SELF-DECEPTION, IMPRESSION MANAGEMENT, AND MOTIVATING CONTEXTS: REVISITING THE SUBSTANCE VERSUS STYLE DEBATE

by

LAWRENCE C. MICHELS

(Under the Direction of Gary J. Lautenschlager)

ABSTRACT

Researchers have claimed respondent faking does not adversely affect the criterion related validity of personality measures used for employee selection. Evidence for this claim has been based on studies indicating the criterion related validity of personality scales remains unimproved when scores on unidimensional social desirability scales are controlled. Personality researchers have suggested that social desirability may be better characterized as a multidimensional construct consisting of an impression management and self-deceptive enhancement component. The former dimension is purported to detect deliberate faking, while the latter is purported to detect unconscious distortion. We experimentally elicited faking from respondents by having them complete a personality inventory in either a condition where they believed there was nothing to lose or gain based on their responses, or where they believed that a desirable outcome was predicted on their responses. Results suggest impression management items detect response distortion, whereas self-deceptive enhancement functions as a personality dimension.

INDEX WORDS: Response distortion, Personality, Employee selection, Impression management, Self-deceptive enhancement

SELF-DECEPTION, IMPRESSION MANAGEMENT, AND MOTIVATING CONTEXTS: REVISITING THE SUBSTANCE VERSUS STYLE DEBATE

by

LAWRENCE C. MICHELS

B.S., Kennesaw State University, 2000

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2004

SELF-DECEPTION, IMPRESSION MANAGEMENT, AND MOTIVATING CONTEXTS: REVISITING THE SUBSTANCE VERSUS STYLE DEBATE

by

LAWRENCE C. MICHELS

Major Professor:	Gary J. Lautenschlager
Committee:	Garnett S. Stokes Karl W. Kuhnert

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2004

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Purpose of the Study

Traditional Industrial/Organizational (I/O) approaches to personnel selection were developed for large, stable, vertical, bureaucratic organizations. The modern organization is becoming less structured, with boundaries between jobs constantly shifting (Schmidt, 1993). The increased reliance on autonomous work teams within the current service oriented economy has led organizations to place a premium on contextual performance, social skills, motivation, and related constructs that are commonly presumed to be personality variables (Casio, 1998). This shift in relevant job performance criteria in the modern organization has given rise to a rebirth of interest in personality variables as predictors of job performance (Guion, 1993).

Many authors have referred to an emerging consensus in the literature surrounding the impact of Social Desirability (SD) on the predictive validity of "Big Five" personality inventories used in selection contexts (Smith & Ellingson, 2002; Ellingson, Ones & Viswesvaran, 1998; Barrick & Mount, 1996; Ones, Viswesvaran, Reiss, 1996; and Moorman & Podsakoff, 1992). The preferred stance seems to be that Social Desirability is a benign construct with little potential to adversely influence the predictive validity of personality measures, and that future exploration into the relationship between SD and the predictive validity of personality measures is belaboring a moot point.

Other researchers have refused to nail the proverbial coffin of Social Desirability and "Faking" shut by continuing to question the influence of SD on actual selection outcomes. Mueller-

Hanson, Heggestad, & Thornton (2003) found that individuals who were presumably faking were more likely to be selected and had lower mean performance than individuals who were presumably honest. They also found more error in performance prediction within samples of respondents who were presumably faking. Rosse, Stecher, Miller, & Levin (1998) posit that the deleterious effects of SD may manifest in contexts where applicants are selected on a top-down basis from scores on personality inventory dimensions (primarily Conscientiousness) in which some applicants distort their responses. The subsequent change in rank ordering, particularly in a context with a low selection ratio, can dramatically affect who is hired without significantly altering the predictive validity of the measure. The preservation of predictive validity, or a predictor's ability to distinguish good from poor performers in the population of job applicants, is clearly a central issue in I/O Psychology (Ones & Viswesvaran, 1998). However, to the extent that Human Resource (HR) professionals are concerned with maximizing the utility of their selection systems and identifying the "best" from a pool of applicants, research concerning the relationship between SD and personality is still relevant and needed.

Before consensus can be reached regarding the extent to which social desirability "matters" in a selection context, researchers would do well to first seek consensus regarding several matters essential to resolving the former issue. First, what exactly is "Social Desirability" as a construct in its own right? Second, how is SD related to various "Big Five" personality dimensions? Third, does the context in which an individual's personality is assessed influence the observed relationship between SD and personality dimensions?

