was performed. There was a significant difference between these two conditions at $p < .01$ for conscientiousness, extroversion, neuroticism, and agreeableness. The difference between the conditions on openness was significant at $p < .07$. Thus, the honest response and desirable response conditions did differ in their responses to the personality inventory.

Additionally, two ANOVAs were performed to examine the effects of response instructions on the BIDR and MCSD. If the manipulation had an effect on the scores of the BIDR and MCSD, the results of the ANOVAs would be significant. In the first ANOVA, the response instruction condition was the independent variable and the total score on the BIDR was the dependent variable. The effect was significant, $F(2, 340) = 14.84$, $p < .001$, partial $\eta^2 = .08$, indicating the response instruction conditions differed in their mean responses. The mean score for the desirable response instruction condition was the largest, the mean score for the no instructions response condition was the smallest, and the mean score for the honest response instruction condition was in between the other two conditions (see Table 20 for means and variances). Additionally, a planned contrast between the honest and desirable response instruction conditions was performed. There was a significant difference between these two conditions at $p < .01$. Thus, there was a difference in the responses of those in the honest response and desirable response conditions to the items on the BIDR.

In the second ANOVA, the response instruction manipulation was the independent variable and the total score on the MCSD was the dependent variable. The effect was significant, $F(2, 308) = 3.304$, $p < .05$, partial $\eta^2 = .021$, indicating the response instruction conditions differed in their mean responses. Surprisingly, the mean score for the honest condition response instruction condition was the largest, the mean score for the desirable condition was the smallest, and the mean score for the no instructions response condition was in between the other two conditions (see Table 20 for the means and variances). This is the opposite of what should occur. A planned contrast between the honest and desirable response instruction conditions revealed the difference between these conditions was significant $p < .05$. Thus, there was a difference in the responses of those in the honest response and desirable response conditions to the items on the MCSD. However, the direction of the difference is opposite of what was expected.

In sum, these results indicate that the response instruction manipulation did lead to differences in the participant's responses in the correct direction on the personality inventory and the BIDR, but not the MCSD. The failure of the MCSD was likely a result of several factors. First is the transparent nature of the items and the instructions to response in a subtle manner. Second was the order of the measure. The MCSD was always last and therefore may have been subject to fatigue on the part of the participants. However, the order of the measures was not manipulated, this possibility could not be tested. Third was the low level of reliability. Because the scores on the MCSD failed to correctly distinguish between the participants responding honestly and desirably, it was not included in the subsequent analyses.

Tests for sex effects. To determine if the sex of the participant was affecting the responses to the personality items, a MANOVA was performed. Sex was entered as the independent variable and the five personality factor scores as the dependent variables. These analyses were performed within each response instruction condition. If the sex of the participants was not exerting an influence on the responses, the overall multivariate effect and the effect from each of five univariate ANOVAs should not be significant. However, there is some evidence that men and women do differ in their mean levels on some of the personality factors. Thus, significant results would not be completely surprising.

There was a significant multivariate effect in the honest response condition, Pillai's trace = .161, $F(5, 123) = 4.706$, $p < .001$, $\eta^2 = .161$. The subsequent univariate F tests revealed that the men and women were significantly different on conscientiousness, $M_M = 205.75$, $M_F = 217.34$, $F(1, 127) = 4.57$, $p < .05$, $\eta^2 = .035$, openness, $M_M = 198.39$, $M_F = 211.44$, $F(1, 127) = 6.47$, $p < .05$, $\eta^2 = .048$, and agreeableness, $M_M = 199.11$, $M_F =$

217.34, $F(1, 127) = 10.93$, $p < .01$, $\eta^2 = .079$. Women scored higher than the men on all three factors. No significant differences were found for the other personality factors (Extroversion, $M_M = 206.50$, $M_F = 216.35$; Neuroticism, $M_M = 163.89$, $M_F = 169.57$) .

Within the desirable response instruction condition, there was also a significant multivariate effect for sex, Pillai's trace = .202, $F(5, 119) = 6.03$, $p < .001$, $\eta^2 = .202$. The subsequent univariate F tests again revealed that the men and women were significantly different on agreeableness, $M_M = 212.18$, $M_F = 230.08$, $F(1, 123) = 16.68$, $p < .001$, $\eta^2 = .119$. The women scored more highly than the men. No significant differences were found for the other personality factors (Conscientiousness, $M_M = 228.97$, $M_F = 237.24$; Extroversion, $M_M = 215.21$, $M_F = 225.52$; Neuroticism, $M_M = 140.76$, $M_F = 141.05$; Openness, $M_M = 208.50$, $M_F = 215.89$).

There was also significant multivariate effect for sex in the no response instructions condition, Pillai's trace = .212, $F(5, 114) = 6.14$, $p < .001$, $\eta^2 = .212$. The subsequent univariate F tests revealed that the men and women were significantly different on extroversion, $M_M = 208.63$, $M_F = 224.18$, $F(1, 118) = 13.50$, $p < .001$, $\eta^2 = .103$, openness, $M_M = 200.33$, $M_F = 214.39$, $F(1, 118) = 13.11$, $p < .001$, $\eta^2 = .100$, and agreeableness, $M_M = 206.82$, $M_F = 222.46$, $F(1, 118) = 9.36$, $p < .01$, $\eta^2 = .074$. Women scored higher than the men on all three factors. No significant differences were found for the other personality factors (Conscientiousness, $M_M = 215.70$, $M_F = 222.08$; Neuroticism, $M_M = 156.32$, $M_F = 160.23$).

Again, the results for agreeableness in all three response instruction conditions and for extroversion in the no response instructions condition are not surprising given previous research demonstrating that women tend to score higher on these factors than men (Costa & McCrae, 1992). The difference between men and women on conscientiousness in the honest response condition and openness in the no response instructions condition, on the other hand, were not expected. However, the sizes of these effects were small.

Additionally, an ANOVA were performed to examine the effects of sex on the BIDR within each response instruction condition. If no sex differences were present, the results of the ANOVA would be non-significant. In the honest response condition, the results were significant, $M_M = 3.85$, $M_F = 5.48$, $F(1, 118) = 4.81$, $p < .05$, $\eta^2 = .039$, with the women obtaining higher scores than men. The results were non-significant in the desirable response instructions and no response instructions conditions (Desirable, $M_M = 6.22$, $M_F = 7.53$; No Response Instructions, $M_M = 4.77$, $M_F = 5.20$).

As was demonstrated here, the differences between men and women on measures of the five-factor model of personality tend to be small (Mount & Barrick, 1995). The magnitude of the changes within each sex between the conditions was small (men, partial $\eta^2 = 0.082$; women, partial $\eta^2 = 0.089$). More importantly for this study, there are no sex differences in the scores on the response distortion scale among those who are responding desirably. These results are consistent with meta-analytic studies examining response distortion (e.g., Ones & Viswesvaran, 1998). Further, because decisions about the presence of desirable responding are made at the individual level, any potential sex differences have no impact on the decision about a given individual. Lastly, because there were three to four times more females than males in these analyses, the presence of any sex differences needs to be interpreted with caution.

Principle components analysis. To test whether the unidimensionality assumption was met, a principle components analysis (PCA) was performed. If this assumption is met, only five factors should emerge and the items should load on the correct factor. The PCA was performed within each factor. To determine the number factors that were present in the data, modified parallel analysis (MPA; Drasgow & Lissak, 1983) was used.

The results of the principle components analysis indicated that a dominant factor emerged for each of the personality factors within each response condition. The first two eigenvalues and the percent of variance the accounted for are presented in Table 22 for each personality factor within each response instruction condition. The first component from the PCA analyses for conscientiousness, agreeableness, and neuroticism in each response instruction condition accounted for more than 20% of the variance, as recommended by Reckase (1979). The first component for openness was less than 20% for each response instruction condition. Although in each case, the percent of variance account for was above fifteen percent. For extroversion, the first factor accounted for twenty percent of the variance in the desirable response instructions condition, but not the other two conditions.

The comparison of the eigenvalues resulting from the PCA and the modified parallel analysis (MPA) using the responses from the honest response instruction condition are presented in Figures 15 through 19 for conscientiousness, extroversion, neuroticism, openness, and agreeableness, respectively. In each of these figures, the unidimensionality criterion is not met. The second eigenvalue from the observed data is greater than the second eigenvalue from the simulated data. The comparison of the eigenvalues resulting from the PCA and the MPA using the responses from the desirable response instructions condition are presented in Figures 20 through 24 for conscientiousness, extroversion, neuroticism, openness, and agreeableness, respectively. Again, the unidimensionality criterion is not met. The second eigenvalue from the observed data is greater than the second eigenvalue from the simulated data. As with the other conditions, the unidimensionality criterion is not met in the no-response instructions condition. The comparisons of the eigenvalues resulting from the PCA and the MPA are available upon request.

Although, each personality trait possess one clear dominant factor and most accounted for more than 20% of the variance, the results of the MPA indicate that strict unidimensionality is not observed. Given, the findings concerning the robustness of IRT model to multidimensionality described above and the moderate degree to which this assumption is violated in this study, the parameter estimation is unlikely to be severely compromised. Moreover, the dimensionality of the data seen in this analysis is comparable to studies with larger sample sizes examining the IPIP with IRT (e.g., Trippe & Harvey, 2003).

Again, the factor analysis also allows the stability of the factor structure to be examined. To examine the stability of the factor structure, a principle components analysis with a promax rotation was performed for each of the three response instruction conditions using the scores of the sub-factors. If the factor structure is stable, the resulting factor structure should be the same for each condition. Parallel analysis was used to determine the number of factors to retain.

The eigenvalues from the honest condition, desirable condition, no response instructions condition and the parallel analysis are presented in Table 23. Because the eigenvalue for the sixth factor from the observed data was less than the eigenvalue for the sixth factor in the random data for each condition, only five factor exist in the data.

As can be seen in Table 24, the rotated factor structure matrix for the honest response condition conforms reasonably well to the five-factor model. The majority of the sub-scales load on the correct factor. The exceptions are the third and fourth sub-scales of extroversion and the third sub-scale of conscientiousness. The third sub-scale is labeled "assertiveness". The assertiveness items loaded on conscientiousness. This sub-factor also loaded highly on conscientiousness in Study 1. The fourth sub-scale of extroversion is labeled "activity". The activity items loaded on the factor for conscientiousness. This was the same result that was found in Study 1. The third sub-scale of conscientiousness is labeled "dutifulness". These items loaded on the factor for agreeableness. However, the difference between this loading and the loading on the conscientiousness factor was small. As was true of Study 1, this pattern of loading is not abnormal. In factor loading matrix presented by Costa & McCrae (1992) for the NEO-PI, the loadings for the dutifulness on agreeableness and the activity items on conscientiousness were both greater than 0.29.

Interestingly, the pattern of factor loadings for the desirable response condition was scattered (see Table 24). Although there were five factors, extroversion and neuroticism loaded on the same factor, albeit in different directions. Thus, there were really only four factors in the desirable responding condition. This finding could be interpreted as the opposite of Schmit & Ryan (1993). Instead of an expanded factor structure emerging, a compressed structure emerged.

Additionally, one of the sub-scales from consciousness, extroversion, and agreeableness loaded on wrong factor. For consciousness, sub-factor number three again loaded on the same factor as agreeableness. For extroversion, sub-factors number four (activity) loaded on the same factor as conscientiousness. For agreeableness, sub-factors number one (trust) loaded on the same factor as extroversion. Additionally, openness loaded on three different factors. Four of the openness items were the only items loading on these factors. Again, the pattern of loadings for activity is consistent with evidence presented by Costa & McCrae (1992). Also, the loadings for openness (sub-factors 1-4) are in line with the evidence presented in Costa and McCrae (1992).

The factor structure of the no response instructions condition more closely resembled the structure of the honest response condition. However, the loadings were less clean. The fourth sub-factor of extroversion again loaded on conscientiousness factor and the first sub-factor loaded on agreeableness. Sub-factor four of neuroticism loaded on the conscientiousness factor and sub-factor five of neuroticism loaded on the extroversion factor. The third sub-factor of openness loaded on the neuroticism factor and the sixth sub-factor of openness loaded on the conscientiousness factor. Lastly, the fifth sub-factor of agreeableness loaded on the extroversion factor.

In summary, it appears there are clear differences in the factor structures between the honest responding, desirable responding, and no response instruction conditions. However, the presence of the desirable and no response condition did not completely destroy the five factor structure. Only the openness factor in the desirable response condition was completely compromised. For the other factors, the majority of the sub-scales loaded on the correct factor. These results are not surprising given the lack of strict unidimensionality in the data. However, the inconsistency in the factor structure between the response instruction conditions is not a problem for the subsequent analyses because the analyses are performed at the item level, not the factor level. Thus, differing patterns of interrelations between specific items does not pose a limitation for the response distortion detection methods utilized below.

Parameter estimation and model fit. The MULTILOG program (version 7; Theissen, 2003) was used to compute the maximum likelihood estimates of the item parameters and the expected a posteriori estimates of the theta parameters in Samejima's (1969) graded response model. The BILOG program (version 2; Mislevy & Bock, 1991) was used to compute the maximum likelihood estimates of the item parameters and the expected a posteriori estimates of the theta parameters in the 2-PL model. For both models, the parameters were estimated for each of the response instruction groups and each of the five personality factors separately. Thus, fifteen analyses were performed. The program defaults were used in both cases. The results of the IRT analyses for the polytomously and dichotomously scored items for each individual and for each item are available upon request.

Model fit was assessed using Drasgow, et al.'s (1995) chi-square tests and the results of the factor analyses. The MODFIT program was used to perform these analyses. Again, because only a limited number of techniques were used to establish model fit, and the same sample was used to calibrate the parameters and evaluate differential responding, the results should be view as preliminary.

The results of the χ^2 tests of fit for the 2-PL model are presented in Table 25 for the honest response condition, Table 26 for the desirable response conditions, and Table 27 for the no response instructions condition. Again, the fit of the IRT was modest in all three conditions. Many of the χ^2 values were above 3 for the singles, doubles and triples.

The results of the χ^2 tests for Samejima's graded response model are presented in Table 28 for the honest response condition, Table 29 for the desirable response condition, and in Table 30 for the no response instructions condition. As can be seen in the table, the fit in all three conditions was good. The majority of the χ^2 values were below 3 for the singles, doubles and triples. Again, there is a modest degree of confidence in the stability of the parameters from the 2-PL model. The degree of confidence for the graded response model is higher, but caution in the interpretation of the parameters is still warranted.

Differential person functioning. To determine if the participants were differentially responding over the two groups of items (i.e., distortable vs. non-distortable) identified in Study 1, a differential person functioning analysis was performed. The analysis was performed using the Mantel-Haenszel and Mantel procedures. The analyses were performed using the all the items from the IPIP.

With the dichotomously scored data, the traditional Mantel-Haenszel procedure was used. In this analysis, the item group (distortable vs. non-distortable) and response (0 vs. 1) were crossed to form a 2 x 2 frequency table. The mean item score over persons was used to form the levels of the stratification variable. The stratification variable was collapsed to produce samples sizes of at least 5 in each response instruction condition at each level of the stratification variable. This resulted in six categories of total item scores. These categories were used as the stratification variable in both the Mantel-Haenszel and Mantel procedures described below.

In the Mantel-Haenszel procedure, the relationship between item group and response is assessed through an odds ratio and the odds ratio over all of the levels of the stratification variable is a measure of the effect size.

As with Study 1, the Educational Testing Service's delta scale and three categories of DIF were used. However, in this case, individuals are classified into one of the three categories, instead of classifying items. Thus, individuals that have a non-significant chi-square and an absolute value of D that is less than one are called "A persons". These people are not demonstrating differential responding over the two groups of items. Participants that have a significant chi-square and the absolute value of D that is greater than 1, but less than 1.5 are called "B persons". These participants are showing a moderate degree of differential responding over the item groups. Participants that have a significant chi-square and an absolute value of D that is greater than 1.5 are called "C persons". These participants are demonstrating a large degree of differential responding. For the purposes of classification in this study, participants falling into category B or C were identified as functioning differentially. Because, the purpose of this analysis is to identify individuals who are responding differentially by selecting the options that represent higher levels of the personality trait on the items that can be distorted, only participants with negative values of D were classified as responding differentially.

The results of the differential person functioning analysis using the Mantel-Haenszel procedure are presented in Table 31. As can be seen in the Table, 38 of the 373 participants (10.2%) were identified as differentially functioning using the B and C categories of Educational Testing Service's classification system and a negative value of D. Of these 38 participants, 16 were in the desirable response instructions condition (42%), 7 were in the honest response instructions condition (18%), and 15 were in the no response instructions condition (39%).

With the polytomous scored data, the Mantel procedure was used. In this analysis, the item group and responses (1 to 5) were crossed to form a 2 x 5 frequency table. The mean item score over persons was used to form the groups for the stratification variable. Again, the stratification variable was collapsed into six categories of total item scores.

The standard mean difference (SMD) was again computed as a descriptive statistics of the degree of differential responding to accompany the Mantel procedure. Values of SMD can range from –1.0 to 1.0. Negative values indicate that the participant has a higher mean score on the items that cannot be distorted than on the items that can be distorted even after controlling for the mean item score. Positive values indicate that the participant has a higher mean score on the items that can be distorted than on the items that cannot be distorted even after controlling for the mean item score. Millsap and Everson (1993) argue that SMD values between –0.05 to 0.05 are considered as demonstrating insignificant degrees of differential responding, values between ±0.05 and ±0.10 are considered as demonstrating moderate degrees of differential responding, and values greater than ±0.10 are considered as demonstrating substantial degrees of differential responding. Again, because the interest of this study is identifying individuals who are distorting their response, only individuals with SMD values that are positive were identified as distorting their responses.

The results of the Mantel procedure are presented in Table 32. As can be seen in the Table, 45 of the 373 participants (12.1%) were identified as responding differentially. Of these 45 participants, 19 were in the desirable response instructions condition (42.2%), 14 were in the honest response instructions condition (31.1%), and 12 were in the no response instructions condition (26%).

Classification decisions using the BIDR. To classify the participants as either distorting or not distorting their responses, the total score on the BIDR was first transformed into a standardized score (i.e., a z-score). The BIDR z-score was compared to the z-score that corresponds to the top 5% (i.e., z = 1.64), top 10% (i.e., z = 1.28) and 15% (i.e., z = 1.03) of a standard normal distribution. Participants with z-score that exceeded these values were classified as distorting their responses and those with z-scores that were below these values were classified as not distorting their responses.

Across the three response instruction conditions, 6.7% (25 participants) of the sample was identified as distorting their responses using the top 5% of the scores on the BIDR, 8.8% (33 participants) using the top 10% of the scores and 12.6% (47 participants) using the top 15% of scores. Of the 25 participants in the top 5% of BIDR scores, 14 were in the desirable response instructions condition (56%), 6 were in the honest response instructions condition (24%), and 5 were in the no response instructions condition (20%). Of the 33 participants in the top 10% of BIDR scores, 20 were in the desirable response instructions condition (60%), 7 were in the honest response instructions condition (21%), and 6 were in the no response instructions condition (18%). Of the 47 participants in the top 15% of BIDR scores, 25 were in the desirable response instructions condition (53%), 10 were in the honest response instructions condition (21%), and 12 were in the no response instructions condition (25%).

As noted above, no classification decisions were made using the MCSD because the participants in the honest response condition scored higher on this scale then those in the desirable response condition. This is the opposite of what should occur.

Classification decision using the person-fit index. The person-fit index, l_z, was computed for each individuals using the WPerfit program (Ferrando & Lorenzo, 2000).

Wperfit computes the l_z index for estimates derived from logistic IRT models. Individuals whose estimates of the latent trait do not fit their response pattern were identified using the guidelines described above (i.e., $l_z < -2.0$).

The results of the person-fit analyses are summarized for each personality factor in Table 33. As can bee seen in the Table, there are several participants with values of l_z that are less than –2.0. On the conscientiousness factor, 37.9% of the participants had values of l_z that were less than –2.0. On the extroversion factor, 44.8% of the participants had values of l_z that were less than –2.0. On the neuroticism factor, 33.6% of the participants had values of l_z that were less than –2.0. On the openness factor, 34.4% of the participants had values of l_z that were less than –2.0. On the agreeableness factor, 38.4% of the participants had values of l_z that were less than –2.0.

Approximately 20% of the participants did not have values of l_z that were less than –2.0 on any of the personality factors. Values l_z that were less than –2.0 were demonstrated on one personality factor by 23.7% if the participants, on two personality factors by 21.3% of the participants, on three personality factors by 18.9% of the participants, on four personality factors by 11.7% of the participants, and on five of the personality factors by 3.7% of the participants. Participants were classified as distorting their responses if they demonstrated values of l_z that were less than –2.0 on three of the five personality factors. This resulted in 129 participants (34.3%) being classified as distorting their responses on the personality inventory. Of the 129, 46 were in the honest response instructions condition (35%), 39 were in the desirable response instructions condition (30%), and 44 were in the no response instructions condition (34%).

Comparison of the classification consistency between DPF, the person-fit index, and the BIDR. After using all four measures to make a classification decision about each individual, the consistency of those decisions was explored. That is, the agreement among the four distortion detection techniques in their decisions about the presence of response distortion was examined. If the four different methods result in the same decision, then the consistency between the decisions will be high. However, if the different methods lead to different decisions, the consistency between these decisions will be low. However, these analyses do not test if the decisions are correct. All of the measures could consistently provide the wrong decision. Tests of the accuracy of these methods were examined in the subsequent section.

Two indices of consistency were utilized. Each index is based on a decision theoretic framework (see Figure 25). First, the estimated probability of a consistent classification was computed. This analysis compares the decisions made (e.g., distorting vs. not distorting) by two different methods (e.g., DPF vs. l_z). The results of this analysis provide evidence about the likelihood of arriving at the same decision using two different classification methods. If the resulting probability is high, it is very likely that the same decision will be reached using either method. If the probability is low, it is much less likely that the same decision will be reached.

The first index is the estimated probability of a consistent decision. This estimate is a sum of the probabilities for a distortion decision on each measure and a non-distortion decision on each measure. Mathematically, this is represented as follows,

$$\hat{P} = \hat{P}_{11} + \hat{P}_{00} \tag{19}$$

where $\hat{P}_{11}$ is the estimated probability of a distortion decision on each measure and $\hat{P}_{00}$ is the estimated probability of a non-distortion decision on each measure (see Figure 25). Values close to 1.0 indicate a high degree of consistency. Values close to 0.0 indicate a low degree of consistency.

All possible comparisons between the four response distortion detection methods were made. The analyses involving the response distortion scale were performed for each of the decisions made using the three different cut-off scores on the scale.

Second, Cohen's Kappa was computed. This index of consistency is similar to $\hat{P}$, but it takes into account the impact of consistency due to chance. Kappa compares the decisions made (e.g., distorting vs. not distorting) by two different methods (e.g., DPF vs. l_z) after removing the consistency attributable to chance alone. If the resulting probability is high, it is very likely that the same decision will be reached using either method even after adjusting for chance. If the probability is low, it indicates that much of the consistency in the decisions may be due to chance.

Mathematically, Kappa is represented as follows,

$$\kappa = \frac{P - P_c}{1 - P_c} \tag{20}$$

where P is the probability of a consistent decision as defined in equation 19 and P_c is the probability of chance consistency. P_c is computed as,

$$P_c = P_{1.}P_{.1} + P_{0.}P_{.0} \tag{21}$$

where $P_{1.}, P_{.1}, P_{0.}$ and $P_{.0}$ represent the column and row totals in Figure 25. All possible comparisons between the four response distortion detection methods were made. The resulting values indicate the percent agreement above chance. For example a value of .40 indicates that the consistency was 40% above chance.

These analyses were first performed using all of the participants. Next, these analyses were performed within each of the three response instruction conditions to determine if consistency varies as function of the type of responding. All possible comparisons between the four response distortion detection methods were made. The analyses involving the response distortion scale were performed for each of the decisions made using the three different cut-off scores on the scales.

The results of the classification consistency (i.e., agreement) analyses across all of the response instruction conditions are presented in Table 34. As can be in the Table, the highest levels of consistency were found between the Mantel-Haenszel procedure and the BIDR at all cut off levels and the Mantel-Haenszel procedure and the Mantel procedure. Although the consistency above chance was appreciable between the Mantel and the Mantel-Haenszel (37.7%), this was not the case for the consistency between the Mantel Haenszel and the BIDR. The above chance consistency ranged from 5% to less than 1%. In part, this is low consistency is a due to the fact that few individuals (< 2% at all cut off values) were classified as distorting by the BIDR and the Mantel-Haenszel. The majority of the consistency was achieved by similar classifications of individuals as not distorting their responses. Not surprisingly, the consistency of the BIDR with itself at the different cut off levels was very high using P and Kappa.

Much less consistency was found between the Mantel procedure and the BIDR. The values for P ranged between .818 and .765, but the values for Kappa were all negative. This indicates that most of the observed agreement may be due to chance alone. Again, the decisions were more consistent for the classification of not distorting. Less than 1% of the cases were classified as distorting by both techniques at all three cut offs scores of the BIDR.

The person fit index demonstrated little consistency with any of the response distortion detection techniques. In fact, Kappa was negative for every comparison. Thus, the observed agreement between the l_z and the other techniques may be due to chance.

To determine if the agreement varied between the response instruction conditions, the consistency of the decisions was examined within each condition. The results for the honest response condition are presented in Table 35. In this condition, the consistency between the Mantel-Haenszel and Mantel procedures, and the Mantel-Haenszel procedure and the BIDR with the top 5% of scores were higher than the overall consistency. Moreover, with the exception of the consistency of the Mantel and BIDR, the classification consistency was higher for all of the other measures. Thus, some of the measures tended to provide reasonably consistent decisions about the participants responding when classifying individuals who are responding honestly. The other measures still failed to provide decisions that agreed at above chance levels. The improvement in agreement is likely a result of each method classifying a larger number for individuals as not distorting their response across the conditions and that the majority of these individuals are concentrated in this condition.

The results of the classification consistency for the desirable response condition are summarized in Table 36. The consistency in this response condition was worse than the overall consistency for most comparisons. With the exceptions of the Mantel and BIDR, BIDR 15% and l_z, l_z and Mantel-Haenszel, and the 5% BIDR and 10% BIDR, the consistency was less. Thus, with a small exception, the different methods came to different conclusions about how the participants were responding when the participants are responding dishonestly. Therefore, the conclusion reached about any participant is dependent on which response distortion detection method is used and how the participant is responding.

Interestingly, in the no response instructions condition, the classification consistency improved for the Mantel procedure and the l_z index (see Table 37 for a summary). In particular, the Mantel procedure had much better agreement with the BIDR at all cut off level. The l_z index had above chance classification consistency for three of the five comparisons. The classification agreement between the other response distortion detection methods was actually less than the overall agreement.

In summary, the classification consistency between the different response distortion detection techniques was modest. No participants were identified as distorting their response by all six of the classification procedures. Only one participant was classified as distorting by 5 of the 6 methods, and only 9 were classified as distorting by 4 of the 6 methods. In the desirable response condition 62.6% of the participants were identified as distorting by at least one of the methods. In the no response instructions condition, 51.2% of the participants were identified as distorting their response by at least of the methods. Lastly, there were 172 participants that were not classified as distorting their responses by any of the methods.

Classification accuracy of DPF. Next, the decision accuracy of the classification decisions that were made using the DPF techniques was examined. If the DPF techniques are accurately classifying the participants, most of the individuals in the desirable response instruction condition should be identified as responding differentially across the items that can and cannot be distorted. Further, very few of the individuals in the honest response instruction group should be identified as responding differentially.

To determine the decision accuracy, the estimated probability of a consistent decision (see Equation 19) and Kappa were used. For this analysis a 2x2 contingency table was created with the decision for DPF (distorting vs. not distorting) versus the experimental response instructions (e.g., respond honestly vs. respond desirably). If the decisions from DPF the techniques match the known response instruction condition, the resulting probability should be close to 1.0. If the decisions do not match, the resulting probability should be close to 0.0.

The overall probability of correct classification using the Mantel-Haenszel procedure was .535. The probability of correctly classifying a participant as not distorting their responses was .476 and the probability of correctly classifying a participant as distorting their responses was .059. The false positive rate (3.1%) was relatively small, but the false negative rate (43.3%) was particularly high for the Mantel-Haenszel procedure. In fact, 88% of the participants in the desirable response condition were not identified as distorting their responses. Cohen's Kappa was .059 indicating the 5.9% increase in accuracy over chance when using the Mantel-Haenszel procedure to identify individuals who are distorting their responses.

The overall probability of correct classification using the Mantel procedure was .520. The probability of correctly classifying a participant as not distorting their responses was .445 and the probability of correctly classifying a participant as distorting their responses was .075. The false positive rate (6.3%) was relatively low and the false negative rate (41.7%) was particularly high. Again, over 80% of the participants that were in the desirable response condition were not identified as distorting their responses. Cohen's Kappa was .028 indicating the 2.8% increase in accuracy over chance when using the Mantel procedure to identify individuals who are distorting their responses.

Classification accuracy of the BIDR . The accuracy of the decisions of the BIDR was examined. The decision from the BIDR was compared with the honest and desirable response instruction conditions in a 2x2 contingency table. Again, each of the decisions made using the three different cut-off scores on the response distortion scales were used in these analyses. The decision accuracy was estimated using equation 19 and Kappa.

The overall probability of correctly classifying participants using the top 5% of BIDR score is .539. The probability of correctly classifying a participant as not distorting their responses was .476 and the probability of correctly classifying a participant as distorting their responses was .059. The false positive rate (2.4%) was low and the false negative rate (43.7%) was much higher. Once again, over 80% of the participants in the desirable response instructions condition were not identified as responding differentially on the items that could be distorted. Cohen's Kappa was .066 indicating the 6.6% increase in accuracy over chance when using the top 5% of scores on the BIDR to identify individuals who are distorting their responses.

Using the top 10% of the scores, the overall probability of a correct classification was .559. The probability of correctly classifying a participant as not distorting their responses was .480 and the probability of correctly classifying a participant, as distorting their responses, was .079. The false positive rate (2.8%) was low and the false negative rate (41.3%) was much higher. Over 80% of the participants in the desirable response instructions condition were not identified as responding differentially on the items that could be distorted. Cohen's Kappa was .107 indicating the 10.7% increase in accuracy over chance when using the top 10% of scores on the BIDR to identify individuals who are distorting their responses.

Using the top 15% of the scores, the overall probability of a correct classification increased to .567. The probability of correctly classifying a participant as not distorting their responses was .469 and the probability of correctly classifying a participant as distorting their responses was .098. The false positive rate (3.9%) was low and the false negative rate (39.4%) was much higher. Although the false negative rate was lower, 80% of the participants in the desirable response instructions condition were still not identified as responding differentially on the items that could be distorted. Cohen's Kappa was .124 indicating the 12.4% increase in accuracy over chance when using the top 15% of scores on the BIDR to identify individuals who are distorting their responses.

Classification accuracy of the person-fit index. The accuracy of the decisions of the person-fit index was examined. The decisions from the person-fit index were compared with the honest and desirable response instruction conditions in a 2x2 contingency table. The decision accuracy was estimated using equation 19 and Kappa.

The overall probability of correctly classifying participants when using a large negative l_z index on three on the five personality factor as a criterion for response distortion is .481. The probability of correctly classifying a participant as not distorting their responses was .327 and the probability of correctly classifying a participant as distorting their responses was .154. The false positive rate (18.4%) and the false negative rate (33.9%) were much high. Approximately 69% of the participants in the desirable response instructions condition were still identified as responding differentially on the items that could be distorted. Cohen's Kappa was -.045 indicating the classification accuracy was less than chance when using the l_z index.

Comparison of the classification accuracy of DPF, the person-fit index, and the BIDR. The accuracy of the decisions from each of the four response distortion detection techniques was compared with the results of probability of an accurate classification, Kappa, the true positive, true negative, false positive, and false negative rates.

The results are summarized in Table 38. As can be seen in the table, the technique with the highest accuracy is the BIDR when using the top 15% of the scores as a cut off for identifying those who are and are not distorting their responses. The next two most accurate methods were the BIDR at 10% and 5%. The differential person function analyses (i.e., the Mantel and Mantel-Haenszel) came in second in accuracy behind the BIDR. However, the differences were not dramatic, especially when using the top 5% of the scores on the BIDR. Interestingly, the decisions of the differential person functioning using the dichotomously scored data (i.e., the Mantel-Haenszel procedure) were more accurate than the decisions using the polytomously scored data (i.e., the Mantel procedure). Regardless of the distortion detection measure used, the accuracy above and beyond chance was not overwhelming. However, these type of results have been found in both lab and field studies (e.g., Brown & Harvey, 2003; Ferrando & Chico, 2001; Reise, 1995; Robie, et al., 2001; Zickar & Drasgow, 1996)

Additionally in Table 38, the true positive, false positive, true negative, and false negative rates for each method are presented. Interestingly, the least accurate measure, the l_z index, has the highest true positive. That is, it had the highest rate of correctly identified the individuals who distorted their responses. However, the high true positive rate comes at the cost of a high false positive rate. This method identifies many individuals as distorting their responses, but it is wrong about as much as it right. The true positive rate for the other methods ranged from approximately 5 to 10 percent. Thus, very few of the individuals that were distorting their response were identified as such. With the exception of the Mantel procedure the false positive rates for the remaining methods were between 2 and 4 percent.

What is also apparent is that all of the procedures failed to classify a large percentage of individuals as distorting their responses. For 5 of 6 methods, the false negative rate was about 40% or greater. Thus, there were many individuals who were distorting their responses, but slipped by the detection methods. However, 5 of the 6 methods did correctly classify over 40% of the participants as not distorting their response when they were in fact responding honestly. Therefore, when the result is negative (i.e., not distorting) the probability that an individual is or is not responding dishonestly is approximately equal.

Study 2 Discussion

The purpose of the second study was to determine if the differential person functioning technique (DPF) could be used to identify individuals who are distorting their responses on personality inventories. The DPF techniques were able to correctly identify those that were and were not distorting their response on the personality inventory. The accuracy of the DPF techniques was approximately equal to that of the response distortion scale, but the accuracy was still low. Further, the response distortion decisions from the DPF technique and the other techniques were not often in agreement.

What was clear is that distortion decisions depend on the response distortion detection method used and what is being detected. The classification consistency between the different methods was slightly above chance in several instances. Some of the inconsistency is likely a result of differences in how response distortion is assessed by the different methods. The BIDR is a special scale that does not use responses on the personality inventory to make a decision about an individual. Individuals who score more highly on this scale are believed to responding dishonestly on the personality inventory. Whereas the DPF techniques and the l_z index use only the responses and properties of the items on the personality inventory. Therefore, an individual's pattern of responding is criteria for determining whether or not that individual is responding dishonestly.

What was surprising was that there was more agreement between the BIDR and the DPF procedures than the DPF and the l_z index. One potential reason for the higher levels of agreement is that individuals who are responding differentially on the IPIP are also responding differentially on the BIDR. Although the BIDR was not included in the DIF analyses in Study 1, it is reasonable to speculate based on previous research with response distortion detection scales (e.g., Stark, et al., 2001) that some of the items on the BIDR could be distorted. If the same individuals were distorting their responses on the BIDR and the IPIP, a higher degree of consistency could be expected between these different types of detection methods than with more similar types of detection methods.

In this study, l_z index performed very poorly. It identified the largest number of individuals as distorting their response, but it was wrong much of the time. This could be potentially a result of low statistical power. The number of items needed for the l_z index to accurately detect response distortion, is often larger than the number of items on a typical personality inventory. Thus, statistical power may be relatively low (Reise 1995). Therefore, it is difficult to determine if the modest findings in this study are a result of low statistical power or a lack of sensitivity of the l_z index.

The lack of agreement is particularly troubling given that typically only one method is used in employment contexts to identify dishonest responders. A completely different decision about the same applicant can be made if one uses a different distortion detection method. With the potential need for justifying the use of one employment-screening device over another, the lack of consistency is a reasonable concern of organizational users of response distortion detection methods. Given that the methods are not interchangeable in terms of their outcomes, the question then becomes what method should be used?

Unfortunately, the results of the accuracy analyses do not provide a clear choice for a method of response distortion detection. From the results of this study, the choice of methods is between the BIDR and the DPF. The BIDR demonstrated slightly higher levels when the top 5% scores were classified as distorting their responses. Higher accuracy was achieved by lowering the cut off on the BIDR. However, many users would be reluctant to use cut off scores that are not very extreme. If being classified as dishonestly responding entails substantial consequences for applicants (e.g., being removed from the applicant pool or being denied employment), organizations will likely only target the most egregious offenders for fear of legal ramifications.

Although the DPF did not outperform the BIDR in terms of overall accuracy, it does offer several advantages over the BIDR that potentially outweigh the difference in accuracy. First, there is no additional administration time or cost to use the DPF techniques. Once properties of the items are known, the costs to apply the procedure are very minimal. It also avoids the addition of another selection instrument and the associated validation of the predictor in terms of its psychometric properties. In assessment contexts where time is a concern, avoiding an additional predictor is desirable, especially given that the administration of a personality inventory such as the IPIP or NEO-PI is time consuming. Further, the DPF has the advantage of being able to specify the items on which the individual was responding dishonestly. No other response distortion technique has this capability.

Also, the scores on the BIDR are substantially correlated with the scores on the personality inventory, regardless of the response instruction condition. Further, the correlations are in the desirable direction. For example, the correlation is positive with conscientiousness, but negative with neuroticism. Similar correlations have been demonstrated using different response distortion scales. Because these scales are correlated with the positive traits, it can be argued that they are too confounded to be a justifiable technique for identifying response distortion (Stark, et al. 2001). In sum, the DPF techniques are a very viable alternative to more traditional methods of detecting response distortion.

However, any potential use of the DPF techniques or other techniques in employment contexts are bound by the potentially low levels of accuracy. There are several potential reasons for the low levels of accuracy demonstrated by the methods in this study. The first is the manipulation used in this study was designed to induce more sophisticated response distortion that will be more similar to the type of response distortion used by job applicants. Therefore, the methods used in this study, including the

DPF, may not be sensitive enough to detect this type of response distortion. Studies using applicants and incumbents (e.g., Robbie, et al., 2001) or subtle response distortion instruction manipulations (e.g., Brown & Harvey, 2003; Paulhus, et al., 1995) have also found low levels of accuracy. Monte Carlo studies and studies using maximal response distortion instructions, have found much higher levels of accuracy (e.g., Zickar, Rosse, Levin, & Hulin, 1997). Thus, high levels of accuracy may only be achieved with these methods when a large degree of response distortion is occurring. Future experimental studies that systematically vary the degree of response distortion are needed to test the veracity of this proposition.

A second reason that is specific to the DPF is statistical power. In the differential person functioning analyses, the sample size is the number of items. In the analyses performed here, there was a sufficient number of items on the personality inventory, but the items were split very unequally between the two item groups (i.e., can or cannot be distorted). Indeed, there were about five times as many items that were classified as cannot be distorted than were classified as can be distorted. Ideally, one would like an equal number of items in each group.

The small number of items constrained the possibility of identifying individuals as distorting their responses. To be identified as distorting, individuals needed to demonstrate substantial differences in responding on the few items that were classified as can be distorted. Identifying individuals that distorted their response to a more moderate degree but on more items was not possible in this analysis. Thus, the results of the DPF techniques may be an underestimate of the actual degree of response distortion.

Regardless of the cause of the lack of consistency and accuracy of the response distortion detection techniques, the fact that they are not uniformly high does pose problems for their use as a screening device in employment and other contexts. However, these methods do have several implications for personality assessment and personnel decisions. Each is discussed below.

General Discussion

In this paper, two studies were conducted that examined intentional response distortion on a measure of the five-factor model of personality. The results of the studies indicate those who are responding dishonestly do respond to some items on this inventory differentially and that these individuals can be detected, albeit at low levels of accuracy. More central to the point of these studies is that the differential person functioning (DPF) techniques detected individuals who are responding dishonestly and did so at levels of accuracy that are comparable to other response distortion detection techniques. Further, these studies extended the use of DPF to situations involving polytomously scored data. The results using the polytomously score data were similar to those using dichotomously score data. The discussion of these studies focuses on the implications of these findings and the differential person functioning technique for personality assessment and the use of personality in personnel decisions.

Implications for Personality Assessment and Response Distortion Detection

The results of these studies have several implications for the assessment of personality. This study demonstrates that individuals when motivated to do so can respond dishonestly to the items on the IPIP. Seventeen percent of the items demonstrated DIF. Although the stability of the particular items demonstrating DIF needs cross validation, the results provide some initial indication of the items on the IPIP that may be improved. Ultimately, the best technique to prevent response distortion on measures of personality is to create items that are resistant to response distortion (Mount & Barrick, 1995). Even though some items may need to be improved, the majority of the items on this personality inventory did not function differentially when individuals are intentionally responding dishonestly. Thus, one could conclude that this personality inventory functions very well when individuals are intentionally responding dishonestly to the items on the inventory.

Moreover, these results were produced under conditions of “realistic” response distortion (Brown & Harvey, 2003; Paulhus, et al., 1995). In these studies, a very mild and more sophisticated form of response distortion was induced by the manipulation. This manipulation is a more ecologically valid replication of nature of response distortion that occurs in employment contexts than manipulations that induce maximal levels of response distortion. Thus, one caveat is that these results may or may not hold under maximal response distortion conditions or some other condition (Stark, et al., 2001). Specifically, many more items could potentially be identified as functioning differentially. In particular, this study found that the most important personality characteristic for predicting work behavior (i.e., conscientiousness) did not demonstrate the highest levels of response distortion. As noted above, this may be a result of a sophisticated type of response distortion. To date, there has only been one study that has attempted inducing differing levels of response distortion (Paulhus, et al., 1995). However, this study only examined mean difference between response instruction conditions. Clearly, other studies that use more extreme and varied forms of response distortion are needed to determine which item function differentially and under what conditions they do so. These studies would provide a better foundation for establishing the generalizability of the DIF results on the IPIP.

Additionally, a very conservative criterion was used to identify DIF. Only items that demonstrated DIF on 3 of 4 different methods were classified as functioning differentially. Because of the small sample size in Study 1, and the fact that it was the basis of the DPF analyses in Study 2, a more conservative criterion was appropriate. However, a similar number of items would be identified as functioning differentially if a more liberal criterion or only one measure of DIF was used. When considering the associated cost with modifying established measures of personality in terms of psychometric re-validation and potential consequences to individuals in employment contexts from the results of these analyses, conservative approaches to identify DIF in personality assessment, such as those used here, are advisable.