The present study attempts to address each of these issues, in turn, by providing theoretical rationale for employing a bifurcated conceptualization of Socially Desirable Responding (SDR) in selection contexts; examining the possible moderating effect of

motivational context on the relationship between SDR and Conscientiousness, the personality dimension most relevant to personnel decision makers; examining the mechanism by which SDR scales "detect" faking in selection contexts; and discussing the implications for using SDR scales to detect, and correct for response distortion on personality measures.

CHAPTER 2

BACKGROUND

Big Five Personality Factor Structure

The discovery of the current dominant personality taxonomy was precipitated by numerous attempts to develop a scientific model of individual differences by addressing two primary issues: First, how to obtain a representative, if not comprehensive, set of such attributes; and second, how to classify or categorize those attributes into a structural model (Goldberg, 1995). Credit for the resolution of the first issue has traditionally been granted to Sir Francis Galton and his "lexical hypothesis." The crux of Galton's hypothesis was that the identification of personality-descriptive terms and the examination of the shared meaning between these terms could provide a means of simplifying individual differences into broad categories.

L.L. Thurstone, considered by some to be the "father of factor analysis," is credited with being the first individual to test Galton's hypothesis by applying data reduction techniques to lexical material. Thurstone administered his 60-item Adjective Check List to 1300 individuals and upon examining the intercorrelations of items, identified five broad factors.

There is on going debate among personality researchers as to whom credit should be given for first identifying the "Big Five." Goldberg (1995) credits Fiske (1949) with the honor of first discovery. Judge and Bono (2000) give the honor to Tupes and Christal (1961). Barrick and Mount (1991) credit Norman (1963) with solidifying the "Big Five" taxonomy and establishing the dimension labels that are currently used today. Despite the lack of consensus regarding their precise origins, personality is nonetheless believed to be comprised of five robust

factors. *Extroversion* refers to the tendency to be sociable, gregarious, assertive, and active. *Emotional Stability (Neuroticism)* refers to the extent to which an individual tends to be anxious, depressed, angry, embarrassed, emotional, worried, and insecure. *Agreeableness* refers to an individual's tendency to be curious, flexible, trusting, good-natured, cooperative, forgiving, soft-hearted, and tolerant. *Conscientiousness* refers to the tendency to be dependable, thorough, responsible, organized, hard working, and achievement-oriented. *Openness to Experience* refers to the tendency to be imaginative, cultured, curious, original, broad-minded, intelligent, and artistic.

Of particular relevance to I/O Psychologists is the extent to which these dimensions of personality are related to job performance. During the past decade, it has been demonstrated that all five factors can predict one or more aspects of job performance (Klimoski, 1993). Extraversion has been found to be a valid predictor of managerial performance whereas Openness to experience has been found to be a valid predictor of training proficiency (Barrick and Mount, 1991). Agreeableness was found to be predictive of transformational leadership (Judge & Bono, 2000).

There have been mixed findings surrounding all but one of the Big Five personality dimensions. Unequivocally, the conscientiousness dimension is the most consistent predictor of job performance criteria across job categories. Aggregating across 117 studies in their landmark meta-analysis, Barrick and Mount (1991) found that Conscientiousness predicted three job performance criteria (job proficiency, training proficiency, and personnel data) for five occupational groups (professionals, police, managers, sales, and skilled/semi-skilled workers; Rho ranges from .20 to .23). Researchers believe that there are two mechanisms by which Conscientiousness affects job performance. First, it improves performance in training programs,

thereby increasing job knowledge, which ultimately leads to better job performance. Second, it affects job performance directly, as conscientious individuals simply are apt to do a better job (Behling, 1998).

It is now generally accepted among HR professionals that non-clinical measures of personality, specifically Conscientiousness, can predict occupational outcomes fairly well while providing incremental predictive validity over traditional cognitive ability tests (Cascio 1998). Perhaps even more importantly, this gain in predictive validity is made without causing adverse impact against protected groups or violating the terms of the Americans with Disabilities Act (Hogan, Hogan, & Roberts, 1996).

Despite the apparent benefits and demonstrated effectiveness of using personality measures in selection contexts, even proponents of their use traditionally reference response distortion as being an inherent limitation. As one authority on the subject states: "This [response distortion] is obviously a problem in all work situations, but especially selection, when subjects know that their careers are at stake. Thus few serious applicants for a sales position would admit to being shy or nervous with people (Kline, 1993)." This unquestioning acceptance of the susceptibility of objective personality measures to faking has ended. Several researchers have provided evidence that "Faking does not matter" (Ones et al., 1996). Before delving into the current debate surrounding the role that response distortion plays in the use of personality measures in the selection context, the supposed source of response distortion, namely, socially desirable responding must first be discussed.