This study also provides additional support of the value of item response theory (IRT) in the assessment and measurement of personality. However, it also demonstrates that IRT may only be useful in certain respects. Specifically, IRT analyses are very useful for identifying the items that are functioning differentially when individuals are distorting their responses. The results from the IRT analyses demonstrated a substantial amount of agreement with more traditional methods of detecting response distortion. Because the IRT methods have a potentially better estimate of the latent trait than the traditional methods (Potenza & Dorans, 1995), they may be more sensitive to difference between groups in their responding to particular items. The drawback is that is an effect size measure for the IRT methods used here (e.g., DFIT) has not been developed. Thus, the degree of DIF cannot be estimated using the IRT methods.

Interestingly, the results of the IRT methods do not appear to depend on the IRT model used. There was substantial agreement between the results from logistic and graded response model. Despite the agreement, the graded response models were superior in model fit. This is counter to other studies that have found that logistic models fit personality data better than the graded response models (e.g., Chernyshenko, et al., 2001). However, none of these studies used the IPIP, so it is difficult to determine if the results of this study in this regard are unique to the IPIP.

Despite the potential value of IRT analyses in identifying differentially functioning items, the IRT analyses appear to have less value in identifying individuals that are responding dishonestly. In this study, as has been found in other studies (e.g., Ferrando & Chico, 2001; Reise, 1995; Robie, et al., 2001), the person-fit index, l_z, functioned exceptionally poorly. The accuracy of this index was less than chance. Moreover, it agreed with the other response distortion detection methods at less than chance levels. One counter argument to this conclusion is that it was an unfair comparison (Coopers & Richardson, 1986). Specifically, the criterion for distortion classifications was too stringent. Only individuals that demonstrated misfit on three of the five personality factors were classified as distorting. The other techniques did not need to demonstrate response distortion on multiple tests.

To determine if the criterion for l_z was causing the low levels of accuracy, the accuracy was estimated using a criterion of one personality factor. The probability of an accuracy decision was still low (54.4%). The true positive rate increased (39.8%), but so did the false positive rate (36.2%). These are essentially the same results as when using a criterion of three of five personality factors showing large negative values for l_z. Thus, this criterion was not resulting in unfair comparisons.

The differential person function approach has several implications of the assessment of personality and detection of response distortion, as described above. Additionally differential person functioning can contribute to some of the vexing issues in the response distortion literature. Namely, identifying individual difference variables related to the occurrence of response distortion, the ability to engage in response distortion, and the motivation to respond dishonestly. To date, clear and confound free relationships between individual difference variables and response distortion have yet to be established (McFarland & Ryan, 2000).

A problem with some of the studies investigating individual differences in response distortion is that the hypothesized individual difference variables (e.g., self-monitoring) are often confounded with the response distortion scale that was used to classify individuals as distorting their responses. The response distortion scales are, in turn, confounded to some degree with the positive personality factors. Alternatively, one could use differential person functioning to identify individuals who are distorting their response on a personality inventory and then investigate situational (e.g., employment status) and stable (e.g., ethnicity) differences between people that are predictive of the classification on the DPF.

For example in this study, one potential individual difference variable is the participants' sex. To determine if men or women were more likely to be identified as distorting their response, the proportion of men and women identified by the Mantel-

Haenszel and Mantel procedures was examined. On both procedures, about 10% of the men and 10% of the women were identified as responding differentially. Thus, sex was not an individual difference variable, in this study, that could distinguish between those responding honest versus dishonestly. However, other individual difference variable such as age, education, and possibly intelligence may distinguish between the individuals identified as distorting or not distorting tier responses. Future work identifying these individual difference variables is sorely needed.

Implications for Personnel Decisions

The results of these studies have several implications for the uses of response distortion techniques as part of personnel decisions. The detection rates of individuals who were distorting their responses were low for all four different methods of detection. The low detection rate has been replicated in several other studies (e.g., Robie, et al., 2001; Zickar & Drasgow, 1996). Because the detection rates are consistently low, using these methods as an exclusionary criterion in personnel and employment decisions is inappropriate. Because so many individuals would be missed using these methods, reacting in the extreme for the few individuals who would be identified may not be justifiable.

Although these response distortion detection methods should not be used to screen out individuals, they still have value in personnel and employment decisions. Differential person functioning, in particular, has many potential uses. First, even though the overall accuracy rate is too low to exclude individuals, the conditional probabilities among the participants responding in a desirable manner (i.e., the proportion of individuals in the desirable response condition that were correctly identified) were high enough to use them in situations where it is almost certain that most individuals are engaging in response distortion. Individuals who are identified can be required to retake the inventory or provide additional evidence validating their responses. Additionally, these techniques can be used with other screening devices such as biodata, interviews, or work samples to determine if individuals are responding differentially over different types of items.

Additionally differential person functioning has potential uses for other personnel functions. Specifically, differential person functioning can be applied to the evaluation of training programs and performance assessment. In terms of training, differential person functioning can be used to determine if individuals have differentially mastered aspects of training the content. For example, at the end of an employee orientation training program, it might be found that trainees are more successful on the items about the organization's history than on the items about the standard operating procedures. If the evaluation of proficiency is made on the total test score, then the lack of proficiency in a particular content domain will be overlooked. The results of DPF analyses can be used to target individuals for remedial training in particular areas or redesign portions of the training program to enhance the learning outcomes. Additionally, differential person functioning can be applied to performance assessment. For example, the performance ratings of supervisors could be examined to determine if they evaluate men and women differently. Alternatively, may be used to identify rater-training needs. For example, if a supervisor is rating high and low performing employees, differential person functioning across the two groups of employees should be occurring. A lack of differential person functioning may indicate an inability to distinguish between good and poor job performance.

Potential Limitations and Future Directions

Of course, this study has potential limitations. The most notable is that the sample size in both studies precluded the cross validation of the item parameter estimates. Therefore, the stability of these estimates cannot be empirically verified. However, the test of the assumptions made by IRT did not reveal serious violations of the assumptions and the results obtained here are comparable to large-scale studies using IRT with the IPIP (e.g., Trippe & Harvey, 2003). Thus, some degree of confidence in the stability is appropriate. Still, replication of this study with larger samples sizes is needed.

A second limitation was that only one personality inventory was used to investigate the consistency and accuracy of the response distortion detection methods. Thus, it is not possible to determine if the low level of accuracy and consistency of the DPF are unique to the IPIP. For example, the IPIP may contain many more or less items that are resistant to response distortion. Thus, it may be the case that accuracy and consistency would be higher or lower on other personality inventories that are much more or less resistant to response distortion.

A related concern is how well the results of this study generalize to actual applicants in employment settings. To increase the generalizability of the results of these studies, a more realistic form of response distortion was induced through the experimental manipulations. Yet, the equivalence of this type of response distortion to the response distortion, which occurs in employment contexts, is unknown. However, the purpose of these studies was not to establish generalizability. It was to establish the feasibility of differential person functioning as a method to detect response distortion, and therefore, generalizability is a lesser concern (Mook, 1983). Nonetheless, future work should be directed at examining differential person functioning using other samples such as applicants and incumbents or individuals with normal and abnormal personality.

A major assumption of all studies that use experimental manipulations of response instructions is that the participants actually did or able to follow the instructions. Outside of the significant differences between the response instruction conditions, an independent verification of the nature of the participants' responding was not collected. This would be particularly useful information given that there were several participants in the honest response condition that were classified as distorting their responses in Study 2. An independent verification would provide an opportunity to test if particular subjects failed to follow the instruction and the impact of this on the accuracy and consistency of the response distortion detection methods. However, even if this was collected, it is likely to be problematic. If individuals respond dishonestly to the personality items regardless of the response instruction, they may also respond dishonestly to questions about how they were responding. Thus, studies using subtle manipulation checks to ferret out individuals that are not responding in the instructed direction are needed.

A related issue is that instructions provided in manipulations did not provide the participants with contextual information (i.e., type of job, organizational culture) about the job they were applying for. Thus, the participants needed to rely on their own conceptions of what was an applicant that "looked good". That is, the applicants were not given a specific personality profile to use when responding to the items. Although excluding contextual information is common practice in response distortion research, a potential criticism is that each individual will have a different conception of what "looking good" means, which would undermine the approach used here.

However, this is generally not a concern when using the five-factor model of personality. Many authors argue that the five-factor model describes the ideal employee (e.g., Hogan & Roberts, 2001; McAdams, 1992). Moreover, there is meta-analytic evidence that differences in the effects of response distortion on the five personality factors are very small. That is, individuals responding dishonestly inflate their score by the same amount on each personality factors. This indicates that individuals either have a similar conceptualization of what "looking good" means or that it does not matter because the individuals are using similar response strategies (i.e., inflate all of my scores).

Further, the results indicate that the participants in this study had very similar conceptions of "looking good" as other samples of students and applicants. The rank order of the five personality factor scores is almost identical to other studies using subtle response distortion manipulations (e.g., Paulhus, et al., 1995), maximal response distortion manipulations (e.g., McFarland & Ryan, 2000), and applicants (e.g., Barrick & Mount, 1996). The rank order also agrees with Schmit and Ryan's (1993) pattern of factor loadings for the factor they labeled the "ideal employee" factor. Given the data and the operationalization of personality, the lack of contextual information is not a limitation of the results or response distortion detection approach advocated here.

Another potential limitation is that it is not possible to examine the quality of the response distortion. Some participants may not have been very good at distorting their response or responding honestly. Thus, for some participants it might be the classification based on the response instruction manipulation that is inaccurate and not the detection methods. On the other hand, it may be the individuals who are not very good at distorting their responses are the ones being identified by the response distortion detection methods. Given that the instructions asked the participants to engage in a particular type of response distortion (i.e., subtle distortion), it is likely that some participants may have been unable to do this very well. However, as noted above there are no individual difference variables that are consistently related to the ability to respond dishonestly or do it well. Thus, future work aimed at identifying individual difference variables related to the ability to engage in response distortion is clearly needed.

There are several additional directions in which the DPF techniques could be applied. The differential person functioning analyses in Study 2 used differentially functioning items as the basis of the item groups in the analyses. However, other item groupings are possible. The perception of whether an item can be distorted is one possible alternative. It may be the case that the perception of whether an item can be distorted is more indicative of distortion behavior than differentially functioning items. Another possibility is grouping items by their verifiability. For example, on measures of biodata, differential person functioning could be used to determine if individuals are responding differentially on the verifiable versus non-verifiable items.

Conclusions

Response distortion on measures of personality has been a long-standing concern of personal specialists and organizational users of personality inventories (Mount & Barrick, 1995; Whyte, 1956; Zickar & Drasgow, 1996). Correspondingly, several approaches have been developed to identify individuals who were distorting their responses on personality inventories in employment context. However, these approaches have not been overwhelmingly successful in identifying these individuals. This study is an initial attempt to introduce differential person functioning as an alternative method that can provide organizational decision makers with a set of tools to understand and manage response distortion on their personality based selection devices.

References

Allport, G. W. (1937). *Personality*. New York: Holt.

Alsmadi, A. A., & Alnabhan, M. (2000). Detecting differential person functioning using Mantel-Haenszel and logistic regression techniques. *Dirasat: Educational Sciences, 27*, 208-217.

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

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261-272.

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

Bartlett, C. J., & Doorley, R. (1967). Social desirability response differences under research, simulated selection, and faking instructional sets. *Personnel Psychology, 20*, 281-288.

Bass, B. M. (1957). Faking by sales applicants of a forced choice personality inventory. *Journal of Applied Psychology, 41*, 403-404.

Behling, O. (1998). Employee selection: Will intelligence and conscientiousness do the job? *Academy of Management Executive 12*, 77-86.

Bingham, W. V. (1922). Measures of occupational success. *Harvard Business Review, 5*, 1-10.

Birenbaum, M. (1985). Comparing the effectiveness of several IRT based appropriateness measures in detecting unusual response patterns. *Educational & Psychological Measurement, 45*, 523-534.

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

Borislow, B. (1958). The Edwards Personal Preference Schedule (EPPS) and fakability. *Journal of Applied Psychology, 42*, 22-27.

Brown, R. D., & Harvey, R. J. (2003). *Detecting personality test faking with appropriateness measurement: Fact of Fantasy?* Poster presented at the annual conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Camilli, G., & Shepard, L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.

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

Cattell, R. B. (1965). *The scientific analysis of personality*. London: Penguin.

Chernyshenko, O. S., Stark, S., Chan, K.-Y., Drasgow, F., & Williams, B. (2001). Fitting item response theory models to two personality inventories: Issues and insights. *Multivariate Behavioral Research, 36*, 523-562.

Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47*, 847-860.

Cohen, A. S., & Kim, S. (1998). An investigation of linking methods under the graded response model. *Applied Psychological Measurement, 22*, 116-130.

Cohen, A. S., Kim, S., & Baker, F. B. (1993). Detection of differntial item functioning in the graded response model. *Applied Psychological Measurement, 17*, 335-350.

Conn, S. R., & Reike, M. L. (eds.) (1994). *The 16PF fifth edition technical manual.* Champaign, IL: Institute for Personality and Ability Testing.

Cooper, W. H., & Richardson, A. J. (1986). Unfair comparisons. *Journal of Applied Psychology, 71*, 179-184.

Costa, P. T. J. (1996). Work and personality: Use of the NEO-PI-R in industrial/organisational psychology. *Applied Psychology: An International Review, 45*, 225-241.

Costa, P. T., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five-factor model. *Journal of Personality & Social Psychology*, 55, 258-265.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, and Winston.

Cronbach, L. J. (1946). Response sets and test validity. *Educational & Psychological Measurement, 6*, 475-494.

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

Crowne, D. P., & Marlow, D. (1964). *The approval motive*. New York: Wiley.

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

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35-66). Hillsdale, NJ: Erlbaum.

Dorans, N. J., & Schmitt, A. P. (1993). Constructed response and differential item functioning: A pragmatic perspective. In R. E. Bennett & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 135-166). Hillsdale, NJ: Erlbaum.

Dorans, N. J., & Schmitt, A. P., & Bleistein, C. A. (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement, 29*, 309-319.

Drasgow, F. (1982). Choice of test model for appropriateness measurement. *Applied Psychological Measurement, 6*, 297-308.

Drasgow, F., & Guertler, E. (1987). A decision-theoretic approach to the use of appropriateness measurement for detecting invalid test and scale scores. *Journal of Applied Psychology, 72*, 10-18.

Drasgow, F., & Hulin, C. L. (1990). Item response theory. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 577-636). Palo Alto, CA: Consulting Psychologists' Press.

Drasgow, F., & Lissak, R. (1983). Modified parallel analysis: A procedure fro examining the latent dimensionality of dichotomously scored item responses. *Journal of Applied Psychology, 68*, 363-373.

Drasgow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement, 7*, 189-199.

Drasgow, F., Levine, M. V., & McLaughlin, M. E. (1991). Appropriateness measurement for some multidimensional test batteries. *Applied Psychological Measurement, 15*, 171-191.

Drasgow, F., Levine, M. V., Tsien, S., Williams, B., Mead, A. (1995). Fitting Polytomous item response theory models to multiple choice tests. *Applied Psychological Measurement, 19*, 143-165.

Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behavior on a forced-choice self-description checklist. *Personnel Psychology, 15*, 13-24.

Dwight, S. A., & Alliger, G. M. (1997). Reactions to overt integrity test items. *Educational & Psychological Measurement, 57*, 937-948.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* Ft Worth, TX: Dryden Press.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*, 155-166.

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

Psychology, 86, 122-133.

Ellis, A. (1946). The validity of personality questionnaires. *Psychological Bulletin, 43*, 385-440.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, NJ: Lawrence Erlbaum.

Eysenck, H. (1970). *The structure of human personality* (3rd ed.). London: Meuthuen.

Ferrando, P. J., & Chico, E. (2001). Detecting dissimulation in personality test scores: A comparison between person-fit indices and detection scales. *Educational & Psychological Measurement, 61*, 997-1012.

Ferrando, P. J., & Lorenzo, U. (2000). Wperfit: A program for computing parametric person-fit statistics and plotting person response curves. *Eductional and Psychological Measurement*, 60, 479-487.

Flowers, C. P., Oshima, T. C., Raju, N. (1999). A Description and Demonstration of the Polytomous-DFIT Framework. *Applied Psychological Measurement, 23*, 309–326.

Funder, D. C. (2000). Personality. *Annual Review of Psychology, 52*, 197-221.

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

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

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I.

Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions.* Mahwah, NJ: Lawrence Erlbaum.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota multiphasic personality inventory (Rev. ed., 2nd printing).* Minneapolis, MN: University of Minnesota Press.

Hogan, R. T. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2 ed., Vol. 2, pp. 873-919). Palo Alto, CA: Consulting Psychologists Press.

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

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

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity.* (pp. 129-145).

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis.

Psychometrika, 30, 179-185.

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

Hough, L. M. (1997). The millennium for personality psychology: New horizons or good ole daze. *Applied Psychology: An International Review, 47*, 233-261.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11*, 209-244.

Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future--remembering the past. *Annual Review of Psychology, 51*, 631-664.

Hough, L. M., & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. In K. R. Murphy (Ed.), Individual differences and behavior in organizations (pp. 31-88). San Francisco: Jossey-Bass.

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

Humphreys, L. G., & Montanelli, R. G. (1973). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research, 10*, 193-205.

Johanson, G., & Alsmadi, A. (2002). Differential person functioning. *Educational & Psychological Measurement, 62*, 435-443.

Johanson, G. A., & Osborn, C. J. (2000). *Acquiescence as Differential Person Functioning.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions*. New York: Springer-Verlag.

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

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*, 530-541.

Kaufman, J. D., & Dunlap, W. P. (2000). Determining the number of factors to retain: A Windows-based FORTRAN-IMSL program for parallel analysis. *Journal of Behavior Research Methods, Instruments & Computers*, 32, 389-395.

Kanfer, R., & Ackerman, P. L. (2000). Individual differences in work motivation: Further explorations of a trait framework. *Applied Psychology: An International Review, 49*, 470-482.

Kelly, E. L., Miles, C. C., & Terman, L. M. (1936). Ability to influence one's score on a typical pencil-and-paper test of personality. *Character & Personality; A Quarterly for Psychodiagnostic & Allied Studies, 4*, 206-215.

Kirchner, W. K., Dunnette, M. D., & Mousley, N. (1960). Use of the Edwards Personal Preference Schedule in the selection of salesmen. *Personnel Psychology, 13*, 421-424.

Kiriscki, L., Hsu, T., & Yu, L. (2002). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied*

Psychological Measurement, 25, 146-162.

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

Kroger, R. O., & Turnbull, W. (1975). Invalidity of validity scales: The case of the MMPI. *Journal of Consulting & Clinical Psychology, 43*, 48-55.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin, 107*, 33-47.

Levine, M. V., & Drasgow, F. (1988). Optimal appropriateness measurement. *Psychometrika, 53*, 161-176.

Levine, M. V., & Rubin, D. B. (1979). Measuring the appropriateness of multiple-choice scores. *Journal of Educational statistics, 4*, 269-290.

Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.

MacKinnon, D. W. (1944) The structure of personality. In J. McVicker Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1., pp. 3-48). New York; Ronald Press.

Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58, 690-700.

Mantel, N., & Haenszel, W. (1958). Statistical aspects of the analysis of data from retrospective studies. *Journal of the National Cancer Institute, 22*, 719-748.

Marco, G. L. (1977). Item characteristic curve solutions to three intractable testing problems. *Journal of Educational Measurement, 14*, 139-160.

Matthews-López, J. L., Larkin, K., & Johanson, G. (2002). Using differential person functioning to detect aberrant response patterns in a standard-setting session for teacher licensure. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

McAdams, D. P. (1992). The five-factor model of personality: A critical appraisal. *Journal of Personality, 60*, 329-361.

McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting & Clinical Psychology, 51*, 882-888.

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

McDougall, W. (1932). On the words character and personality. *Character and Personality*, 1, 3-16.

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

Michaelis, W., & Eysenck, H. J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real-life motivation. *Journal of Genetic Psychology, Vol. 118*, 223-234.

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

Mislevy, R. J., & Bock, R. D. (1991). BILOG users' guide. Chicago: Scientific Software.

Montag, I., & Comrey, A. L. (1990). Stability of major personality factors under changing motivational conditions. *Journal of Social Behavior & Personality, 5*, 265-274.

Montag, I., & Levin, J. (1994). The five-factor personality model in applied settings. *European Journal of Personality, 8*, 1-11.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*, 379-387.

Mount, M. K., & Barrick, M. R. (1995). The big-five personality dimensions: Implications for research and practice in human resource management. In G. R. Ferris (Ed.), *Research in Personnel and Human Resources Management* (Vol. 13, pp. 153-200). Greenwich, CN: JAI Press Inc.

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

Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11*, 245-269.

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

Orpen, C. (1971). The fakability of the Edwards Personal Preference schedule in personnel selection. *Personnel Psychology, 24*, 1-4.

Parsons, C. K., & Hulin, C. L. (1982). An empirical comparison of item response theory and hierarchical factor analysis in applications to job satisfaction. *Journal of Applied Psychology, 67*, 826-834.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality & Social Psychology, 46*, 598-609.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, & P. R. Shaver (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1., pp. 17-59).

Paulhus, D. L., Bruce, M. N., & Trapnell, P. D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality & Social Psychology Bulletin, 21*, 100-108.

Raju, N. S. (1998). DFIT4: A fortran program for calculating DIF/DTF [computer program]. Chicago: Illinois Institute of Technology.

Raju, N. S., van der Linden, W. J., & Fleer, P. F. (1995). IRT-based internal measure of differential functioning of items and tests. *Applied Psychological Measurement, 19,* 353–368.

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics, 4*, 207-230.

Reise, S. P. (1995). Scoring method and the detection of person misfit in a personality assessment context. *Applied Psychological Measurement, 19*, 213-229.

Reisse, S. P., & Waller, N. G. (1993). Traitedness and the assessment of response pattern scalability. *Journal of Personality and Social Psychology*, 65, 143-151.

Reisse, S. P., & Yu, J. (1990). Parameter recovery in the graded response model using MULTILOG. *Journal of Educational Measurement*, 27, 133-144.

Robie, C., Zickar, M. J., & Schmit, M. J. (2001). Measurement equivalence between applicant and incumbent groups: An IRT analysis of personality scales. *Human Performance, 14*, 187-207.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology, 83*, 634-644.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist, 56*, 302-318.

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

Salgado, J. F. (1998). Big Five personality dimensions and job performance in Army and civil occupations: A European perspective. *Human Performance, 11*, 271-288.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement, Vol. 34*, 100.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

Schmit, M. J., & Ryan, A. M. (1992). Test-taking dispositions: A missing link? *Journal of Applied Psychology, 77*, 629-637.

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology, 78*, 966-974.

Schmitt, N., & Chan, D. (1998). *Personnel selection: A theoretical approach.* Thousand Oaks, CA: Sage.

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

Schmitt, N., Chan, D., Sacco, J. M., McFarland, L. A., & Jennings, D. (1999). Correlates of person fit and effect of person fit on test validity. *Applied Psychological Measurement, 23*, 41-53.

Schwab, D. P. (1971). Issues in response distortion studies of personality inventories: A critique and replicated study. *Personnel Psychology, Vol. 24*, 637-647.

Skrzypek, G. J., & Wiggins, J. S. (1966). Contrasted groups versus repeated measurement designs in the evaluation of social desirability scales. *Educational & Psychological Measurement, 26*, 131-138.

Smith, L. L. (2002). On the usefulness of item bias analysis to personality psychology. *Personality & Social Psychology Bulletin, 28*, 754-763.

Somes, G. W. (1986). Te generalized Mantel-Haenszel statistic. *The American Statistician*, 40, 106-108.

Stark, S., Chernyshenko, O. S., Chan, K.-Y., Lee, W. C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of*

Applied Psychology, 86, 943-953.

Stricker, L. J. (1963). Acquiescence and social desirability response styles, item characteristics, and conformity. *Psychological Reports, 12*, 319-341.

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

Thissen, D. (2003). MULTILOG users' guide. Chicago: Scientific Software.

Thissen, D., & Steinberg, L. (1988). Data analysis using item response theory. *Psychological Bulletin, 104*, 385-395.

Trippe, D. M., & Harvey, R. J. (2003). *Item response theory of the IPIP big-five scales*. Poster presented at the annual conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings*. United States Air Force ASD Technical Report 61-97.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational & Psychological Measurement, 59*, 197-210.

Waters, L. K. (1965). A note on the "fakability" of forced-choice scales. *Personnel Psychology, 18*, 187-191.

White, L. A., Nord, R. D., Mael, F. A., Young, M. C. (1993). The assessment of Background and Life Experiences (ABLE). In T. Trent & J. H. Laurence (Eds.), Adaptability screening for the armed forces (pp. 101-162). Washington, D. C.: Office of the Assistant Secretary of Defense.

Whyte, W. H. (1956). *The organization man*. New York: Simon and Schuster.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review, 12*, 250-264.

Zickar, M. J. (2000). Modeling faking on personality tests. In D. R. Ilgen, & C. L. Hulin, (Eds.), *Computational modeling of behavior in organizations: The third scientific discipline.* (pp. 95-113). Washington, D.C.: American Psychological Association.

Zickar, M. J. (2001). Conquering the next frontier: Modeling personality data with item response theory. In B. W. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 141-160). Washington, DC: American Psychological Association.

Zickar, M. J., & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement, 20*, 71-87.

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

Zickar, M. J., & Ury, K. L. (2002). Developing an interpretation of item parameters for personality items: Content correlates of parameter estimates. *Educational & Psychological Measurement, 62*, 19-31.

Zickar, M. J., Rosse, J., Levin, R., & Hulin, C. L. (1997). *Modeling the effects of faking on a personality test*. Unpublished manuscript.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement,* 30, 233–251.

Appendix A

The International Personality Item Pool

On the following pages, there are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes ***you***. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then fill in the bubble that corresponds to the number on the scale.

Response Options

1: Very Inaccurate **2**: Moderately Inaccurate **3**: Neither Inaccurate nor Accurate
4: Moderately Accurate **5**: Very Accurate

1) Excel in what I do.
2) Become overwhelmed by events.
3) Can handle a lot of information.
4) Stick to the rules.
5) Sympathize with the homeless.
6) Believe in one true religion.
7) Leave my belongings around.
8) Make people feel welcome.
9) Distrust people.
10) Cheer people up.
11) Love surprise parties.
12) Jump into things without thinking.
13) Believe that others have good intentions.
14) Cheat to get ahead.
15) Have a good word for everyone.
16) Believe that we should be tough on crime.
17) Am not easily affected by my emotions.

18) Often feel uncomfortable around others.

19) Love flowers.

20) Am indifferent to the feelings of others.

21) Can manage many things at the same time.

22) Am not highly motivated to succeed.

23) Turn my back on others.

24) Try to lead others.

25) Have a high opinion of myself.

26) Try not to think about the needy.

27) Am easy to satisfy.

28) Like to visit new places.

29) Experience very few emotional highs and lows.

30) Can handle complex problems.

31) Believe that people are basically moral.

32) Do things I later regret.

33) Want everything to be "just right."

34) Handle tasks smoothly.

35) Am able to control my cravings.

36) Feel that my life lacks direction.

37) Use others for my own ends.

38) Love action.

39) Don't like crowded events.

40) Panic easily.

41) Love life.

42) Trust others.

43) Am easily intimidated.

44) Don't know why I do some of the things I do.

45) Demand quality.

46) Stick to my chosen path.

47) Pay my bills on time.

48) Rarely overindulge.

49) Am very pleased with myself.

50) Carry out my plans.

51) Interested in many things.

52) Enjoy being part of a group.

53) Have a low opinion of myself.

54) See beauty in things that others might not notice.

55) Willing to try anything once.

56) Like order.

57) Break my promises.

58) Rarely notice my emotional reactions.

59) Avoid mistakes.

60) Seldom get mad.

61) Remain calm under pressure.

62) Am relaxed most of the time.

63) Am not really interested in others.

64) Often eat too much.

65) Prefer variety to routine.

66) Avoid philosophical discussions.

67) Keep others at a distance.

68) Believe that I am better than others.

69) Dislike changes.

70) Take no time for others.

71) Love to help others.

72) Often feel blue.

73) Can't stand weak people.

74) Don't like to draw attention to myself.

75) Do not like concerts.

76) Do a lot in my spare time.

77) Have difficulty understanding abstract ideas.

78) Am a creature of habit.

79) Feel desperate.

80) Am not bothered by messy people.

81) Look at the bright side of life.

82) Get overwhelmed by emotions.

83) Am not interested in other people's problems.

84) Use flattery to get ahead.

85) Spend time reflecting on things.

86) Believe in the importance of art.

87) Believe laws should be strictly enforced.

88) Do just enough work to get by.

89) Indulge in my fantasies.

90) Like to begin new things.

91) Get caught up in my problems.

92) Wait for others to lead the way.

93) Anticipate the needs of others.

94) Have a vivid imagination.

95) Have a lot of fun.

96) Turn plans into actions.

97) Find it difficult to get down to work.

98) Tell the truth.

99) Feel comfortable with myself.

100) Get angry easily.

101) Get chores done right away.

102) Waste my time.

103) Enjoy the beauty of nature.

104) Feel others' emotions.

105) Like a leisurely lifestyle.

106) Do things according to a plan.

107) Seldom feel blue.

108) Rush into things.

109) Can't stand confrontations.

110) Radiate joy.

111) Seek to influence others.

112) Have little to say.

113) Am passionate about causes.

114) Am wary of others.

115) Get others to do my duties.

116) Like to act on a whim.

117) Know how to cope.

118) Believe in human goodness.

119) Dislike myself.

120) Try to follow the rules.

121) Act without thinking.

122) Find it difficult to approach others.

123) Am not bothered by difficult social situations.

124) React quickly.

125) Experience my emotions intensely.

126) Obstruct others' plans.

127) Pretend to be concerned for others.

128) Listen to my conscience.

129) Get back at others.

130) Would never cheat on my taxes.

131) Dislike talking about myself.

132) Trust what people say.

133) Easily resist temptations.

134) Am not easily disturbed by events.

135) Am calm even in tense situations.

136) Am hard to get to know.

137) Rarely get irritated.

138) Don't worry about things that have already happened.

139) Am always prepared.

140) Enjoy wild flights of fantasy.

141) Have a rich vocabulary.

142) Enjoy examining myself and my life.

143) Contradict others.

144) Do not enjoy going to art museums.

145) Like to solve complex problems.

146) Seldom joke around.

147) Love a good fight.

148) Avoid crowds.

149) Seldom toot my own horn.

150) Dislike loud music.

151) Set high standards for myself and others.

152) Let things proceed at their own pace.

153) Am afraid that I will do the wrong thing.

154) Insult people.

155) Can talk others into doing things.

156) Fear for the worst.

157) Come up with good solutions.

158) Choose my words with care.

159) Complete tasks successfully.

160) Have difficulty imagining things.

161) Adapt easily to new situations.

162) Avoid contacts with others.

163) Often forget to put things back in their proper place.

164) Am always on the go.

165) Believe that people are essentially evil.

166) Can't make up my mind.

167) Love order and regularity.

168) Believe people should fend for themselves.

169) Am not easily bothered by things.

170) Do not like poetry.

171) Love large parties.

172) Enjoy being part of a loud crowd.

173) Stumble over my words.

174) Don't like the idea of change.

175) Believe that we coddle criminals too much.

176) Have little to contribute.

177) Don't understand things.

178) Laugh my way through life.

179) Consider myself an average person.

180) Seldom get emotional.

181) Act comfortably with others.

182) Keep my cool.

183) Break rules.

184) Am often down in the dumps.

185) Am always busy.

186) Misrepresent the facts.

187) Express childlike joy.

188) Get stressed out easily.

189) Take advantage of others.

190) Hold back my opinions.

191) Know how to get around the rules.

192) Get irritated easily.

193) Would never go hang gliding or bungee jumping.

194) Avoid difficult reading material.

195) Prefer to stick with things that I know.

196) Make friends easily.

197) Like to take my time.

198) Am not easily amused.

199) Am concerned about others.

200) Laugh aloud.

201) Do more than what's expected of me.

202) Tend to vote for liberal political candidates.

203) Am afraid to draw attention to myself.

204) Make rash decisions.

205) Warm up quickly to others.

206) Rarely complain.

207) Don't see the consequences of things.

208) Like to stand during the national anthem.

209) Feel comfortable around people.

210) Know how to get things done.

211) Love to daydream.

212) Amuse my friends.

213) Postpone decisions.

214) Put little time and effort into my work.

215) Seldom daydream.

216) Lose my temper.

217) Never spend more than I can afford.

218) Have frequent mood swings.

219) Leave a mess in my room.

220) Love excitement.

221) Believe that criminals should receive help rather than punishment.

222) Go on binges.

223) Try to understand myself.

224) Love to read challenging material.

225) Dislike being the center of attention.

226) Love to eat.

227) Yell at people.

228) Do not like art.

229) Prefer to be alone.

230) Enjoy being reckless.

231) Go straight for the goal.

232) Tend to dislike soft-hearted people.

233) Never splurge.

234) Have difficulty starting tasks.

235) Am not interested in theoretical discussions.

236) Know the answers to many questions.

237) Seek quiet.

238) Hold a grudge.

239) Do not have a good imagination.

240) Make myself the center of attention.

241) Get upset easily.

242) Hate to seem pushy.

243) Only feel comfortable with friends.

244) Like to take it easy.

245) Am not bothered by disorder.

246) Talk to a lot of different people at parties.

247) Am not embarrassed easily.

248) Like to tidy up.

249) Boast about my virtues.

250) Believe in an eye for an eye.

251) Suspect hidden motives in others.

252) Am not interested in abstract ideas.

253) Work hard.

254) Am comfortable in unfamiliar situations.

255) Readily overcome setbacks.

256) Get to work at once.

257) Seek danger.

258) Tend to vote for conservative political candidates.

259) Think that all will be well.

260) Believe that there is no absolute right or wrong.

261) Have a sharp tongue.

262) Think highly of myself.

263) Am afraid of many things.

264) Believe that too much tax money goes to support artists.

265) Am often in a bad mood.

266) Want to be left alone.

267) Keep my promises.

268) Don't understand people who get emotional.

269) Make people feel uncomfortable.

270) Seek adventure.

271) Do the opposite of what is asked.

272) React slowly.

273) Keep in the background.

274) Plunge into tasks with all my heart.

275) Am attached to conventional ways.

276) Suffer from others' sorrows.

277) Dislike new foods.

278) Feel that I'm unable to deal with things.

279) Value cooperation over competition.

280) Look down on others.

281) Am not easily annoyed.

282) Seldom get lost in thought.

283) Enjoy thinking about things.

284) Put people under pressure.

285) Like music.

286) Misjudge situations.

287) Worry about things.

288) Involve others in what I am doing.

289) Take charge.

290) Often make last-minute plans.

291) Like to get lost in thought.

292) Am able to stand up for myself.

293) Do not enjoy watching dance performances.

294) Do crazy things.

295) Need a push to get started.

296) Am sure of my ground.

297) Feel sympathy for those who are worse off than myself.

298) Act wild and crazy.

299) Start tasks right away.

300) Take control of things.

Appendix B

Experimental Protocol for Session 1 of Study 1

Honest Condition

Today, you will complete a personality inventory. This is a two-part study and you will need to sign-up for a second session to receive your second point. When you complete the inventory, I will schedule you for your second session.

Please mark your responses only on the answer forms. You will mark your responses to question 1-200 on the first form and responses to 201-300 on the second form. Do not answer any of the items after 300.

On the first form only, please indicate you gender and year of birth. In the place for your grade, please mark "1" if you have ever held a job and "0" if you have never held job. In the section for special codes please put a **1** in the box labeled "M" on both forms. Please put the identification number on the back of your experiment card so that you can use it again for the second session. It is very important that you write this number down so you do not forget.

Please complete this personality inventory as honestly as you can. The results will be completely anonymous and will be used for research purposes only. It is *very* important that you respond to this survey by describing yourself as you really are and not as you want to be or as you want others to see you. Please look at each side of the page as you fill this out. You can start whenever you are ready. Please let me know if you have any questions.

Desirable Condition

Today, you will complete a personality inventory. This is a two-part study and you will need to sign-up for a second session to receive your second point. When you complete the inventory, I will schedule you for your second session.

Please mark your responses only on the answer forms. You will mark your responses to question 1-200 on the first form and 201-300 on the second form. Do not answer any of the items after 300.

On the first form only, please indicate you gender and year of birth. In the place for your grade, please mark "1" if you have ever held a job and "0" if you have never held job. In the section for special codes please put a **2** in the box labeled "M" on both forms. Please put the identification number on the back of your experiment card so that you can use it again for the second session. It is very important that you write this number down so you do not forget.

Please complete the personality inventory as if you were applying for a job you really want. To increase your chances of being hired, you should respond in way that will make you look good to the organization. But, do not respond in a way that will look like you were obviously faking your responses. Please look at each side of the page as you fill this out. You can start whenever you are ready. Please let me know if you have any questions.

Appendix C

Experimental Protocol for Session 2 of Study 1

Honest Condition

Today, you will complete a personality inventory. Please mark your responses only on the answer forms. You will mark your responses to question 1-200 on the first form and responses to 201-300 on the second form. Do not answer any of the items after 300.

On both forms, please mark the 7-digit identification number that you used in the first session in the boxes for the identification number. You probably recorded this number on the back of your experiment card. Please put the id number on both forms. In the section for special codes please put a **1** in the box labeled "M" on both forms.

Please complete this personality inventory as honestly as you can. The results will be completely anonymous and will be used for research purposes only. It is *very* important that you respond to this survey by describing yourself as you really are and not as you want to be or as you want others to see you.

Please start with the answer form that has 12 in the 1st two boxes of the special codes section. Please look at each side of the page as you fill this out. You can start whenever you are ready. Please let me know if you have any questions.

Desirable Condition

Today, you will complete a personality inventory. Please mark your responses only on the answer forms. You will mark your responses to question 1-200 on the first form and responses to 201-300 on the second form. Do not answer any of the items after 300.

On both forms, please mark the 7-digit identification number that you used in the first session in the boxes for the identification number. You probably recorded this number on the back of your experiment card. Please put the id number on both forms. In the section for special codes please put a **2** in the box labeled "M" on both forms.

Please complete the personality inventory as if you were applying for a job you really want. To increase your chances of being hired, you should respond in way that will make you look good to the organization. But, do not respond in a way that will look like you were obviously faking your responses.

Please start with the answer form that has 12 in the 1st two boxes of the special codes section. Please look at each side of the page as you fill this out. You can start whenever you are ready. Please let me know if you have any questions.

Appendix D

Balanced Inventory of Desirable Responding

Using the scale below as a guide, mark the number on the Scantron for each statement to indicate how much you agree with it.

1	2	3	4	5	6	7
NOT TRUE			SOMEWHAT TRUE			VERY TRUE

1. I sometimes tell lies if I have to.
2. I never cover up my mistakes.
3. There have been occasions when I have taken advantage of someone.
4. I never swear.
5. I sometimes try to get even rather than forgive and forget.
6. I always obey laws, even if I'm unlikely to get caught.
7. I have said something bad about a friend behind his or her back.
8. When I hear people talking privately, I avoid listening.
9. I have received too much change from a salesperson without telling him or her.
10. I always declare everything at customs.
11. When I was young I sometimes stole things.
12. I have never dropped litter on the street.
13. I sometimes drive faster than the speed limit.
14. I never read sexy books or magazines.
15. I have done things that I don't tell other people about.
16. I never take things that don't belong to me.
17. I have taken sick-leave from work or school even though I wasn't really sick.
18. I have never damaged a library book or store merchandise without reporting it.
19. I have some pretty awful habits.
20. I don't gossip about other people's business.

Appendix E

Marlowe-Crowne Social Desirability Scale

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true or false as it pertains to you personally. Please mark a 1 for false and a 2 for true on your Scantron answer form.

1. Before voting I thoroughly investigate the qualifications of all the candidates.
2. I never hesitate to go out of my way to help someone in trouble.
3. It is sometimes hard for me to go on with my work if I am not encouraged.
4. I have never intensely disliked anyone.
5. On occasion I have had doubts about my ability to succeed in life.
6. I sometimes feel resentful when I don't get my way.
7. I am always careful about my manner of dress.
8. My table manners at home are as good as when, I eat out in a restaurant.
9. If I could get into a movie without paying and be sure I was not seen, I would probably do it.
10. On a few occasions, I have given up doing something because I thought too little of my ability.
11. I like to gossip at times.
12. There have been times when I felt like rebelling against people in authority even though I knew they were right.
13. No matter who I'm talking to, I'm always a good listener.
14. I can remember "playing sick" to get out of something.
15. There have been occasions when I took advantage of someone.
16. I'm always willing to admit it when I make a mistake.
17. I always try to practice what I preach.
18. I don't find it particularly difficult to get along with loud mouthed, obnoxious people.
19. I sometimes try to get even, rather than forgive and forget.
20. When I don't know something I don't at all mind admitting it.

21. I am always courteous, even to people who are disagreeable.
22. At times I have really insisted on having things my own way.
23. There have been occasions when I felt like smashing things.
24. I would never think of letting someone else be punished for my wrongdoings.
25. I never resent being asked to return a favor.
26. I have never been irked when people expressed ideas very different from my own.
27. I never make a long trip without checking the safety of my car.
28. There have been times when I was quite jealous of the good fortune of others.
29. I have almost never felt the urge to tell someone off.
30. I am sometimes irritated by people who ask favors of me.
31. I have never felt that I was punished without cause.
32. I sometimes think when people have a misfortune they only got what they deserved.
33. I have never deliberately said something that hurt someone's feelings.

Appendix F

Experimental Protocol for Study 2

Honest Condition

Today, you will complete a personality inventory. Please mark your responses only on the answer forms. You will mark your responses to question 1-200 on the first blue form, 201-320, on the green form, and your responses to the remaining items on the second blue form.

On the first blue form only, please indicate you gender in the section labeled sex. Please bubble in your **actual age** in the boxes for year of birth. In the section labeled grade/EDUC, please mark "1" if you have ever held a job and "0" if you have never held job. In the section for special codes please bubble in the number **1** in the box labeled "L".

Please complete this personality inventory as honestly as you can. The results will be completely anonymous and will be used for research purposes only. It is *very* important that you respond to this survey by describing yourself as you really are and not as you want to be or as you want others to see you. Again, please answer as honestly as you can.

Please start with the first blue answer form. Please look at each side of the page as you fill this out and use both sides of the answer sheet. You can start whenever you are ready. Please let me know if you have any questions.

When you are finished, please check that you have completely erased any stray marks or responses that you have changed.

Desirable Condition

Today, you will complete a personality inventory. Please mark your responses only on the answer forms. . You will mark your responses to question 1-200 on the first blue form, 201-320, on the green form, and your responses to the remaining items on the second blue form.

On the first blue form only, please indicate you gender in the section labeled sex. Please bubble in your **actual age** in the boxes for year of birth. In the section labeled grade/EDUC, please mark "1" if you have ever held a job and "0" if you have never held job. In the section for special codes please bubble in the number **2** in the box labeled "L".