Socially Desirable Responding

Broadly stated, SDR consists of the tendency of individuals to present themselves in a favorable manner with respect to social norms and standards (Zerbe & Paulhus, 1987). The

creators of the most commonly used measure of Social Desirability, the Marlowe-Crowne (MC) scale (Crowne & Marlowe, 1960), conceptualized SDR as a contaminant, or evidence of deliberate "faking." However, they gradually began to revise their conceptualization of SDR suggesting it is a relatively stable personality trait reflecting a need for approval. Since the publication of the Marlowe-Crowne scale, there have been no fewer that 12 additional Social Desirability scales developed (Paulhus, 1984). As these unidimensional SDR scales typically measure the extent to which an individual claims to possess unlikely virtues, or denies possessing likely faults, these scales are frequently referred to as simply Unlikely Virtue scales (UV).

UV scales have been repeatedly treated as measuring equivalent single constructs and have been used interchangeably in the SD literature. Researchers have made policy recommendations based on findings from studies in which meta-analytic techniques were used to aggregate across studies using various SD scales (Ones et al., 1996). Realizing that the unidimensional conceptualization of SD may be somewhat simplistic and limited, Paulhus and his associates developed a two factor model of SDR.

The two factor conceptualization of SDR consists of *Self-Deceptive Enhancement* (SDE) and *Impression Management* (IM). The first instrument developed to measure this bifurcated model of SDR was the Balanced Inventory of Desirable Responding (BIDR) (Paulhus, 1984). The Self-Deception factor is purported to reflect an unconscious tendency to provide inflated, positive self-reports. By unconscious, we mean outside of the awareness of respondents. SDE items consist of statements referring to thoughts and feelings that are judged to be universally true, but psychologically threatening (e.g., "*Worry about what people think of me*"); (Paulhus, 1984). Individuals engaging in SDE actually believe that their responses are indicative of their

actual personality traits and reflect their own conceptualization of themselves (Ellingson et al., 1999).

Impression Management is proposed to represent the conscious presentation of a false front, and an intentional attempt to create a favorable impression. The IM factor is generally thought to reflect deliberate "Faking" and is the aspect of response distortion that unitary conceptualizations of SD have sought to capture (Barrick & Mount, 1996). IM items consist of the extent to which respondents claim to engage in desirable but statistically infrequent behavior (e.g., "*Always admit it when I make a mistake,"* and undesirable but common behavior (e.g., "*Get back at others*"); (Zerbe & Paulhus, 1987). Since the IM items reference specific behaviors rather than thoughts, these items are believed to be minimally subject to Self-Deception.

Researchers have yet to agree on the relative merit of treating SD as a two factor construct. A recent published study using the BIDR impression management scale and the BIDR self-deception scale reported a correlation between the scales of .49 ($p < .01$) (Smith & Ellingson, 2002). This moderate correlation seems to indicate that either the BIDR is a highly unreliable measure of a single underlying SD construct, or that the two-factor model of SDR is a viable alternative to single component models, and that two conceptually distinct SDR facets exist. Given that internal consistency reliabilities for SDE and IM scales are generally fairly high ($\alpha \approx .80$ and .82, respectively), a double correction for attenuation was applied to reveal what the relationship between the two scales would be if both were measured with perfect reliability (Spearman, 1904; Muchinsky, 1996).

$$\frac{r_{xy}}{\sqrt{r_{xx'}}\sqrt{r_{yy'}}} = \frac{.49}{\sqrt{.80}\sqrt{.82}} = .60 \qquad [2\text{-}1]$$

In the present example, we see that the disattenuated correlation between SDE and IM is $\rho = .60$. Because coefficient alpha represents a lower bound estimate of internal consistency

reliability, the dissattenuated population correlation (ρ) is likely an over correction, such that the "true" population correlation is somewhat lower. Therefore, the obtained dissattenuated correlation between SDE and IM is an inflated upper bound estimate of the true population correlation. While significant, it seems that the correlation is low enough to suggest that IM and SDE are distinct constructs and warrant independent examination in future research. The disagreement among researchers regarding the proper conceptualization of SDR will likely continue. Unfortunately, recent research seems to have given up on resolving this issue and has instead shifted its focus onto examining the role that this ill-defined construct plays in the use of objective personality measures.

The Substance vs. Style Debate

For decades, a widely held axiom was that self-report personality inventories were affected by the response style of individuals, and that dishonest responses would compromise the validity of personality measures. Extensive research has established that personality measures are "Fakeable" (Bartlett & Doorley, 1967; Hough, 1990), in that individuals can successfully distort their responses in an instructed manner (e.g., fake good or fake bad). Many techniques have been employed in an attempt to remedy this apparent threat of distortion with varying degrees of success. Use of forced-choice items in which statements are matched according to their level of social desirability and the use of subtle item content to reduce a measures "transparency" have proven to be somewhat useful in reducing distortion (Hough, 1998). In fact, simply warning respondents about the consequences of distorting their responses and telling them that detection methods are in place has been demonstrated to significantly reduce distortion (Schrader & Osburn, 1977). This technique has been deemed effective in reducing faking on biodata questionnaires which by their nature contain items that are verifiable (Lautenschlager,

1994). The effectiveness of a warning has not been sufficiently examined in the context of "Big Five" personality inventories to make a determination as to the extent to which it reduces faking. However, since personality items are generally unverifiable, it seems reasonable that a warning would be, at best, modestly effective at reducing deliberate distortion on personality inventories.

As test developers have become increasingly sophisticated, a popular method for detecting response distortion has been to include a SD or "lie" scale in non-cognitive measures (Ellingson, Smith, & Sackett, 2001). SD scales are used as "lie" scales because of their sensitivity to response distortion. This sensitivity is indicated by virtue of the fact that SD scale scores change more than any other content-oriented personality scale when individuals are instructed to "fake good" on a measure (Ones & Viswesvaran, 1998).

Ganster et al., (1983) proposed three alternative models for measuring the effects of SDR on the predictive validity of non-cognitive measures when using a SDR scale. First, SDR contamination can produce spurious relationships between variables. When SDR is correlated with both a predictor and criterion, any observed predictive relationship may be due to shared variance in SDR rather than shared variance in the construct supposedly being tapped. Second, SDR may function as a suppressor variable by essentially "masking" a true relationship between a predictor and a criterion. When this is the case, controlling for SDR should result in an increase in the observed relationship between the predictor and criterion. Third, SDR may function as a moderator such that the relationship between a predictor and criterion depends on the level of SDR. Ganster's (1983) study was significant, as it outlined the specific mechanism by which the one-factor model of SDR could influence the predictive validity of personality measures, and it demonstrated methods to test for these effects. Additionally, it was the first of

several studies to suggest that SD may be more substantive than previously believed, and that it may be a useful predictor variable in its own right, rather than simply a style of responding.

Several researchers tested Ganster's theory and found that elevated scores on SD scales did not decrease the predictive validity of personality measures (Hough et al., 1990). Ones et al., (1996) went further in their meta-analysis by examining the effect of SD on the predictive validity of personality measures for 1,460 correlations across 409,496 individuals. They found that SD did not function as a suppressor or a mediator, nor did SD appear to cause a spurious relationship between personality dimensions and performance criteria, as it was not significantly correlated with the performance criteria. The authors concluded that SD introduces redundant rather than error variance into the predictive relationship between personality and job performance. They concluded that SD should be viewed as a personality construct related to Conscientiousness and Emotional Stability, rather than evidence of response distortion.

As a follow up study, Ones and Viswesvaran (1998) again demonstrated that respondents did engage in SDR, as they did show elevated scores on SD scales. But again, SDR did not attenuate the predictive validity between personality measures and performance. The authors found that SD appeared to be a substantive construct related to Emotional Stability and Conscientiousness with estimated population correlations of .18 and .13, respectively. It was, however, not useful for predicting job performance by itself. The authors provide an interesting possible explanation for their counterintuitive findings. They report that SDR has essentially no bearing on the predictive validity of personality measures because an individual who decides to create an identity for themselves by distorting their responses will also tend to behave in a manner consistent with that identity (e.g., such as engaging in hard-working and dependable behaviors). Thus, they conclude: "selecting individuals who respond to items on a personality

scale so as to project an identity of a Conscientiousness individual is sensible (Ones et al., 1998)."