Please complete the personality inventory as if you were applying for a job you really want. To increase your chances of being hired, you should respond in way that will make you look good to the organization. But, do not respond in a way that will look like you were obviously faking your responses.

Please start with the first blue answer form. Please look at each side of the page as you fill this out and use both sides of the answer sheet. You can start whenever you are ready. Please let me know if you have any questions.

When you are finished, please check that you have completely erased any stray marks or responses that you have changed.

No Instructions Condition

Today, you will complete a personality inventory. Please mark your responses only on the answer forms. . You will mark your responses to question 1-200 on the first blue form, 201-320, on the green form, and your responses to the remaining items on the second blue form.

On the first blue form only, please indicate you gender in the section labeled sex. Please bubble in your **actual age** in the boxes for year of birth. In the section labeled grade/EDUC, please mark "1" if you have ever held a job and "0" if you have never held job. In the section for special codes please bubble in the number **3** in the box labeled "L".

Please complete the personality inventory as if you were applying for a job.

Please start with the first blue answer form. Please look at each side of the page as you fill this out and use both sides of the answer sheet. You can start whenever you are ready. Please let me know if you have any questions.

When you are finished, please check that you have completely erased any stray marks or responses that you have changed.

Table 1

Internal Consistency Reliability for the Personality Factors and Sub-Facets on the IPIP from both Administrations of the Personality Inventory and Both Response Instruction Conditions

Factor	Sub-factor	First	Second	Honest	Desirable
Conscientiousness		**0.949**	**0.961**	**0.939**	**0.958**
	1	0.799	0.844	0.770	0.833
	2	0.789	0.806	0.796	0.788
	3	0.800	0.871	0.825	0.840
	4	0.836	0.867	0.835	0.849
	5	0.889	0.907	0.867	0.915
	6	0.806	0.826	0.774	0.815
Extroversion		**0.910**	**0.922**	**0.923**	**0.900**
	1	0.826	0.878	0.866	0.840
	2	0.768	0.831	0.816	0.799
	3	0.781	0.783	0.799	0.778
	4	0.731	0.438	0.557	0.553
	5	0.678	0.724	0.735	0.700
	6	0.746	0.787	0.802	0.712
Neuroticism		**0.956**	**0.954**	**0.941**	**0.954**
	1	0.859	0.849	0.833	0.842
	2	0.906	0.880	0.892	0.882
	3	0.853	0.888	0.864	0.878
	4	0.842	0.825	0.834	0.819
	5	0.752	0.776	0.729	0.754
	6	0.857	0.837	0.823	0.847
Openness		**0.863**	**0.836**	**0.871**	**0.832**
	1	0.784	0.793	0.825	0.757
	2	0.817	0.816	0.842	0.806
	3	0.772	0.737	0.769	0.724
	4	0.664	0.687	0.654	0.717
	5	0.781	0.783	0.761	0.791
	6	0.669	0.653	0.696	0.608

Agreeableness		**0.903**	**0.918**	**0.916**	**0.904**
	1	0.822	0.852	0.850	0.840
	2	0.815	0.831	0.817	0.810
	3	0.754	0.822	0.792	0.786
	4	0.676	0.721	0.701	0.669
	5	0.705	0.668	0.706	0.643
	6	0.633	0.645	0.688	0.622

Note. First = First Administration; Second = Second Administration; Honest = Honest Response Instruction Condition; Desirable = Desirable Response Instructions Condition. All reliabilities were computed using Cronbach's alpha. Reliabilities in boldface are the factor level reliabilities.

Table 2

Means and Variances in the Honest and Desirable Response Instruction Conditions for Each Item using the Polytomously Scored Data

		H		D	
Item	Factor	*M*	S^2	*M*	S^2
1	C	4.09	0.47	4.48	0.38
7	C	3.06	1.38	3.71	1.19
12	C	3.12	0.90	3.63	0.91
22	C	4.07	1.37	4.40	1.25
33	C	3.85	0.90	3.99	0.71
34	C	3.64	0.66	4.22	0.45
45	C	3.83	0.56	4.11	0.51
46	C	3.30	1.03	3.62	0.70
47	C	4.08	0.92	4.57	0.52
50	C	3.93	0.47	4.24	0.52
56	C	3.72	0.91	4.05	0.78
57	C	4.10	0.73	4.52	0.45
59	C	3.32	0.89	3.88	0.54
80	C	3.14	1.60	3.38	1.46
88	C	3.41	1.33	4.04	1.00
96	C	3.93	0.60	4.28	0.44
97	C	3.04	1.26	3.88	1.18
98	C	4.05	0.70	4.49	0.45
101	C	3.11	1.14	3.80	1.12
102	C	3.29	1.19	3.99	1.03
106	C	3.71	0.67	3.93	0.59
108	C	2.98	1.01	3.45	0.82
115	C	3.95	0.77	4.25	0.69
116	C	2.94	1.08	3.23	1.11
120	C	4.03	0.50	4.32	0.41
121	C	3.23	1.14	3.79	1.00
128	C	3.90	0.68	4.12	0.54
139	C	3.38	0.79	3.91	0.82
151	C	4.12	0.72	4.38	0.43
157	C	3.88	0.44	4.16	0.51
158	C	3.34	1.05	3.96	0.75

159	C	3.99	0.50	4.32	0.57
163	C	3.41	1.35	3.91	1.05
167	C	3.40	0.99	3.73	0.93
176	C	4.10	0.64	4.42	0.62
177	C	3.71	0.86	4.14	0.62
183	C	3.50	0.93	3.99	1.10
186	C	3.69	0.74	4.06	0.74
201	C	3.71	0.85	4.17	0.69
204	C	3.20	0.99	3.54	1.13
207	C	3.64	1.10	4.01	0.81
210	C	4.05	0.50	4.40	0.48
213	C	2.78	1.09	3.53	1.29
214	C	3.91	0.90	4.36	0.83
219	C	3.10	1.79	3.66	1.51
231	C	3.74	0.84	4.15	0.74
234	C	3.14	1.25	3.79	1.15
245	C	3.03	1.29	3.21	1.13
248	C	3.55	1.15	3.77	1.08
253	C	4.14	0.62	4.58	0.36
256	C	3.09	1.02	3.83	1.24
267	C	4.03	0.74	4.49	0.43
271	C	4.01	0.82	4.39	0.63
274	C	3.45	0.81	3.99	0.79
286	C	3.22	0.86	3.84	0.69
290	C	2.54	1.17	2.94	1.42
294	C	2.54	1.13	2.93	1.31
295	C	3.07	1.15	3.73	1.04
296	C	3.71	0.85	3.98	0.54
299	C	3.11	0.97	3.80	1.10
10	E	4.12	0.54	4.35	0.48
11	E	3.95	1.13	3.99	1.05
18	E	3.63	1.04	4.02	0.99
21	E	3.78	0.82	4.26	0.55
24	E	3.84	0.56	4.17	0.47
38	E	3.89	0.66	4.06	0.56
39	E	3.61	1.24	3.74	1.25
41	E	4.41	0.61	4.63	0.36

52	E	4.22	0.82	4.39	0.51
55	E	3.77	1.05	3.94	0.92
63	E	4.23	0.85	4.30	0.58
67	E	3.74	1.19	4.03	0.86
74	E	2.99	1.17	2.95	1.14
76	E	3.59	1.14	3.85	1.05
81	E	4.10	0.69	4.35	0.55
92	E	3.35	1.07	3.81	0.91
95	E	4.41	0.71	4.61	0.38
105	E	2.05	0.95	2.15	0.85
110	E	3.68	0.85	4.01	0.67
111	E	3.83	0.60	3.95	0.61
112	E	3.81	1.11	4.03	0.67
124	E	3.58	0.88	3.83	0.72
136	E	3.47	1.78	3.86	1.25
146	E	4.44	0.69	4.21	0.94
148	E	3.75	1.22	3.99	0.98
150	E	4.11	0.99	3.94	1.02
152	E	2.49	0.76	2.52	0.77
155	E	3.58	0.77	3.58	0.73
162	E	4.00	0.94	4.25	0.67
164	E	3.82	1.00	3.91	0.90
171	E	3.97	1.20	4.00	1.06
172	E	3.63	1.14	3.70	1.15
178	E	3.42	1.38	3.26	1.21
181	E	3.92	0.74	4.23	0.60
185	E	3.82	1.01	3.94	0.72
187	E	3.75	0.85	3.61	0.86
190	E	3.31	1.13	3.65	1.05
193	E	3.76	1.77	3.88	1.53
196	E	3.99	0.97	4.25	0.70
197	E	2.31	0.91	2.28	0.68
198	E	4.01	1.03	4.15	0.70
200	E	4.44	0.68	4.39	0.54
205	E	3.84	0.73	4.17	0.47
209	E	3.99	0.89	4.40	0.45
212	E	4.21	0.65	4.35	0.37
220	E	4.28	0.72	4.34	0.49

229	E	3.46	1.37	3.73	1.02
230	E	2.50	1.21	2.17	1.22
237	E	3.01	0.90	2.98	1.05
244	E	2.09	0.91	2.22	0.63
246	E	3.52	1.34	3.92	0.98
257	E	2.21	0.96	2.20	1.10
266	E	3.65	1.13	3.94	0.88
270	E	3.68	1.02	3.90	0.79
272	E	3.69	0.89	3.96	0.85
273	E	3.42	1.09	3.75	0.95
288	E	3.88	0.51	4.06	0.46
289	E	3.56	0.88	3.95	0.76
298	E	3.18	1.09	2.84	1.40
300	E	3.72	0.78	4.09	0.66
2	N	2.90	1.18	2.23	0.79
30	N	2.37	0.91	1.88	0.61
32	N	3.01	1.06	2.30	0.89
35	N	2.80	0.88	2.26	0.78
36	N	2.34	1.33	1.74	0.70
40	N	2.22	1.20	1.75	0.86
43	N	2.57	1.18	2.17	1.00
44	N	3.17	1.13	2.51	1.30
48	N	2.88	1.04	2.55	0.90
49	N	2.19	0.82	1.85	0.65
53	N	2.01	1.25	1.61	0.69
60	N	2.78	1.29	2.41	1.01
61	N	2.63	1.04	1.95	0.68
62	N	2.40	1.03	1.90	0.67
64	N	2.81	1.32	2.41	1.18
72	N	2.29	1.19	1.92	0.81
79	N	1.96	1.00	1.62	0.60
82	N	2.92	1.29	2.30	1.17
91	N	2.79	1.12	2.24	1.13
99	N	2.05	0.98	1.62	0.59
100	N	2.40	1.37	1.85	0.92
107	N	2.48	1.19	2.29	1.13
117	N	2.25	0.80	1.94	0.48

119	N	1.71	0.78	1.47	0.60
122	N	2.75	1.51	2.27	1.26
123	N	2.99	1.44	2.75	1.46
133	N	3.09	0.96	2.65	0.96
134	N	3.02	1.16	2.70	1.00
135	N	2.81	1.07	2.33	0.86
137	N	3.05	1.50	2.55	1.38
138	N	3.40	1.38	2.97	1.30
153	N	3.41	1.08	3.02	1.21
156	N	2.77	1.44	2.32	1.34
161	N	2.33	0.90	1.93	0.63
166	N	3.21	1.24	2.42	1.41
169	N	2.86	1.31	2.53	1.10
173	N	2.80	1.10	2.28	1.02
182	N	2.05	0.70	1.69	0.57
184	N	2.12	1.04	1.75	0.73
188	N	3.14	1.60	2.51	1.38
192	N	2.85	1.36	2.10	1.18
203	N	2.69	1.25	2.40	1.18
206	N	3.05	1.23	2.43	1.08
216	N	2.19	1.17	1.81	0.85
217	N	2.71	1.90	2.50	1.62
218	N	2.41	1.23	1.89	0.98
222	N	2.61	1.19	2.15	1.17
226	N	3.91	0.95	3.71	0.90
233	N	3.70	0.93	3.33	1.05
241	N	2.61	1.47	1.95	0.91
243	N	2.73	1.29	2.34	1.09
247	N	3.14	1.41	2.78	1.44
254	N	2.71	1.21	2.38	1.09
255	N	2.53	0.94	2.21	0.93
263	N	2.53	1.06	2.11	0.82
265	N	2.05	0.92	1.69	0.62
278	N	2.27	1.05	1.86	0.60
281	N	2.97	1.42	2.58	1.43
287	N	3.54	1.17	3.01	1.41
292	N	2.22	0.95	1.80	0.74

3	O	3.68	0.66	4.15	0.46
6	O	2.63	1.92	2.60	1.79
16	O	2.08	0.61	1.84	0.62
17	O	3.49	1.38	3.21	1.06
19	O	3.72	1.27	3.86	1.26
28	O	4.52	0.60	4.54	0.47
29	O	3.39	1.31	2.97	1.25
51	O	4.40	0.63	4.56	0.31
54	O	3.99	0.68	4.05	0.58
58	O	3.58	1.15	3.53	0.82
65	O	3.18	1.15	3.35	1.10
66	O	3.38	1.47	3.61	1.21
69	O	3.11	1.24	3.48	0.94
75	O	4.39	0.80	4.43	0.62
77	O	3.46	1.15	3.85	0.84
78	O	2.70	0.86	2.77	0.89
85	O	3.83	0.89	3.88	0.73
86	O	3.51	1.60	3.70	1.13
87	O	2.17	0.78	1.99	0.69
89	O	3.16	0.93	3.14	1.04
90	O	3.83	0.67	4.06	0.64
94	O	3.83	0.89	4.01	0.75
103	O	4.04	0.90	4.12	0.73
104	O	3.88	0.64	3.94	0.67
113	O	3.73	0.74	4.01	0.49
125	O	3.27	1.24	3.03	1.09
140	O	3.39	1.04	3.30	1.25
141	O	3.26	1.06	3.74	0.90
142	O	3.61	0.99	3.88	0.77
144	O	3.33	1.77	3.54	1.63
145	O	3.03	1.51	3.48	1.41
160	O	3.95	0.96	4.10	0.87
170	O	3.37	1.80	3.47	1.64
174	O	3.21	1.23	3.62	0.96
175	O	2.83	1.03	2.72	0.97
180	O	3.27	1.38	3.10	1.11
194	O	3.23	1.42	3.73	1.26
195	O	2.57	1.03	3.01	1.29

202	O	3.01	0.95	3.08	0.99
208	O	1.82	1.05	1.55	0.88
211	O	3.99	0.99	3.43	1.60
215	O	3.82	1.10	3.43	1.53
221	O	2.77	1.10	2.75	1.35
223	O	4.03	0.69	4.09	0.71
224	O	2.77	1.32	3.41	1.35
228	O	3.74	1.64	3.88	1.42
235	O	3.35	1.50	3.63	1.22
239	O	3.98	1.08	4.14	1.03
252	O	3.56	1.08	3.83	0.81
258	O	3.08	1.10	3.16	1.04
260	O	3.14	1.14	3.05	1.32
264	O	3.42	1.06	3.48	1.05
268	O	3.87	1.00	3.90	0.82
275	O	2.98	0.83	3.00	0.81
277	O	3.69	1.30	3.81	1.10
282	O	3.45	0.93	3.16	1.16
283	O	4.05	0.85	4.28	0.57
285	O	4.52	0.71	4.61	0.46
291	O	3.48	1.12	3.16	1.20
293	O	3.50	1.74	3.60	1.64
4	A	3.75	0.78	4.30	0.57
5	A	3.51	1.02	3.67	0.98
8	A	4.23	0.63	4.50	0.43
9	A	3.32	0.97	3.71	0.96
13	A	3.72	0.60	3.88	0.54
14	A	4.02	1.02	4.55	0.51
15	A	3.48	0.90	3.88	0.68
20	A	3.79	0.96	4.05	0.76
23	A	4.29	0.79	4.56	0.62
25	A	2.41	1.02	2.16	0.76
26	A	3.53	0.97	3.70	0.76
27	A	3.66	1.20	3.80	0.77
31	A	3.45	0.76	3.70	0.68
37	A	3.92	0.97	4.23	0.75
42	A	3.59	0.80	3.84	0.71

68	A	3.86	1.37	3.88	1.12
70	A	4.19	0.85	4.45	0.43
71	A	4.12	0.80	4.35	0.66
73	A	3.23	1.15	3.16	0.98
83	A	4.04	0.74	4.22	0.64
84	A	3.44	1.33	3.72	1.19
93	A	3.48	0.77	3.69	0.57
109	A	3.38	1.58	3.32	1.56
114	A	3.22	0.93	3.29	1.00
118	A	3.91	0.89	4.10	0.67
126	A	3.94	0.70	4.15	0.72
127	A	4.02	0.96	4.35	0.66
129	A	3.60	1.27	4.09	0.81
130	A	3.92	1.55	4.24	1.25
131	A	3.02	1.34	3.01	1.08
132	A	3.47	0.84	3.58	0.74
143	A	3.31	1.00	3.54	0.83
147	A	3.29	1.60	3.49	1.59
149	A	3.06	1.33	3.14	1.20
154	A	3.82	1.01	4.24	1.00
165	A	4.26	0.86	4.43	0.69
168	A	2.78	1.01	3.00	1.10
179	A	3.42	1.38	3.22	1.59
189	A	4.05	0.77	4.36	0.73
191	A	2.84	1.06	3.32	1.41
199	A	4.17	0.66	4.28	0.54
225	A	2.70	1.13	2.66	1.22
227	A	3.88	1.12	4.21	0.75
232	A	3.88	1.00	4.11	0.79
236	A	2.70	0.90	2.23	0.81
238	A	3.38	1.29	3.86	1.14
240	A	3.43	1.20	3.49	1.31
242	A	3.82	0.98	3.88	1.01
249	A	3.57	0.95	3.58	0.90
250	A	2.91	1.23	3.01	1.22
251	A	3.04	1.29	3.27	1.23
259	A	3.86	0.75	4.02	0.63
261	A	3.16	1.21	3.32	1.20

262	A	2.56	0.95	2.36	1.21
269	A	4.06	0.93	4.30	0.84
276	A	2.82	1.17	2.70	1.18
279	A	3.57	1.10	3.79	1.11
280	A	4.01	0.97	4.31	0.75
284	A	3.82	0.87	3.83	0.82
297	A	3.78	0.83	3.93	0.69

Note. H = Honest Response Condition; D = Desirable Response Condition; C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 3

Means and Variances in the Honest and Desirable Response Instruction Conditions for Each Factor and Sub-Factor using the Polytomously Scored Data

		H		D	
Factor	Sub-factor	*M*	S^2	*M*	S^2
Conscientiousness		**212.81**	**760.09**	**238.18**	**850.85**
	1	38.09	23.91	41.94	23.08
	2	34.06	42.78	37.34	36.26
	3	39.52	31.54	43.14	24.80
	4	38.36	34.09	42.42	30.05
	5	32.12	47.24	38.47	60.70
	6	30.67	37.23	34.88	36.29
Extroversion		**219.44**	**701.28**	**227.72**	**435.80**
	1	39.13	45.42	41.86	28.52
	2	37.08	42.50	38.61	33.01
	3	35.58	35.13	37.85	28.14
	4	31.43	27.09	32.95	16.14
	5	35.13	33.23	34.93	28.04
	6	41.08	32.69	41.52	18.24
Neuroticism		**161.58**	**1065.59**	**134.76**	**962.95**
	1	28.94	57.81	24.27	45.18
	2	26.03	66.09	21.06	48.83
	3	21.69	52.17	17.86	36.80
	4	28.01	53.76	24.22	44.20
	5	30.69	33.07	26.44	33.57
	6	26.23	44.28	20.91	35.38
Openness		**206.46**	**482.77**	**210.66**	**346.20**
	1	36.81	40.23	35.73	35.58
	2	38.12	52.16	39.27	41.87
	3	36.33	38.98	35.60	25.10
	4	34.31	24.18	36.25	24.23
	5	33.90	38.22	37.65	36.12
	6	26.98	30.55	26.15	24.78

Agreeableness		**214.06**	**641.04**	**223.73**	**498.41**
	1	35.97	38.47	37.80	32.46
	2	37.92	41.84	41.00	31.73
	3	39.81	28.54	42.32	21.83
	4	35.53	36.02	37.69	28.00
	5	30.78	35.06	29.64	27.07
	6	34.05	28.09	35.27	21.85

Note. H = Honest Response Condition; D = Desirable Response Condition; C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness. Numbers in Bold are the factor means and variances.

Table 4

Descriptive Statistics and Correlations between Factor Scores on Each Personality Factor

Factor	1	2	3	4	5	6	7	8	9	10
1. C1	--									
2. E1	0.34*	--								
3. N1	-0.31*	-0.42*	--							
4. O1	0.13*	0.35*	0.00	--						
5. A1	0.53*	0.38*	-0.22*	0.39*	--					
6. C2	0.47*	0.20*	-0.21*	0.06	0.35*	--				
7. E2	0.06	0.60*	-0.33*	0.08	0.12	0.46*	--			
8. N2	-0.19*	-0.28*	0.53*	0.04	-0.20*	-0.70*	-0.61*	--		
9. O2	-0.04	0.09	-0.04	0.63*	0.19*	0.20*	0.29*	-.016*	--	
10. A2	0.27*	0.16*	-0.14*	0.19*	0.68*	0.64*	0.35*	-0.43*	0.31*	--
M	211.14	217.37	160.50	205.10	212.52	237.99	227.64	134.21	210.47	223.58
SD	27.85	26.26	31.01	21.80	25.65	29.03	20.86	30.78	18.47	22.19

Note. C1 = Conscientiousness first session; E1 = Extroversion first session; N1 = Neuroticism first session; O1 = Openness first session; A1 = Agreeableness first session; C2 = Conscientiousness second session; E2 = Extroversion second session; N2 = Neuroticism second session; O2 = Openness second session; A2 = Agreeableness second session.

Table 5

Eigenvalues from the Principle Components Analysis of the Observed Data from the Honest Response and Desirable Response Instruction Conditions

Factor	1st eigenvalue	% Variance	2nd eigenvalue	% Variance
		Desirable Response Condition		
C	19.025	31.709	3.195	5.326
E	12.013	20.022	4.446	7.410
N	17.425	29.042	4.063	6.771
O	7.941	13.235	5.134	8.557
A	11.675	19.458	3.376	5.626
		Honest Response Condition		
C	14.443	24.072	4.218	7.029
E	14.041	23.402	4.031	6.718
N	14.947	24.911	4.372	7.286
O	9.919	16.531	4.250	7.083
A	12.237	20.396	4.164	6.939

Note. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 6

Eigenvalues from the Principle Components Analysis of the Observed Sub-Factor Scores from the Honest Response and Desirable Response Instruction Conditions, and the eigenvalues from the Parallel Analysis

	Component					
Condition	1	2	3	4	5	6
Honest	8.584	4.164	3.378	2.320	1.727	1.089
Desirable	11.199	3.196	2.674	1.753	1.573	1.133
Parallel Analysis	1.772	1.666	1.572	1.502	1.436	1.374

Table 7

Factor Loadings of the IPIP Factor Scores for the Honest Response and Desirable Response Instruction Conditions

	Component				
Sub-scale	1	2	3	4	5
			Honest Condition		
C_1	**0.588**				
C_2	**0.666**			0.400	-0.345
C_3	**0.571**	0.505			
C_4	**0.774**				
C_5	**0.853**				
C_6	**0.430**	0.356	-0.532		
E_1			**0.781**		
E_2			**0.822**		
E_3	0.477		**0.604**		
E_4	0.834				
E_5			**0.745**		
E_6		0.376	**0.681**		
N_1				**0.915**	
N_2	0.327	-0.411		**0.820**	
N_3			-0.458	**0.483**	
N_4		0.314	-0.379	**0.553**	
N_5		-0.373	0.439	**0.406**	
N_6				**0.807**	
O_1					**0.640**
O_2					**0.687**
O_3			0.310	0.659	
O_4				-0.320	**0.516**
O_5					**0.881**

O_6	-0.544				
A_1		**0.623**	0.334		
A_2		**0.741**			
A_3		**0.649**	0.364		
A_4		**0.857**			
A_5		**0.623**	-0.447		
A_6		**0.665**			
			Desirable Condition		
C_1	**0.705**				
C_2		0.405		0.613	
C_3	0.390	0.477			
C_4	**0.536**			0.300	
C_5	**0.652**			0.329	
C_6	**0.468**	0.464	-0.461		
E_1			**0.584**		
E_2			**0.706**		
E_3	0.763	-0.430			
E_4	0.476			0.334	
E_5		-0.459	**0.651**		
E_6			**0.767**		
N_1	-0.875				0.374
N_2	-0.526	-0.363		0.323	0.335
N_3	-0.428		-0.410		
N_4	-0.800				
N_5	-0.499	-0.317			
N_6	-0.930				
O_1				-0.535	**0.574**
O_2	0.343			-0.380	**0.709**
O_3	-0.300				**0.789**
O_4	0.531			-0.311	
O_5	1.021			-0.404	

O_6				-0.713	
A_1		**0.640**	0.502		
A_2		**0.736**			
A_3		**0.478**	0.326		
A_4		**0.894**			
A_5	-0.560	**0.834**			
A_6		**0.545**			0.440

Note. Only loading > .30 are displayed. Loadings on the five personality factors are reported in boldface. Subscripts indicate the number of the sub-scale for each factor. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 8

Fit Statistics for the 2-PL Model in the Honest Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	21	10	5	6	3	13	2	2.653	2.383
Doubles	7	17	24	9	3	0	0	2.271	1.038
Triples	0	12	8	0	0	0	0	1.915	0.468

Frequency Table of Chi-Square/DF Ratios for Extroversion									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	15	7	5	10	3	5	15	4.029	3.414
Doubles	6	10	20	9	6	7	2	3.119	2.111
Triples	0	10	6	2	1	1	0	2.415	1.279

Frequency Table of Chi-Square/DF Ratios for Neuroticism									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	15	8	7	4	7	10	9	3.548	2.774
Doubles	3	15	20	15	6	1	0	2.711	1.058
Triples	0	7	12	1	0	0	0	2.15	0.502

Frequency Table of Chi-Square/DF Ratios for Openness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	26	3	4	4	5	6	12	3.559	3.777
Doubles	10	13	14	9	8	2	4	2.925	2.318
Triples	2	7	6	1	2	2	0	2.444	1.4

Frequency Table of Chi-Square/DF Ratios for Agreeableness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	19	11	5	6	5	8	6	2.99	2.823
Doubles	7	14	18	14	5	2	0	2.552	1.253
Triples	1	9	8	2	0	0	0	2.137	0.723

Table 9

Fit Statistics for the 2-PL Model in the Desirable Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	15	12	2	4	5	15	7	3.431	2.667
Doubles	0	17	20	18	4	1	0	2.665	0.948
Triples	0	6	11	2	1	0	0	2.302	0.623

Frequency Table of Chi-Square/DF Ratios for Extroversion									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	5	4	10	4	6	10	21	5.481	3.369
Doubles	0	7	12	13	15	7	6	4.047	1.905
Triples	0	1	11	2	5	1	0	3.217	1.273

Frequency Table of Chi-Square/DF Ratios for Neuroticism									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	5	1	6	3	10	20	15	5.276	2.449
Doubles	0	2	11	27	14	5	1	3.699	1.13
Triples	0	1	6	10	3	0	0	3.241	0.712

Frequency Table of Chi-Square/DF Ratios for Openness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	19	6	6	1	2	11	15	4.349	3.946
Doubles	4	8	16	18	6	6	2	3.399	2.019
Triples	0	7	6	4	2	1	0	2.835	1.162

Frequency Table of Chi-Square/DF Ratios for Agreeableness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	16	6	3	4	4	14	13	4.336	3.518
Doubles	5	7	15	19	9	2	3	3.259	1.756
Triples	0	6	10	1	2	1	0	2.629	1.175

Table 10

Fit Statistics for Samejima's Graded Response Model in the Honest Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	19	24	14	2	0	0	1	3.315	13.857
Doubles	18	32	6	2	0	2	0	1.523	0.887
Triples	5	10	3	2	0	0	0	1.597	0.858

Frequency Table of Chi-Square/DF Ratios for Extroversion

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	7	12	6	7	9	14	5	6.357	20.932
Doubles	3	16	26	8	4	3	0	-3.189	37.156
Triples	2	0	7	6	2	2	1	1.659	9.488

Frequency Table of Chi-Square/DF Ratios for Neuroticism

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	60	0	0	0	0	0	0	0.388	0.231
Doubles	38	22	0	0	0	0	0	0.844	0.373
Triples	8	10	0	1	1	0	0	1.228	1.011

Frequency Table of Chi-Square/DF Ratios for Openness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	43	13	3	1	0	0	0	0.816	0.808
Doubles	25	28	6	1	0	0	0	1.218	0.642
Triples	7	8	3	1	0	0	1	1.802	1.659

Frequency Table of Chi-Square/DF Ratios for Agreeableness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	19	21	9	9	2	0	0	1.692	1.2
Doubles	14	40	6	0	0	0	0	1.392	0.537
Triples	2	11	4	0	1	0	0	1.776	1.343

Table 11

Fit Statistics for Samejima's Graded Response in the Desirable Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	0	0	1	0	2	2	55	27.849	22.234
Doubles	0	0	1	0	4	9	46	12.159	7.531
Triples	1	3	1	7	2	1	5	5.126	4.325

Frequency Table of Chi-Square/DF Ratios for Extroversion									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	2	2	4	1	3	9	39	13.059	17.445
Doubles	3	1	4	6	10	16	20	3.406	15.288
Triples	2	1	6	2	2	4	3	4.076	9.943

Frequency Table of Chi-Square/DF Ratios for Neuroticism									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	0	0	1	2	5	4	48	11.933	7.644
Doubles	0	2	1	8	16	21	12	5.447	1.803
Triples	1	1	2	3	4	2	7	5.796	3.831

Frequency Table of Chi-Square/DF Ratios for Openness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	6	8	9	4	9	11	13	5.27	4.464
Doubles	2	8	22	11	5	8	4	3.517	1.909
Triples	6	2	5	3	1	2	1	2.711	2.044

Frequency Table of Chi-Square/DF Ratios for Agreeableness									
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	1	4	6	4	4	10	31	9.331	6.873
Doubles	1	4	4	7	12	16	16	5.646	2.766
Triples	2	2	4	3	1	4	4	5.712	6.669

Table 12

Results of the Differential Functioning Item Analysis using the Mantel-Haenszel Procedure on the Dichotomously Scored Items

Item	Factor	χ^2_{MH}	Ω	|Delta|	Group
1	C	0.051	1.296	0.6087	
2	**N**	**8.098**	**0.328**	**2.62**	**D**
3	**O**	**11.361 *****	**2.941**	**2.535**	**D**
4	**A**	**6.998 *****	**2.291**	**1.948**	**D**
5	A	0.072	0.908	0.227	
6	O	0.19	0.862	0.349	
7	C	0.457	1.252	0.5288	
8	A	1.286	1.847	1.442	
9	A	1.66	1.43	0.841	
10	E	0.881	1.480	0.921	
11	E	0.950	0.743	0.6981	
12	C	0.263	1.206	0.4395	
13	A	0.795	1.358	0.719	
14	**A**	**11.364 *****	**3.498**	**2.943**	**D**
15	**A**	**3.675 ***	**1.682**	**1.222**	**D**
16	O	0.05	0.59	1.24	
17	**O**	**9.542 *****	**0.456**	**1.845**	**H**
18	**E**	**3.145 ***	**1.666**	**1.200**	**D**
19	O	0.254	1.168	0.365	
20	A	0.247	1.179	0.387	
21	**E**	**7.721 *****	**2.643**	**2.284**	**D**
22	C	0.022	1.114	0.2538	
23	A	0.247	1.302	0.62	
24	**E**	**7.849 *****	**2.678**	**2.315**	**D**
25	**A**	**4.881 ****	**0.399**	**2.159**	**H**
26	A	0.071	0.903	0.24	
27	A	0.556	1.249	0.523	
28	O	0.04	1.245	0.515	
29	**O**	**14.129 *****	**0.383**	**2.255**	**H**
30	N	0.709	0.589	1.244	
31	A	1.256	1.391	0.776	
32	**N**	**3.25 ***	**0.522**	**1.528**	**D**

33	C	0.219	0.827	0.4442	
34	**C**	**7.052 *****	**2.796**	**2.4158**	**D**
35	N	1.323	0.626	1.101	
36	**N**	**4.97 ****	**0.31**	**2.752**	**D**
37	A	0.439	1.269	0.56	
38	**E**	**3.020 ***	**1.662**	**1.193**	**D**
39	E	0.059	0.938	0.150	
40	N	0.021	1.027	0.063	
41	E	2.220	2.163	1.813	
42	A	0.605	1.301	0.618	
43	N	0.192	1.29	0.598	
44	N	0.003	0.946	0.13	
45	C	0.976	1.477	0.9165	
46	C	0.494	0.791	0.5499	
47	C	2.135	1.944	1.5628	
48	N	0.021	0.997	0.007	
49	N	0.001	0.85	0.382	
50	C	0.255	0.766	0.6275	
51	**O**	**3.14 ***	**3.582**	**2.998**	**D**
52	E	0.133	1.147	0.322	
53	N	0.12	0.748	0.682	
54	O	0.022	0.906	0.232	
55	E	0.604	1.222	0.472	
56	C	0.241	0.832	0.4324	
57	C	0.103	1.248	0.5217	
58	O	1.892	0.693	0.862	
59	**C**	**2.929 ***	**1.675**	**1.2126**	**D**
60	N	0.037	0.895	0.261	
61	**N**	**4.205 ****	**0.402**	**2.142**	**D**
62	N	1.401	0.508	1.592	
63	E	0.235	1.193	0.414	
64	N	0.603	0.77	0.614	
65	O	0.491	1.219	0.465	
66	O	0.03	1.078	0.177	
67	E	0.718	1.279	0.579	
68	**A**	**4.759 ****	**0.525**	**1.514**	**H**
69	**O**	**5.083 ****	**1.759**	**1.327**	**D**
70	**A**	**3.779 ***	**2.815**	**2.432**	**D**

71	A	0.515	1.381	0.759	
72	N	0	0.913	0.214	
73	**A**	**4.262 ****	**0.579**	**1.284**	**H**
74	E	2.732	0.662	0.971	
75	O	0.009	1.042	0.097	
76	E	2.442	1.529	0.997	
77	**O**	**4.202 ****	**1.766**	**1.336**	**D**
78	O	1.131	1.417	0.819	
79	N	0.013	0.796	0.536	
80	C	0.001	1.041	0.0940	
81	E	0.256	1.207	0.442	
82	N	0.401	0.749	0.679	
83	A	0	0.939	0.148	
84	A	0.038	0.924	0.186	
85	O	0.268	0.822	0.461	
86	O	0.056	1.128	0.283	
87	O	0.882	0.55	1.405	
88	C	0.023	1.102	0.2280	
89	O	0.031	0.931	0.168	
90	O	1.743	1.527	0.995	
91	N	0.133	1.214	0.456	
92	**E**	**6.325 ****	**1.959**	**1.581**	**D**
93	A	0.026	1.075	0.17	
94	O	1.837	1.565	1.053	
95	E	1.511	2.010	1.641	
96	C	0.327	1.380	0.7567	
97	**C**	**3.103 ***	**1.773**	**1.3442**	**D**
98	C	0.035	0.990	0.0235	
99	N	0.718	0.528	1.501	
100	N	0.251	0.757	0.654	
101	C	0.145	1.171	0.3713	
102	C	0.014	0.987	0.0306	
103	O	0.555	1.356	0.716	
104	O	0.068	0.889	0.276	
105	E	0.160	0.856	0.365	
106	C	0.739	0.745	0.6933	
107	N	0.422	1.332	0.674	
108	C	0.182	0.846	0.3925	

109	A	2.176	0.668	0.948	
110	E	1.298	1.391	0.775	
111	E	0.052	0.936	0.155	
112	E	1.693	1.491	0.938	
113	**O**	**4.217 ** **	**1.839**	**1.432**	**D**
114	A	0.144	0.882	0.295	
115	C	1.532	0.636	1.0646	
116	C	0.004	1.019	0.0447	
117	N	0.207	0.665	0.959	
118	A	0.004	1.094	0.211	
119	N	0	0.779	0.587	
120	C	0.180	0.774	0.6040	
121	C	0.212	1.198	0.4230	
122	N	0.519	1.365	0.731	
123	N	0.656	1.302	0.62	
124	E	1.145	1.341	0.689	
125	**O**	**8.455 *** **	**0.473**	**1.759**	**H**
126	A	0.001	1.062	0.141	
127	A	1.867	1.692	1.236	
128	C	0.017	1.109	0.2444	
129	A	0.934	1.399	0.789	
130	A	1.686	1.515	0.976	
131	A	2.695	0.638	1.056	
132	A	0.001	0.958	0.101	
133	N	0.009	1.017	0.04	
134	N	0.018	0.997	0.007	
135	N	0.592	0.735	0.724	
136	E	0.457	1.197	0.422	
137	N	0	0.955	0.108	
138	N	0.002	1.023	0.053	
139	C	0.008	0.980	0.0470	
140	O	1.037	0.756	0.657	
141	**O**	**6.659**	**1.976**	**1.601**	**D**
142	O	0.783	1.303	0.622	
143	A	0.553	1.245	0.515	
144	O	0.053	1.106	0.237	
145	**O**	**7.086 *** **	**1.931**	**1.546**	**D**
146	**E**	**8.227 *** **	**0.347**	**2.491**	**H**

147	A	0.278	1.171	0.371	
148	E	0.054	1.071	0.161	
149	A	0.699	0.778	0.59	
150	**E**	**5.818 ****	**0.506**	**1.602**	**H**
151	C	0.486	1.490	0.9353	
152	E	0.034	0.942	0.140	
153	N	1.861	1.532	1.002	
154	**A**	**3.087 ***	**1.86**	**1.458**	**D**
155	E	4.322	0.577	1.291	
156	N	1.299	1.572	1.063	
157	C	0.000	0.928	0.1739	
158	C	0.188	1.170	0.3690	
159	**C**	**3.120 ***	**0.434**	**1.9623**	**H**
160	O	0	1.045	0.103	
161	N	0.038	0.987	0.031	
162	E	1.663	1.566	1.054	
163	C	0.051	1.126	0.2797	
164	E	0.661	0.809	0.500	
165	A	0.236	0.729	0.743	
166	**N**	**3.784 ***	**0.551**	**1.401**	**D**
167	C	0.731	0.770	0.6134	
168	A	0.715	1.294	0.606	
169	N	0.396	1.318	0.649	
170	O	0.253	0.844	0.399	
171	E	2.602	0.630	1.085	
172	E	2.401	0.662	0.970	
173	N	1.208	0.681	0.903	
174	**O**	**10.776 *****	**2.25**	**1.906**	**D**
175	O	0.504	0.791	0.551	
176	C	0.106	1.233	0.4935	
177	C	0.256	1.218	0.4630	
178	**E**	**6.664 ****	**0.539**	**1.452**	**H**
179	**A**	**5.665 ****	**0.54**	**1.448**	**H**
180	**O**	**3.09 ***	**0.633**	**1.075**	**H**
181	**E**	**2.968 ***	**1.935**	**1.551**	**D**
182	N	0.054	0.971	0.069	
183	C	0.077	0.875	0.3126	
184	N	0.024	1.047	0.108	

185	E	0.235	1.130	0.286	
186	C	1.233	0.660	0.9776	
187	**E**	**4.699 ****	**0.556**	**1.382**	**H**
188	N	0.177	1.216	0.46	
189	A	0.912	1.491	0.939	
190	E	1.439	1.353	0.711	
191	**A**	**3.999 ****	**1.78**	**1.355**	**D**
192	N	0.507	0.735	0.724	
193	E	1.466	0.724	0.758	
194	**O**	**8.851 *****	**2.212**	**1.866**	**D**
195	**O**	**11.804 *****	**2.639**	**2.28**	**D**
196	E	1.077	1.500	0.952	
197	E	2.388	0.561	1.358	
198	E	0.923	1.374	0.746	
199	A	0	0.921	0.193	
200	E	1.382	0.595	1.222	
201	C	0.000	0.942	0.1410	
202	O	0.014	0.993	0.017	
203	N	0.973	1.407	0.802	
204	C	0.313	0.829	0.4395	
205	**E**	**3.978 ****	**2.026**	**1.659**	**D**
206	**N**	**3.861 ****	**0.527**	**1.505**	**D**
207	C	0.093	1.146	0.3220	
208	O	0.373	0.706	0.818	
209	**E**	**6.811 *****	**2.935**	**2.531**	**D**
210	C	0.096	0.787	0.5640	
211	**O**	**22.495 *****	**0.264**	**3.13**	**H**
212	E	1.336	1.659	1.189	
213	**C**	**4.068 ****	**1.969**	**1.5910**	**D**
214	C	0.031	1.002	0.0047	
215	**O**	**7.552 *****	**0.482**	**1.715**	**H**
216	N	0.264	0.746	0.689	
217	N	0.037	0.915	0.209	
218	N	0.192	0.774	0.602	
219	C	0.140	0.862	0.3478	
220	E	0.727	0.700	0.839	
221	O	0.002	0.953	0.113	
222	N	0.05	0.892	0.269	

223	O	0.019	0.903	0.24	
224	**O**	**16.219 *****	**2.91**	**2.51**	**D**
225	A	0.007	0.942	0.14	
226	N	0.654	0.768	0.62	
227	A	0.88	1.383	0.762	
228	**O**	**3.76 ***	**0.493**	**1.662**	**H**
229	E	0.097	1.085	0.192	
230	**E**	**5.406 ****	**0.483**	**1.712**	**H**
231	C	0.024	1.115	0.2562	
232	A	1.543	1.503	0.958	
233	**N**	**3.235 ***	**0.607**	**1.173**	**D**
234	C	0.116	1.165	0.3596	
235	O	0.415	1.228	0.483	
236	**A**	**5.38 ****	**0.422**	**2.027**	**H**
237	**E**	**3.221 ***	**0.619**	**1.127**	**H**
238	A	2.636	1.568	1.057	
239	O	0.058	0.876	0.311	
240	A	2.295	0.658	0.984	
241	N	2.221	0.562	1.354	
242	A	1.793	0.647	1.023	
243	N	0.171	0.829	0.441	
244	E	0.900	0.668	0.947	
245	**C**	**2.867 ***	**0.597**	**1.2103**	**H**
246	E	2.096	1.512	0.971	
247	N	0.254	1.199	0.426	
248	**C**	**2.784 ***	**0.609**	**1.1633**	**H**
249	A	1.18	0.734	0.727	
250	A	0.184	0.863	0.346	
251	A	0.012	1.064	0.146	
252	**O**	**4.515 ****	**1.916**	**1.528**	**D**
253	C	0.044	1.242	0.5076	
254	N	0.455	1.318	0.649	
255	N	0.07	1.174	0.377	
256	C	1.828	1.549	1.0293	
257	E	0.001	0.990	0.024	
258	O	0	1.038	0.088	
259	A	0.059	0.884	0.29	
260	O	0.352	0.848	0.387	

261	A	1.037	0.745	0.692	
262	A	0.619	0.756	0.657	
263	N	0.311	1.35	0.705	
264	O	0.003	1.016	0.037	
265	N	0.18	0.689	0.875	
266	E	0.010	1.030	0.069	
267	C	0.000	1.082	0.1833	
268	O	0.01	0.99	0.024	
269	A	0.09	1.172	0.373	
270	E	0.320	1.169	0.366	
271	C	0.119	1.175	0.3784	
272	E	2.472	1.550	1.029	
273	E	3.036	1.603	1.109	
274	C	0.490	1.287	0.5946	
275	O	0.14	0.873	0.319	
276	**A**	**3.872 ****	**0.58**	**1.28**	**H**
277	O	0.009	0.988	0.028	
278	**N**	**4.898 ****	**0.286**	**2.942**	**D**
279	A	0.052	0.907	0.229	
280	A	0.099	0.832	0.432	
281	N	0.328	1.23	0.486	
282	**O**	**7.288 *****	**0.497**	**1.643**	**H**
283	O	0.346	1.336	0.681	
284	A	1.879	0.651	1.009	
285	O	0.393	1.458	0.886	
286	**C**	**4.072 ****	**1.746**	**1.3113**	**D**
287	N	0.005	1.067	0.152	
288	E	0.077	0.904	0.238	
289	**E**	**7.296 *****	**2.100**	**1.743**	**D**
290	C	0.010	0.925	0.1833	
291	**O**	**13.531 *****	**0.386**	**2.237**	**H**
292	N	0.012	0.854	0.371	
293	O	0.001	0.963	0.089	
294	C	0.028	0.910	0.2209	
295	C	0.966	1.397	0.7873	
296	C	0.180	0.845	0.3972	
297	A	0.013	1.074	0.168	
298	**E**	**10.652 *****	**0.441**	**1.925**	**H**

299	C	0.548	1.280	0.5805	
300	**E**	**8.894** ***	**2.473**	**2.127**	**D**

Note. *p < .10, **p < .05, ***p < .01. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; χ^2_{MH} = Mantel-Haenszel Chi-square; Ω = Odds Ratio; |Delta| = MH D-DIF; D = Desirable Response Condition; H = Honest Response Condition. Items in boldface are identified as functioning differentially.

Table 13

Results of the Differential Functioning Item Analysis using the Mantel Procedure on the Polytomously Scored Items

Item	Factor	M	SMD	Group
1	**C**	**4.801** **	**0.136**	**D**
2	**N**	**4.659** **	**-0.134**	**D**
3	**O**	**22.426** ***	**0.389**	**D**
4	**A**	**21.448** ***	**0.351**	**D**
5	A	0.096	0.032	
6	O	0.322	-0.118	
7	C	2.617	0.174	
8	**A**	**5.210** **	**0.110**	**D**
9	**A**	**5.181** **	**0.191**	**D**
10	E	2.089	0.064	
11	E	1.687	-0.089	
12	C	0.420	0.087	
13	A	0.009	0.024	
14	**A**	**17.354** ***	**0.300**	**D**
15	**A**	**7.008** ***	**0.178**	**D**
16	**O**	**6.414** **	**-0.178**	**H**
17	**O**	**6.326** **	**-0.222**	**H**
18	**E**	**3.387** *	**0.159**	**D**
19	O	0.153	0.058	
20	A	1.575	0.081	
21	**E**	**17.298** ***	**0.348**	**D**
22	C	0.023	-0.033	
23	A	2.096	0.042	
24	**E**	**8.262** ***	**0.206**	**D**
25	**A**	**5.719** **	**-0.196**	**H**
26	A	0.044	0.036	
27	A	0.142	0.072	
28	O	0.502	-0.062	
29	**O**	**14.902** ***	**-0.374**	**H**
30	N	1.457	-0.134	
31	A	0.780	0.102	
32	**N**	**8.092** ***	**-0.215**	**D**

33	C	0.245	-0.033	
34	**C**	**8.229** ***	**0.144**	**D**
35	**N**	**4.861** **	**-0.235**	**D**
36	**N**	**5.067** **	**-0.208**	**D**
37	A	1.204	0.052	
38	E	1.189	0.090	
39	E	0.475	-0.055	
40	N	0.024	0.044	
41	E	2.158	0.107	
42	A	0.990	0.058	
43	N	1.062	0.047	
44	N	2.080	-0.246	
45	C	0.017	0.046	
46	C	0.001	0.045	
47	**C**	**3.599** *	**0.132**	**D**
48	N	0.345	-0.084	
49	N	0.002	0.000	
50	C	0.002	-0.043	
51	O	1.317	0.063	
52	E	0.119	0.016	
53	N	0.011	-0.047	
54	O	0.648	-0.069	
55	E	0.504	0.176	
56	C	0.205	0.064	
57	C	2.243	0.086	
58	O	1.692	-0.148	
59	**C**	**8.193** ***	**0.161**	**D**
60	N	0.251	0.181	
61	**N**	**4.792** **	**-0.112**	**D**
62	N	1.281	-0.037	
63	E	0.473	-0.059	
64	N	0.816	-0.209	
65	O	0.621	0.120	
66	O	0.786	0.124	
67	E	0.914	0.113	
68	**A**	**5.304** **	**-0.222**	**H**
69	**O**	**7.498** ***	**0.295**	**D**
70	A	2.076	0.056	

71	A	0.629	0.027	
72	N	0.394	0.050	
73	**A**	**3.627 ***	**-0.192**	**H**
74	E	2.188	-0.135	
75	O	0.936	-0.099	
76	E	0.215	0.066	
77	**O**	**7.002 *****	**0.295**	**D**
78	O	0.778	0.070	
79	N	0.070	-0.064	
80	C	0.057	0.080	
81	E	1.070	0.028	
82	N	0.248	-0.155	
83	A	0.005	0.026	
84	A	0.604	0.049	
85	O	0.412	-0.030	
86	O	0.045	0.027	
87	**O**	**2.849 ***	**-0.148**	**H**
88	C	0.165	-0.017	
89	O	0.615	-0.086	
90	**O**	**2.922 ***	**0.129**	**D**
91	N	0.000	-0.024	
92	**E**	**7.087 *****	**0.251**	**D**
93	A	1.474	0.083	
94	O	0.452	0.081	
95	E	0.537	0.041	
96	C	0.856	0.030	
97	**C**	**4.559 ****	**0.099**	**D**
98	C	1.912	0.096	
99	N	1.625	-0.127	
100	N	0.271	-0.016	
101	C	2.267	0.126	
102	C	0.751	0.073	
103	O	0.251	-0.020	
104	O	0.070	-0.016	
105	E	1.104	0.155	
106	C	0.768	-0.046	
107	**N**	**3.036 ***	**0.178**	**H**
108	C	0.134	0.045	

109	A	1.574	-0.102	
110	**E**	**3.178 ***	**0.119**	**D**
111	E	0.015	0.044	
112	E	0.064	0.005	
113	**O**	**5.870 ****	**0.186**	**D**
114	A	0.092	-0.040	
115	C	2.330	-0.113	
116	C	0.274	0.019	
117	N	0.000	0.057	
118	A	0.027	0.047	
119	N	0.867	0.063	
120	C	0.086	0.004	
121	C	0.037	0.008	
122	N	0.216	0.023	
123	N	0.302	0.227	
124	E	1.792	0.107	
125	**O**	**5.605 ****	**-0.171**	**H**
126	A	0.280	0.026	
127	A	2.628	0.088	
128	C	0.631	-0.022	
129	**A**	**7.308 *****	**0.174**	**D**
130	A	0.947	0.114	
131	A	0.769	-0.020	
132	A	0.182	-0.018	
133	N	0.884	-0.191	
134	N	0.164	-0.028	
135	N	0.075	0.099	
136	E	1.787	0.116	
137	N	0.323	0.182	
138	N	0.002	-0.072	
139	C	1.157	0.096	
140	O	1.903	-0.134	
141	**O**	**11.993 *****	**0.304**	**D**
142	**O**	**3.136 ***	**0.174**	**D**
143	A	0.346	0.086	
144	O	0.032	-0.030	
145	**O**	**8.828 *****	**0.367**	**D**
146	**E**	**12.018 *****	**-0.284**	**H**

147	A	0.053	-0.005	
148	E	0.122	0.047	
149	A	0.364	-0.007	
150	**E**	**5.488 ****	**-0.226**	**H**
151	C	0.045	0.006	
152	E	0.185	-0.024	
153	N	1.101	0.060	
154	**A**	**4.019 ****	**0.143**	**D**
155	E	0.674	-0.072	
156	N	0.234	0.025	
157	C	0.009	0.017	
158	**C**	**3.642 ***	**0.150**	**D**
159	C	0.714	-0.020	
160	O	0.001	-0.028	
161	N	0.123	-0.001	
162	E	0.504	0.042	
163	C	0.060	0.015	
164	E	0.329	-0.090	
165	A	0.141	-0.019	
166	**N**	**5.860 ****	**-0.344**	**D**
167	C	0.015	0.012	
168	A	1.470	0.156	
169	N	1.973	0.146	
170	O	0.309	-0.078	
171	**E**	**3.670 ***	**-0.174**	**H**
172	E	1.672	-0.157	
173	N	0.898	-0.058	
174	**O**	**8.987 *****	**0.293**	**D**
175	O	1.436	-0.176	
176	C	0.342	0.028	
177	C	1.688	0.094	
178	**E**	**3.247 ***	**-0.169**	**H**
179	**A**	**4.702 ****	**-0.207**	**H**
180	**O**	**3.289 ***	**-0.197**	**H**
181	**E**	**3.023 ***	**0.092**	**D**
182	N	0.127	0.034	
183	C	0.007	0.011	
184	N	0.230	0.074	

185	E	0.040	-0.111	
186	C	0.712	-0.038	
187	**E**	**5.583 ****	**-0.214**	**H**
188	N	0.451	0.035	
189	A	1.886	0.073	
190	**E**	**3.126 ***	**0.115**	**D**
191	**A**	**5.304 ****	**0.167**	**D**
192	N	2.639	-0.173	
193	E	0.018	-0.003	
194	**O**	**9.310 *****	**0.345**	**D**
195	**O**	**9.795 *****	**0.275**	**D**
196	E	0.378	0.025	
197	E	0.068	-0.066	
198	E	0.009	0.007	
199	A	0.438	-0.062	
200	**E**	**4.018 ****	**-0.140**	**H**
201	C	0.163	0.022	
202	O	0.018	-0.053	
203	N	0.240	0.005	
204	C	0.033	-0.108	
205	**E**	**5.366 ****	**0.136**	**D**
206	**N**	**3.899 ****	**-0.297**	**D**
207	C	0.092	0.047	
208	**O**	**3.705 ***	**-0.261**	**H**
209	**E**	**7.086 *****	**0.177**	**D**
210	C	0.011	-0.011	
211	**O**	**26.274 *****	**-0.578**	**H**
212	E	0.008	-0.063	
213	C	2.322	0.129	
214	C	0.002	0.008	
215	**O**	**13.478 *****	**-0.369**	**H**
216	N	0.073	0.022	
217	N	0.000	0.041	
218	N	0.101	-0.009	
219	C	0.004	0.026	
220	E	1.280	-0.082	
221	O	0.352	-0.105	
222	N	0.832	-0.114	

223	O	0.355	-0.040	
224	**O**	**18.382 *****	**0.455**	**D**
225	A	0.431	-0.077	
226	N	0.502	-0.044	
227	A	2.096	0.123	
228	O	0.605	-0.075	
229	E	0.724	0.077	
230	**E**	**5.497 ****	**-0.194**	**H**
231	C	0.338	0.001	
232	A	0.616	0.019	
233	**N**	**4.190 ****	**-0.162**	**D**
234	C	0.267	0.013	
235	O	1.401	0.164	
236	**A**	**16.757 *****	**-0.288**	**H**
237	E	2.298	-0.187	
238	**A**	**6.217 ****	**0.257**	**D**
239	O	0.001	0.013	
240	A	0.853	-0.111	
241	N	2.419	-0.083	
242	A	0.597	-0.038	
243	N	0.134	0.038	
244	E	1.666	0.162	
245	C	0.741	-0.102	
246	E	1.648	0.026	
247	N	0.570	0.041	
248	**C**	**2.736 ***	**-0.138**	**H**
249	A	2.121	-0.131	
250	A	0.046	-0.009	
251	A	0.079	0.070	
252	O	2.548	0.144	
253	C	1.508	0.074	
254	N	1.552	0.141	
255	N	0.009	0.006	
256	C	1.667	0.122	
257	E	0.154	0.096	
258	O	0.010	-0.048	
259	A	0.418	0.038	
260	O	0.891	-0.158	

261	A	0.151	-0.041	
262	**A**	**4.010** **	**-0.128**	**H**
263	N	0.187	0.036	
264	O	0.029	0.014	
265	N	0.423	0.078	
266	E	0.483	0.078	
267	**C**	**2.721** *	**0.092**	**D**
268	O	0.119	0.027	
269	A	0.603	-0.011	
270	E	1.108	0.041	
271	C	0.000	-0.004	
272	E	2.128	0.075	
273	E	0.764	0.073	
274	C	0.288	0.051	
275	O	0.042	-0.044	
276	A	1.927	-0.090	
277	O	0.024	0.023	
278	N	0.019	0.034	
279	A	0.075	0.021	
280	A	0.529	0.055	
281	N	1.205	0.056	
282	**O**	**10.493** ***	**-0.269**	**H**
283	O	2.307	0.186	
284	**A**	**3.739** *	**-0.143**	**H**
285	O	0.000	-0.032	
286	**C**	**8.792** ***	**0.185**	**D**
287	N	0.013	-0.113	
288	E	0.530	0.009	
289	**E**	**6.651** **	**0.162**	**D**
290	C	0.161	-0.005	
291	**O**	**13.105** ***	**-0.327**	**H**
292	N	2.404	-0.238	
293	O	0.023	0.044	
294	C	0.045	0.000	
295	C	0.599	0.047	
296	C	0.223	-0.072	
297	A	0.196	0.066	
298	**E**	**11.386** ***	**-0.371**	**H**

299	C	1.385	0.088	
300	**E**	**7.586*****	**0.125**	**D**

Note. *p < .10, **p < .05, ***p < .01. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; χ^2_M = Mantel Chi-Square; SMD = Standardized Mean Difference; D = Desirable Response Condition; H = Honest Response Condition. Items in boldface are identified as functioning differentially.

Table 14

Linking Coefficients from the Mean-Sigma Method for the Dichotomously and Polytomously Scored Data

	Dichotomous Data		Polytomous Data	
Factor	α	β	α	β
C	1.064	0.821	1.099	0.281
E	0.905	0.273	1.200	0.637
N	0.894	-0.595	1.143	-0.435
O	0.892	0.178	0.909	0.206
A	0.758	0.190	0.888	-0.113

Note. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 15

Results of the Differential Functioning Item Analysis using the DFIT Procedure on the Dichotomously Scored Items

Item	Factor	NCDIF	χ^2
1	C	0.0000	444.130 *
2	**N**	**0.0200**	**738.180** *
3	**O**	**0.0140**	**12277.090** *
4	**A**	**0.0160**	**595.090** *
5	A	0.0030	156.240
6	O	0.0010	1614.070 *
7	C	0.0010	20904.440 *
8	A	0.0040	392.710 *
9	**A**	**0.0080**	**283.140** *
10	E	0.0020	298.040 *
11	E	0.0020	9009.610 *
12	C	0.0030	155.240
13	A	0.0020	1345.700 *
14	**A**	**0.0180**	**3827.020** *
15	**A**	**0.0130**	**275.720** *
16	O	0.0000	10579.430 *
17	**O**	**0.0120**	**2528.650** *
18	E	0.0040	5638.450 *
19	O	0.0000	378.890 *
20	A	0.0010	1541.850 *
21	**E**	**0.0100**	**15767.100** *
22	C	0.0030	154.440
23	A	0.0020	206.790 *
24	**E**	**0.0110**	**3925.190** *
25	**A**	**0.0100**	**105130.600** *
26	A	0.0020	154.320
27	A	0.0030	278.050 *
28	O	0.0000	222.640 *
29	**O**	**0.0190**	**2600.050** *
30	N	0.0040	535.040 *
31	A	0.0240	177.440
32	**N**	**0.0080**	**3209.280** *

33	C	0.0020	330.500 *
34	**C**	**0.0200**	**436.270 ***
35	N	0.0050	348.720 *
36	**N**	**0.0150**	**916.860 ***
37	A	0.0030	319.420 *
38	E	0.0030	13800.870 *
39	E	0.0010	172.170
40	N	0.0000	159.900
41	E	0.0020	1680.180 *
42	**A**	**0.0070**	**224.460 ***
43	N	0.0090	156.980
44	N	0.0040	154.370
45	C	0.0020	2933.040 *
46	C	0.0010	704.390 *
47	C	0.0030	3599.260 *
48	N	0.0010	162.060
49	N	0.0000	154.920
50	C	0.0030	1451.590 *
51	**O**	**0.0080**	**374.510 ***
52	E	0.0020	176.740
53	N	0.0010	262.120 *
54	O	0.0040	239.120 *
55	E	0.0000	7690.650 *
56	C	0.0010	541.670 *
57	C	0.0000	3365.440 *
58	O	0.0050	684.800 *
59	**C**	**0.0070**	**6400.540 ***
60	N	0.0020	215.780 *
61	**N**	**0.0090**	**828.860 ***
62	N	0.0020	1250.130 *
63	E	0.0010	154.650
64	N	0.0020	270.200 *
65	O	0.0010	372.970 *
66	O	0.0000	233.530 *
67	E	0.0030	252.050 *
68	A	0.0050	964.320 *
69	O	0.0050	447.530 *
70	**A**	**0.0070**	**2553.600 ***

71	A	0.0010	402.440 *
72	N	0.0000	165.010
73	A	0.0040	11628.290 *
74	E	0.0030	375.970 *
75	O	0.0000	243.410 *
76	**E**	**0.0130**	**206.710** *
77	O	0.0050	673.400 *
78	O	0.0010	261.040 *
79	N	0.0000	332.850 *
80	C	0.0010	177.920
81	E	0.0020	226.820 *
82	N	0.0030	430.730 *
83	A	0.0020	161.010
84	A	0.0000	274.820 *
85	O	0.0020	8065.340 *
86	O	0.0020	160.220
87	O	0.0030	22031.770 *
88	C	0.0040	166.520
89	O	0.0000	154.010
90	O	0.0020	2950.620 *
91	N	0.0010	168.660
92	**E**	**0.0100**	**610.960** *
93	A	0.0010	541.220 *
94	O	0.0030	522.930 *
95	E	0.0020	1095.670 *
96	C	0.0030	178.930
97	**C**	**0.0060**	**2958.380** *
98	C	0.0010	377.440 *
99	N	0.0010	1268.440 *
100	N	0.0030	273.650 *
101	C	0.0070	154.000
102	C	0.0030	155.510
103	O	0.0010	413.070 *
104	O	0.0000	5102.600 *
105	E	0.0020	1051.750 *
106	C	0.0020	799.380 *
107	N	0.0010	5675.260 *
108	C	0.0000	227.890 *

109	A	0.0020	475.170 *
110	E	0.0020	4397.240 *
111	E	0.0000	469.920 *
112	E	0.0020	2624.270 *
113	**O**	**0.0060**	**774.220** *
114	A	0.0000	155.650
115	**C**	**0.0060**	**1245.120** *
116	C	0.0020	160.200
117	N	0.0010	216.940 *
118	A	0.0040	159.810
119	N	0.0040	160.580
120	C	0.0020	3012.160 *
121	C	0.0040	165.290
122	N	0.0000	980.320 *
123	N	0.0010	22505.300 *
124	E	0.0040	270.150 *
125	**O**	**0.0100**	**15615.540** *
126	A	0.0000	155.890
127	A	0.0030	2593.050 *
128	C	0.0000	154.020
129	A	0.0030	463.430 *
130	A	0.0040	789.530 *
131	A	0.0040	978.270 *
132	A	0.0040	159.520
133	N	0.0020	169.690
134	N	0.0010	177.200
135	N	0.0020	1403.780 *
136	E	0.0010	640.850 *
137	N	0.0060	174.600
138	N	0.0000	157.220
139	C	0.0010	154.380
140	O	0.0010	269.290 *
141	**O**	**0.0140**	**275.020** *
142	O	0.0010	1058.070 *
143	**A**	**0.0060**	**210.190** *
144	O	0.0010	154.290
145	**O**	**0.0080**	**511.160** *
146	**E**	**0.0140**	**3721.800** *

147	A	0.0050	188.810
148	E	0.0010	179.570
149	A	0.0030	187.610
150	**E**	**0.0090**	**20829.080** *
151	C	0.0020	206.690 *
152	E	0.0030	326.910 *
153	N	0.0020	567.060 *
154	**A**	**0.0060**	**2414.360** *
155	**E**	**0.0060**	**774.010** *
156	N	0.0020	1070.090 *
157	C	0.0000	365.150 *
158	C	0.0040	159.690
159	**C**	**0.0150**	**1128.100** *
160	O	0.0000	243.190 *
161	N	0.0050	156.580
162	E	0.0020	2284.710 *
163	C	0.0000	1219.860 *
164	E	0.0010	594.380 *
165	A	0.0040	173.340
166	**N**	**0.0110**	**386.080** *
167	C	0.0020	221.270 *
168	A	0.0030	294.130 *
169	N	0.0010	196.510
170	O	0.0020	466.940 *
171	E	0.0040	3815.680 *
172	E	0.0040	540.340 *
173	N	0.0030	418.910 *
174	**O**	**0.0110**	**1574.120** *
175	O	0.0020	1025.730 *
176	C	0.0000	156.450
177	C	0.0000	256.970 *
178	**E**	**0.0090**	**16482.290** *
179	**A**	**0.0060**	**35458.570** *
180	**O**	**0.0060**	**693.530** *
181	E	0.0050	1303.400 *
182	N	0.0050	179.550
183	C	0.0030	184.400
184	N	0.0010	502.130 *

185	E	0.0000	53656.010 *
186	**C**	**0.0080**	**244.190** *
187	**E**	**0.0070**	**2384.030** *
188	N	0.0020	163.340
189	A	0.0020	371.120 *
190	E	0.0030	326.300 *
191	**A**	**0.0150**	**291.550** *
192	N	0.0020	441.950 *
193	E	0.0020	3044.950 *
194	**O**	**0.0120**	**638.250** *
195	**O**	**0.0180**	**9404.390** *
196	E	0.0010	1305.720 *
197	**E**	**0.0060**	**8251.420** *
198	E	0.0010	1058.100 *
199	A	0.0010	275.140 *
200	**E**	**0.0080**	**641.990** *
201	C	0.0010	198.890 *
202	O	0.0000	156.560
203	N	0.0030	322.850 *
204	C	0.0010	4930.390 *
205	**E**	**0.0140**	**323.940** *
206	**N**	**0.0080**	**6434.030** *
207	C	0.0000	228.560 *
208	O	0.0020	2404.600 *
209	**E**	**0.0100**	**400.580** *
210	C	0.0020	1235.540 *
211	**O**	**0.0320**	**2165.100** *
212	E	0.0020	1340.340 *
213	**C**	**0.0090**	**353.190** *
214	C	0.0000	174.810
215	**O**	**0.0140**	**786.640** *
216	N	0.0010	901.430 *
217	N	0.0000	154.630
218	N	0.0010	272.950 *
219	**C**	**0.0090**	**196.330** *
220	**E**	**0.0070**	**332.250** *
221	O	0.0000	293.230 *
222	N	0.0050	156.970

223	O	0.0020	167.550
224	**O**	**0.0230**	**413.640** *
225	A	0.0000	302.910 *
226	N	0.0010	1746.500 *
227	**A**	**0.0070**	**229.980** *
228	**O**	**0.0110**	**525.710** *
229	E	0.0010	216.370 *
230	**E**	**0.0090**	**4549.100** *
231	C	0.0000	154.200
232	A	0.0030	8240.410 *
233	N	0.0040	1891.670 *
234	C	0.0010	1088.450 *
235	O	0.0010	524.060 *
236	**A**	**0.0130**	**4380.780** *
237	E	0.0050	9850.490 *
238	**A**	**0.0180**	**202.740** *
239	O	0.0020	545.950 *
240	A	0.0040	284.910 *
241	**N**	**0.0070**	**425.250** *
242	A	0.0020	1401.720 *
243	N	0.0010	373.730 *
244	E	0.0050	1129.600 *
245	**C**	**0.0060**	**341.970** *
246	E	0.0050	363.090 *
247	N	0.0020	212.730 *
248	**C**	**0.0060**	**4582.160** *
249	A	0.0010	495.200 *
250	A	0.0010	161.200
251	A	0.0050	172.020
252	**O**	**0.0060**	**540.310** *
253	C	0.0000	184.700
254	N	0.0010	295.650 *
255	N	0.0010	1105.790 *
256	C	0.0040	392.810 *
257	E	0.0000	422.530 *
258	O	0.0000	156.260
259	A	0.0020	169.240
260	O	0.0010	513.810 *

261	A	0.0000	570.080 *
262	A	0.0010	590.890 *
263	N	0.0040	194.670
264	O	0.0010	161.650
265	N	0.0020	284.670 *
266	E	0.0040	156.000
267	C	0.0000	238.250 *
268	O	0.0010	199.600 *
269	A	0.0020	182.790
270	E	0.0040	166.350
271	C	0.0010	165.370
272	**E**	**0.0080**	**251.070 ***
273	**E**	**0.0080**	**311.560 ***
274	C	0.0020	288.320 *
275	O	0.0020	335.460 *
276	A	0.0050	7495.990 *
277	O	0.0010	161.240
278	**N**	**0.0130**	**686.780 ***
279	A	0.0060	154.260
280	A	0.0020	174.210
281	N	0.0010	234.240 *
282	**O**	**0.0130**	**491.740 ***
283	O	0.0010	356.330 *
284	**A**	**0.0070**	**310.440 ***
285	O	0.0020	314.920 *
286	**C**	**0.0070**	**663.510 ***
287	N	0.0000	170.970
288	E	0.0000	293.160 *
289	**E**	**0.0140**	**488.680 ***
290	C	0.0010	330.940 *
291	**O**	**0.0200**	**589.160 ***
292	N	0.0000	876.780 *
293	O	0.0020	166.240
294	C	0.0010	423.680 *
295	C	0.0020	465.390 *
296	C	0.0010	1341.520 *
297	A	0.0010	217.150 *
298	**E**	**0.0140**	**1241.650 ***

299	C	0.0020	232.170 *
300	**E**	**0.0180**	**563.610 ***

Note. *p < .01. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; NCDIF = Non-compensatory Differential Item Functioning. Items in boldface are identified as functioning differentially.

Table 16

Results of the Differential Functioning Item Analysis using the DFIT Procedure on the Polytomously Scored Items

Item	Factor	NCDIF	χ^2
1	C	0.027	10406.660 *
2	N	0.083	472.040 *
3	**O**	**0.126**	**16457.670** *
4	**A**	**0.366**	**7905.900** *
5	A	0.054	1407.120 *
6	O	0.010	5414.690 *
7	C	0.082	5488.450 *
8	A	0.089	15259.630 *
9	**A**	**0.193**	**13248.980** *
10	**E**	**4.600**	**8435.610** *
11	**E**	**1.710**	**6705.590** *
12	C	0.032	369.790 *
13	A	0.032	10295.970 *
14	**A**	**0.249**	**3073.100** *
15	**A**	**0.166**	**103769.800** *
16	O	0.073	728675.300 *
17	**O**	**0.146**	**22209.130** *
18	E	0.069	29974.150 *
19	O	0.002	7313.550 *
20	A	0.060	7746.310 *
21	E	0.037	105020.100 *
22	C	0.008	1069.290 *
23	A	0.086	1920.650 *
24	E	0.005	169.050
25	A	0.062	NR *
26	A	0.035	21609.300 *
27	A	0.022	220872.800 *
28	O	0.002	650.190 *
29	**O**	**0.285**	**11607.120** *
30	N	0.005	551.840 *
31	**A**	**0.110**	**3533.130** *
32	**N**	**0.111**	**23502.590** *

33	C	0.005	312.320 *
34	C	0.056	14543.330 *
35	N	0.056	2933.910 *
36	N	0.040	16141.160 *
37	**A**	**0.106**	**6236.870** *
38	E	0.010	624.960 *
39	**E**	**0.320**	**32298.230** *
40	N	0.003	3952.880 *
41	E	0.034	695.800 *
42	A	0.086	23817.920 *
43	N	0.025	2762.950 *
44	N	0.016	325.480 *
45	C	0.002	1106.720 *
46	C	0.002	5149.010 *
47	C	0.041	4963.690 *
48	N	0.014	439.320 *
49	N	0.001	916.640 *
50	C	0.011	159.920
51	O	0.008	195.360
52	**E**	**0.105**	**6279.050** *
53	N	0.002	635.030 *
54	O	0.022	1537.090 *
55	E	0.016	378.620 *
56	C	0.003	169.350
57	C	0.021	392.440 *
58	O	0.014	3992.920 *
59	**C**	**0.097**	**22083.840** *
60	N	0.038	173.190
61	N	0.029	492.810 *
62	N	0.033	157.370
63	**E**	**0.181**	**4313.490** *
64	N	0.019	444.860 *
65	O	0.007	358.300 *
66	O	0.007	233.880 *
67	**E**	**0.105**	**8094.250** *
68	A	0.001	154.260
69	**O**	**0.096**	**137198.200** *
70	A	0.058	5009.880 *

71	A	0.057	6792.920 *
72	N	0.037	961.250 *
73	A	0.005	154068.800 *
74	**E**	**0.136**	**660.860** *
75	O	0.001	2190.840 *
76	E	0.085	2952.270 *
77	O	0.061	93955.450 *
78	O	0.003	1390.910 *
79	N	0.004	922.520 *
80	C	0.001	275.500 *
81	E	0.089	24224.740 *
82	N	0.001	340.430 *
83	A	0.035	10907.320 *
84	**A**	**0.133**	**6932.980** *
85	O	0.016	2270.330 *
86	O	0.020	1910.260 *
87	O	0.042	225159.900 *
88	C	0.001	154.970
89	O	0.014	1892.760 *
90	O	0.019	2261.910 *
91	N	0.002	373.190 *
92	E	0.017	8668.420 *
93	A	0.045	6219.060 *
94	O	0.003	154.690
95	E	0.061	923.290 *
96	C	0.004	736.030 *
97	C	0.078	3062.940 *
98	C	0.020	371.110 *
99	N	0.012	646.960 *
100	N	0.002	175.710
101	C	0.034	1038.790 *
102	C	0.014	17231.830 *
103	O	0.016	4222.640 *
104	O	0.001	6161.270 *
105	E	0.002	320.580 *
106	C	0.006	1720.180 *
107	**N**	**0.101**	**2322.990** *
108	C	0.012	17464.350 *

109	A	0.003	1224.230 *
110	E	0.089	3701.190 *
111	E	0.081	4403.310 *
112	**E**	**0.171**	**2558.880 ***
113	O	0.023	2202.760 *
114	A	0.012	2735.750 *
115	C	0.015	3691.860 *
116	C	0.007	3559.590 *
117	N	0.011	248.840 *
118	A	0.037	17213.790 *
119	N	0.012	1807.340 *
120	C	0.014	340.370 *
121	C	0.037	219.290 *
122	N	0.014	3614.600 *
123	N	0.020	624.200 *
124	E	0.014	1924.270 *
125	**O**	**0.103**	**139187.000 ***
126	A	0.073	4705.920 *
127	**A**	**0.103**	**1848.880 ***
128	C	0.015	213.940 *
129	**A**	**0.262**	**38381.530 ***
130	**A**	**0.109**	**6627.530 ***
131	A	0.000	516.230 *
132	A	0.021	3770.740 *
133	N	0.019	277.440 *
134	N	0.003	16597.930 *
135	N	0.015	204.160 *
136	**E**	**0.155**	**4320.790 ***
137	N	0.037	160.420
138	N	0.002	176.410
139	C	0.021	203.430 *
140	O	0.045	1090.200 *
141	**O**	**0.161**	**6345.880 ***
142	O	0.021	6620.250 *
143	A	0.092	8167.600 *
144	O	0.003	158.880
145	**O**	**0.161**	**4843.180 ***
146	**E**	**0.293**	**44012.640 ***

147	**A**	**0.117**	**1254.800** *
148	**E**	**0.243**	**10069.700** *
149	A	0.007	2360.920 *
150	**E**	**0.216**	**4118.090** *
151	C	0.003	1725.390 *
152	E	0.010	2740.460 *
153	N	0.034	842.540 *
154	**A**	**0.268**	**4058.800** *
155	E	0.067	1507.710 *
156	N	0.019	1677.540 *
157	C	0.007	167.910
158	C	0.040	620.890 *
159	C	0.021	170.840
160	O	0.001	237.700 *
161	N	0.003	200.730 *
162	**E**	**0.181**	**14278.620** *
163	C	0.014	4242.290 *
164	E	0.085	26791.640 *
165	A	0.023	897.910 *
166	**N**	**0.122**	**512.750** *
167	C	0.001	158.520
168	A	0.061	18287.000 *
169	N	0.043	645.640 *
170	O	0.025	8901.660 *
171	**E**	**0.442**	**1471.330** *
172	**E**	**0.350**	**5471.820** *
173	N	0.005	346.260 *
174	**O**	**0.144**	**178803.000** *
175	O	0.020	1332.660 *
176	C	0.011	191.950
177	C	0.016	269.170 *
178	**E**	**0.139**	**NR** *
179	A	0.032	2365.890 *
180	O	0.077	11016.070 *
181	E	0.090	7128.670 *
182	N	0.003	221.970 *
183	C	0.032	157.550
184	N	0.026	763.960 *

185	E	0.021	580.540 *
186	C	0.016	183.160
187	**E**	**0.172**	**1170.010** *
188	N	0.029	595.780 *
189	**A**	**0.121**	**2644.150** *
190	E	0.038	322.540 *
191	**A**	**0.521**	**1184.760** *
192	N	0.008	305.790 *
193	E	0.073	5115.350 *
194	**O**	**0.130**	**19212.160** *
195	**O**	**0.130**	**79311.130** *
196	**E**	**0.244**	**7817.940** *
197	E	0.005	366832.000 *
198	E	0.067	1175.400 *
199	A	0.009	803.060 *
200	**E**	**0.246**	**2858.040** *
201	C	0.002	174.790
202	O	0.000	1718.210 *
203	N	0.023	1859.720 *
204	C	0.004	176.760
205	E	0.057	5481.540 *
206	N	0.035	14845.170 *
207	C	0.002	585.900 *
208	O	0.064	163383.200 *
209	E	0.084	762.160 *
210	C	0.003	156.140
211	**O**	**0.458**	**12750.660** *
212	E	0.083	1155.710 *
213	C	0.086	336.300
214	C	0.039	154.200
215	**O**	**0.346**	**2359.110** *
216	N	0.009	613.750 *
217	N	0.002	154.630
218	N	0.013	2323.490 *
219	C	0.008	222.100 *
220	**E**	**0.218**	**2729.240** *
221	O	0.004	1241.550 *
222	N	0.028	205.370 *

223	O	0.005	68081.640	*
224	**O**	**0.363**	**11784.360**	*
225	A	0.001	423.300	*
226	N	0.002	574.310	*
227	**A**	**0.140**	**14204.900**	*
228	O	0.054	3258.050	*
229	**E**	**0.117**	**784.220**	*
230	**E**	**0.165**	**65050.140**	*
231	C	0.001	2073.810	*
232	A	0.084	2509.350	*
233	N	0.032	93064.410	*
234	C	0.011	197.610	*
235	O	0.007	165.900	
236	**A**	**0.212**	**246092.200**	*
237	**E**	**0.130**	**5845.740**	*
238	**A**	**0.401**	**5413.270**	*
239	O	0.007	1783.930	*
240	A	0.056	381.740	*
241	N	0.030	207.220	*
242	A	0.003	780.600	*
243	N	0.016	2388.150	*
244	E	0.001	4813.970	*
245	C	0.004	1797.430	*
246	**E**	**0.177**	**47267.610**	*
247	N	0.011	15103.730	*
248	C	0.048	3337.890	*
249	A	0.000	609.620	*
250	A	0.015	10830.400	*
251	A	0.060	6948.650	*
252	O	0.006	390.660	*
253	C	0.015	393.160	*
254	N	0.032	220437.300	*
255	N	0.006	412.630	*
256	C	0.059	288.860	*
257	E	0.002	7901.000	*
258	O	0.000	154.360	
259	A	0.026	243708.400	*
260	O	0.019	1458.100	*

261	A	0.026	898.370	*
262	A	0.028	143480.700	*
263	N	0.008	505.020	*
264	O	0.008	178.090	
265	N	0.039	7347.840	*
266	**E**	**0.173**	**14862.730**	*
267	C	0.014	6371.620	*
268	O	0.007	21821.520	*
269	**A**	**0.148**	**2321.780**	*
270	E	0.036	234.280	*
271	C	0.001	157.120	
272	E	0.057	422.030	*
273	**E**	**0.172**	**23079.280**	*
274	C	0.005	796.310	*
275	O	0.002	451.320	*
276	A	0.017	2238.190	*
277	O	0.002	160.420	
278	N	0.003	280.960	*
279	A	0.085	1641.410	*
280	**A**	**0.144**	**2830.040**	*
281	N	0.032	40995.250	*
282	**O**	**0.215**	**5980.970**	*
283	O	0.013	167.700	
284	A	0.006	189.120	
285	O	0.002	154.040	
286	C	0.085	2691.140	*
287	N	0.001	4197.370	*
288	E	0.047	10618.320	*
289	E	0.008	172.710	
290	C	0.010	166.580	
291	**O**	**0.330**	**2183.750**	*
292	N	0.012	3166.700	*
293	O	0.000	474.970	*
294	C	0.033	262.500	*
295	C	0.016	2339.170	*
296	C	0.004	18337.300	*
297	A	0.017	789.620	*
298	**E**	**0.398**	**4548.470**	*

299	C	0.023	393.420 *
300	E	0.002	175.790

Note. *p < .01. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness, NR = Not Reported by DFIT; NCDIF = Non-compensatory Differential Item Functioning. Items in boldface are identified as functioning differentially.