Both of the Ones et al. studies (1996, 1998) were very direct and vehement in their position that SDR and faking do not matter in prediction for personnel selection. Despite, or perhaps because of, the fervor with which they declare "Faking" to be a non-issue, there have been some criticisms of methodological aspects of their studies. First, the Ones & Viswesvaran (1996) meta-analysis used a rather loosely knit sample that consisted of an aggregation of studies that employed over 17 different measures of SD (all of which were unidimensional conceptualizations of SD). Using single factor conceptualizations to make statements regarding the futility of attempting to remove the effects of response distortion from personality measures may have been misleading. If the two factor conceptualization of SD is accurate, it stands to reason that only the IM factor would affect the predictive validity of personality scales, since it is the factor believed to capture deliberate faking. Therefore, before discounting the merit of detecting and removing SD from personality scale scores, the authors could attempt to examine the effect of IM and SDE separately.

Second, they did not differentiate between the instruction sets used in each study. It has been demonstrated that the instruction set under which SDR occurs influences the degree of SDR respondents engage in (Smith & Ellingson, 2002). Some of the research designs included in the Ones et al. (1996) meta-analysis employed "fake good" instruction sets in which participants were instructed to present themselves in the most favorable light, while other studies had "honest" instruction sets in which participants were instructed to respond honestly. Even these "honest" instruction sets can be further divided into studies in which respondents had nothing to be lost or gained as a function of scores on the measure and studies that created a motivational

context in which participants perceived that there was something to be gained by presenting themselves in a positive light, or something to be lost by presenting themselves negatively.

It seems plausible that the SDR occurring in contexts with "fake good" instruction sets would be quite different both in terms of magnitude and in relation to other constructs, than SDR occurring in contexts with "honest" instructions absent any motivating context. The former condition seems to tap an individual's "ability" to distort, while the latter taps an individual's "propensity" to distort. Until the effect of motivational context is examined further by assessing the relationship between SDR and personality dimensions in contexts which are equivalent to those encountered in practice (e.g., "honest" instruction sets with considerable reason to distort), researchers should not accept the notion that SDR simply does not "matter."

SDR and Hiring Decisions

With the accumulation of research indicating that SD does not lead to the degradation of predictive validity of personality measures, several researchers have instead turned to the effect of SD on actual hiring decisions (Christiansen et al., 1994; Rosse et al., 1998). They point out that SD, whether it is defined as response distortion or redundant personality variance, can have a dramatic effect on who is hired without having a detectable effect on predictive validity. Essentially, the composition of the upper limits of the distribution of personality scores changes as a function of increased SDR, resulting in a change in rank ordering of applicants. Christiansen et al. (1994) used a sample of 495 respondents and found that individuals tended to shift an average of 23 (Standard Deviation = 36) rank order positions when personality scores were corrected for SD.

Depending on the selection ratio used, this shift and subsequent change in rank ordering resulted in discrepant hires (i.e., individuals who were not hired on the basis of their uncorrected

scores, but who would have been hired had personality scores been corrected for SD contamination). The magnitude of this discrepancy was only 0.4 % in high selection ratios (.95) but was considerably higher in lower selection ratios (16% and 12% for selection ratios of .15 and .05 respectively). This finding suggests that the job performance of individuals scoring lower on SD scales may be under-predicted by their scores on uncorrected personality scales. Perhaps a greater issue identified by the authors involves the potential for legal implications when an "honest" person loses a job to an applicant who "fakes" responses. The authors make a case that correcting personality scores for SD may increase the utility of using personality measures in a selection context particularly when used in conjunction with a low selection ratio.

Correction Techniques

Several correction procedures have been identified as having some merit for obtaining "true" personality scores from respondents who, presumably, are motivated to distort their responses. Hough et al. (1998) described two techniques that resulted in significant changes in hiring outcomes. The first technique involves obtaining a sample of SD scores given by job incumbents. The use of job incumbents as a sample population for obtaining distributions of "true" SD scores absent response distortion has been a commonly employed technique in the SD literature (Rosse et al., 1998). Incumbents are believed to have little motivation to distort their responses on personality measures, as it is reasoned that they have nothing to gain from presenting themselves in a positive light. Therefore, it is argued that their obtained SD scores represent a component of SD that reflects something other than "distortion," such as redundant personality variance. There seem to be several issues inherent in presuming that incumbent and applicant samples are equivalent in all respects other than their motivation to distort. These issues are addressed in the current study.