Table 17

Summary of the Decisions about Differential Item Functioning from the Observed Score and Latent Parameter Analyses

Item	Factor	MH	M	DFIT D	DFIT P
1	C	No	Yes	No	No
2	**N**	**Yes**	**Yes**	**Yes**	**No**
3	**O**	**Yes**	**Yes**	**Yes**	**Yes**
4	**A**	**Yes**	**Yes**	**Yes**	**Yes**
5	A	No	No	No	No
6	O	No	No	No	No
7	C	No	No	No	No
8	A	No	Yes	No	No
9	**A**	**No**	**Yes**	**Yes**	**Yes**
10	E	No	No	No	Yes
11	E	No	No	No	Yes
12	C	No	No	No	No
13	A	No	No	No	No
14	**A**	**Yes**	**Yes**	**Yes**	**Yes**
15	**A**	**Yes**	**Yes**	**Yes**	**Yes**
16	O	No	Yes	No	No
17	**O**	**Yes**	**Yes**	**Yes**	**Yes**
18	E	Yes	Yes	No	No
19	O	No	No	No	No
20	A	No	No	No	No
21	**E**	**Yes**	**Yes**	**Yes**	**No**
22	C	No	No	No	No
23	A	No	No	No	No
24	**E**	**Yes**	**Yes**	**Yes**	**No**
25	**A**	**Yes**	**Yes**	**Yes**	**No**
26	A	No	No	No	No
27	A	No	No	No	No
28	O	No	No	No	No
29	**O**	**Yes**	**Yes**	**Yes**	**Yes**
30	N	No	No	No	No
31	A	No	No	No	Yes
32	**N**	**Yes**	**Yes**	**Yes**	**Yes**

33	C	No	No	No	No
34	**C**	**Yes**	**Yes**	**Yes**	**No**
35	N	No	Yes	No	No
36	**N**	**Yes**	**Yes**	**Yes**	**No**
37	A	No	No	No	Yes
38	E	Yes	No	No	No
39	E	No	No	No	Yes
40	N	No	No	No	No
41	E	No	No	No	No
42	A	No	No	Yes	No
43	N	No	No	No	No
44	N	No	No	No	No
45	C	No	No	No	No
46	C	No	No	No	No
47	C	No	Yes	No	No
48	N	No	No	No	No
49	N	No	No	No	No
50	C	No	No	No	No
51	O	Yes	No	Yes	No
52	E	No	No	No	Yes
53	N	No	No	No	No
54	O	No	No	No	No
55	E	No	No	No	No
56	C	No	No	No	No
57	C	No	No	No	No
58	O	No	No	No	No
59	**C**	**Yes**	**Yes**	**Yes**	**Yes**
60	N	No	No	No	No
61	**N**	**Yes**	**Yes**	**Yes**	**No**
62	N	No	No	No	No
63	E	No	No	No	Yes
64	N	No	No	No	No
65	O	No	No	No	No
66	O	No	No	No	No
67	E	No	No	No	Yes
68	A	Yes	Yes	No	No
69	**O**	**Yes**	**Yes**	**No**	**Yes**
70	A	Yes	No	Yes	No

71	A	No	No	No	No
72	N	No	No	No	No
73	A	Yes	Yes	No	No
74	E	No	No	No	Yes
75	O	No	No	No	No
76	E	No	No	Yes	No
77	O	Yes	Yes	No	No
78	O	No	No	No	No
79	N	No	No	No	No
80	C	No	No	No	No
81	E	No	No	No	No
82	N	No	No	No	No
83	A	No	No	No	No
84	A	No	No	No	Yes
85	O	No	No	No	No
86	O	No	No	No	No
87	O	No	Yes	No	No
88	C	No	No	No	No
89	O	No	No	No	No
90	O	No	Yes	No	No
91	N	No	No	No	No
92	**E**	**Yes**	**Yes**	**Yes**	**No**
93	A	No	No	No	No
94	O	No	No	No	No
95	E	No	No	No	No
96	C	No	No	No	No
97	**C**	**Yes**	**Yes**	**Yes**	**No**
98	C	No	No	No	No
99	N	No	No	No	No
100	N	No	No	No	No
101	C	No	No	No	No
102	C	No	No	No	No
103	O	No	No	No	No
104	O	No	No	No	No
105	E	No	No	No	No
106	C	No	No	No	No
107	N	No	Yes	No	Yes
108	C	No	No	No	No

109	A	No	No	No	No
110	E	No	Yes	No	No
111	E	No	No	No	No
112	E	No	No	No	Yes
113	**O**	**Yes**	**Yes**	**Yes**	**No**
114	A	No	No	No	No
115	C	No	No	Yes	No
116	C	No	No	No	No
117	N	No	No	No	No
118	A	No	No	No	No
119	N	No	No	No	No
120	C	No	No	No	No
121	C	No	No	No	No
122	N	No	No	No	No
123	N	No	No	No	No
124	E	No	No	No	No
125	**O**	**Yes**	**Yes**	**Yes**	**Yes**
126	A	No	No	No	No
127	A	No	No	No	Yes
128	C	No	No	No	No
129	A	No	Yes	No	Yes
130	A	No	No	No	Yes
131	A	No	No	No	No
132	A	No	No	No	No
133	N	No	No	No	No
134	N	No	No	No	No
135	N	No	No	No	No
136	E	No	No	No	Yes
137	N	No	No	No	No
138	N	No	No	No	No
139	C	No	No	No	No
140	O	No	No	No	No
141	**O**	**Yes**	**Yes**	**Yes**	**Yes**
142	O	No	Yes	No	No
143	A	No	No	Yes	Yes
144	O	No	No	No	No
145	**O**	**Yes**	**Yes**	**Yes**	**Yes**
146	**E**	**Yes**	**Yes**	**Yes**	**Yes**

147	A	No	No	No	Yes
148	E	No	No	No	Yes
149	A	No	No	No	No
150	**E**	**Yes**	**Yes**	**Yes**	**Yes**
151	C	No	No	No	No
152	E	No	No	No	No
153	N	No	No	No	No
154	**A**	**Yes**	**Yes**	**Yes**	**Yes**
155	E	No	No	Yes	No
156	N	No	No	No	No
157	C	No	No	No	No
158	C	No	Yes	No	No
159	C	Yes	No	Yes	No
160	O	No	No	No	No
161	N	No	No	No	No
162	E	No	No	No	Yes
163	C	No	No	No	No
164	E	No	No	No	No
165	A	No	No	No	No
166	**N**	**Yes**	**Yes**	**Yes**	**Yes**
167	C	No	No	No	No
168	A	No	No	No	No
169	N	No	No	No	No
170	O	No	No	No	No
171	E	No	Yes	No	Yes
172	E	No	No	No	Yes
173	N	No	No	No	No
174	**O**	**Yes**	**Yes**	**Yes**	**Yes**
175	O	No	No	No	No
176	C	No	No	No	No
177	C	No	No	No	No
178	**E**	**Yes**	**Yes**	**Yes**	**Yes**
179	**A**	**Yes**	**Yes**	**Yes**	**No**
180	**O**	**Yes**	**Yes**	**Yes**	**No**
181	E	Yes	Yes	No	No
182	N	No	No	No	No
183	C	No	No	No	No
184	N	No	No	No	No

185	E	No	No	No	No
186	C	No	No	Yes	No
187	**E**	**Yes**	**Yes**	**Yes**	**Yes**
188	N	No	No	No	No
189	A	No	No	No	Yes
190	E	No	Yes	No	No
191	**A**	**Yes**	**Yes**	**Yes**	**Yes**
192	N	No	No	No	No
193	E	No	No	No	No
194	**O**	**Yes**	**Yes**	**Yes**	**Yes**
195	**O**	**Yes**	**Yes**	**Yes**	**Yes**
196	E	No	No	No	Yes
197	E	No	No	Yes	No
198	E	No	No	No	No
199	A	No	No	No	No
200	**E**	**No**	**Yes**	**Yes**	**Yes**
201	C	No	No	No	No
202	O	No	No	No	No
203	N	No	No	No	No
204	C	No	No	No	No
205	**E**	**Yes**	**Yes**	**Yes**	**No**
206	**N**	**Yes**	**Yes**	**Yes**	**No**
207	C	No	No	No	No
208	O	No	Yes	No	No
209	**E**	**Yes**	**Yes**	**Yes**	**No**
210	C	No	No	No	No
211	**O**	**Yes**	**Yes**	**Yes**	**Yes**
212	E	No	No	No	No
213	C	Yes	No	Yes	No
214	C	No	No	No	No
215	**O**	**Yes**	**Yes**	**Yes**	**Yes**
216	N	No	No	No	No
217	N	No	No	No	No
218	N	No	No	No	No
219	C	No	No	Yes	No
220	E	No	No	Yes	Yes
221	O	No	No	No	No
222	N	No	No	No	No

223	O	No	No	No	No
224	**O**	**Yes**	**Yes**	**Yes**	**Yes**
225	A	No	No	No	No
226	N	No	No	No	No
227	A	No	No	Yes	Yes
228	O	Yes	No	Yes	No
229	E	No	No	No	Yes
230	**E**	**Yes**	**Yes**	**Yes**	**Yes**
231	C	No	No	No	No
232	A	No	No	No	No
233	N	Yes	Yes	No	No
234	C	No	No	No	No
235	O	No	No	No	No
236	**A**	**Yes**	**Yes**	**Yes**	**Yes**
237	E	Yes	No	No	Yes
238	**A**	**No**	**Yes**	**Yes**	**Yes**
239	O	No	No	No	No
240	A	No	No	No	No
241	N	No	No	Yes	No
242	A	No	No	No	No
243	N	No	No	No	No
244	E	No	No	No	No
245	C	Yes	No	Yes	No
246	E	No	No	No	Yes
247	N	No	No	No	No
248	**C**	**Yes**	**Yes**	**Yes**	**No**
249	A	No	No	No	No
250	A	No	No	No	No
251	A	No	No	No	No
252	O	Yes	No	Yes	No
253	C	No	No	No	No
254	N	No	No	No	No
255	N	No	No	No	No
256	C	No	No	No	No
257	E	No	No	No	No
258	O	No	No	No	No
259	A	No	No	No	No
260	O	No	No	No	No

261	A	No	No	No	No
262	A	No	Yes	No	No
263	N	No	No	No	No
264	O	No	No	No	No
265	N	No	No	No	No
266	E	No	No	No	Yes
267	C	No	Yes	No	No
268	O	No	No	No	No
269	A	No	No	No	Yes
270	E	No	No	No	No
271	C	No	No	No	No
272	E	No	No	Yes	No
273	E	No	No	Yes	Yes
274	C	No	No	No	No
275	O	No	No	No	No
276	A	Yes	No	No	No
277	O	No	No	No	No
278	N	Yes	No	Yes	No
279	A	No	No	No	No
280	A	No	No	No	Yes
281	N	No	No	No	No
282	**O**	**Yes**	**Yes**	**Yes**	**Yes**
283	O	No	No	No	No
284	A	No	Yes	Yes	No
285	O	No	No	No	No
286	**C**	**Yes**	**Yes**	**Yes**	**No**
287	N	No	No	No	No
288	E	No	No	No	No
289	**E**	**Yes**	**Yes**	**Yes**	**No**
290	C	No	No	No	No
291	**O**	**Yes**	**Yes**	**Yes**	**Yes**
292	N	No	No	No	No
293	O	No	No	No	No
294	C	No	No	No	No
295	C	No	No	No	No
296	C	No	No	No	No
297	A	No	No	No	No
298	**E**	**Yes**	**Yes**	**Yes**	**Yes**

299	C	No	No	No	No
300	**E**	**Yes**	**Yes**	**Yes**	**No**

Note. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; MH = Mantel-Haenszel Procedure; M = Mantel Procedure; DFIT D = DFIT Procedure with Dichotomous Data; DFIT P = DFIT Procedure with Polytomous Data. Items in boldface are classified as capable of being distorted.

Table 18

Internal Consistency Reliability for the Personality Factors and Sub-Facets on the IPIP for Each Response Instruction Condition

Factor	Sub Factor	Honest Condition	Desirable Condition	No Instructions Condition
C		**0.928**	**0.947**	**0.853**
	1	0.758	0.845	0.688
	2	0.869	0.824	0.347
	3	0.748	0.846	0.822
	4	0.769	0.793	0.805
	5	0.837	0.874	0.878
	6	0.767	0.786	0.798
E		**0.908**	**0.939**	**0.910**
	1	0.813	0.876	0.853
	2	0.801	0.842	0.781
	3	0.781	0.814	0.713
	4	0.706	0.581	0.263
	5	0.775	0.786	0.755
	6	0.814	0.847	0.807
N		**0.930**	**0.951**	**0.926**
	1	0.781	0.807	0.829
	2	0.861	0.872	0.865
	3	0.879	0.902	0.866
	4	0.770	0.813	0.730
	5	0.739	0.707	0.725
	6	0.772	0.801	0.720
O		**0.899**	**0.890**	**0.875**
	1	0.740	0.758	0.808
	2	0.807	0.837	0.836
	3	0.764	0.742	0.781
	4	0.769	0.770	0.745
	5	0.793	0.796	0.782

	6	0.632	0.668	0.625
A		**0.924**	**0.911**	**0.934**
	1	0.875	0.829	0.850
	2	0.803	0.761	0.829
	3	0.814	0.837	0.839
	4	0.784	0.732	0.805
	5	0.658	0.615	0.718
	6	0.745	0.697	0.739

Note. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness. Reliabilities were computed using Cronbach's alpha. Reliabilities in boldface are for the personality factor.

Table 19

Means and Variances in the Honest, Desirable, and No Response Instructions Conditions for Each Item using the Polytomously Scored Data

			H		D		NI	
Scale	Item	Factor	*M*	S^2	*M*	S^2	*M*	S^2
IPIP	1	C	4.078	0.494	4.480	0.381	4.273	0.350
	7	C	3.171	1.393	3.696	1.310	3.289	1.357
	12	C	3.062	1.152	3.669	0.841	3.215	1.287
	22	C	4.101	1.388	4.512	0.978	4.306	1.081
	33	C	3.938	0.980	4.088	0.581	3.967	0.782
	34	C	3.822	0.538	4.088	0.403	3.843	0.400
	45	C	3.845	0.788	4.064	0.706	3.950	0.581
	46	C	3.240	0.965	3.440	0.813	3.471	0.751
	47	C	4.287	0.878	4.312	0.974	4.223	1.008
	50	C	3.930	0.440	4.176	0.501	4.066	0.496
	56	C	3.938	0.918	4.112	0.713	3.983	0.833
	57	C	4.054	0.802	4.472	0.525	4.207	0.932
	59	C	3.465	1.094	3.726	0.802	3.467	1.007
	80	C	3.411	1.650	3.312	1.555	3.256	1.475
	88	C	3.426	1.184	3.952	0.885	3.504	1.219
	96	C	3.837	0.794	4.200	0.435	4.000	0.567
	97	C	3.132	1.365	3.792	1.166	3.281	1.454
	98	C	4.171	0.580	4.496	0.381	4.264	0.729
	101	C	3.147	1.267	3.696	1.004	3.339	1.076
	102	C	3.240	1.371	3.808	0.898	3.397	1.025
	106	C	3.651	0.916	3.784	0.719	3.783	0.709
	108	C	3.008	1.180	3.528	0.929	3.116	1.037
	115	C	4.132	0.693	4.352	0.649	4.066	0.846
	116	C	3.031	1.046	3.096	1.136	2.983	1.008
	120	C	3.938	0.793	4.128	0.774	4.074	0.669
	121	C	3.357	1.309	3.920	0.881	3.380	1.154
	128	C	3.822	0.741	4.072	0.648	3.777	0.791
	139	C	3.302	1.056	3.832	0.625	3.446	0.933
	151	C	4.256	0.645	4.360	0.668	4.215	0.570
	157	C	3.891	0.410	4.096	0.491	4.058	0.288
	158	C	3.434	0.951	3.887	0.784	3.678	0.820

159	C	4.093	0.523	4.360	0.442	4.215	0.370
163	C	3.550	1.609	3.992	1.234	3.471	1.485
167	C	3.465	1.157	3.536	1.057	3.512	1.202
176	C	4.047	0.779	4.400	0.661	4.149	0.594
177	C	3.798	1.006	4.194	0.661	3.810	0.855
183	C	3.519	1.095	4.056	0.827	3.603	0.991
186	C	3.805	0.851	4.120	0.816	3.711	0.741
201	C	3.674	0.956	4.176	0.582	3.876	0.860
204	C	3.318	1.031	3.371	1.178	3.331	0.890
207	C	3.705	1.038	4.000	0.823	3.686	1.184
210	C	4.163	0.590	4.400	0.419	4.331	0.406
213	C	2.946	1.177	3.600	0.903	3.066	0.962
214	C	4.000	1.039	4.320	0.655	4.074	0.953
219	C	3.194	1.736	3.616	1.545	3.345	1.635
231	C	3.636	0.874	3.984	0.855	3.808	0.677
234	C	3.008	1.252	3.624	1.043	3.339	1.026
245	C	3.225	1.473	3.304	1.181	3.116	1.237
248	C	3.535	1.469	3.792	1.069	3.694	1.164
253	C	4.155	0.773	4.408	0.711	4.231	0.513
256	C	2.961	1.178	3.544	1.169	3.198	1.094
267	C	4.093	0.648	4.256	0.805	4.190	0.772
271	C	4.248	0.657	4.472	0.525	4.182	0.733
274	C	3.349	0.885	3.912	0.823	3.554	0.633
286	C	3.383	0.994	3.832	0.738	3.308	0.904
290	C	2.434	1.185	2.688	1.200	2.413	1.011
294	C	2.713	1.269	2.856	1.334	2.388	1.040
295	C	3.172	1.262	3.742	0.908	3.198	1.160
296	C	3.750	0.709	3.903	0.722	3.777	0.558
299	C	2.930	1.097	3.659	1.063	3.215	1.170
10	E	4.147	0.502	4.352	0.375	4.215	0.487
11	E	3.891	1.269	4.040	1.119	4.083	0.893
18	E	3.250	1.260	3.824	0.985	3.537	0.984
21	E	3.891	0.754	4.144	0.576	4.041	0.690
24	E	3.736	0.696	3.976	0.653	3.942	0.722
38	E	3.915	0.907	4.032	0.773	4.050	0.598
39	E	3.496	1.455	3.408	1.453	3.438	1.248
41	E	4.271	0.605	4.480	0.639	4.331	0.790

52	E	4.171	0.768	4.424	0.569	4.281	0.720
55	E	3.760	1.168	3.792	0.924	3.983	0.766
63	E	4.147	0.845	4.320	0.655	4.190	0.905
67	E	3.752	1.141	4.048	0.853	3.785	1.187
74	E	2.806	1.392	2.824	1.227	2.909	1.267
76	E	3.442	1.264	3.855	0.922	3.595	1.210
81	E	3.659	0.867	4.256	0.595	4.033	0.616
92	E	3.457	1.266	3.704	1.033	3.529	0.951
95	E	4.333	0.568	4.416	0.616	4.421	0.613
105	E	2.023	0.773	2.104	0.787	1.917	0.626
110	E	3.395	0.803	3.792	0.843	3.636	0.867
111	E	3.667	0.771	3.728	0.861	3.645	0.814
112	E	3.798	1.162	3.912	0.871	3.802	1.110
124	E	3.558	0.811	3.616	0.884	3.595	0.626
136	E	3.543	1.531	3.816	1.297	3.711	1.557
146	E	4.357	0.778	4.392	0.660	4.372	0.702
148	E	3.837	1.106	3.776	1.401	3.702	1.394
150	E	4.054	1.099	4.008	1.250	4.066	0.846
152	E	2.581	0.683	2.456	0.750	2.413	0.561
155	E	3.481	0.861	3.576	0.956	3.504	0.835
162	E	4.085	0.938	4.264	0.889	4.033	0.932
164	E	3.690	1.169	3.864	0.877	3.669	0.890
171	E	3.806	1.361	3.824	1.566	3.826	1.295
172	E	3.535	1.407	3.288	1.529	3.488	1.419
178	E	3.349	1.229	3.500	1.130	3.446	1.083
181	E	3.853	0.642	4.056	0.892	3.860	0.688
185	E	3.690	1.122	3.848	0.775	3.529	0.901
187	E	3.566	1.029	3.792	0.940	3.587	0.928
190	E	3.488	1.330	3.544	0.831	3.545	1.333
193	E	3.860	1.684	3.888	1.762	3.983	1.316
196	E	3.884	0.760	4.136	0.715	3.967	0.816
197	E	2.411	1.010	2.344	0.679	2.364	0.817
198	E	4.054	0.974	4.224	0.643	4.107	0.947
200	E	4.344	0.605	4.368	0.863	4.355	0.664
205	E	3.612	1.020	3.880	0.897	3.702	0.944
209	E	3.899	0.685	4.160	0.765	3.959	0.823
212	E	4.217	0.531	4.224	0.514	4.298	0.594
220	E	4.156	0.590	4.360	0.652	4.306	0.664

229	E	3.426	1.184	3.479	1.235	3.592	1.067
230	E	2.372	1.376	2.137	1.290	2.372	1.169
237	E	2.938	1.105	2.888	1.068	3.074	0.919
244	E	2.085	0.641	2.224	0.756	2.083	0.593
246	E	3.434	1.419	3.632	1.605	3.587	1.311
257	E	2.070	1.159	2.064	0.996	2.264	1.213
266	E	3.473	1.157	3.888	1.020	3.752	1.021
270	E	3.651	1.104	3.744	1.224	3.826	1.061
272	E	3.605	1.007	4.040	0.700	3.727	0.933
273	E	3.380	1.066	3.576	1.037	3.455	1.050
288	E	3.767	0.711	4.000	0.645	3.876	0.626
289	E	3.543	0.938	3.808	0.624	3.752	0.805
298	E	3.016	1.433	2.879	1.294	3.350	1.255
300	E	3.643	0.872	3.992	0.549	3.725	0.588
2	N	3.217	0.984	2.472	0.913	2.810	0.955
30	N	2.202	0.693	1.944	0.505	2.174	0.711
32	N	3.023	1.273	2.520	0.897	2.769	1.013
35	N	2.837	1.215	2.232	0.793	2.645	1.148
36	N	2.357	1.372	1.808	0.850	2.223	1.175
40	N	2.256	1.348	1.864	0.973	2.289	1.224
43	N	2.736	1.258	2.384	0.980	2.736	1.079
44	N	3.349	1.385	2.656	1.308	3.041	1.173
48	N	3.016	1.125	2.432	0.876	2.843	0.983
49	N	2.465	1.032	2.040	0.603	2.190	0.872
53	N	2.233	1.352	1.696	0.713	2.050	1.181
60	N	3.047	1.435	2.508	1.049	2.851	1.278
61	N	2.674	1.018	2.104	0.787	2.405	1.043
62	N	2.450	1.015	2.056	0.957	2.364	1.167
64	N	2.984	1.703	2.560	1.168	2.736	1.446
72	N	2.581	1.261	1.952	0.869	2.355	1.164
79	N	2.016	1.031	1.656	0.760	1.893	0.980
82	N	3.101	1.185	2.424	1.004	2.826	1.111
91	N	2.953	1.201	2.328	0.932	2.826	1.145
99	N	2.217	0.906	1.896	0.787	2.008	0.832
100	N	2.612	1.442	1.984	1.000	2.380	1.154
107	N	2.884	1.432	2.464	1.460	2.727	1.200
117	N	2.333	0.943	1.888	0.487	2.074	0.653

119	N	1.969	1.171	1.584	0.761	1.867	0.940
122	N	2.845	1.445	2.400	1.161	2.702	1.377
123	N	3.023	1.429	2.784	1.316	2.702	1.327
133	N	3.217	0.906	2.728	0.958	3.149	0.961
134	N	3.070	1.081	2.888	1.262	3.017	0.966
135	N	2.806	1.048	2.448	0.959	2.785	1.020
137	N	3.341	1.367	2.688	1.184	3.041	1.190
138	N	3.535	1.266	3.128	1.209	3.140	1.288
153	N	3.620	1.066	3.208	1.182	3.628	1.002
156	N	2.992	1.367	2.280	1.106	2.826	1.361
161	N	2.380	1.019	1.976	0.927	2.198	0.777
166	N	3.116	1.400	2.589	1.138	3.066	1.229
169	N	3.271	1.293	2.584	1.132	2.876	1.276
173	N	2.667	1.068	2.400	1.097	2.620	0.938
182	N	2.132	0.725	1.784	0.590	2.174	0.561
184	N	2.318	1.219	1.720	0.816	2.182	1.050
188	N	3.333	1.240	2.568	1.296	3.091	1.350
192	N	2.992	1.289	2.272	1.167	2.645	1.348
203	N	2.852	1.340	2.648	1.149	2.702	1.094
206	N	3.124	1.266	2.416	1.035	2.909	1.183
216	N	2.450	1.499	1.968	1.015	2.107	1.147
217	N	2.791	1.964	2.760	1.668	2.901	1.773
218	N	2.758	1.209	1.992	0.927	2.542	1.427
222	N	2.806	1.642	2.080	1.106	2.628	1.352
226	N	3.845	1.226	3.896	0.852	3.802	1.060
233	N	3.729	1.059	3.653	0.911	3.711	0.724
241	N	2.674	1.393	2.208	1.118	2.579	1.263
243	N	2.814	1.528	2.424	1.149	2.752	1.155
247	N	3.023	1.554	2.912	1.387	3.099	1.257
254	N	2.829	0.971	2.544	1.073	2.612	0.823
255	N	2.481	0.798	2.224	0.933	2.372	0.752
263	N	2.620	1.269	2.144	0.657	2.496	0.852
265	N	2.116	0.822	1.718	0.643	2.165	0.989
278	N	2.256	1.161	1.832	0.899	2.066	0.779
281	N	3.256	1.254	2.576	1.295	2.934	1.346
287	N	3.736	1.149	3.272	1.200	3.653	1.045
292	N	2.054	1.005	1.896	0.804	1.926	0.653

3	O	3.806	0.814	4.056	0.521	3.884	0.603
6	O	2.651	2.057	2.728	2.071	2.488	1.902
16	O	1.969	0.608	1.832	0.657	1.909	0.633
17	O	3.868	1.068	3.344	1.131	3.496	1.285
19	O	3.930	1.628	3.904	1.313	3.818	1.300
28	O	4.457	0.719	4.576	0.504	4.554	0.566
29	O	3.558	1.405	3.096	1.168	3.620	0.904
51	O	4.357	0.606	4.440	0.652	4.529	0.418
54	O	3.977	0.773	4.032	0.644	4.033	0.749
58	O	3.713	1.175	3.782	0.855	3.744	0.925
65	O	3.240	1.403	3.392	1.160	3.372	1.336
66	O	3.326	1.737	3.632	1.202	3.364	1.333
69	O	3.147	1.423	3.600	1.081	3.223	1.375
75	O	4.535	0.766	4.408	0.873	4.458	0.755
77	O	3.488	1.127	3.816	0.797	3.545	1.250
78	O	2.775	1.098	2.896	0.965	2.645	0.898
85	O	3.922	0.947	4.024	0.620	3.975	0.791
86	O	3.473	1.376	3.816	1.184	3.612	1.556
87	O	2.380	1.112	2.224	0.869	2.281	0.837
89	O	3.240	0.903	2.992	0.976	3.215	0.853
90	O	3.868	0.584	4.016	0.629	3.950	0.631
94	O	3.845	0.976	4.088	0.758	4.058	0.805
103	O	4.062	0.871	4.336	0.709	4.165	0.656
104	O	3.891	0.801	4.008	0.524	3.967	0.766
113	O	3.682	0.734	3.904	0.813	3.769	0.729
125	O	3.496	1.127	3.112	0.842	3.405	1.110
140	O	3.388	0.958	3.432	0.941	3.463	1.051
141	O	3.295	1.303	3.544	1.040	3.446	1.166
142	O	3.550	0.953	3.792	0.876	3.562	0.998
144	O	3.240	1.731	3.616	1.771	3.347	1.612
145	O	3.209	1.292	3.568	1.215	3.198	1.327
160	O	3.946	0.927	4.226	0.697	4.158	0.689
170	O	3.395	1.850	3.613	1.573	3.463	1.451
174	O	3.256	1.192	3.608	0.918	3.372	1.119
175	O	2.891	0.879	2.792	0.811	2.916	0.891
180	O	3.388	1.270	3.089	1.203	3.364	1.217
194	O	3.209	2.011	3.776	1.288	3.322	1.420
195	O	2.814	1.137	3.112	1.116	2.736	1.146

202	O	3.078	0.994	3.137	1.176	2.950	1.114
208	O	1.876	1.297	1.608	1.031	1.686	0.834
211	O	4.101	0.966	3.736	1.212	4.041	0.923
215	O	3.837	1.481	3.584	1.229	3.810	1.422
221	O	2.829	1.330	2.744	1.176	2.612	1.223
223	O	3.961	0.959	4.152	0.791	3.950	0.848
224	O	2.829	1.752	3.216	1.542	3.041	1.323
228	O	3.705	1.881	4.048	1.417	3.851	1.444
235	O	3.349	1.354	3.456	1.315	3.430	1.214
239	O	3.930	1.050	4.168	0.867	4.116	0.870
252	O	3.364	1.187	3.776	1.175	3.579	0.996
258	O	3.217	1.156	3.280	1.461	3.174	1.245
260	O	3.054	1.349	3.000	1.419	3.050	1.214
264	O	3.388	0.989	3.680	1.106	3.430	0.830
268	O	3.899	1.091	3.944	0.940	3.843	1.000
275	O	3.039	0.959	3.264	0.841	3.157	0.717
277	O	3.752	1.000	3.752	1.188	3.826	1.078
282	O	3.341	1.195	3.240	0.974	3.413	1.078
283	O	4.109	0.629	4.256	0.515	4.041	0.707
285	O	4.667	0.396	4.656	0.631	4.545	0.633
291	O	3.473	1.298	3.248	1.398	3.471	1.301
293	O	3.566	1.873	3.856	1.640	3.521	1.635
4	A	3.783	0.953	4.296	0.807	4.132	0.732
5	A	3.736	0.774	3.824	0.807	3.818	0.900
8	A	4.178	0.601	4.536	0.525	4.405	0.443
9	A	3.256	1.036	3.584	0.842	3.339	1.243
13	A	3.512	0.721	3.904	0.491	3.752	0.638
14	A	4.155	0.929	4.488	0.784	4.198	0.877
15	A	3.318	0.875	3.680	0.639	3.496	0.802
20	A	3.907	0.944	4.032	0.926	3.975	0.708
23	A	4.380	0.722	4.584	0.503	4.512	0.719
25	A	2.527	0.954	2.192	0.673	2.421	0.996
26	A	3.651	0.854	3.816	0.780	3.653	0.862
27	A	3.426	1.059	3.728	0.764	3.521	1.002
31	A	3.391	0.870	3.632	0.605	3.372	0.952
37	A	3.760	1.200	4.216	0.751	4.000	0.800
42	A	3.496	1.158	3.752	0.720	3.504	1.002

68	A	3.860	1.262	3.992	1.121	3.901	1.107
70	A	4.240	0.856	4.456	0.508	4.190	1.089
71	A	4.109	0.910	4.240	0.861	4.174	0.811
73	A	3.147	1.392	3.312	1.152	3.107	1.263
83	A	4.016	0.765	4.240	0.684	4.066	0.929
84	A	3.791	1.276	3.824	1.098	3.793	1.015
93	A	3.488	0.986	3.680	0.606	3.537	0.684
109	A	3.240	1.731	3.136	1.280	3.289	1.507
114	A	2.907	0.913	3.336	1.096	3.240	0.900
118	A	3.899	0.748	4.144	0.544	3.983	0.850
126	A	4.000	0.656	4.192	0.640	4.033	0.716
127	A	4.163	1.012	4.200	0.887	4.198	0.844
129	A	3.457	1.453	3.944	1.086	3.694	1.214
130	A	4.062	1.512	4.306	1.287	4.132	1.549
131	A	3.178	1.413	3.056	1.037	3.124	1.193
132	A	3.302	1.103	3.576	0.875	3.298	1.044
143	A	3.225	1.082	3.492	0.821	3.347	1.095
147	A	3.380	1.941	3.696	1.375	3.463	1.334
149	A	3.000	1.156	3.288	1.336	3.380	1.021
154	A	3.729	1.309	4.232	0.712	3.843	1.167
165	A	4.240	0.840	4.384	0.787	4.240	1.084
168	A	2.953	0.935	3.056	1.198	2.893	1.030
179	A	3.352	1.364	3.169	1.410	3.458	0.973
189	A	4.070	1.019	4.392	0.724	4.289	0.774
191	A	2.938	1.293	3.440	1.135	3.017	1.316
199	A	4.070	0.847	4.272	0.764	4.174	0.695
225	A	2.853	1.533	2.784	0.993	2.917	1.410
227	A	3.791	1.229	4.160	0.781	3.992	1.008
232	A	3.829	1.002	4.080	1.010	3.901	1.123
236	A	2.659	0.898	2.524	0.853	2.545	0.867
238	A	3.124	1.734	3.648	1.375	3.372	1.302
240	A	3.783	1.046	3.648	1.117	3.479	1.418
242	A	3.760	0.996	3.944	1.053	3.909	1.083
249	A	3.729	0.996	3.742	0.908	3.570	0.980
250	A	3.008	1.602	3.176	1.372	2.909	1.333
251	A	3.062	1.184	3.488	1.333	3.124	1.143
259	A	3.581	0.917	3.800	0.919	3.595	1.043
261	A	3.085	1.563	3.360	1.361	2.959	1.257

	262	A	2.690	1.137	2.584	1.067	2.727	1.250
	269	A	4.023	1.289	4.360	0.716	4.099	1.023
	276	A	2.868	1.240	3.032	0.999	2.901	1.023
	279	A	3.597	1.117	3.744	1.208	3.537	1.067
	280	A	3.938	0.887	4.248	0.865	4.132	0.816
	284	A	3.798	1.022	3.864	0.925	3.678	0.904
	297	A	3.961	0.866	3.903	1.031	3.826	0.978
BIDR	1		4.344	2.615	4.924	2.138	4.188	1.895
	2		3.424	2.117	3.803	2.746	3.444	2.025
	3		4.629	3.162	5.319	1.784	4.726	2.390
	4		1.824	2.275	2.526	3.191	1.931	1.995
	5		4.864	3.457	5.560	1.744	4.821	2.269
	6		3.912	3.258	4.569	3.030	4.353	2.631
	7		3.016	2.629	3.759	2.515	3.259	2.489
	8		3.176	2.017	3.922	2.455	3.509	2.426
	9		4.240	4.103	4.974	4.269	5.103	3.763
	10		4.331	3.605	5.122	3.090	4.776	3.828
	11		5.260	4.637	5.647	3.239	4.974	4.478
	12		2.504	3.203	3.112	4.257	2.578	3.481
	13		1.951	2.047	2.336	2.469	2.207	2.026
	14		3.154	4.246	3.733	4.615	3.431	4.265
	15		2.680	2.947	2.887	2.891	2.930	3.083
	16		4.869	3.569	5.439	2.886	4.730	3.269
	17		2.697	4.378	3.826	4.847	3.165	4.139
	18		4.835	4.389	5.372	3.075	4.600	4.382
	19		4.598	2.970	5.412	2.067	4.974	2.394
	20		3.248	2.988	4.246	2.541	3.896	2.743
MCSD	1		0.452	0.452	0.383	0.238	0.513	0.252
	2		0.597	0.597	0.650	0.229	0.696	0.213
	3		1.000	1.000	1.000	0.000	1.000	0.000
	4		0.385	0.385	0.342	0.227	0.295	0.210
	5		1.000	1.000	1.000	0.000	1.000	0.000
	6		1.000	1.000	1.000	0.000	1.000	0.000
	7		0.574	0.574	0.627	0.236	0.622	0.237
	8		0.549	0.549	0.547	0.250	0.432	0.248
	9		1.000	1.000	1.000	0.000	1.000	0.000

10	1.000	1.000	1.000	0.000	1.000	0.000
11	1.000	1.000	1.000	0.000	1.000	0.000
12	1.000	1.000	1.000	0.000	1.000	0.000
13	0.664	0.664	0.652	0.229	0.688	0.217
14	1.000	1.000	1.000	0.000	1.000	0.000
15	1.000	1.000	1.000	0.000	1.000	0.000
16	0.661	0.661	0.435	0.248	0.486	0.252
17	0.771	0.771	0.765	0.181	0.817	0.151
18	0.415	0.415	0.452	0.250	0.330	0.223
19	1.000	1.000	1.000	0.000	1.000	0.000
20	0.667	0.667	0.609	0.240	0.667	0.224
21	0.658	0.658	0.583	0.245	0.654	0.228
22	1.000	1.000	1.000	0.000	1.000	0.000
23	1.000	1.000	1.000	0.000	1.000	0.000
24	0.786	0.786	0.688	0.217	0.792	0.166
25	0.786	0.786	0.759	0.185	0.819	0.150
26	0.487	0.487	0.464	0.251	0.371	0.236
27	0.618	0.618	0.528	0.252	0.446	0.250
28	1.000	1.000	1.000	0.000	1.000	0.000
29	0.318	0.318	0.215	0.170	0.190	0.155
30	1.000	1.000	1.000	0.000	1.000	0.000
31	0.495	0.495	0.411	0.244	0.361	0.233
32	1.000	1.000	1.000	0.000	1.000	0.000
33	0.422	0.422	0.400	0.242	0.433	0.248

Note. H = Honest Response Condition; D = Desirable Response Condition; NI = No Response Instructions Condition; C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; MCSD = Marlowe-Crowne Social desirability Scale; BIDR = Balanced Index of Desirable Responding.

Table 20

Means and Variances in the Honest, Desirable, and No Response Instructions Conditions for Each Factor and Sub-factor using the Polytomously Scored Data

		H		D		NI	
Factor	Sub-factor	*M*	S^2	*M*	S^2	*M*	S^2
C		**214.82**	**661.49**	**234.99**	**724.98**	**219.47**	**724.62**
	1	38.67	21.52	41.69	24.60	39.42	15.40
	2	35.08	60.98	37.23	42.37	35.33	49.39
	3	40.04	24.10	42.74	29.00	40.30	31.61
	4	38.25	30.30	41.89	25.46	39.49	28.02
	5	31.72	45.91	37.38	43.93	33.55	49.45
	6	31.06	36.06	34.06	33.24	31.39	35.22
E		**214.21**	**560.90**	**222.71**	**694.76**	**219.10**	**524.99**
	1	38.15	34.39	40.86	39.29	38.96	40.16
	2	36.24	41.32	37.25	46.82	37.18	35.27
	3	35.00	34.80	36.58	32.49	35.78	26.91
	4	30.98	25.30	32.46	16.06	30.93	10.28
	5	34.33	38.83	34.15	39.16	35.66	32.26
	6	39.51	29.86	41.42	31.41	40.59	28.46
N		**168.34**	**891.68**	**140.98**	**905.46**	**158.83**	**716.09**
	1	30.34	40.09	25.22	39.01	28.49	44.19
	2	27.74	55.58	22.09	46.15	25.79	51.82
	3	23.78	58.13	18.81	45.38	21.98	48.33
	4	28.44	41.36	25.60	42.16	27.48	31.23
	5	31.60	40.24	27.49	28.90	30.22	33.47
	6	26.44	34.61	21.77	30.97	24.87	26.95
O		**208.60**	**601.27**	**213.88**	**485.01**	**209.84**	**441.88**
	1	37.02	32.04	36.70	30.42	37.69	35.77
	2	38.55	47.95	40.26	47.48	38.78	47.24
	3	37.01	33.84	36.17	27.87	36.72	32.94
	4	34.71	32.88	36.66	29.45	35.36	28.20
	5	33.98	46.14	37.10	37.46	34.85	38.28
	6	27.33	27.32	27.00	30.39	26.45	24.52

A		**213.38**	**718.07**	**225.21**	**535.65**	**217.11**	**741.58**
	1	34.62	44.41	37.60	32.31	35.45	42.07
	2	38.52	39.16	41.18	28.55	39.47	37.50
	3	39.65	33.34	42.06	28.09	40.69	31.85
	4	34.22	47.87	37.31	30.89	35.39	43.44
	5	31.60	28.82	30.90	23.59	31.50	31.72
	6	34.77	32.04	36.15	27.42	34.61	31.34
BIDR		**5.12**	**11.95**	**7.21**	**12.97**	**5.05**	**9.16**
MCSD		**25.52**	**12.81**	**24.43**	**9.82**	**24.68**	**9.07**

Note. H = Honest Response Condition; D = Desirable Response Condition; NI = No Response Instructions Condition; C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; MCSD = Marlowe-Crowne Social desirability Scale; BIDR = Balanced Index of Desirable Responding

Table 21

Descriptive Statistics and Correlations for the Personality Factor Scores in the Honest, Desirable and No Response Instructions Conditions

				Honest Response Condition			
Factor	1	2	3	4	5	6	7
1. C	--						
2. E	0.16	--					
3. N	-0.39*	-0.30*	--				
4. O	0.05	0.38*	-0.11	--			
5. A	0.44*	0.19*	-0.28*	0.34*	--		
6. BIDR	0.509*	-0.07	-0.37*	0.09	0.51*	--	
7. MCSD	-0.23*	-0.09	0.18	-0.002	0.09	0.01	--
M	214.82	214.21	168.34	208.60	213.38	5.12	25.52
SD	25.72	23.68	29.86	24.52	26.80	3.46	3.58

				Desirable Response Condition			
Factor	1	2	3	4	5	6	7
1. C	--						
2. E	0.40*	--					
3. N	-0.65*	-0.56*	--				
4. O	0.34*	0.35*	-0.26*	--			
5. A	0.63*	0.36*	-0.54*	0.42*	--		
6. BIDR	0.48*	0.07	-0.49*	0.11	0.48*	--	
7. MCSD	0.11	0.31*	-0.20*	0.11	0.18	-0.001	--
M	234.99	222.71	140.98	213.88	225.21	7.21	24.43
SD	26.93	26.36	30.09	22.02	23.14	3.60	3.13

				No Response Instructions Condition			
Factor	1	2	3	4	5	6	7
1. C	--						
2. E	0.23*	--					
3. N	-0.41*	-0.41*	--				

4. O	0.07	0.44*	-0.09	--			
5. A	0.50*	0.41*	-0.37*	0.35*	--		
6. BIDR	0.42*	0.16	-0.46*	0.05	0.48*	--	
7. MCSD	-0.05	-0.07	0.21*	0.05	-0.16	-0.04	--
M	219.47	219.10	158.83	209.84	217.11	5.05	24.68
SD	26.92	22.91	26.76	21.02	27.23	3.03	3.01

Note. $p < .05$. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness; MCSD = Marlowe-Crowne Social Desirability Scale; BIDR = Balanced Index of Desirable Responding.

Table 22

Eigenvalues from the Principle Components Analysis of the Observed Data from the Honest Response, Desirable Response, and No Response Instructions Condition

Factor	1st eigenvalue	% Variance		2nd eigenvalue	% Variance
			Desirable Response Condition		
C	12.071	20.119		5.093	8.488
E	11.278	18.797		4.108	6.847
N	12.913	21.522		4.081	6.802
O	10.128	16.879		4.330	7.216
A	12.465	20.775		4.060	6.766
			Honest Response Condition		
C	16.446	27.410		4.072	6.787
E	16.244	27.073		4.021	6.702
N	16.860	28.099		4.049	6.749
O	10.364	17.274		4.337	7.228
A	12.252	20.419		3.852	6.420
			No Response Instructions Condition		
C	14.026	23.377		4.625	7.709
E	13.248	13.248		3.868	3.868
N	12.798	21.330		4.211	7.018
O	9.196	15.327		4.969	8.282
A	14.315	23.859		3.519	5.864

Note. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 23

Eigenvalues from the Principle Components Analysis of the Observed Sub-Factor Scores from the Honest Response, and Desirable Response, and No Response Instructions Conditions, and the eigenvalues from the Parallel Analyses

	Component					
Condition	1	2	3	4	5	6
Honest	7.36	4.18	3.31	2.57	1.99	1.22
Parallel Analysis	1.86	1.74	1.64	1.57	1.47	1.41
Desirable	10.88	3.63	2.84	1.79	1.71	1.02
Parallel Analysis	1.87	1.74	1.65	1.56	1.49	1.42
No Instructions	8.29	4.14	3.27	2.54	1.78	1.20
Parallel Analysis	1.90	1.76	1.66	1.57	1.48	1.41

Table 24

Factor Loadings of the IPIP Factor Scores for the Honest Response, Desirable Response, and No Response Instruction Conditions.

	Component				
Sub-scale	1	2	3	4	5
			Honest Condition		
C_1		**0.514**	-0.312		0.389
C_2		**0.779**			-0.384
C_3	0.523	0.512			
C_4		**0.809**			
C_5		**0.740**			
C_6		**0.553**		-0.393	
E_1				**0.816**	
E_2				**0.841**	
E_3	-0.569	0.373		0.338	
E_4	-0.344	0.670			
E_5		-0.324		**0.595**	
E_6				**0.801**	
N_1			**0.928**		
N_2	-0.534		**0.693**		
N_3			**0.649**		
N_4	0.366		**0.544**	-0.310	
N_5			**0.697**	0.439	
N_6			**0.752**		
O_1					**0.563**
O_2					**0.721**
O_3			0.459		**0.626**
O_4			-0.314		**0.661**
O_5					**0.904**

O_6		-0.456			**0.616**
A_1	**0.501**			0.518	
A_2	**0.546**	0.355			
A_3	**0.578**			0.431	
A_4	**0.875**				
A_5	**0.728**				
A_6	**0.646**				0.347
			Desirable Condition		
C_1	0.373		**0.382**		
C_2	-0.346		**0.892**		
C_3		0.645	0.371		
C_4			**0.601**	0.358	
C_5			**0.790**		
C_6	-0.317	0.576	0.398		
E_1	**0.786**				
E_2	**0.921**				-0.302
E_3	**0.638**	-0.424	0.316		
E_4		-0.415	0.682		
E_5	**0.819**	-0.489			
E_6	**0.581**			0.352	
N_1	**-0.705**			0.452	
N_2	-0.429	-0.604			
N_3	**-0.729**				
N_4	**-0.914**				
N_5		-0.434			
N_6	**-0.551**				
O_1				0.860	0.494
O_2				0.771	0.363
O_3				0.819	
O_4	0.702				0.307
O_5				0.401	**0.849**

O_6	-0.330			0.385	**0.717**
A_1	0.564	0.451			
A_2		**0.796**			
A_3	0.394	**0.480**		0.350	
A_4		**0.864**			
A_5	-0.407	**0.827**			
A_6		**0.582**		0.553	

No Response
Instructions Condition

C_1		**0.574**			0.303
C_2		**0.797**			
C_3	0.510	**0.511**			
C_4		**0.731**			
C_5		**0.858**			
C_6		**0.673**	-0.419		
E_1	0.576		0.542		
E_2	0.592		**0.645**		
E_3			**0.777**		
E_4		0.517	0.464		
E_5		-0.340	**0.646**		
E_6	0.615		0.355		
N_1				**0.888**	
N_2	-0.488			**0.751**	
N_3		-0.344		**0.490**	
N_4			-0.699	0.300	
N_5		-0.601	0.333		
N_6				**0.842**	
O_1					**0.794**
O_2					**0.696**
O_3	0.509			0.657	
O_4		-0.349			**0.362**
O_5					**0.733**

O_6		-0.558			
A_1	**0.725**				-0.308
A_2	**0.553**	0.394			
A_3	**0.845**				
A_4	**0.943**				
A_5	0.479			-0.529	
A_6	**0.822**				

Note. Only loading > .30 are displayed. Loadings on the five personality factors are reported in boldface. Subscripts indicate the number of the sub-scale for each factor. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 25

Fit Statistics for the 2-PL Model in the Honest Response Instruction Condition

	Frequency Table of Chi-Square/DF Ratios for Conscientiousness								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	22	7	4	6	4	14	3	2.823	2.551
Doubles	6	21	21	9	3	0	0	2.2	1.011
Triples	0	13	7	0	0	0	0	1.837	0.497

	Frequency Table of Chi-Square/DF Ratios for Extroversion								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	20	10	5	6	2	6	11	3.197	3.119
Doubles	9	16	19	8	3	2	3	2.689	2.079
Triples	0	13	3	2	1	1	0	2.21	1.136

	Frequency Table of Chi-Square/DF Ratios for Neuroticism								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	24	6	9	3	5	8	5	2.691	2.496
Doubles	7	17	23	9	4	0	0	2.29	1.055
Triples	0	12	7	1	0	0	0	2.034	0.552

	Frequency Table of Chi-Square/DF Ratios for Openness								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	25	1	8	7	8	3	8	2.879	2.904
Doubles	9	15	15	18	2	0	1	2.36	1.312
Triples	0	11	8	0	1	0	0	2.012	0.705

	Frequency Table of Chi-Square/DF Ratios for Agreeableness								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	20	11	11	4	5	8	1	2.366	2.07
Doubles	7	25	17	8	1	2	0	2.107	1.064
Triples	2	11	6	0	1	0	0	1.851	0.855

Table 26

Fit Statistics for the 2-PL Model in the Desirable Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	15	6	10	9	7	12	1	2.973	2.02
Doubles	1	15	29	12	3	0	0	2.403	0.84
Triples	0	7	12	1	0	0	0	2.198	0.391

Frequency Table of Chi-Square/DF Ratios for Extroversion

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	9	10	9	11	6	4	11	3.67	2.715
Doubles	4	16	14	15	4	2	5	3.146	2.152
Triples	0	10	3	5	0	2	0	2.721	1.335

Frequency Table of Chi-Square/DF Ratios for Neuroticism

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	5	7	4	9	16	16	3	3.959	1.901
Doubles	0	7	28	15	7	3	0	2.99	0.974
Triples	0	2	12	6	0	0	0	2.646	0.493

Frequency Table of Chi-Square/DF Ratios for Openness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	18	10	7	3	2	10	10	3.288	2.999
Doubles	7	13	22	7	6	3	2	2.8	1.902
Triples	1	7	9	0	2	1	0	2.427	1.236

Frequency Table of Chi-Square/DF Ratios for Agreeableness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	17	8	9	2	7	13	4	3.072	2.575
Doubles	6	18	25	6	2	2	1	2.403	1.499
Triples	1	12	4	2	1	0	0	1.991	0.928

Table 27

Fit Statistics for the 2-PL Model in the No Response Instructions Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	24	3	6	5	4	17	1	2.869	2.642
Doubles	7	14	25	10	3	0	1	2.336	1.289
Triples	0	10	8	2	0	0	0	2.114	0.696

Frequency Table of Chi-Square/DF Ratios for Extroversion

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	20	8	6	5	6	8	7	3.081	2.904
Doubles	9	17	15	9	4	3	3	2.66	2.068
Triples	2	10	4	2	1	1	0	2.167	1.259

Frequency Table of Chi-Square/DF Ratios for Neuroticism

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	13	9	10	6	6	8	8	3.403	2.55
Doubles	3	16	22	15	3	1	0	2.567	1.014
Triples	1	7	10	2	0	0	0	2.128	0.586

Frequency Table of Chi-Square/DF Ratios for Openness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	24	6	3	4	9	6	8	3.065	3.069
Doubles	10	15	15	11	7	0	2	2.583	1.794
Triples	1	8	8	2	0	1	0	2.162	1.09

Frequency Table of Chi-Square/DF Ratios for Agreeableness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	20	10	6	5	8	9	2	2.54	2.21
Doubles	11	19	21	5	1	2	1	2.14	1.253
Triples	1	14	2	2	1	0	0	1.847	0.815

Table 28

Fit Statistics for Samejima's Graded Response Model in the Honest Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	47	13	0	0	0	0	0	0.777	0.319
Doubles	34	24	1	1	0	0	0	1.015	0.466
Triples	6	8	6	0	0	0	0	1.508	0.711

Frequency Table of Chi-Square/DF Ratios for Extroversion

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	34	20	6	0	0	0	0	0.997	0.656
Doubles	30	24	6	0	0	0	0	1.154	0.561
Triples	10	5	4	1	0	0	0	1.342	0.821

Frequency Table of Chi-Square/DF Ratios for Neuroticism

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	60	0	0	0	0	0	0	0.351	0.186
Doubles	42	17	1	0	0	0	0	0.868	0.428
Triples	8	8	0	1	1	0	1	1.826	2.715

Frequency Table of Chi-Square/DF Ratios for Openness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	34	20	5	1	0	0	0	1.015	0.747
Doubles	21	33	6	0	0	0	0	1.246	0.557
Triples	10	7	2	0	0	0	0	1.072	0.607

Frequency Table of Chi-Square/DF Ratios for Agreeableness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	23	29	8	0	0	0	0	1.237	0.677
Doubles	25	33	1	1	0	0	0	1.132	0.479
Triples	6	6	5	2	0	0	0	1.653	0.973

Table 29

Fit Statistics for Samejima's Graded Response Model in the Desirable Response Instruction Condition

Frequency Table of Chi-Square/DF Ratios for Conscientiousness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	3	2	3	0	5	17	30	6.516	2.511
Doubles	2	5	14	10	13	12	4	3.996	1.991
Triples	2	5	1	5	2	4	1	0.822	11.233

Frequency Table of Chi-Square/DF Ratios for Extroversion

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	7	3	6	7	7	17	13	4.942	2.856
Doubles	2	11	20	5	12	8	2	3.375	1.73
Triples	2	4	3	4	2	4	1	3.395	2.01

Frequency Table of Chi-Square/DF Ratios for Neuroticism

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	7	11	20	15	6	1	0	2.629	1.142
Doubles	5	35	13	6	1	0	0	1.885	0.816
Triples	5	7	4	2	1	0	0	1.853	1.023

Frequency Table of Chi-Square/DF Ratios for Openness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	30	19	7	3	1	0	0	1.174	0.979
Doubles	18	34	8	0	0	0	0	1.329	0.5
Triples	4	10	3	1	1	0	0	1.618	1.015

Frequency Table of Chi-Square/DF Ratios for Agreeableness

	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	12	8	9	10	7	12	2	3.328	2.209
Doubles	6	18	17	15	3	1	0	2.461	1.164
Triples	1	4	5	1	3	4	1	3.668	2.585

Table 30

Fit Statistics for Samejima's Graded Response Model in the No Response Instructions Condition

	Frequency Table of Chi-Square/DF Ratios for Conscientiousness								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	23	25	12	0	0	0	0	1.348	0.569
Doubles	14	41	4	1	0	0	0	1.375	0.564
Triples	7	10	2	1	0	0	0	1.409	0.735

	Frequency Table of Chi-Square/DF Ratios for Extroversion								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	10	10	17	8	11	4	0	2.689	1.617
Doubles	5	25	22	6	2	0	0	2.142	0.824
Triples	2	6	6	5	0	0	0	2.215	0.973

	Frequency Table of Chi-Square/DF Ratios for Neuroticism								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	44	16	0	0	0	0	0	0.712	0.481
Doubles	33	24	3	0	0	0	0	1.039	0.522
Triples	5	8	2	0	1	1	0	1.704	1.583

	Frequency Table of Chi-Square/DF Ratios for Openness								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	40	13	4	2	1	0	0	0.937	0.898
Doubles	33	22	5	0	0	0	0	1.095	0.566
Triples	8	6	2	1	0	1	0	-7.826	40.197

	Frequency Table of Chi-Square/DF Ratios for Agreeableness								
	<1	1<2	2<3	3<4	4<5	5<7	>7	Mean	SD
Singles	23	22	10	5	0	0	0	1.409	0.924
Doubles	19	32	9	0	0	0	0	1.402	0.524
Triples	6	5	4	1	1	0	1	2.215	2.496

Table 31

Results of the Differential Peron Functioning Analysis using the Mantel-Haenszel Procedure on the Dichotomously Scored Items

Participant	Condition	χ^2_{MH}	Ω	Delta
1112220	**D**	**3.227** *	**2.147**	**-1.796**
1112221	D	2.808 *	0.487	1.691
1122330	H	0.463	1.387	-0.769
1122331	H	**2.835** *	**2.024**	**-1.657**
1122332	H	0.925	1.568	-1.057
1122333	H	0.725	0.716	0.785
1122334	H	1.118	1.466	-0.899
3100000	NI	0.025	0.988	0.029
3100001	NI	1.692	1.624	-1.139
3100002	NI	0.953	1.415	-0.816
3100003	NI	0.208	1.211	-0.450
3100004	**NI**	**18.263** ***	**5.415**	**-3.970**
3100006	NI	0.010	0.968	0.076
3100007	NI	0.027	1.131	-0.289
3100008	NI	0.001	0.953	0.113
3100009	D	1.841	1.768	-1.339
3100010	D	0.943	1.920	-1.533
3100011	D	0.009	0.888	0.279
3100012	NI	0.001	1.047	-0.108
3100013	D	0.000	0.909	0.224
3100014	NI	1.211	1.630	-1.148
3100015	NI	0.658	1.503	-0.958
3100016	D	0.000	0.933	0.163
3100017	NI	1.840	0.617	1.135
3100018	NI	0.402	0.787	0.563
3100019	NI	2.487	1.907	-1.517
3100020	H	1.197	0.627	1.097
3100021	NI	0.018	1.146	-0.320
3100022	NI	1.273	1.574	-1.066
3100023	NI	0.572	0.731	0.736
3100024	NI	2.451	1.868	-1.468
3100025	D	0.025	0.836	0.421

3100026	NI	1.354	1.636	-1.157
3100027	D	0.001	0.911	0.219
3100028	**D**	**3.074 ***	**2.442**	**-2.098**
3100029	D	2.171	0.513	1.569
3100030	D	1.024	1.720	-1.274
3100031	D	0.021	0.856	0.365
3100032	**D**	**5.819 ****	**2.203**	**-1.856**
3100033	NI	0.292	0.788	0.560
3100034	D	0.174	1.280	-0.580
3100035	**H**	**3.143 ***	**2.289**	**-1.946**
3100036	H	0.866	0.694	0.858
3100037	H	0.001	1.048	-0.110
3100038	H	0.593	1.372	-0.743
3100039	D	0.330	0.743	0.698
3100040	D	0.769	1.643	-1.167
3100041	D	0.028	1.158	-0.345
3100042	H	1.102	1.558	-1.042
3100043	H	5.489 **	0.443	1.913
3100044	H	0.768	0.713	0.795
3100045	H	2.138	0.537	1.461
3100046	D	1.044	1.578	-1.072
3100047	D	0.253	0.766	0.626
3100048	H	0.002	0.957	0.103
3100049	H	0.548	1.371	-0.742
3100050	D	0.030	0.888	0.279
3100051	D	0.015	1.345	-0.697
3100052	D	0.002	0.934	0.160
3100053	**D**	**3.520 ***	**1.913**	**-1.524**
3100055	D	0.000	0.947	0.128
3100056	D	0.131	0.787	0.563
3100057	D	1.855	0.497	1.643
3100058	D	1.571	0.613	1.150
3100059	NI	2.463	0.569	1.325
3100060	D	0.805	1.537	-1.010
3100061	**NI**	**3.766 ***	**2.254**	**-1.910**
3100062	NI	0.003	1.075	-0.170
3100063	H	0.007	1.026	-0.060
3100064	H	0.026	0.894	0.263

3100065	NI	1.826	0.641	1.045
3100066	NI	0.409	0.760	0.645
3100067	NI	0.259	1.266	-0.554
3100068	D	0.000	0.943	0.138
3100069	D	0.255	0.723	0.762
3100070	**D**	**4.267 ****	**2.278**	**-1.935**
3100071	H	0.026	0.999	0.002
3100072	H	2.546	1.836	-1.428
3100073	H	0.056	0.868	0.333
3100074	H	3.459 *	0.482	1.715
3100075	H	0.003	1.075	-0.170
3100076	H	0.200	1.237	-0.500
3100077	D	0.000	0.889	0.276
3100078	NI	0.870	1.407	-0.802
3100079	D	0.110	0.842	0.404
3100080	D	0.497	1.350	-0.705
3100081	D	0.503	1.394	-0.781
3100082	D	0.043	1.156	-0.341
3100083	D	0.016	0.989	0.026
3100084	D	5.077 **	0.432	1.972
3100085	NI	1.608	1.602	-1.107
3100086	NI	0.257	1.270	-0.562
3100087	NI	1.613	1.716	-1.269
3100089	NI	0.074	0.858	0.360
3100090	NI	1.877	1.736	-1.296
3100091	NI	1.050	1.491	-0.939
3100092	NI	0.405	1.394	-0.781
3100093	NI	0.043	0.879	0.303
3100094	NI	0.001	0.959	0.098
3100095	NI	0.000	0.945	0.133
3100096	D	0.142	1.304	-0.624
3100097	D	0.942	0.542	1.439
3100098	D	0.104	1.234	-0.494
3100099	D	0.505	0.681	0.903
3100100	D	0.797	1.593	-1.094
3100101	D	2.203	1.924	-1.538
3100102	D	0.010	1.089	-0.200
3100103	D	3.609 *	0.410	2.095

3100104	NI	0.023	0.898	0.253
3100105	NI	0.174	0.808	0.501
3100106	NI	0.449	1.302	-0.620
3100107	NI	0.000	1.051	-0.117
3100108	NI	0.015	1.102	-0.228
3100109	NI	2.639	0.574	1.305
3100111	NI	0.097	0.846	0.393
3100113	H	0.395	1.340	-0.688
3100114	H	2.744 *	0.516	1.555
3100115	H	0.027	1.005	-0.012
3100116	H	8.782 **	0.302	2.814
3100117	H	0.683	0.707	0.815
3100118	H	0.016	1.128	-0.283
3100119	H	0.155	1.223	-0.473
3100120	H	0.519	1.364	-0.729
3334444	D	0.099	0.865	0.341
3355335	D	0.270	1.300	-0.617
4433300	NI	0.107	1.173	-0.375
4443333	D	0.203	1.384	-0.764
4443334	D	1.327	2.234	-1.889
4443335	D	1.012	1.580	-1.075
4443336	**D**	**5.181 ****	**2.923**	**-2.521**
4443337	D	0.035	0.856	0.365
4443338	D	0.055	0.867	0.335
4443339	D	2.522	0.460	1.825
4455550	**D**	**3.126 ***	**2.209**	**-1.862**
4455551	D	2.071	0.531	1.488
4455552	**NI**	**3.146 ***	**2.250**	**-1.906**
4455553	NI	0.791	1.448	-0.870
4455554	NI	0.028	1.109	-0.243
4455555	NI	0.997	1.463	-0.894
4455556	NI	0.002	1.039	-0.090
4455660	NI	0.781	1.432	-0.844
4455661	D	0.102	1.212	-0.452
4455662	H	10.736 ***	0.284	2.958
5200000	H	0.140	1.176	-0.381
5200001	H	0.012	0.902	0.242
5200002	D	0.414	0.650	1.012

5200003	D	0.013	1.018	-0.042
5200004	D	0.407	1.321	-0.654
5200005	**D**	**3.871 ****	**3.571**	**-2.991**
5200006	D	0.020	1.159	-0.347
5200007	D	0.000	1.055	-0.126
5200008	D	0.108	0.803	0.516
5200009	D	0.001	1.074	-0.168
5200010	D	0.142	0.824	0.455
5200011	D	0.005	1.092	-0.207
5200012	NI	0.474	1.360	-0.723
5200013	D	0.309	0.788	0.560
5200014	D	0.358	0.759	0.648
5200015	D	2.600	1.766	-1.336
5200016	D	1.310	0.650	1.012
5200017	D	0.238	1.274	-0.569
5200018	**D**	**4.768 ****	**3.538**	**-2.969**
5200018	H	1.716	0.617	1.135
5200019	D	2.177	0.586	1.256
5200020	D	0.747	0.686	0.886
5200023	D	0.648	1.367	-0.735
5200024	H	0.514	1.346	-0.698
5200025	H	0.035	1.133	-0.293
5200026	H	0.000	1.066	-0.150
5200027	H	0.008	0.976	0.057
5200028	H	3.832 *	0.502	1.620
5200029	H	0.397	1.366	-0.733
5200030	H	0.078	1.179	-0.387
5200031	H	1.676	1.735	-1.295
5200032	H	0.293	0.786	0.566
5200033	H	0.472	1.365	-0.731
5200035	D	0.921	0.561	1.358
5200036	D	0.253	1.335	-0.679
5200037	H	0.137	0.828	0.444
5200038	D	1.154	0.588	1.248
5200039	D	0.003	0.923	0.188
5200040	**H**	**5.345 ****	**2.133**	**-1.780**
5200041	H	2.178	1.696	-1.241
5200042	**D**	**3.797 ***	**1.975**	**-1.599**

5200043	D	0.001	0.912	0.216
5200044	D	0.145	1.184	-0.397
5200045	D	0.001	0.893	0.266
5200046	D	3.008 *	0.552	1.396
5200047	D	0.000	0.905	0.235
5200048	**D**	**10.724 *****	**3.501**	**-2.945**
5200049	D	0.010	1.047	-0.108
5200050	D	3.153 *	0.455	1.851
5200051	**NI**	**3.239 ***	**2.309**	**-1.967**
5200052	D	0.005	0.963	0.089
5200053	NI	0.159	0.802	0.519
5200055	D	0.070	0.880	0.300
5200056	D	0.011	1.030	-0.069
5200057	D	0.897	0.660	0.976
5200058	D	0.052	1.000	0.000
5200059	D	1.037	1.570	-1.060
5200060	D	0.087	1.159	-0.347
5200061	NI	0.046	1.157	-0.343
5200062	NI	0.000	1.062	-0.141
5200063	H	1.497	0.580	1.280
5200064	H	0.086	1.206	-0.440
5200065	H	0.522	0.717	0.782
5200066	H	0.748	1.391	-0.776
5200067	**H**	**5.396 ****	**2.898**	**-2.500**
5200068	H	2.436	1.847	-1.442
5200069	H	1.150	0.658	0.984
5200070	H	1.266	1.489	-0.936
5200071	H	0.019	1.097	-0.218
5200072	H	0.023	0.994	0.014
5200073	NI	0.008	0.902	0.242
5200074	H	0.176	1.208	-0.444
5200075	H	0.043	0.868	0.333
5200076	H	0.286	1.278	-0.576
5200077	H	0.795	1.531	-1.001
5200078	H	0.417	1.313	-0.640
5200079	H	0.131	1.202	-0.432
5200080	H	0.010	1.020	-0.047
5200082	H	0.018	1.104	-0.233

5200083	H	0.003	0.917	0.204
5200085	NI	0.200	1.209	-0.446
5200086	NI	0.639	1.380	-0.757
5200087	H	0.193	0.816	0.478
5200088	H	0.008	1.115	-0.256
5200089	H	0.785	0.685	0.889
5200090	H	0.002	1.041	-0.094
5200091	H	0.909	1.603	-1.109
5200092	H	2.540	1.851	-1.447
5200093	D	2.815 *	0.446	1.897
5200094	D	0.367	1.344	-0.695
5200095	D	0.478	1.525	-0.992
5200096	**D**	**2.703** *	**2.068**	**-1.707**
5200097	NI	0.369	0.782	0.578
5200098	NI	0.474	1.347	-0.700
5200099	NI	0.000	1.055	-0.126
5200100	NI	0.000	0.955	0.108
5200101	NI	0.103	1.193	-0.415
5200102	NI	1.324	0.634	1.071
5200103	NI	0.014	0.974	0.062
5200104	H	2.453	0.565	1.342
5200105	H	0.418	0.776	0.596
5200106	H	1.918	0.574	1.305
5200107	H	0.008	1.079	-0.179
5200108	H	0.992	1.446	-0.867
5200109	H	0.120	1.170	-0.369
5200110	H	0.011	1.020	-0.047
5200111	H	0.063	0.849	0.385
5200112	H	4.922 **	0.461	1.820
5200113	H	0.018	0.983	0.040
5200114	H	0.005	0.967	0.079
5200115	H	0.000	0.952	0.116
5200116	H	0.911	0.683	0.896
5200128	D	0.856	0.646	1.027
5200130	D	0.004	0.946	0.130
5200131	D	0.039	1.189	-0.407
5200132	D	0.644	1.358	-0.719
5200135	NI	0.453	0.750	0.676

5200137	NI	0.537	1.356	-0.716
5200138	NI	0.145	0.815	0.481
5200140	**NI**	**5.251 ****	**2.298**	**-1.955**
5200141	NI	1.416	1.692	-1.236
5200142	NI	1.620	0.643	1.038
5200143	NI	1.225	1.560	-1.045
5200144	**NI**	**4.048 ****	**2.050**	**-1.687**
5200145	NI	0.018	1.096	-0.215
5200146	H	0.024	0.897	0.255
5200147	NI	0.231	0.802	0.519
5200148	**NI**	**5.572 ****	**2.536**	**-2.187**
5200149	NI	0.404	1.342	-0.691
5200150	**NI**	**6.873 *****	**2.563**	**-2.212**
5200151	NI	0.596	1.382	-0.760
5200152	NI	0.043	1.129	-0.285
5200153	NI	1.003	1.472	-0.909
5200154	NI	0.336	0.789	0.557
5200156	NI	1.729	0.538	1.457
5200158	NI	0.004	0.963	0.089
5200159	**NI**	**3.630 ***	**2.093**	**-1.736**
5200160	NI	2.611	1.778	-1.352
5200161	NI	0.136	0.789	0.557
5200162	NI	1.059	1.491	-0.939
5200163	NI	0.299	0.787	0.563
5200164	NI	1.988	0.616	1.139
5200165	NI	0.000	0.936	0.155
5200166	**NI**	**3.636 ***	**2.011**	**-1.642**
5200167	NI	0.475	1.305	-0.626
5200168	NI	0.010	1.019	-0.044
5200169	NI	0.174	1.241	-0.507
5200170	D	0.165	1.251	-0.526
5200171	D	0.121	0.857	0.363
5200172	D	0.182	1.215	-0.458
5200173	H	2.086	1.861	-1.460
5200174	**H**	**8.245 ****	**2.696**	**-2.331**
5200175	H	0.615	0.709	0.808
5200176	H	0.002	0.910	0.222
5200177	H	0.034	0.883	0.292

5200178	H	0.040	1.160	-0.349
5200179	H	0.288	0.802	0.519
5200180	**NI**	**3.676** *	**2.001**	**-1.630**
5200181	H	0.495	1.357	-0.717
5200182	**NI**	**6.163** **	**2.422**	**-2.079**
5200183	**NI**	**7.084** ***	**3.077**	**-2.641**
5200184	NI	1.072	1.451	-0.875
5200185	**D**	**3.185** *	**2.314**	**-1.972**
5200186	D	0.005	1.049	-0.112
5200187	NI	0.022	1.119	-0.264
5200188	NI	0.857	0.680	0.906
5200189	NI	0.682	1.378	-0.753
5200190	NI	0.080	1.151	-0.330
5200191	NI	0.165	0.838	0.415
5200192	NI	0.316	1.355	-0.714
5200193	NI	0.248	1.242	-0.509
5200194	NI	0.441	1.383	-0.762
5200195	NI	0.155	1.244	-0.513
5200196	H	0.285	1.276	-0.573
5200197	**H**	**4.301** **	**2.103**	**-1.747**
5200198	H	2.476	0.449	1.882
5200199	H	0.025	1.003	-0.007
5200200	D	0.700	1.443	-0.862
5200201	D	1.005	1.598	-1.102
5200202	D	0.811	1.565	-1.053
5200203	H	0.401	0.775	0.599
5200204	H	4.893 **	0.420	2.039
5200205	H	0.124	0.836	0.421
5200207	H	0.090	0.842	0.404
5200208	H	2.037	0.609	1.165
5200209	H	0.054	0.878	0.306
5200211	H	0.143	1.260	-0.543
5200213	H	0.857	0.659	0.980
5200214	H	0.005	1.040	-0.092
5200215	H	0.024	0.996	0.009
5200216	H	0.373	1.276	-0.573
5200217	H	1.608	1.615	-1.126
5200218	H	6.448 **	0.364	2.375

5200219	H	1.799	1.726	-1.283
5200220	**H**	**3.159 ***	**2.002**	**-1.631**
5200225	**D**	**4.935 ****	**3.807**	**-3.142**
5200226	D	0.005	0.879	0.303
5200227	NI	1.338	1.532	-1.002
5290129	NI	2.322	1.830	-1.420
5790210	NI	0.354	1.269	-0.560
7104088	D	1.133	1.654	-1.183
8877801	**D**	**5.282 ****	**2.377**	**-2.035**
8877802	**D**	**3.777 ***	**2.387**	**-2.045**
8877804	NI	0.085	1.178	-0.385
8877805	H	1.164	0.616	1.139
8877806	H	0.026	1.113	-0.252
8877807	H	0.289	0.799	0.527
8877808	H	0.006	0.964	0.086
8877809	H	0.161	1.188	-0.405
8877810	**NI**	**7.891 *****	**2.949**	**-2.541**
8877811	NI	0.004	0.972	0.067
8877812	H	7.663 ***	0.349	2.474
8877813	NI	1.399	0.657	0.987
8877814	NI	1.480	1.671	-1.207
8877815	NI	0.261	1.234	-0.494
8877817	NI	0.002	1.106	-0.237
8877818	**NI**	**3.644 ***	**2.099**	**-1.742**
8877819	D	0.015	1.124	-0.275
8877820	D	0.002	1.035	-0.081
8877821	D	0.090	1.187	-0.403
8877822	D	0.654	1.360	-0.723
8877823	D	0.005	0.920	0.196
8877824	NI	0.812	1.558	-1.042
8877825	H	0.015	1.015	-0.035
8877826	D	0.310	1.412	-0.811
8877827	H	0.029	1.170	-0.369
8877828	H	0.251	1.239	-0.504
8877829	H	0.475	0.750	0.676
8877830	H	0.008	0.915	0.209
8877831	H	15.235 ***	0.205	3.724
8877832	H	0.203	1.223	-0.473

8877833	H	0.002	1.075	-0.170

Note. *p < .10, **p < .05, ***p < .01. χ^2_{MH} = Mantel-Haenszel Chi-Square; Ω = Odds Ratio; Delta = MH D-DIF; D = Desirable Response Condition; H = Honest Response Condition; NI = No Response Instructions Condition. Participants in boldface are identified as responding differentially.

Table 32

Results of the Differential Person Functioning Analysis using the Mantel Procedure on the Polytomously Scored Items

Participant	Condition	χ^2_M	SMD
1112220	D	0.409	0.038
1112221	D	2.310	-0.048
1122330	H	0.544	0.029
1122331	H	2.545	0.041
1122332	H	1.167	0.021
1122333	H	0.391	-0.017
1122334	**H**	**3.174** *	**0.057**
3100000	NI	3.201 *	-0.039
3100001	NI	0.001	0.005
3100002	NI	2.328	0.060
3100003	NI	0.201	0.021
3100004	**NI**	**13.971** ***	**0.122**
3100006	NI	1.799	-0.035
3100007	NI	1.494	0.036
3100008	NI	0.372	-0.023
3100009	**D**	**4.437** **	**0.076**
3100010	D	0.002	-0.006
3100011	D	0.289	0.039
3100012	**NI**	**6.003** **	**0.072**
3100013	D	0.361	-0.005
3100014	NI	0.708	-0.026
3100015	NI	0.271	-0.020
3100016	D	0.541	-0.016
3100017	NI	0.438	0.027
3100018	NI	2.200	-0.053
3100019	**NI**	**3.011** *	**0.046**
3100020	H	1.455	-0.030
3100021	NI	0.335	0.025
3100022	NI	0.625	0.030
3100023	NI	0.649	-0.029
3100024	**NI**	**7.048** **	**0.085**
3100025	D	0.135	0.002

3100026	**NI**	**3.417** *	**0.064**
3100027	D	0.190	0.030
3100028	D	0.404	0.019
3100029	D	1.014	-0.031
3100030	D	0.473	0.022
3100031	D	0.146	-0.021
3100032	**D**	**4.165** **	**0.073**
3100033	NI	1.262	0.034
3100034	D	1.389	0.062
3100035	H	1.974	0.048
3100036	H	5.065 **	-0.050
3100037	H	0.063	-0.008
3100038	H	0.903	0.041
3100039	D	1.479	-0.029
3100040	D	0.945	-0.032
3100041	D	0.023	-0.004
3100042	H	1.755	0.034
3100043	H	13.864 ***	-0.108
3100044	H	0.557	-0.026
3100045	H	2.635	-0.059
3100046	**D**	**7.880** **	**0.088**
3100047	D	0.023	-0.013
3100048	H	0.173	-0.022
3100049	H	3.009	0.049
3100050	D	0.020	0.012
3100051	D	0.282	0.014
3100052	D	0.008	0.003
3100053	D	0.684	0.021
3100055	D	1.320	0.054
3100056	D	0.457	-0.021
3100057	D	2.902 *	-0.070
3100058	D	1.303	-0.039
3100059	NI	1.399	-0.028
3100060	**D**	**5.420** **	**0.049**
3100061	NI	2.391	0.042
3100062	NI	0.039	-0.008
3100063	H	0.821	0.009
3100064	H	0.604	-0.025

3100065	NI	0.870	-0.009
3100066	NI	0.102	0.003
3100067	NI	1.478	0.047
3100068	D	2.645	-0.012
3100069	D	0.007	-0.014
3100070	D	1.695	0.049
3100071	H	0.141	0.012
3100072	**H**	**6.222** **	**0.096**
3100073	H	0.663	0.006
3100074	H	1.997	-0.033
3100075	H	0.003	-0.009
3100076	H	1.757	0.044
3100077	D	0.546	0.025
3100078	NI	0.002	-0.007
3100079	D	1.018	-0.016
3100080	D	0.176	0.010
3100081	D	0.438	0.030
3100082	D	1.664	0.033
3100083	D	0.399	0.023
3100084	D	7.768 **	-0.082
3100085	NI	0.009	0.008
3100086	NI	2.076	0.044
3100087	NI	0.431	-0.018
3100089	NI	4.445 **	-0.048
3100090	NI	0.002	0.016
3100091	NI	0.285	-0.013
3100092	NI	0.696	0.020
3100093	NI	4.969 **	-0.059
3100094	NI	0.284	-0.021
3100095	NI	0.021	-0.001
3100096	D	0.245	0.010
3100097	D	2.866 *	-0.042
3100098	D	0.605	0.019
3100099	D	0.797	-0.028
3100100	D	0.001	-0.003
3100101	D	0.809	0.027
3100102	**D**	**4.906** **	**0.074**
3100103	D	1.457	-0.041

3100104	NI	0.058	-0.010
3100105	NI	1.405	-0.020
3100106	NI	1.681	0.040
3100107	NI	0.437	-0.018
3100108	NI	0.849	-0.026
3100109	NI	12.841 ***	-0.099
3100111	NI	0.257	0.009
3100113	H	0.510	0.024
3100114	H	1.871	-0.036
3100115	H	0.585	0.022
3100116	H	1.889	-0.025
3100117	H	0.376	-0.035
3100118	H	0.201	0.009
3100119	H	0.056	0.012
3100120	H	1.806	0.049
3334444	D	0.028	-0.022
3355335	**D**	**11.270** ***	**0.129**
4433300	NI	0.743	0.017
4443333	**D**	**4.915** **	**0.042**
4443334	D	1.760	-0.038
4443335	D	0.020	-0.004
4443336	D	1.970	0.040
4443337	D	0.071	0.003
4443338	D	0.002	0.013
4443339	D	5.302 **	-0.064
4455550	**D**	**4.941** **	**0.040**
4455551	D	2.654	-0.037
4455552	NI	0.193	-0.004
4455553	**NI**	**3.320** *	**0.049**
4455554	NI	0.587	0.025
4455555	**NI**	**3.362** *	**0.054**
4455556	NI	0.231	0.007
4455660	NI	2.326	0.038
4455661	**D**	**3.210** *	**0.059**
4455662	H	5.691 **	-0.055
5200000	H	0.495	0.001
5200001	H	0.010	-0.006
5200002	D	0.043	-0.007

5200003	D	1.086	0.027
5200004	D	2.359	0.047
5200005	D	1.648	0.045
5200006	D	0.907	0.033
5200007	D	0.176	0.009
5200008	D	0.020	-0.006
5200009	D	0.433	-0.014
5200010	D	1.647	-0.048
5200011	D	1.636	-0.047
5200012	NI	1.976	-0.058
5200013	D	0.567	-0.021
5200014	D	1.215	-0.032
5200015	**D**	**3.439** *	**0.072**
5200016	D	0.281	0.009
5200017	D	0.046	0.011
5200018	**D**	**12.612** ***	**0.083**
5200018	H	0.372	-0.016
5200019	D	3.208 *	-0.036
5200020	D	2.321	-0.046
5200023	D	0.184	0.019
5200024	H	0.909	0.032
5200025	H	0.022	0.001
5200026	H	0.092	0.013
5200027	H	0.937	0.025
5200028	H	2.392	-0.034
5200029	H	0.845	0.030
5200030	H	0.038	0.008
5200031	H	1.125	0.051
5200032	H	2.164	-0.028
5200033	H	0.485	0.033
5200035	D	0.927	-0.025
5200036	D	1.574	0.026
5200037	H	1.011	-0.043
5200038	D	0.431	-0.013
5200039	D	0.058	-0.013
5200040	**H**	**6.569** **	**0.091**
5200041	**H**	**3.491** *	**0.072**
5200042	**D**	**3.530** *	**0.056**

5200043	D	0.259	-0.009
5200044	D	**7.003** **	**0.068**
5200045	D	0.643	0.015
5200046	D	0.681	-0.037
5200047	D	1.557	-0.022
5200048	**D**	**5.375** **	**0.051**
5200049	D	0.777	0.036
5200050	D	0.058	-0.013
5200051	NI	0.657	-0.013
5200052	D	0.055	-0.002
5200053	NI	0.070	0.012
5200055	D	0.704	0.041
5200056	D	1.000	0.032
5200057	D	0.016	0.011
5200058	D	2.374	0.046
5200059	D	1.916	0.032
5200060	D	0.690	0.014
5200061	NI	0.563	-0.036
5200062	NI	0.612	-0.009
5200063	H	1.299	-0.016
5200064	H	0.228	0.009
5200065	H	1.158	-0.023
5200066	H	1.367	0.033
5200067	**H**	**4.356** **	**0.075**
5200068	H	0.389	0.022
5200069	H	0.799	-0.030
5200070	H	2.538	0.073
5200071	H	0.090	0.000
5200072	H	0.310	0.011
5200073	NI	0.616	-0.015
5200074	H	2.197	0.040
5200075	H	1.924	-0.069
5200076	H	0.527	0.022
5200077	**H**	**5.831** **	**0.064**
5200078	H	0.040	-0.016
5200079	H	1.935	0.029
5200080	H	0.015	-0.003
5200082	H	0.043	0.014

5200083	H	0.179	-0.022
5200085	NI	2.720 *	-0.049
5200086	NI	0.242	0.014
5200087	H	0.011	0.003
5200088	H	0.888	0.017
5200089	H	1.370	-0.041
5200090	H	0.032	-0.005
5200091	H	1.158	0.037
5200092	**H**	**3.093** *	**0.031**
5200093	D	1.486	-0.041
5200094	D	0.740	0.033
5200095	D	0.322	0.019
5200096	**D**	**7.572** **	**0.069**
5200097	NI	4.497 **	-0.082
5200098	NI	0.282	-0.017
5200099	NI	1.918	-0.033
5200100	NI	0.228	0.012
5200101	NI	1.151	0.036
5200102	NI	5.781 **	-0.066
5200103	NI	1.278	-0.041
5200104	H	1.768	-0.022
5200105	H	0.006	-0.020
5200106	H	0.452	-0.012
5200107	H	0.221	0.031
5200108	H	0.570	0.026
5200109	H	0.055	-0.020
5200110	H	0.858	0.025
5200111	H	0.080	-0.003
5200112	H	5.736 **	-0.057
5200113	H	0.126	-0.025
5200114	H	0.062	0.008
5200115	H	0.025	-0.007
5200116	H	0.650	-0.024
5200128	D	0.002	-0.009
5200130	D	0.003	-0.009
5200131	D	0.561	-0.022
5200132	**D**	**4.063** **	**0.071**
5200135	NI	0.089	-0.004

5200137	NI	2.047	0.039
5200138	NI	2.648	-0.027
5200140	NI	0.494	0.020
5200141	NI	0.016	0.002
5200142	NI	2.255	-0.031
5200143	NI	0.099	0.027
5200144	**NI**	**3.973** **	**0.051**
5200145	NI	1.905	0.053
5200146	H	0.081	0.013
5200147	NI	0.712	-0.014
5200148	NI	0.011	-0.002
5200149	NI	0.132	0.003
5200150	**NI**	**2.847** *	**0.048**
5200151	NI	0.604	-0.025
5200152	NI	0.482	0.028
5200153	NI	1.777	0.047
5200154	NI	1.496	-0.046
5200156	NI	8.196 **	-0.069
5200158	NI	0.749	-0.012
5200159	**NI**	**6.067** **	**0.070**
5200160	NI	0.620	0.020
5200161	NI	0.097	-0.002
5200162	NI	0.226	0.016
5200163	NI	0.139	0.018
5200164	NI	0.057	0.007
5200165	NI	0.293	-0.007
5200166	NI	0.886	0.014
5200167	**NI**	**3.179** *	**0.066**
5200168	NI	0.104	-0.002
5200169	NI	0.004	-0.002
5200170	D	0.039	0.002
5200171	D	0.307	-0.014
5200172	D	0.655	0.019
5200173	**H**	**4.102** **	**0.071**
5200174	**H**	**6.610** **	**0.064**
5200175	H	0.640	-0.019
5200176	H	3.068 *	-0.060
5200177	H	0.231	-0.010

5200178	H	0.524	0.024
5200179	H	0.422	-0.034
5200180	NI	1.225	0.040
5200181	H	0.892	0.025
5200182	NI	2.195	0.028
5200183	**NI**	**10.080** **	**0.087**
5200184	NI	0.429	0.006
5200185	D	1.712	0.025
5200186	D	0.039	-0.001
5200187	NI	0.003	-0.006
5200188	NI	0.268	-0.020
5200189	NI	1.000	0.031
5200190	NI	0.639	-0.035
5200191	NI	0.630	0.029
5200192	NI	0.175	-0.012
5200193	NI	1.568	-0.036
5200194	NI	1.338	-0.033
5200195	NI	2.016	-0.035
5200196	**H**	**4.835** **	**0.048**
5200197	**H**	**2.942** *	**0.066**
5200198	H	5.787 **	-0.066
5200199	H	0.263	-0.010
5200200	D	1.281	0.068
5200201	D	0.093	0.043
5200202	D	2.177	0.036
5200203	H	0.083	-0.023
5200204	H	4.935 **	-0.076
5200205	H	0.280	-0.024
5200207	H	0.842	-0.011
5200208	H	1.059	-0.011
5200209	H	0.773	0.021
5200211	H	1.489	0.044
5200213	H	0.006	0.005
5200214	H	0.679	0.010
5200215	H	0.012	0.000
5200216	H	2.302	0.080
5200217	**H**	**2.804** *	**0.053**
5200218	H	15.989 ***	-0.125

5200219	**H**	**2.798** *	**0.053**
5200220	**H**	**5.589** **	**0.050**
5200225	**D**	**4.416** **	**0.059**
5200226	D	0.985	-0.020
5200227	NI	0.630	0.024
5290129	NI	1.180	0.040
5790210	NI	0.012	0.001
7104088	D	1.674	0.042
8877801	**D**	**3.101** *	**0.050**
8877802	**D**	**5.142** **	**0.053**
8877804	NI	0.864	-0.011
8877805	H	3.116 *	-0.042
8877806	H	0.667	-0.023
8877807	H	0.007	0.004
8877808	H	0.013	0.009
8877809	H	0.592	0.041
8877810	NI	0.303	0.014
8877811	NI	0.955	-0.027
8877812	H	3.205 *	-0.023
8877813	NI	0.662	-0.017
8877814	NI	0.135	0.016
8877815	NI	0.121	0.006
8877817	NI	0.026	0.006
8877818	NI	0.229	-0.014
8877819	D	1.625	0.027
8877820	D	1.290	0.034
8877821	D	0.001	-0.007
8877822	D	0.118	0.024
8877823	D	0.176	0.025
8877824	NI	1.817	-0.040
8877825	H	0.718	-0.019
8877826	**D**	**5.827** **	**0.063**
8877827	H	2.007	0.049
8877828	H	1.192	0.035
8877829	H	0.311	-0.014
8877830	H	0.047	0.003
8877831	H	5.981 **	-0.057
8877832	H	0.459	0.021

8877833	H	2.432	0.036

Note. *p < .10, **p < .05, ***p < .01. χ^2_M = Mantel Chi-Square; SMD = Standardize Mean Difference; D = Desirable Response Condition; H = Honest Response Condition; NI = No Response Instructions Condition. Participants in boldface are identified as responding differentially.