Hough's correction procedure recommends determining the mean and standard deviation for SD scores for the incumbent sample, and making the following adjustments to applicant scores. First, if an applicant scores three or more standard deviations higher than the incumbent SD scale mean, that applicant's score on the content scale (Conscientiousness) is reduced by an equivalent of two standard deviations (based on incumbent scores). Second, if an applicant scores more than two, but less than three standard deviations above the incumbent SD scale mean, that applicant's score on the content scale is reduced by an equivalent of one standard deviation on the content scale. Another possibility is to simply identify a SD scale "cut score," above which an applicant will be excluded from the selection proceedings. Hough recommends using a 5% cut score that refers to the SD score which only 5% of the presumably "honest" incumbent sample score at or above. Either of these techniques is effective in changing the rank ordering of applicants and thus hiring decisions as well as eliminating the effects of extreme distortion. However, the extent to which, if any, these techniques improve the predictive validity of personality measures has not been determined. Nor has it been unequivocally demonstrated that these techniques are useful for removing applicants who are faking.

Another approach for generating corrected personality scores free from the effects of SD is the partialling approach. This approach consists of removing the variance in each personality scale that can be accounted for by SD scales. The process consists of residualizing obtained personality scale scores on SD scale scores. The resultant residual value is used as a corrected score reflecting the portion of Conscientiousness not shared with SD. This procedure results in a semipartial correlation between the Conscientiousness scale scores and the performance criteria in the absence of SDR.

Ultimately, while we know that correcting does change the rank ordering of personality scores in a distribution, none of these correction procedures have demonstrated their ability to increase the predictive validity of personality scale scores. Several researchers have actually advised against using these approaches to correct for SD because they believe that full partialling of SD from personality may remove the substantive variance in personality scale scores shared with SD (Costa & McCrae, 1997). While the ideal technique for correcting personality scale scores for response distortion has yet to be discovered, it seems the goal should be one in which variance attributable to deliberate "faking" is removed without "overcorrecting" and removing relevant personality information. If such a technique is to be developed, it seems that Paulhus' fifteen year old model has the potential to enable the researcher to accomplish just such an objective.

CHAPTER 3

THE PRESENT STUDY

Hypotheses

Given the lack of agreed upon operationalizations of constructs such as Social Desirability, Faking, and Response Distortion, it seems necessary to take a step back from the effect of SD on the predictive validity of personality scales, and to instead focus on clarifying SD as a construct and identifying its relationship to relevant personality dimensions. In an effort to determine the nature of SDR and definitively state whether it is "style" or "substance," it is necessary to study the phenomenon in a setting in which it could be expected to naturally occur. Using an experimental setting will allow for comparison of respondents who will be motivated to present themselves in the best possible light to a comparison group comprised of similar individuals without such motivation.

While until recently little attention had been paid to motivational contexts, a recent study examined the effect of motivating context on SDR by comparing applicant groups to incumbent groups (Rosse et al., 1998). It was assumed that applicants would feel pressure to engage in response distortion characterized by deliberate faking or Impression Management (IM) while incumbents would feel no such pressure. There are several problems with using this paradigm. First, it may not be appropriate to assume that job incumbents are an impartial group who are not motivated to present themselves in a positive light. Individuals may perceive that their job is in danger and distort their responses to present themselves in a positive light. Hence, treating scores obtained from incumbent samples as an "honest" baseline may be inappropriate. Second,

by virtue of the fact that incumbents are, in fact, incumbents, it stands to reason that their distribution of scores on relevant personality dimensions would be negatively skewed, and thus inappropriate to use as an "honest" comparison group. It seems reasonable that they would possess higher levels of Conscientiousness by virtue of the fact that they were hired and still presumably effective employees.

To properly assess the effect of motivating contexts on SDR, it is necessary to compare mean scores between groups of individuals equivalent in all respects except on the motivational context to which they are exposed. To accomplish this level of control, it is essential to utilize the lab environment where proper control can be exercised. Motivational context from here on will be referred to in terms of degrees of "IM pressure." High IM pressure refers to a condition in which a respondent perceives that there is something to be lost or gained as a function of their responses on a measure. Low IM pressure will refer to a condition in which a respondent perceives that there is nothing to be lost or gained as a function of their scores on a measure.

Based on prior research indicating that individuals can, and do, distort their responses when given adequate reason to, and using the two factor model of SD, the following hypotheses will be tested:

H1a: Respondents subjected to high levels of IM pressure will yield significantly higher Conscientiousness scale scores than respondents subjected to low levels of IM pressure.

This hypothesis serves as a manipulation check, of sorts, as the successful manipulation of "High IM pressure" will be evidenced by inflated, or positively biased self reports. Because of the previously observed relationship between Conscientiousness and SDE, we expect SDE to be somewhat effected by IM pressure such that:

H1b: Respondents subjected to high levels of IM pressure will yield significantly higher Self-Deceptive Enhancement scale scores than respondents subjected to low levels of IM pressure.

It must be noted that this hypothesized increase in SDE is believed to be a manifestation of respondents positively endorsing items throughout the entire scale. Or, to be more specific, the increase is due to respondents engaging in impression management and endorsing items in such a way so as to maximize scores on socially desirable dimensions (e.g., Conscientiousness and Agreeableness) and minimize scores on socially undesirable dimensions (e.g., Neuroticism). It does not, however, suggest that the underlying construct, Self-Deception, is actually increasing in the presence of IM pressure. If that were the case, it would be tantamount to saying that an individual's conscious desire to positively self present is influencing their unconscious tendency to self deceive. This is a proposition that neither Freud, Jung, Paulhus, nor the current authors would endorse. Again, because SDE items refer to cognitions rather than overt behaviors it has been proposed that SDE items detect an unconscious tendency to present oneself in a positive light, and not deliberate distortion. Therefore in the High IM pressure condition where respondents are presumably engaging in deliberate distortion, we expect the SDE scale to be elevated along with the substantive personality scale Conscientiousness, as both scales contain items with socially desirable content. We contend that the elevation in the SDE scale is simply evidence that the IM pressure manipulation is, in fact, inducing respondents to inflate their responses.

An important element in the present study is the extent to which IM scales detect deliberate faking. When provided with the impetus to engage in deliberate self enhancement, we expect that individuals will engage in deliberate distortion, and this distortion will be "detected" by the IM scale such that:

H1c: Respondents subjected to high levels of IM pressure will yield significantly higher Impression Management scale scores than respondents subjected to low levels of IM pressure.

To determine the extent to and manner in which the two factor conceptualization of SD is useful for detecting response distortion it is necessary to examine the two components in the presence of IM pressure rather than the induced "fake good" conditions in which they have previously been examined. It seems plausible that if Self-Deceptive Enhancement is, in fact, an unconscious tendency to have an overly positive view of oneself, then it's relationship to other relevant personality dimensions should be unaffected by motivational context. Therefore, although we expect mean differences in SDE between groups subjected to High vs. Low "IM pressure," we predict that the linear relationship between SDE and Conscientiousness will remain the same across both the High and Low IM pressure conditions.

To clarify, it seems plausible that individuals who are deliberately distorting their responses may endorse all items in an overly favorable manner. Consequently, as they inflate their scores on Conscientiousness scales, they will also inflate their scores on SDE scales because the items in both scales are relatively transparent and susceptible to distorting. This inflation should happen in tandem such that the linear relationship between SDE and Conscientiousness remains intact, while the mean scores of each scale increase due to the respondent engaging in IM. Therefore, we hypothesize:

H2a: Level of "IM pressure" should not affect the linear relationship between Self-Deceptive Enhancement scale scores and Conscientiousness scale scores.

Unlike SDE, if IM scales detect deliberate "Faking," we would expect the relationship between IM scale scores and Conscientiousness scale scores to change as a function of IM pressure. In the Low IM pressure condition, individuals respond to items in a context in which

they do not feel they have anything on the line. We expect that the distortion occurring in this Low IM pressure condition will be trait related and detected primarily by SDE scales. When confronted with High IM pressure we expect the ensuing deliberate distortion to be detected by the IM scale. We believe that respondents will positively endorse items on the IM scale as they artificially inflate all other items in the measure. Therefore, the following hypothesis will be tested:

H2b: Level of IM pressure will moderate the relationship between IM scale scores and Conscientiousness scale scores such that there will be a weak relation between IM and Conscientiousness in the low IM pressure condition, and a strong positive relation between IM and Conscientiousness in the high IM pressure condition, where respondents believe there is something to be gained by presenting themselves in a positive light.

It seems plausible that previous research which characterized SD as a substantive personality dimension related to Conscientiousness and Neuroticism based these assertions on relationships observed between personality and Self-Deceptive Enhancement exclusively (not Impression Management). However, since the preponderance of SDR research over the last decade used unidimensional conceptualizations of SD that used items referencing both cognitions as well as overt behaviors (for exception see Barrick & Mount, 1996), the exclusive nature of this relationship may have gone undetected.