Table 33

Person-Fit Statistic, l_z, for Each Participant on Each Personality Factor

Participant	C	E	N	O	A
1112220	-0.056	0.609	-0.075	-1.018	0.028
1112221	0.193	-3.689	-0.331	0.515	-0.771
1122330	1.362	0.603	0.544	1.668	2.255
1122331	-1.086	-0.042	-0.616	0.505	-1.635
1122332	0.760	0.908	1.448	0.796	-0.322
1122333	-1.642	-0.695	-1.497	-1.266	1.215
1122334	-1.390	-1.500	-3.791	-4.743	-3.810
3100000	-0.658	0.934	0.243	0.226	1.012
3100001	-0.888	-1.722	-1.997	-2.020	-0.698
3100002	1.152	-5.300	-3.414	-5.499	-0.422
3100003	-2.238	-8.473	-2.857	-6.275	-1.914
3100004	0.985	1.140	1.420	-0.818	-0.436
3100006	0.697	0.638	1.195	0.891	1.194
3100007	0.085	0.955	0.159	0.189	0.262
3100008	-0.242	-2.135	-0.026	0.377	0.337
3100009	2.137	-2.284	-0.494	0.327	0.914
3100010	1.034	0.995	0.830	1.123	-0.353
3100011	0.133	1.182	0.591	0.971	-0.023
3100012	0.339	-1.116	1.136	-0.083	-1.370
3100013	0.659	0.452	-0.013	-0.123	0.690
3100014	0.710	1.096	0.525	2.047	-0.173
3100015	2.067	0.744	-0.275	-0.135	1.126
3100016	0.371	-0.869	0.523	-19.234	-0.760
3100017	-1.030	-7.664	-5.673	-5.642	-3.736
3100018	-0.099	-2.365	-1.299	-0.550	-1.681
3100019	0.028	0.852	2.625	1.791	1.194
3100020	-0.146	0.935	1.966	2.316	-0.329
3100021	-5.544	-0.522	-0.184	-4.716	-8.903
3100022	-0.612	-11.550	0.507	-5.990	-5.393
3100023	-4.119	-2.195	-13.193	-2.640	-2.078
3100024	-0.926	1.885	-3.062	1.532	-2.367
3100025	0.431	-0.490	0.125	0.447	1.082
3100026	0.618	-0.842	-2.427	0.556	-4.963
3100027	0.486	1.264	0.148	-1.692	1.021

3100028	0.754	-1.755	0.823	0.932	1.324
3100029	0.117	-1.700	-0.534	-3.089	0.501
3100030	1.060	0.018	-0.180	0.064	0.650
3100031	-0.828	-0.172	0.113	-13.278	0.220
3100032	0.782	-3.782	-1.335	-1.179	-4.870
3100033	-0.868	0.387	1.842	-3.183	-0.249
3100034	0.100	1.343	-0.277	0.559	0.442
3100035	0.132	2.011	0.411	0.924	0.505
3100036	-0.262	0.769	0.449	-1.212	-0.549
3100037	0.220	-0.971	1.483	-1.751	0.264
3100038	-3.575	-1.578	-1.060	-3.290	-0.166
3100039	-0.465	0.585	0.363	1.555	0.074
3100040	0.371	0.664	0.575	1.449	1.044
3100041	0.028	-3.203	-0.157	1.544	-0.155
3100042	0.960	-0.261	-0.483	-0.428	0.748
3100043	-1.350	-3.818	1.489	-0.990	-0.181
3100044	-1.415	-1.597	-1.766	-4.008	-0.721
3100045	-0.612	0.471	0.320	0.319	1.485
3100046	2.141	-1.260	-0.798	1.181	1.540
3100047	-1.245	-3.913	-1.115	-0.485	1.644
3100048	-1.107	-0.256	0.194	1.823	0.180
3100049	2.242	0.760	0.435	0.046	-0.086
3100050	-14.900	-13.134	2.941	-2.579	-11.653
3100051	-0.609	-4.512	0.993	-1.868	-9.723
3100052	-33.971	-12.402	23.349	-8.047	-11.142
3100053	-15.209	-5.570	0.989	-14.020	-2.731
3100055	2.233	2.352	2.421	-1.253	-0.283
3100056	-3.686	-5.538	0.167	0.579	-1.830
3100057	-0.709	-5.075	-7.013	0.987	-0.281
3100058	-7.840	-8.119	1.397	0.780	-3.019
3100059	-0.412	0.123	-13.734	0.236	0.150
3100060	0.829	-2.675	0.926	-7.224	-0.220
3100061	1.447	-8.639	-1.644	-11.713	-0.306
3100062	-6.279	-1.051	1.692	-0.167	-1.149
3100063	-0.740	0.644	-0.937	1.559	-0.788
3100064	0.049	0.988	-2.290	-0.315	-3.886
3100065	0.811	0.990	1.081	0.178	0.936
3100066	-3.696	1.998	-1.901	2.188	-3.599

3100067	-1.484	-16.942	-4.660	-4.960	-5.275
3100068	1.314	-2.543	-4.552	1.704	1.405
3100069	-12.283	-2.057	1.174	0.974	1.399
3100070	2.340	1.221	0.521	1.482	-1.367
3100071	-1.014	0.886	-0.762	1.209	-0.704
3100072	2.297	0.249	2.115	0.591	0.225
3100073	0.966	1.490	0.409	-0.124	0.871
3100074	0.366	-0.718	0.345	0.329	0.461
3100075	1.114	-0.942	-0.501	-0.617	0.019
3100076	1.577	0.767	0.336	0.653	-0.845
3100077	-2.482	-2.706	0.823	-7.495	-6.716
3100078	1.118	-2.763	0.101	-0.246	-10.904
3100079	-2.097	-5.315	-1.902	-2.087	0.240
3100080	1.506	1.085	1.484	1.124	0.468
3100081	-14.406	-9.695	-5.915	-1.743	-4.087
3100082	-0.706	-9.329	-2.278	-3.162	-3.245
3100083	-10.876	-0.485	2.603	-0.773	-4.875
3100084	2.458	-1.271	1.799	1.838	-5.090
3100085	-0.260	-7.091	-7.249	-0.754	-2.361
3100086	-6.340	1.795	1.691	-1.060	1.123
3100087	-0.207	-6.626	0.273	-1.584	0.974
3100089	-6.686	1.408	-2.931	-3.033	-0.141
3100090	0.192	-13.360	-2.124	0.766	0.887
3100091	1.678	2.523	1.468	-0.394	0.066
3100092	0.440	-6.571	0.239	-5.899	-0.035
3100093	-9.024	-3.601	-3.428	-1.447	-2.801
3100094	0.668	0.736	-2.965	1.449	-7.195
3100095	0.818	-1.559	0.264	-3.208	-2.823
3100096	-0.520	-0.113	1.043	0.953	2.358
3100097	-2.117	-4.481	-0.812	1.537	-0.846
3100098	-0.934	1.737	-0.575	1.339	-0.245
3100099	0.527	-1.859	-0.225	-0.593	-5.568
3100100	-11.001	-17.441	15.407	-9.458	-23.530
3100101	-2.652	-3.134	-0.105	0.214	0.762
3100102	-0.503	2.367	-2.334	-2.308	-11.961
3100103	-3.433	-1.374	14.921	-1.861	0.667
3100104	-1.926	-3.133	0.210	-1.551	-3.640
3100105	-1.262	-3.433	-10.858	-1.441	0.363

3100106	-0.422	1.183	-0.104	-4.196	1.979
3100107	-0.246	-4.430	1.215	0.053	-2.203
3100108	0.919	0.077	-0.432	0.920	-4.496
3100109	0.341	-1.083	-9.769	2.108	-2.439
3100111	-7.345	0.067	-0.396	-1.359	1.997
3100113	1.903	0.980	0.268	-1.249	1.355
3100114	0.147	0.503	2.026	0.162	1.480
3100115	-0.716	0.536	-1.864	1.510	-0.123
3100116	0.679	0.560	1.159	2.327	0.153
3100117	0.650	0.265	0.091	0.340	-1.396
3100118	0.739	-0.242	0.448	0.596	0.543
3100119	-0.108	-0.003	2.315	0.105	1.478
3100120	-1.190	-0.502	-0.682	0.570	2.308
3334444	0.905	-18.802	2.836	1.703	-0.100
3355335	2.606	-3.687	1.854	1.478	1.503
4433300	0.670	-0.247	-2.640	1.444	0.152
4443333	-5.676	-0.974	1.225	-9.330	-8.326
4443334	-2.275	-2.720	-0.652	-0.084	-12.493
4443335	0.386	-2.000	0.333	-2.517	-4.502
4443336	1.137	-0.122	-0.175	-5.116	-0.245
4443337	-11.574	1.376	1.432	-2.682	-3.789
4443338	-0.892	1.505	0.629	0.622	-2.806
4443339	0.151	-2.751	-0.214	-1.049	2.890
4455550	-0.571	-0.448	23.083	-5.090	-0.361
4455551	-7.292	1.128	12.719	0.232	-3.702
4455552	0.487	-2.695	-1.515	-3.695	-6.444
4455553	-4.383	-2.451	-1.283	-1.125	-3.134
4455554	0.921	0.108	-0.634	-8.203	0.417
4455555	-3.569	-4.827	-1.139	-5.417	-0.779
4455556	-4.179	-0.494	0.561	-0.177	0.667
4455660	-1.307	-5.040	1.463	-0.576	-1.835
4455661	1.817	2.122	2.733	-2.894	1.102
4455662	0.321	-0.158	0.453	-1.478	0.279
5200000	-0.789	-0.514	-0.319	-1.126	1.006
5200001	1.507	0.154	-1.169	-0.541	0.274
5200002	-0.315	-1.043	-2.236	-1.558	-0.169
5200003	-5.355	-9.882	-0.810	-14.609	-1.755
5200004	0.613	-1.844	1.584	0.938	0.530

5200005	-0.312	-6.982	-9.257	0.193	-4.719
5200006	-1.700	1.017	1.469	-5.314	1.220
5200007	-0.877	-1.608	1.506	-0.743	-0.288
5200008	-2.210	-7.134	-5.836	-0.097	-2.135
5200009	-2.026	0.199	1.187	2.182	-2.039
5200010	-7.125	-8.633	-7.903	-0.374	-5.561
5200011	0.544	-3.486	24.050	-1.465	-3.887
5200012	2.026	1.518	1.662	0.257	1.005
5200013	-0.497	-2.299	1.935	-0.886	0.255
5200014	0.893	1.027	0.095	-2.658	1.839
5200015	-1.783	-2.581	0.658	-1.135	-1.543
5200016	0.968	-3.168	0.974	-2.640	-2.371
5200017	-20.311	-28.041	28.939	-17.288	-35.432
5200018	1.470	-4.835	2.599	-5.363	1.654
5200018	2.147	-13.704	-0.724	0.395	0.736
5200019	2.095	-2.326	2.485	0.499	1.723
5200020	-2.488	1.280	0.026	0.070	0.774
5200023	-5.000	-6.584	0.380	-10.427	-7.219
5200024	-1.010	-8.401	-1.451	-7.390	-8.297
5200025	-0.973	0.580	1.732	-1.695	-4.487
5200026	0.895	-1.084	-4.503	0.458	-1.035
5200027	-0.790	0.230	-0.977	-0.692	-0.721
5200028	1.406	-2.296	-0.487	0.695	-3.462
5200029	1.151	1.436	-0.832	-1.718	0.447
5200030	-8.449	-4.527	-1.416	0.690	-0.480
5200031	-3.326	-4.984	0.362	0.962	-3.592
5200032	-0.864	-0.158	0.449	-3.913	0.778
5200033	-10.241	-5.490	-1.551	0.291	-9.244
5200035	-33.953	-37.094	16.265	-9.000	-11.557
5200036	-8.488	-4.400	23.510	-2.686	-14.259
5200037	1.004	1.055	1.205	1.584	0.625
5200038	0.855	0.929	0.113	0.416	1.091
5200039	-6.634	-8.113	12.262	-9.740	-6.369
5200040	-2.939	0.436	-8.667	-0.029	-2.933
5200041	-4.843	-3.507	-4.023	-5.444	-0.936
5200042	-4.201	1.838	1.397	-6.484	-8.304
5200043	-8.042	-8.224	0.789	-6.252	-1.935
5200044	-0.967	-2.915	0.577	1.236	-5.398

5200045	-2.068	-5.407	0.568	0.415	-1.756
5200046	-0.428	-2.803	-1.619	-0.690	0.181
5200047	-2.262	-3.762	16.501	-1.378	-3.383
5200048	0.661	0.211	-9.671	-7.217	-0.019
5200049	-5.601	-2.235	-0.112	-0.432	0.869
5200050	1.517	0.850	-3.389	1.507	1.096
5200051	-3.447	0.338	-9.817	-7.203	-6.574
5200052	-0.220	0.479	-1.422	0.571	-0.977
5200053	1.967	0.657	-0.871	-1.564	-0.092
5200055	-9.191	-3.910	-3.167	-2.499	-6.798
5200056	-1.518	-3.134	1.665	-13.937	-1.517
5200057	1.268	-1.326	1.988	-1.694	-4.076
5200058	0.554	-0.226	-0.341	-0.235	-0.474
5200059	1.469	1.862	0.358	1.219	0.069
5200060	-6.809	-3.476	1.575	-6.403	0.360
5200061	1.085	-3.711	0.280	0.996	-0.034
5200062	0.214	-0.276	0.293	1.141	-0.512
5200063	-6.708	-1.076	-4.022	-4.624	19.351
5200064	-0.444	0.585	-10.106	-2.052	-1.306
5200065	-2.807	-4.848	-0.949	0.345	-0.439
5200066	-2.633	-10.843	0.642	1.133	-3.491
5200067	-0.124	-11.589	-6.028	-0.275	0.275
5200068	-1.637	-6.834	1.558	-1.623	-5.200
5200069	-3.417	-5.891	-3.455	-13.017	0.767
5200070	-5.904	-0.956	-9.396	0.317	-5.106
5200071	-5.567	-6.304	-2.956	0.548	-7.151
5200072	-0.200	-4.481	-0.788	-3.796	-5.425
5200073	0.047	1.411	0.201	-3.046	-3.496
5200074	-1.914	1.573	2.503	-1.362	-2.908
5200075	-11.612	-0.428	-5.978	-6.291	-6.115
5200076	-4.574	1.235	-5.381	0.745	-5.179
5200077	-2.902	-2.123	-1.833	-2.476	-2.555
5200078	-6.674	0.087	1.864	0.896	0.184
5200079	1.500	-0.970	-0.064	-8.809	2.201
5200080	-0.532	0.736	-2.416	-0.440	0.104
5200082	-4.432	-2.898	2.313	-4.642	1.783
5200083	1.540	1.157	1.333	-0.733	-1.707
5200085	-2.247	0.423	-5.454	-1.998	-3.312

5200086	1.135	-0.477	0.589	-1.411	-1.247
5200087	-3.775	-3.353	-8.736	-2.857	12.806
5200088	0.268	-2.440	-2.808	2.328	0.369
5200089	-0.857	-0.087	-12.078	-0.207	-3.793
5200090	-7.339	-6.497	-2.543	-3.662	-2.775
5200091	-6.379	-3.930	-6.683	-12.065	-9.082
5200092	-1.142	-2.550	0.356	0.241	-6.128
5200093	1.349	-4.615	-2.043	-0.275	0.499
5200094	1.588	0.628	1.897	-0.084	-2.594
5200095	-25.164	-4.102	-7.935	-0.264	-5.077
5200096	0.653	-6.393	1.217	0.996	1.216
5200097	-8.344	0.074	1.752	-1.109	0.024
5200098	-4.294	0.403	1.714	-1.623	-5.209
5200099	-0.288	0.607	-20.631	0.306	-0.419
5200100	0.597	-6.224	0.228	-0.538	-1.278
5200101	-1.533	-12.945	-6.445	-7.685	-1.207
5200102	-4.441	-1.327	-5.530	1.540	0.622
5200103	1.381	-1.699	-7.425	-16.818	-3.324
5200104	-11.461	-5.222	-10.191	-4.579	-6.458
5200105	-0.580	-1.225	-3.813	-1.892	-1.128
5200106	-15.823	-4.269	-7.023	-10.815	-2.723
5200107	-0.371	-3.239	-3.970	-0.258	-2.822
5200108	-1.119	-3.572	-2.015	-0.700	-5.841
5200109	-11.947	-5.224	-3.769	-11.707	-4.772
5200110	-1.515	-3.708	-6.414	-1.591	-7.165
5200111	-0.799	-8.711	-1.902	1.865	-5.779
5200112	-4.035	-0.400	-14.492	-1.771	-1.765
5200113	0.405	-7.439	0.164	-1.459	1.650
5200114	-14.464	-2.258	-1.232	-3.078	-1.833
5200115	1.961	0.658	-7.973	-1.958	1.731
5200116	-6.192	-1.685	-15.194	1.238	26.280
5200128	-10.521	-0.305	-7.633	0.562	-1.288
5200130	-9.711	-2.227	-5.425	-4.932	-0.526
5200131	-3.884	-1.391	-0.266	-8.904	1.157
5200132	-2.840	-2.641	-3.285	-0.019	-0.710
5200135	-8.258	2.466	2.088	0.032	-1.106
5200137	-2.783	1.602	1.521	-10.950	1.018
5200138	-1.128	-2.026	-8.971	-5.931	0.894

5200140	1.412	-2.704	-6.780	-1.553	-0.507
5200141	-8.274	-9.295	-7.308	-0.064	-25.433
5200142	0.469	-1.706	-0.855	-2.119	-1.407
5200143	1.773	-7.931	1.599	1.034	-3.987
5200144	-0.659	-1.474	-3.914	-0.655	-5.840
5200145	-3.106	-1.548	-6.260	-1.367	1.666
5200146	-19.900	1.237	1.804	0.566	12.795
5200147	-0.539	-6.149	-1.923	-1.407	-6.632
5200148	-3.145	1.268	-7.227	-6.185	-2.941
5200149	-9.947	-1.615	-3.660	-8.907	-6.376
5200150	-3.332	-6.090	-2.374	-1.974	-2.385
5200151	0.277	-10.556	-0.847	1.334	0.563
5200152	-3.360	0.141	-1.804	-0.899	-0.428
5200153	-1.930	0.046	-6.219	-0.227	0.900
5200154	-10.917	-0.117	-3.412	-1.099	-2.817
5200156	-24.796	-21.219	-23.305	-7.219	-20.707
5200158	-13.248	-7.707	-0.675	-1.977	-13.177
5200159	-0.974	2.726	-0.612	-2.058	-1.094
5200160	0.396	0.979	-4.799	1.917	-3.768
5200161	-17.030	-11.503	1.434	-5.991	-15.189
5200162	-2.597	-0.198	0.626	-6.241	2.071
5200163	-4.925	-11.153	-5.385	-1.683	-2.308
5200164	1.887	-2.197	-2.027	-12.952	-9.912
5200165	1.374	2.251	2.553	-3.789	0.907
5200166	1.612	-3.970	-12.401	-0.447	-12.549
5200167	0.227	-0.025	1.644	-0.743	-3.614
5200168	-23.466	-2.332	-1.319	-1.570	-20.280
5200169	-12.789	-12.544	-5.362	-16.168	-10.062
5200170	-2.908	1.559	-0.970	1.024	-8.176
5200171	-2.119	-8.363	0.242	-2.456	0.161
5200172	1.695	0.502	1.544	0.036	-1.550
5200173	-7.188	1.990	-13.399	-1.795	1.242
5200174	-0.744	0.714	-5.916	1.156	-6.462
5200175	-6.227	-5.201	-2.159	1.003	-1.123
5200176	-2.227	-3.302	-2.076	-12.362	-0.393
5200177	-0.080	0.888	-1.515	-1.401	1.261
5200178	1.886	-4.700	-1.977	-0.062	-5.020
5200179	-6.525	-4.054	-7.124	-2.491	-3.949

5200180	0.692	-1.584	-2.976	-2.476	0.172
5200181	-2.597	0.956	1.680	0.370	-1.497
5200182	-4.377	-4.124	0.036	-6.124	1.285
5200183	-6.337	-0.425	0.840	0.799	-0.019
5200184	-6.470	-6.023	-0.228	-4.813	-0.787
5200185	2.568	2.049	-1.590	-2.760	2.228
5200186	2.184	-3.076	1.318	-0.860	2.283
5200187	-0.999	-2.934	-11.692	-0.129	-4.149
5200188	-1.317	-12.485	2.837	-11.126	1.245
5200189	0.582	-7.544	-2.130	-30.281	-9.058
5200190	1.309	0.419	-5.613	1.510	2.277
5200191	-4.090	-3.460	-1.019	0.550	-3.040
5200192	-16.195	-1.280	-9.852	-1.065	-10.077
5200193	1.614	1.902	-4.786	-0.478	-3.100
5200194	-3.451	-8.539	-0.010	-0.360	-2.404
5200195	-13.742	-1.319	0.209	1.573	-16.972
5200196	-0.009	-2.699	-3.126	-4.835	-5.250
5200197	-6.777	-6.560	-4.431	-4.449	-1.521
5200198	-10.786	2.299	-9.716	-1.249	-8.237
5200199	-0.474	-2.416	0.141	-19.786	-9.185
5200200	-0.258	1.091	-3.455	-0.772	1.084
5200201	-8.334	-8.482	-3.910	-2.689	-2.980
5200202	-2.481	1.377	0.347	0.793	0.672
5200203	-1.820	0.708	-0.645	-4.483	-1.490
5200204	-0.272	-11.049	-0.882	-2.807	-3.635
5200205	1.215	1.774	-3.092	-0.928	-2.496
5200207	1.470	-2.301	1.196	-4.785	1.681
5200208	0.255	0.727	1.860	-3.528	-4.598
5200209	-0.074	0.065	-5.474	-2.458	-5.271
5200211	-4.406	-12.935	-5.038	0.122	-5.440
5200213	-5.504	1.053	-7.136	1.207	-3.755
5200214	-1.771	-0.887	1.481	1.134	12.048
5200215	-2.424	-7.184	-1.247	-0.499	11.667
5200216	1.792	-6.337	-30.077	-0.738	0.603
5200217	0.400	-10.100	1.531	-10.512	-0.875
5200218	-4.116	0.060	-12.490	-2.133	1.647
5200219	-1.072	0.203	-2.113	0.038	-0.488
5200220	-0.342	-1.948	-26.077	-2.276	-2.503

5200225	-9.849	-2.082	-1.277	-1.509	-0.537
5200226	-21.294	-2.459	24.486	-5.259	-17.738
5200227	-2.922	2.669	-2.103	1.020	-6.351
5290129	-2.161	-3.794	-0.589	-1.359	-7.122
5790210	-2.900	0.737	2.131	-4.048	1.512
7104088	-2.078	-1.544	0.718	-3.004	-0.584
8877801	-0.825	-0.086	-9.333	-1.605	-0.283
8877802	-1.597	-1.617	-3.138	-0.612	0.979
8877804	-0.032	1.275	-0.650	0.744	-0.829
8877805	-5.424	-2.713	1.508	-3.406	0.924
8877806	-3.605	-4.753	-3.482	-19.452	-2.141
8877807	2.540	1.541	1.714	1.051	-5.606
8877808	-3.046	0.299	-11.665	1.031	-2.585
8877809	-1.658	-11.294	-25.144	1.022	-1.499
8877810	-8.974	-2.662	-0.463	-3.518	-3.197
8877811	1.427	-3.694	0.862	1.073	-0.716
8877812	-1.863	-0.249	-9.451	-1.371	1.526
8877813	0.449	0.890	-2.418	-3.519	-9.297
8877814	-2.874	0.869	1.753	0.422	-2.064
8877815	-10.989	-2.832	2.526	-4.754	1.027
8877817	1.588	-1.146	-0.667	0.059	-0.765
8877818	-1.004	-0.570	0.708	-11.841	-0.126
8877819	1.177	-4.956	-2.860	-3.936	-2.242
8877820	-0.364	-3.568	-4.611	-3.545	0.215
8877821	-12.510	-34.224	-8.791	-15.013	-24.677
8877822	-5.753	-7.492	0.813	-7.095	1.437
8877823	0.033	-3.050	-2.848	0.179	1.850
8877824	-6.606	-12.962	-24.471	-12.338	0.810
8877825	2.116	1.002	-3.084	0.630	1.905
8877826	-7.178	-0.480	1.849	-0.770	-2.880
8877827	-1.774	2.557	2.061	-0.637	-1.885
8877828	-11.299	-1.114	-6.549	-3.136	2.362
8877829	-5.364	0.806	-9.726	-2.179	-1.861
8877830	1.021	-3.117	-0.647	-0.274	-0.865
8877831	-4.899	-11.492	-25.770	-2.416	-3.885
8877832	1.584	-4.763	-1.294	-1.607	1.108
8877833	-6.349	-1.330	-5.099	-6.266	-4.706

Note. C = Conscientiousness; E = Extroversion; N = Neuroticism; O = Openness; A = Agreeableness.

Table 34

Results of the Classification Consistency Analyses for the DPF, the Person-Fit Index, and the BIDR across all Response Instruction Conditions

Method		1	2	3	4	5
1. MH	P					
	κ					
2. M	P	0.875				
	κ	0.377				
3. BIDR 5%	P	0.854	0.818			
	κ	0.051	-0.034			
4. BIDR 10%	P	0.832	0.797	0.979		
	κ	0.020	-0.060	0.851		
5. BIDR 15%	P	0.800	0.765	0.942	0.963	
	κ	0.006	-0.070	0.665	0.805	
6. l_z	P	0.624	0.605	0.611	0.600	0.595
	κ	-0.001	-0.030	-0.067	-0.077	-0.058

Note. MH = Mantel-Haenszel Procedure; M = Mantel Procedure; l_z = Person-Fit Index; P = Probability of a Consistent Classification; κ = Cohen's Kappa.

Table 35

Results of the Classification Consistency Analyses for the DPF, the Person-Fit Index, and the BIDR in the Honest Response Instructions Condition

Method		1	2	3	4	5
1. MH	P					
	κ					
2. M	P	0.907				
	κ	0.455				
3. BIDR 5%	P	0.907	0.829			
	κ	0.132	-0.073			
4. BIDR 10%	P	0.899	0.822	0.993		
	κ	0.08	-0.082	0.919		
5. BIDR 15%	P	0.876	0.798	0.969	0.976	
	κ	0.045	-0.105	0.735	0.811	
6. l_z	P	0.628	0.628	0.613	0.605	0.597
	κ	0.006	0.051	-0.048	-0.062	-0.064

Note. MH = Mantel-Haenszel Procedure; M = Mantel Procedure; l_z = Person-Fit Index; P = Probability of a Consistent Classification; κ = Cohen's Kappa.

Table 36

Results of the Classification Consistency Analyses for the DPF, the Person-Fit Index, and the BIDR in the Desirable Response Instructions Condition

Method		1	2	3	4	5
1. MH	P					
	κ					
2. M	P	0.859				
	κ	0.292				
3. BIDR 5%	P	0.835	0.876			
	κ	-0.066	0.063			
4. BIDR 10%	P	0.826	0.868	0.991		
	κ	-0.076	0.048	0.905		
5. BIDR 15%	P	0.793	0.835	0.942	0.951	
	κ	-0.041	0.075	0.563	0.643	
6. l_z	P	0.628	0.587	0.629	0.62	0.637
	κ	0.064	-0.058	0.008	-0.008	0.069

Note. MH = Mantel-Haenszel Procedure; M = Mantel Procedure; l_z = Person-Fit Index; P = Probability of a Consistent Classification; κ = Cohen's Kappa.

Table 37

Results of the Classification Consistency Analysis for the DPF, the Person-Fit Index, and the BIDR in the No Response Instructions Condition

Test		1	2	3	4	5
1. MH	P					
	κ					
2. M	P	0.859				
	κ	0.292				
3. BIDR 5%	P	0.835	0.876			
	κ	-0.066	0.063			
4. BIDR 10%	P	0.826	0.868	0.991		
	κ	-0.076	0.048	0.905		
5. BIDR 15%	P	0.793	0.835	0.942	0.951	
	κ	-0.041	0.075	0.563	0.643	
6. l_z	P	0.628	0.587	0.629	0.620	0.637
	κ	0.064	-0.058	0.008	-0.008	0.069

Note. MH = Mantel-Haenszel Procedure; M = Mantel Procedure; l_z = Person-Fit Index; P = Probability of a Consistent Classification; κ = Cohen’s Kappa.

Table 38

Summary of the Decision Accuracy of Each Response Distortion Detection Technique and the Associated True Positive, True Negative, False Positive, and False Negative Rates

Method	P	κ	TP	FP	TN	FN
MH	0.535	0.059	5.9%	3.1%	47.6%	43.3%
M	0.520	0.028	7.5%	6.3%	44.5%	41.7%
BIDR 5%	0.539	0.066	5.5%	2.4%	48.4%	43.7%
BIDR 10%	0.559	0.107	7.9%	2.8%	48.0%	41.3%
BIDR 15%	0.567	0.124	9.8%	3.9%	46.9%	39.4%
l_z	0.481	-0.045	15.4%	18.1%	32.7%	33.9%

Note. MH = Mantel-Haenszel Procedure; M = Mantel Procedure; l_z = Person-Fit Index; P = Probability of a Accurate Classification; κ = Cohen's Kappa; TP = True Positive; FP = False Negative; TN = True Negative; FN = False Negative.

Figure 1

Item Characteristic Curve

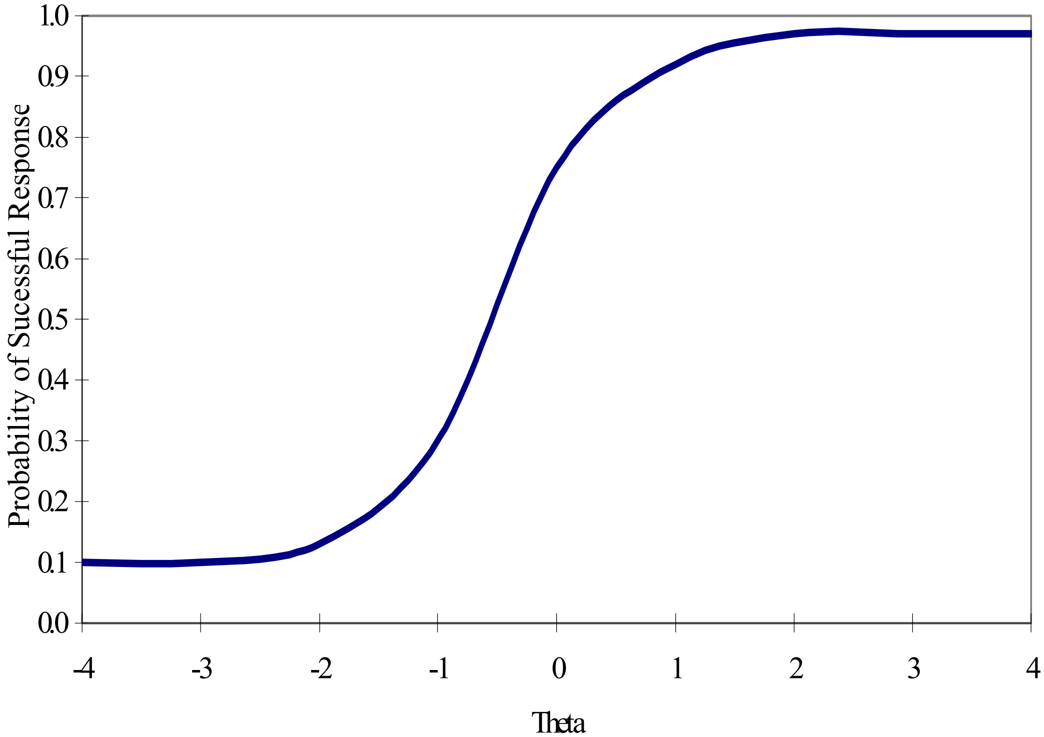


Figure 2

Option Characteristic Curve

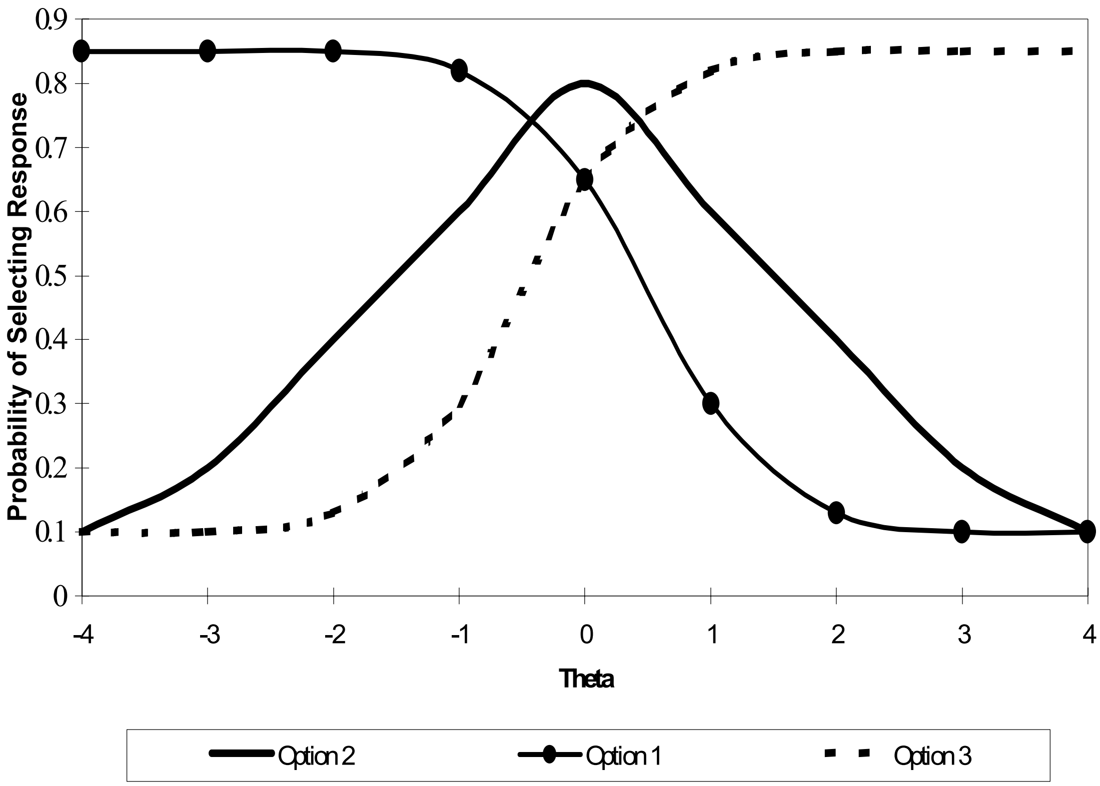


Figure 3

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Conscientiousness

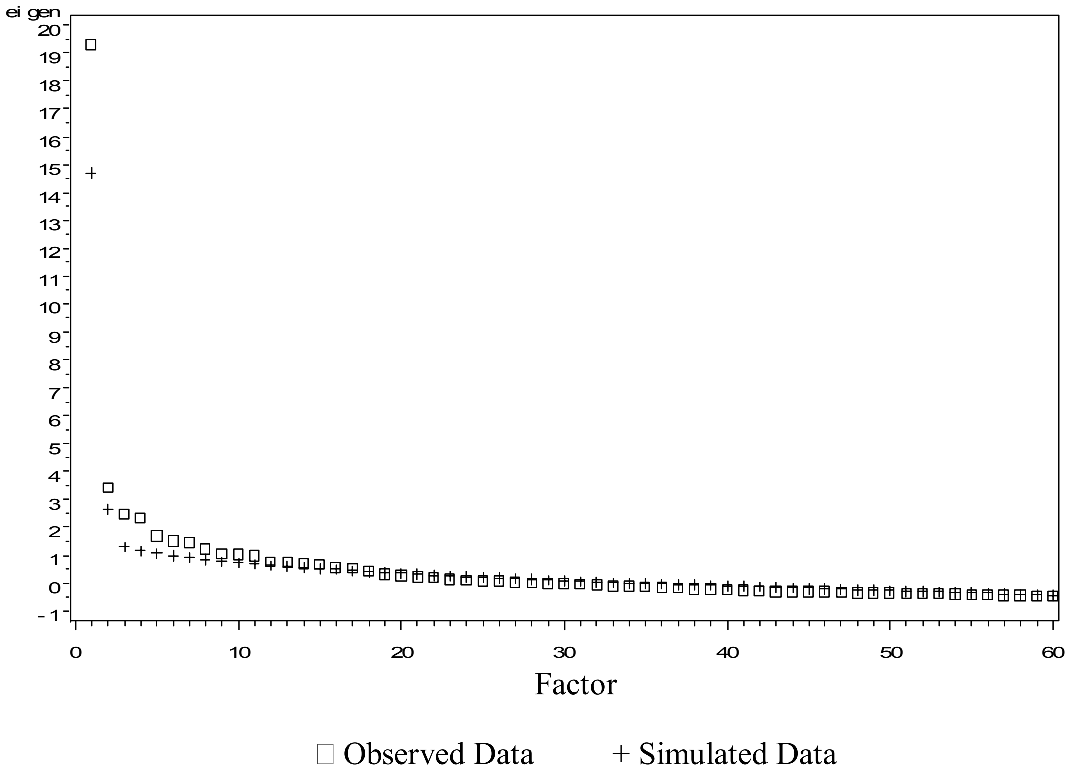
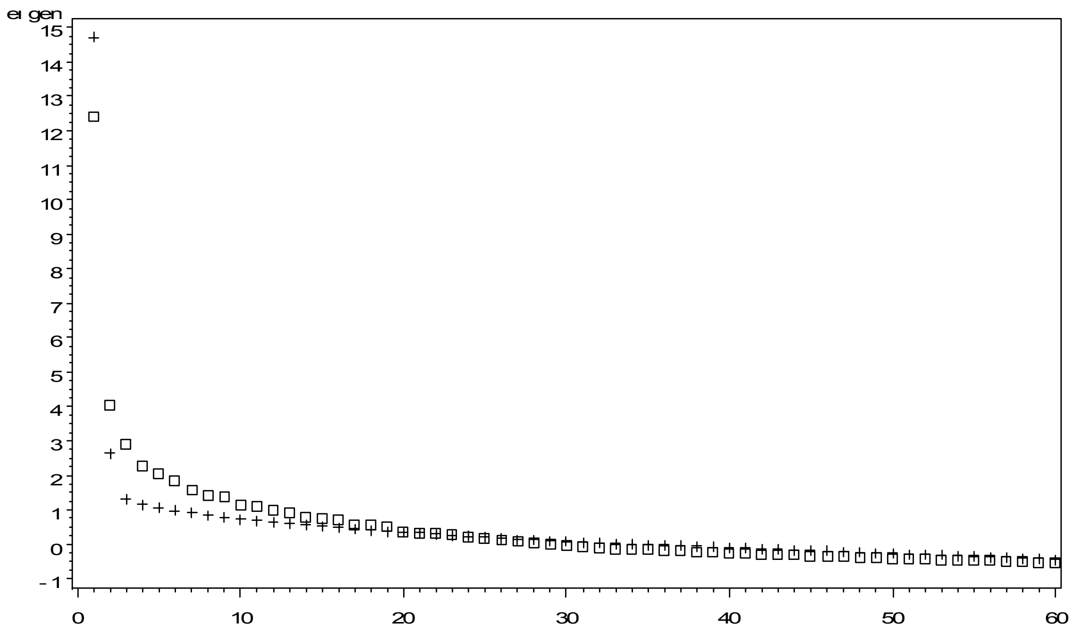


Figure 4

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Extroversion

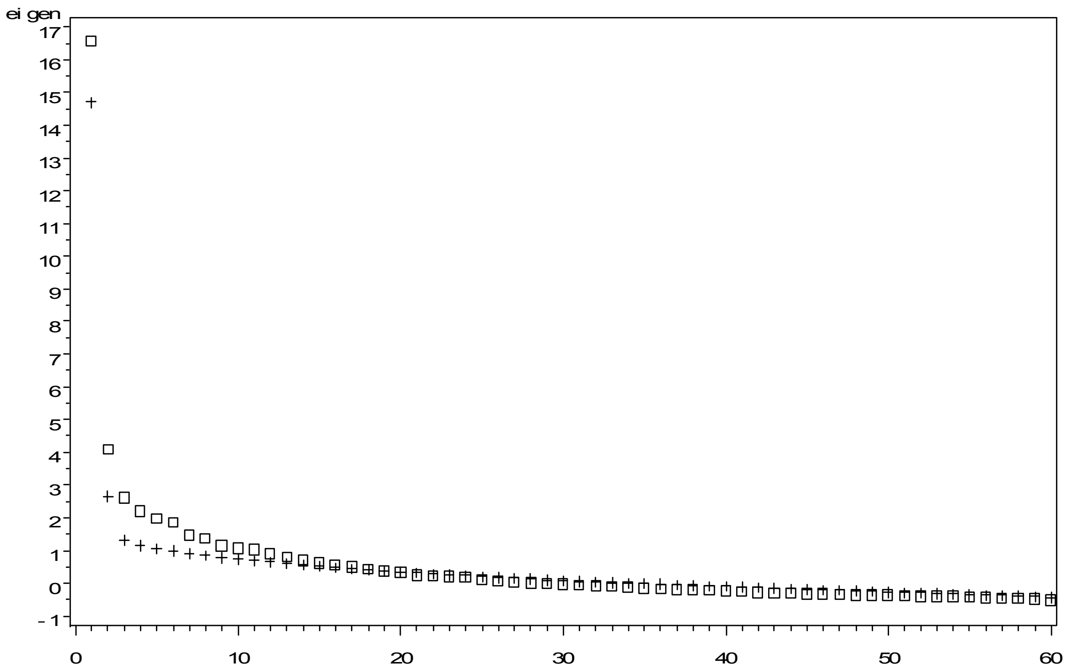


Factor

□ Observed Data + Simulated Data

Figure 5

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Neuroticism



Factor

□ Observed Data + Simulated Data

Figure 6

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Openness

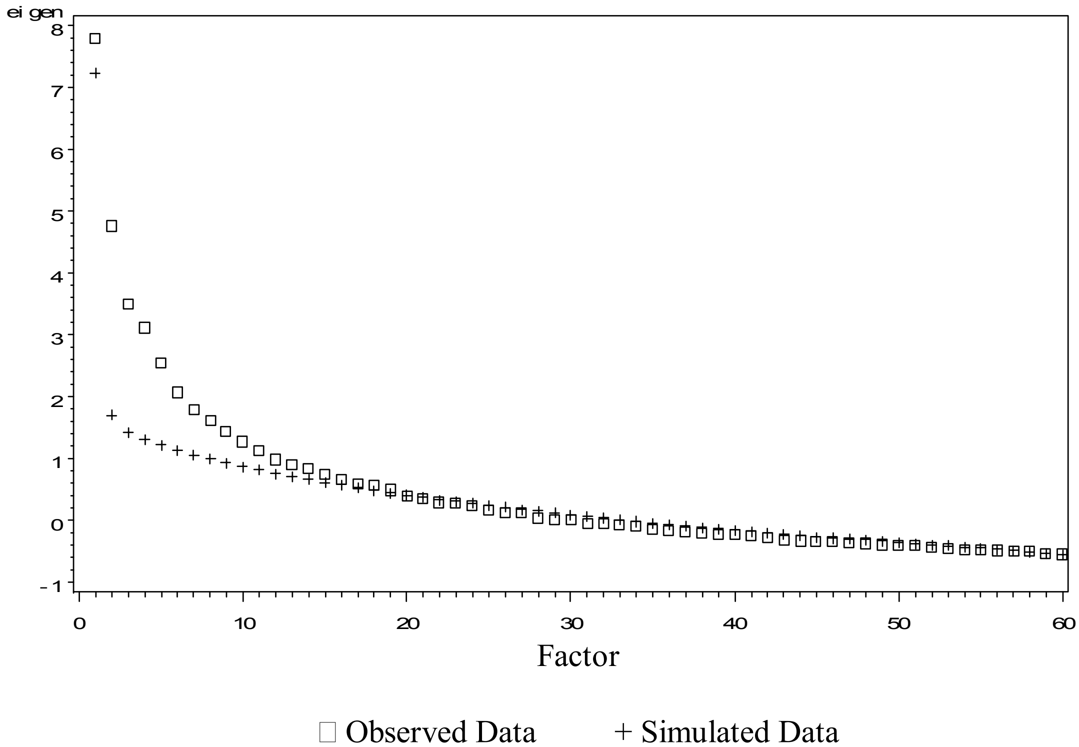


Figure 7

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Agreeableness

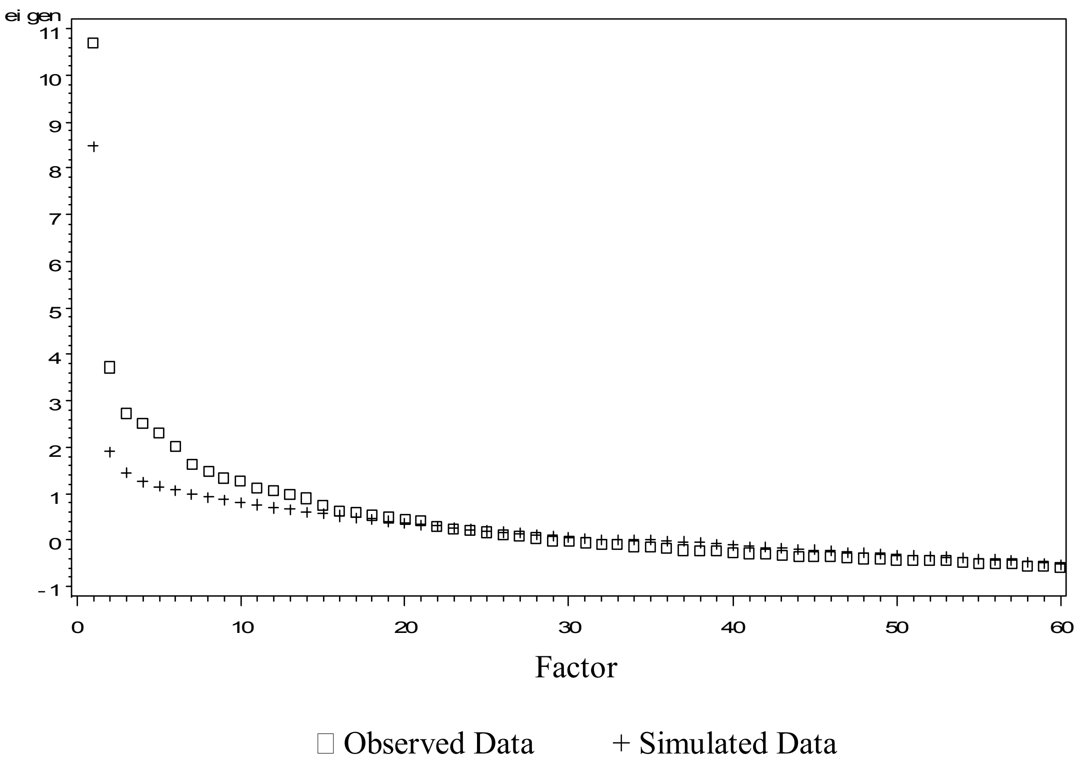


Figure 8

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Conscientiousness

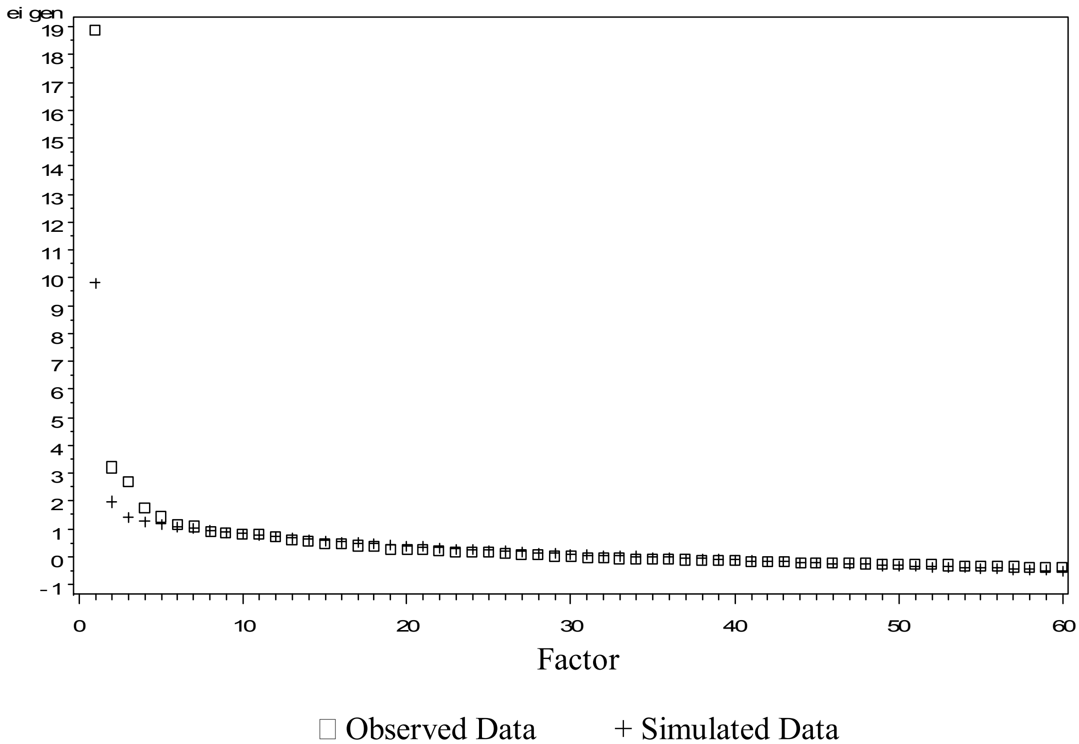
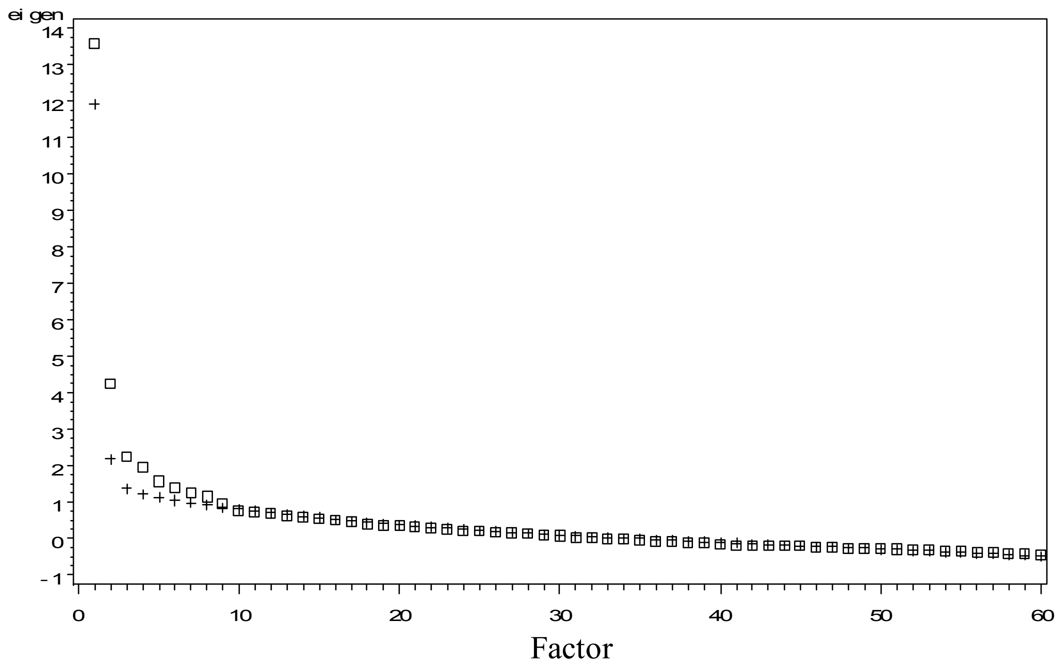


Figure 9

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Extroversion



□ Observed Data + Simulated Data

Figure 10

Modified Parallel Analysis Results for the Honest Response Instructions Condition for

Neuroticism

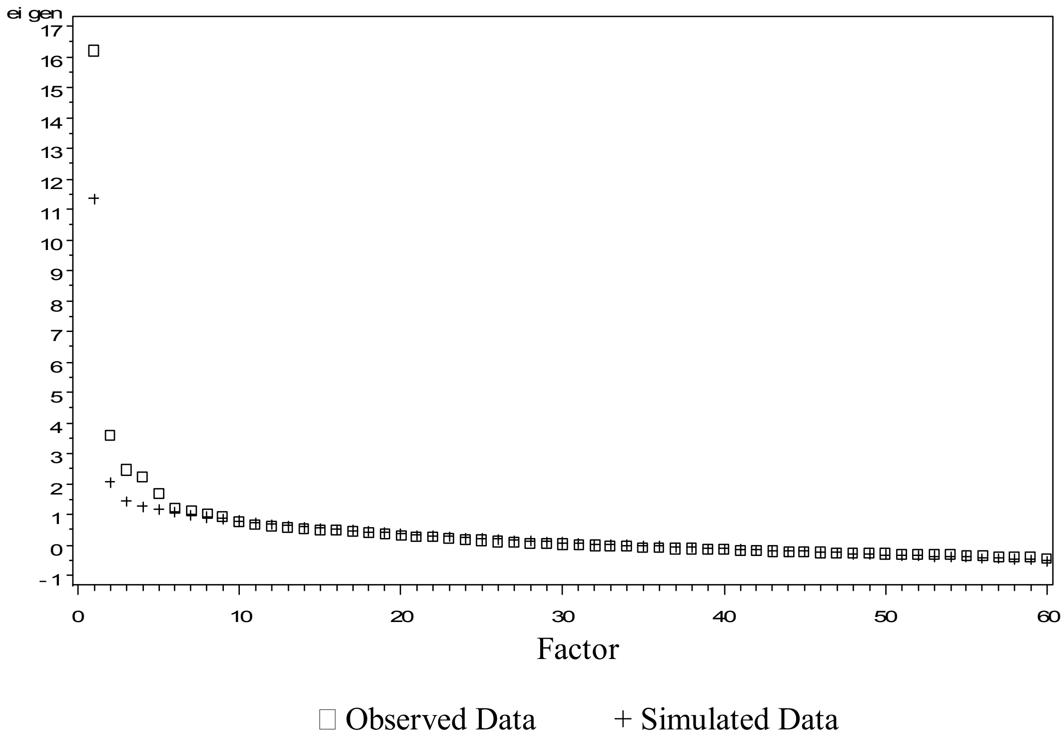


Figure 11

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Openness

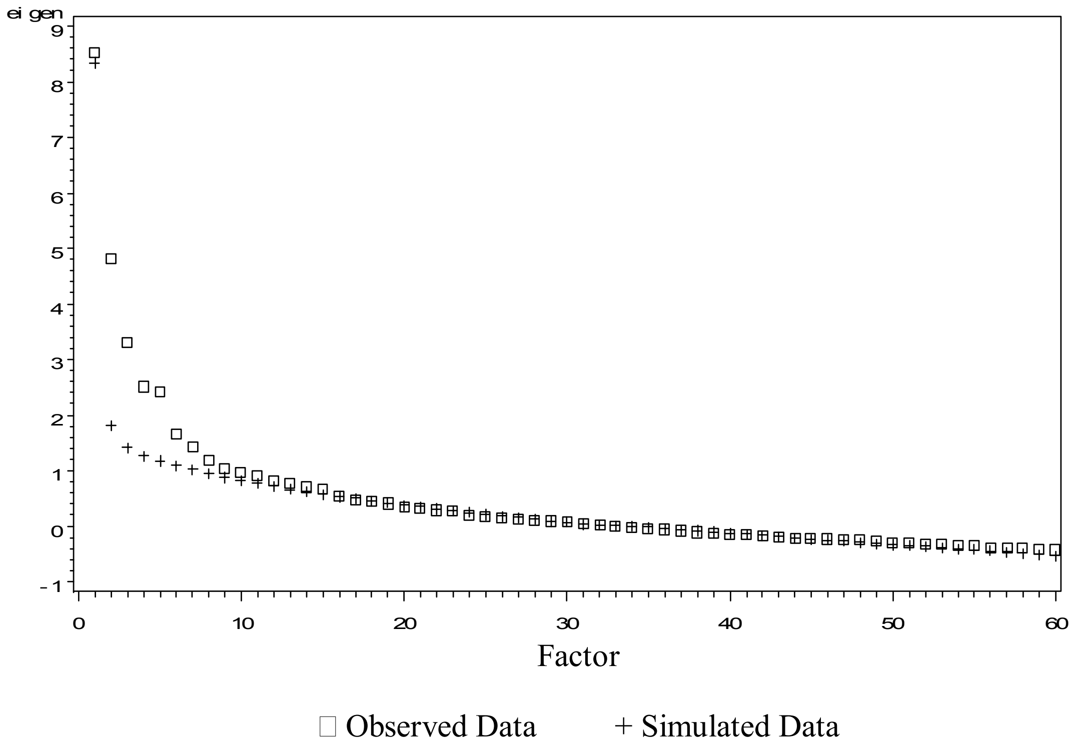
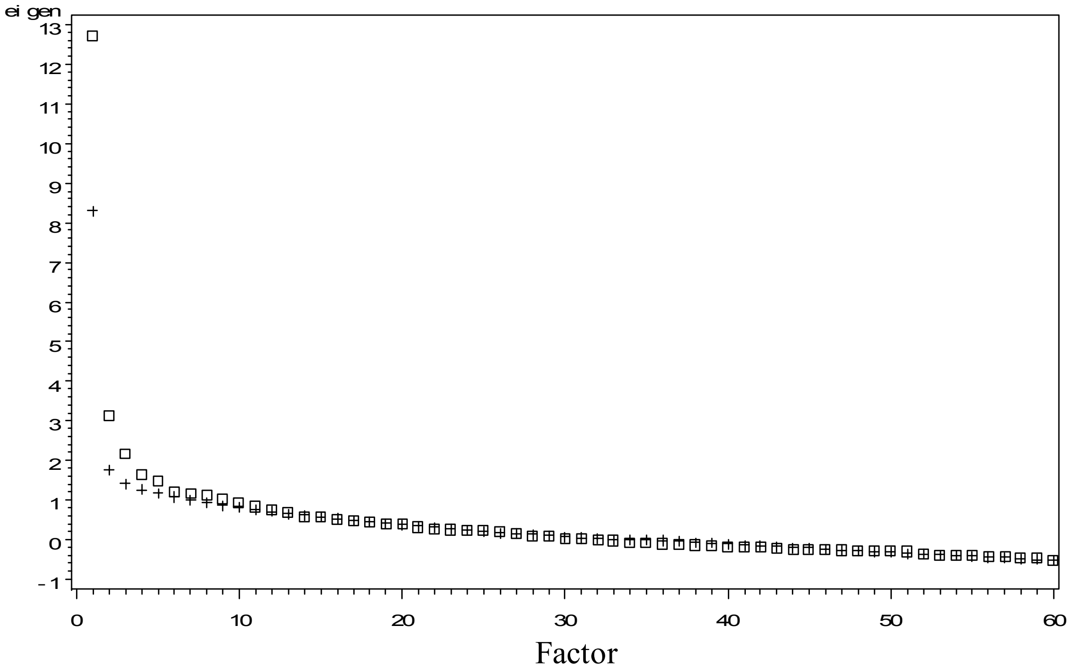


Figure 12

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Agreeableness



□ Observed Data + Simulated Data

Figure 13

Distribution of the Standard Errors of the Item Difficulty Parameter in the Desirable

Response Instructions Condition

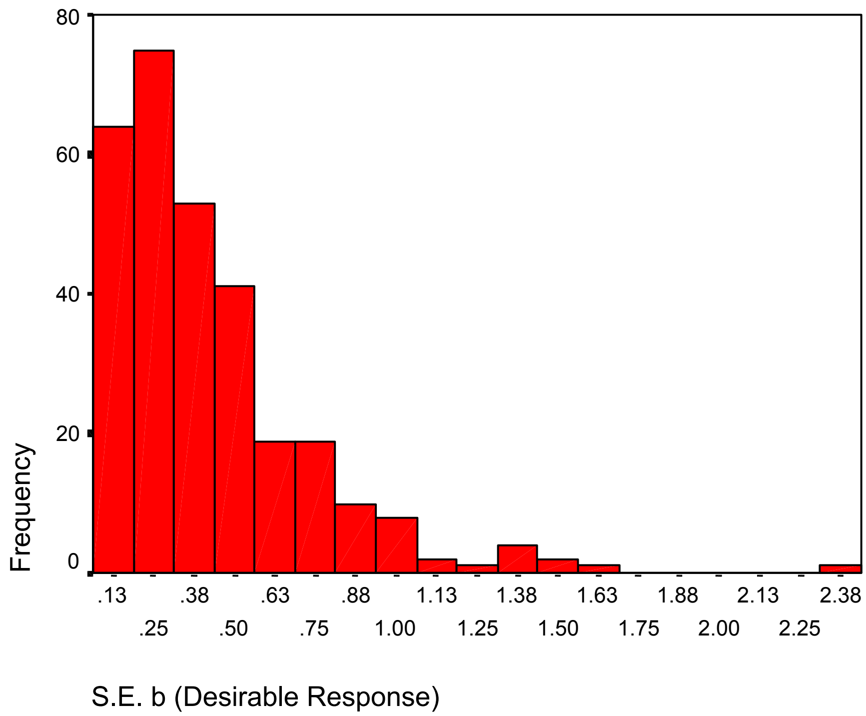


Figure 14

Distribution of the Standard Errors of the Item Difficulty Parameter in the Honest

Response Instructions Condition

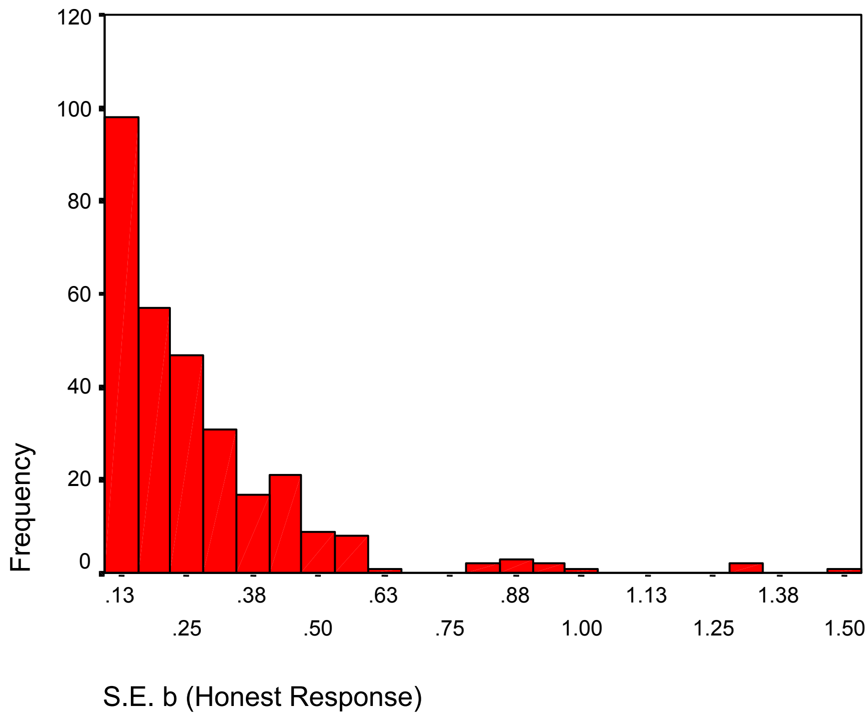
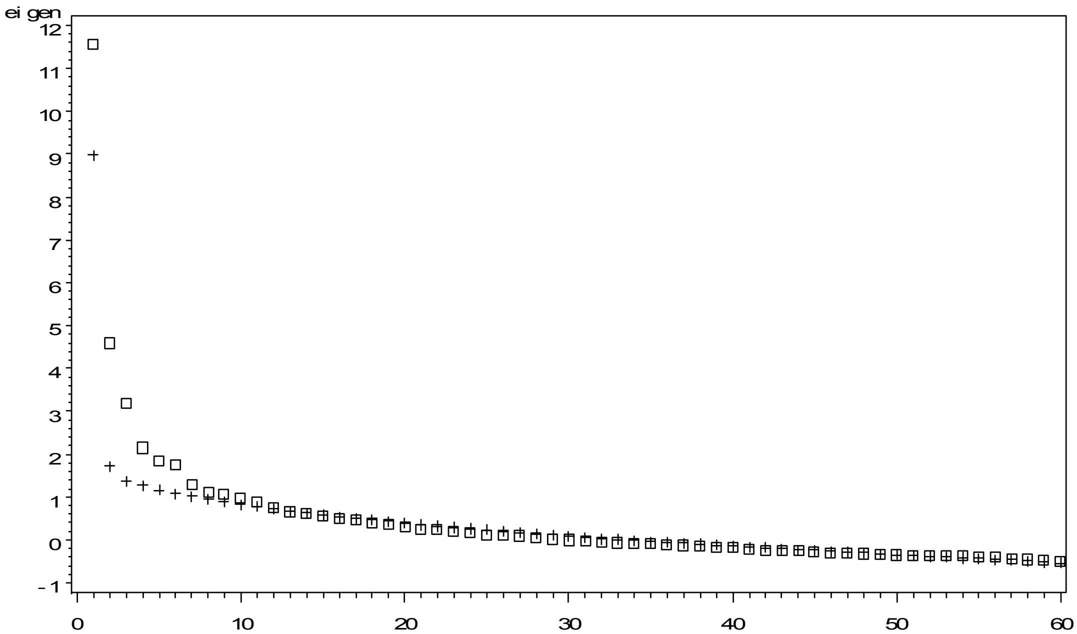


Figure 15

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Conscientiousness

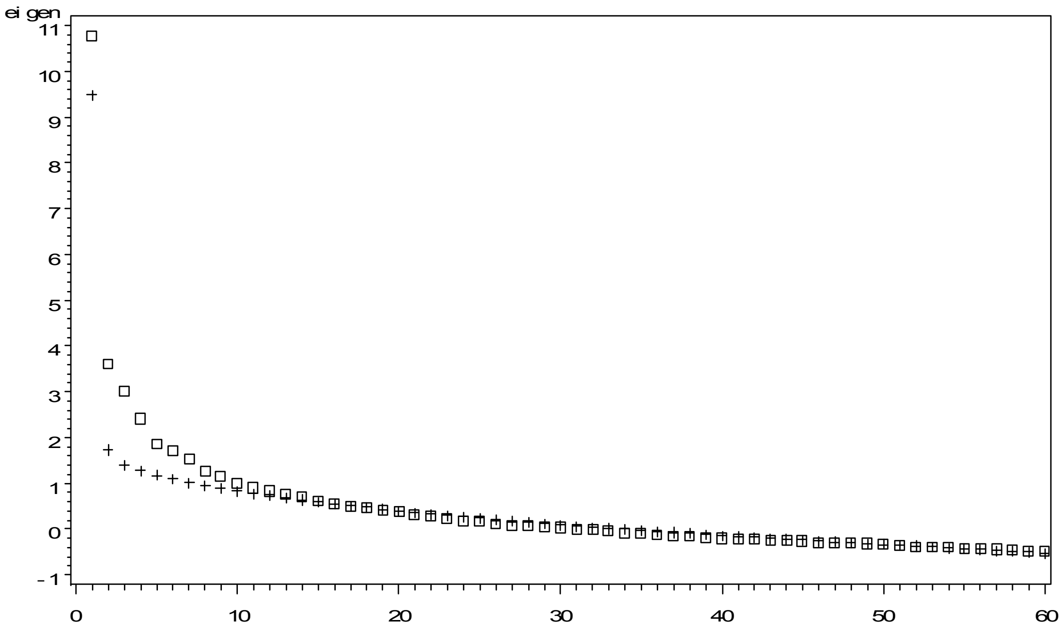


Factor

□ Observed Data + Simulated Data

Figure 16

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Extroversion

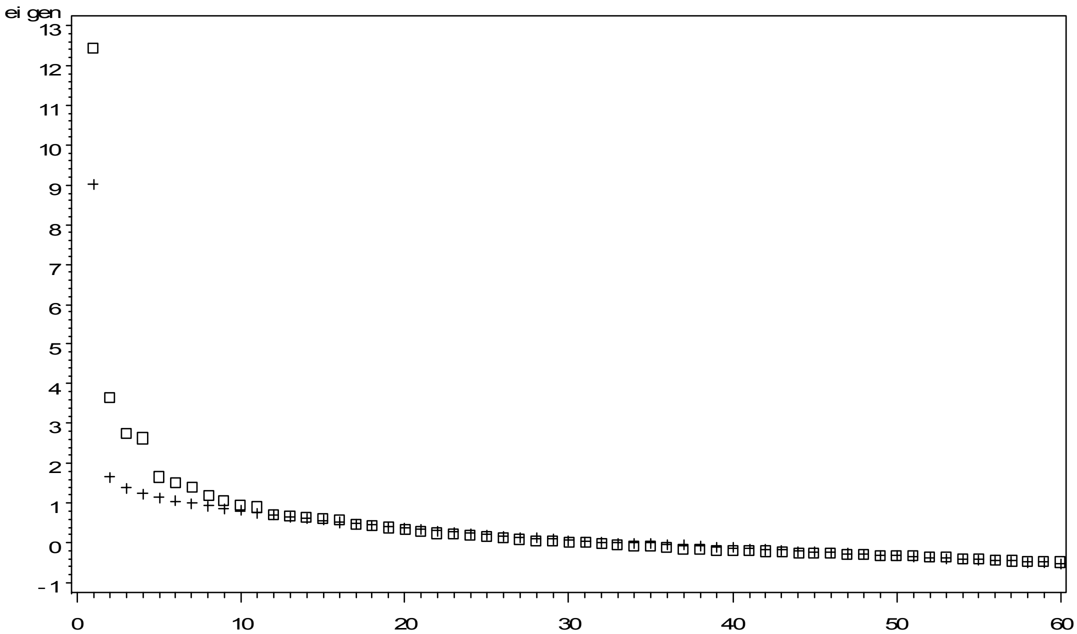


Factor

□ Observed Data + Simulated Data

Figure 17

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Neuroticism

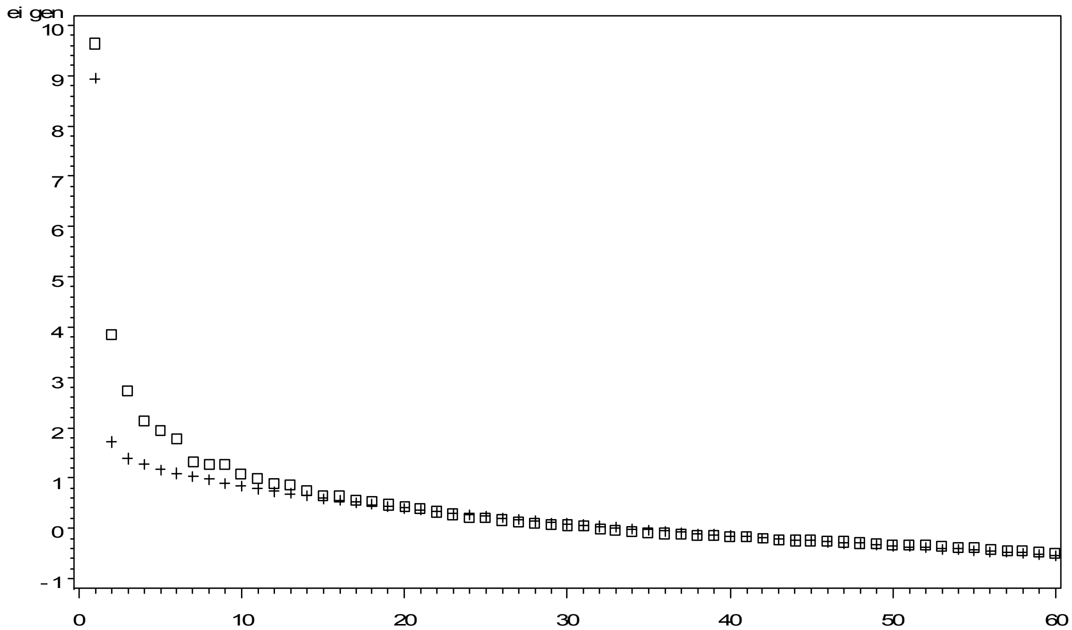


Factor

□ Observed Data + Simulated Data

Figure 18

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Openness

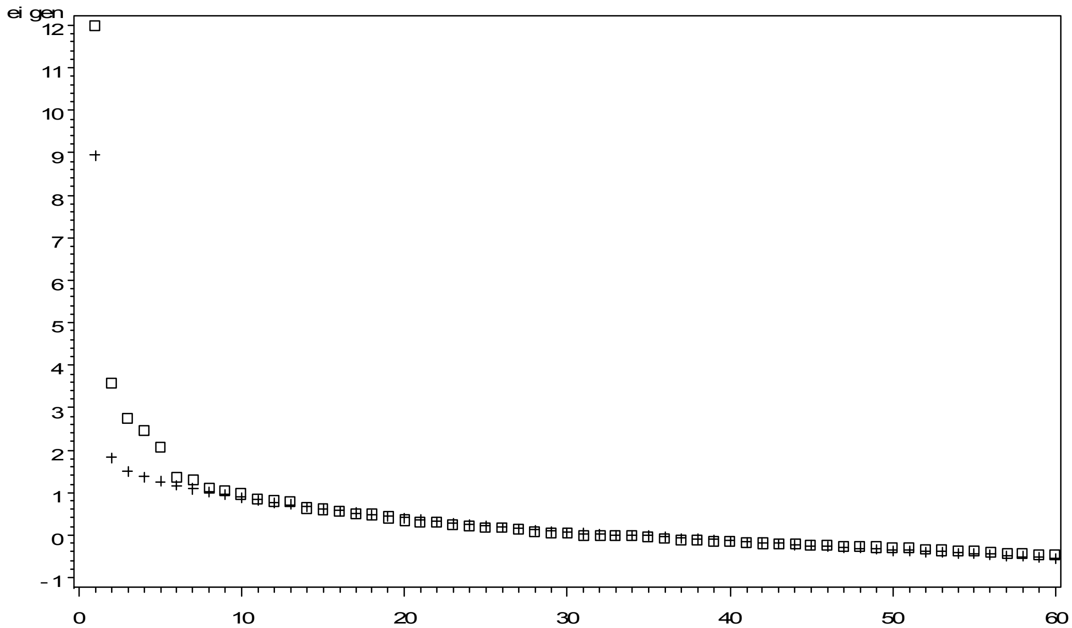


Factor

□ Observed Data + Simulated Data

Figure 19

Modified Parallel Analysis Results for the Honest Response Instructions Condition for Agreeableness

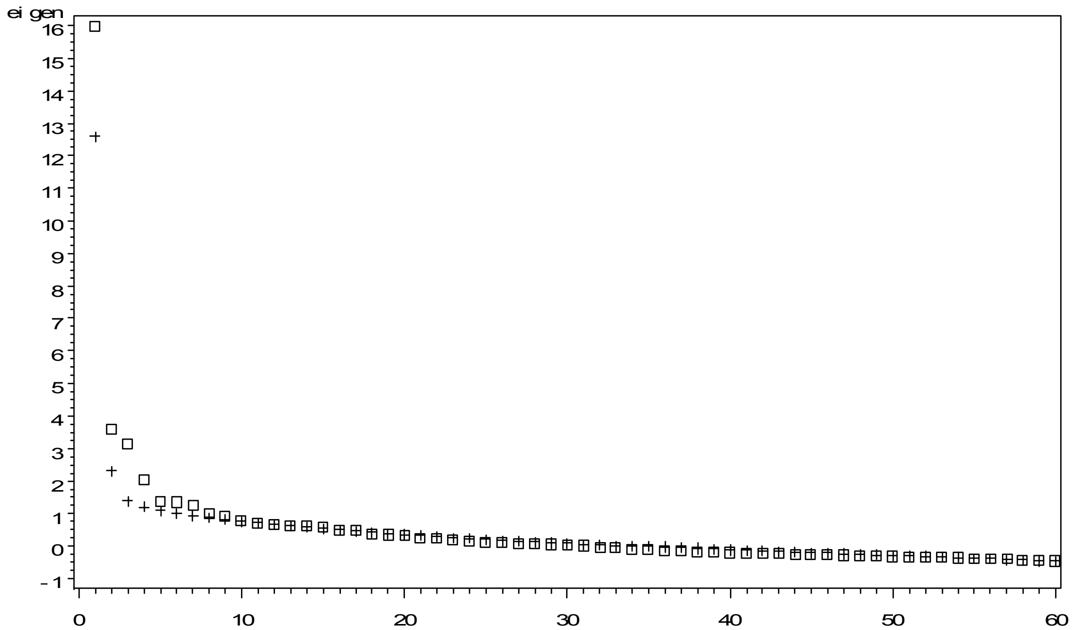


Factor

□ Observed Data + Simulated Data

Figure 20

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Conscientiousness

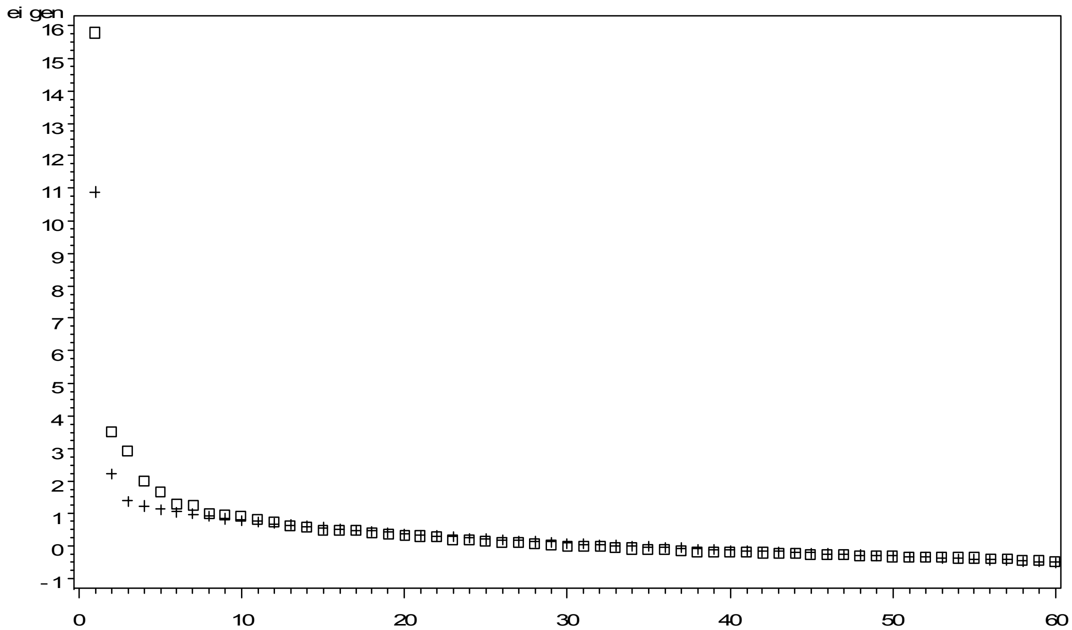


Factor

□ Observed Data + Simulated Data

Figure 21

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Extroversion



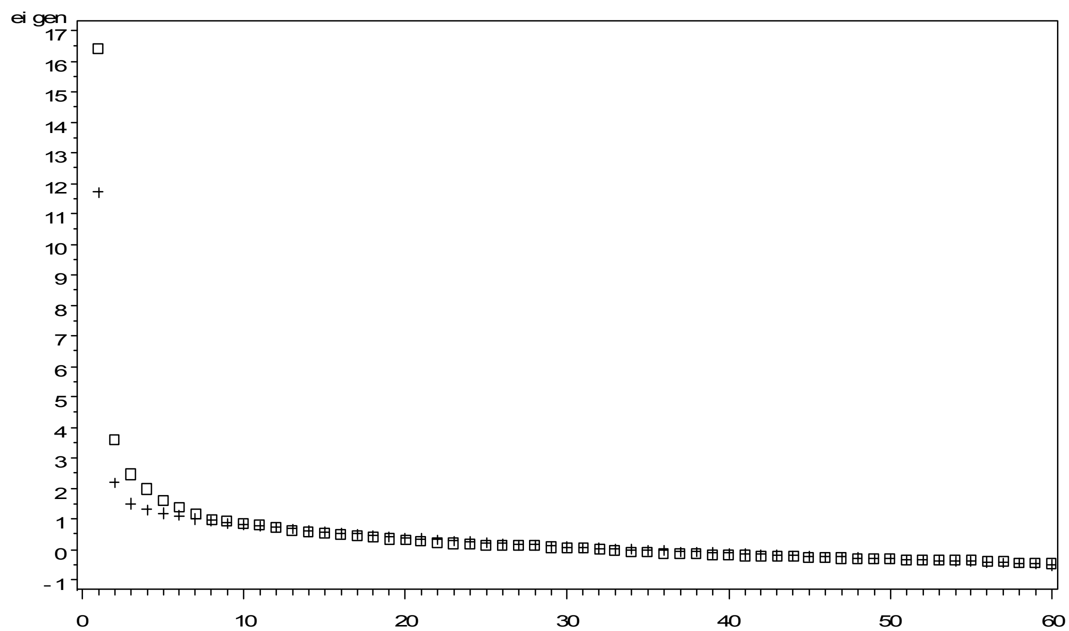
Factor

□ Observed Data + Simulated Data

Figure 22

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for

Neuroticism

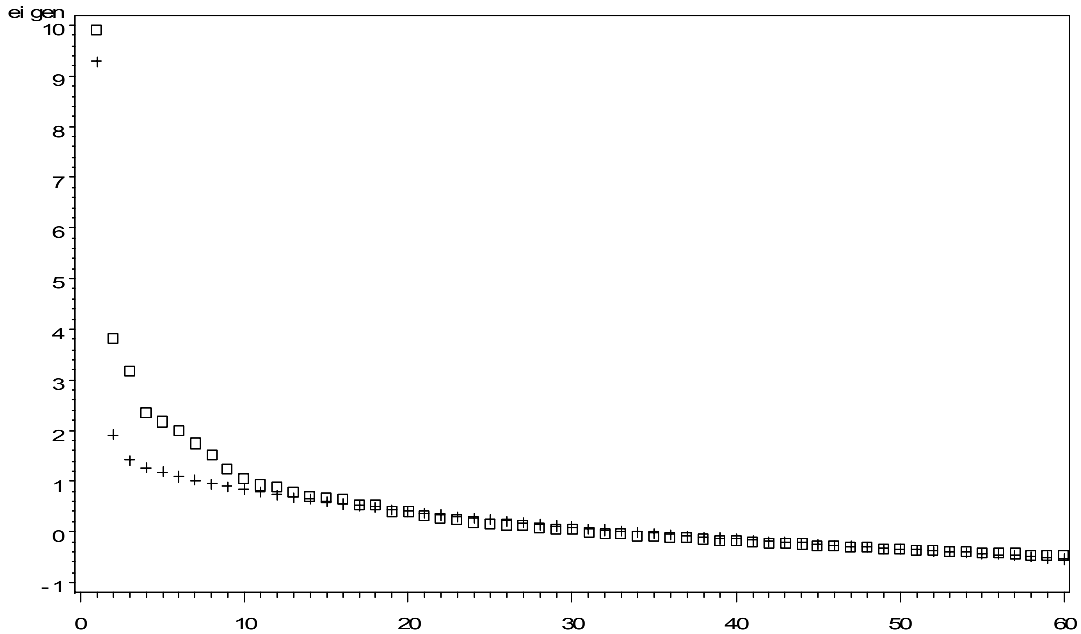


Factor

□ Observed Data + Simulated Data

Figure 23

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Openness

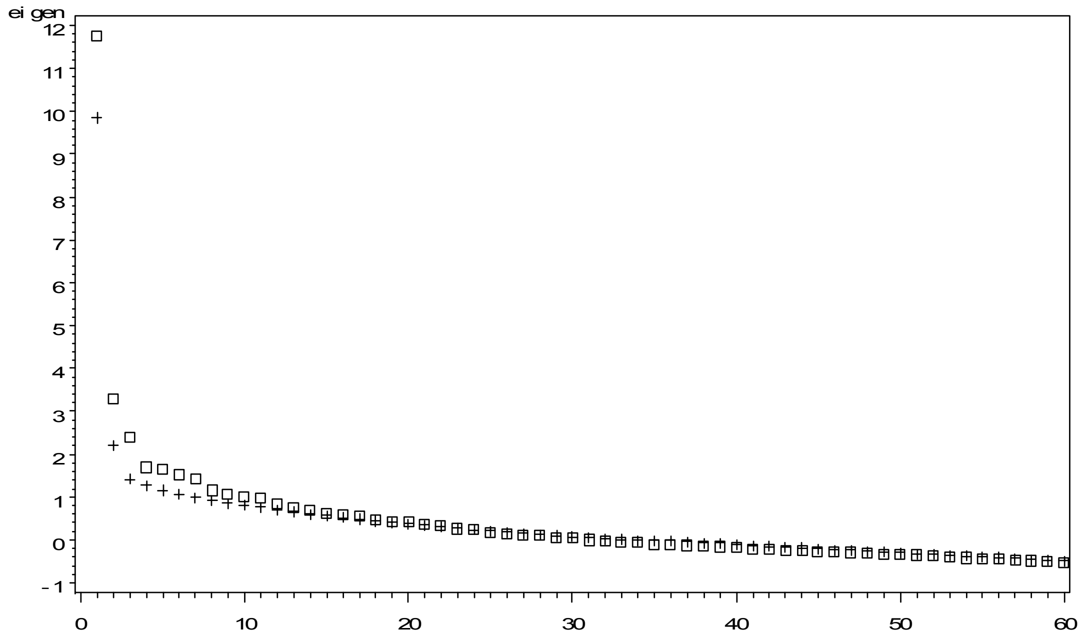


Factor

□ Observed Data + Simulated Data

Figure 24

Modified Parallel Analysis Results for the Desirable Response Instructions Condition for Agreeableness



Factor

□ Observed Data + Simulated Data

Figure 25

Decision Theoretic Classification Table

		Technique 1		
		D	ND	
Technique 2	D	P_{11}	P_{10}	$P_{1.}$
	ND	P_{01}	P_{00}	$P_{0.}$
		$P_{.1}$	$P_{.0}$	

Note. D = Distortion decision, ND = Non-distortion decision.