Prior research has consistently indicated that there is a measurable relation between SD scales and Personality scales, as evidenced by small to moderate correlations. It seems plausible that these previous studies have, in fact, observed two separate relations between personality and SD scales. In contexts absent IM pressure, the relation may be between personality and SDE, while in the presence of IM pressure, the relation may be between personality and IM. Because

most researchers have used a unidimensional conceptualization of SD, which consisted of items referencing both cognitions and overt behaviors, they may have been unable to tease out the source of the covariation and thus not adequately resolve the "substance" versus "style" debate. We propose that the "substance" of SD will be detected by SDE scales while the "style" will detected by IM scales. To confirm this theory, the following hypothesis will be tested:

H3: Self-Deceptive Enhancement scale scores will account for significant variance in Conscientiousness scale scores above and beyond that accounted for by Impression Management for respondents subjected to low levels of IM pressure.

If this is the case, it stands to reason also that if IM is, in fact, akin to deliberate faking, IM scale scores should be related to relevant personality scale scores in the presence of IM pressure when respondents could reasonably be expected to fake. Therefore, the following hypothesis will be tested:

H4: Impression management scale scores will account for significant variance in Conscientiousness scale scores above and beyond that accounted for by Self-Deceptive Enhancement for respondents subjected to high levels of IM pressure.

Summary of Theoretical Rationale

To summarize, researchers have proposed that "Social Desirability," as measured by unidimensional SD scales, is a substantive personality dimension rather than a response style that introduces systematic upward bias in other personality scales. They have taken this position because they observed that the criterion related validity between Conscientiousness and measures of job performance generally does not improve when scores on Social Desirability scales are controlled. These assertions have been made based on studies employing Social Desirability scales that included items that reference both behaviors as well as cognitions. Within

the personality literature, it has been proposed that items referencing cognitions detect an unconscious tendency to positively self-present (SDE), while items referencing behaviors detect deliberate distortion which is the type of distortion that constitutes faking (IM). Attempting to remove the effects of faking by controlling for Social Desirability as measured by unidimensional scales that include both types of items may fail to isolate "Faking," which is purportedly detected only in items referencing specific behaviors. Furthermore, because SDE may be substantively related to personality, controlling for it may remove trait relevant variance. If this is true, then the effects of faking can not be assessed by controlling for Social Desirability as measured by unidimensional scales, because doing so will remove both trait relevant (SDE) as well as contaminant variance (IM).

The goal of this experiment is to determine if items referencing behaviors do, in fact, tend to detect Faking while items referencing cognitions do not. Unlike much of the previous research on faking, we are attempting to elicit faking in the form and magnitude with which it would be expected to occur in an evaluative context such as employee selection. This methodology is in sharp contrast to previous studies that have employed "fake good" instruction sets. Furthermore, we will use a comparison group that is identical to the experimental group in all respects except for the degree to which they feel pressure to distort their responses. This aspect is in contrast to previous studies that have employed control groups whose degree of similarity to the experimental group is questionable (e.g., applicants vs. incumbents).

We predict that individuals will engage in increased levels of deliberate distortion when they are being evaluated as evidenced by significantly higher scores on three personality dimensions (hypotheses 1a, 1b, and 1c). However, we hope to illustrate that the relationship between Conscientiousness and SD items referencing cognitions (SDE) will remain constant

irrespective of whether individuals are distorting their responses (hypotheses 2a). This hypothesis will be supported if the relationship between Conscientiousness and SDE is not moderated by IM pressure. Confirmation of this hypothesis will lend support to the notion that SDE is a substantive personality dimension related to Conscientiousness.

Conversely, we hope to illustrate that the relationship between Conscientiousness and SD items referencing behaviors (IM) will change as a function of IM pressure (hypothesis 2b). The logic behind this hypothesis is that Conscientiousness scores obtained from respondents in the high IM pressure condition will reflect "true" Conscientiousness as well as some systematic upward bias, or faking. This latter component of the Conscientiousness score should covary with the IM scales that are purported to detect faking and this covariation should be greater in the condition where respondents presumably are faking.

Additionally, we predict that the relative strength of the relationship between Conscientiousness and each of the SD dimensions will change depending on the degree to which respondents are faking (hypotheses 3 and 4). When respondents are engaging in minimal faking, we predict that SDE items will demonstrate a stronger relationship with Conscientiousness. When respondents are faking, we predict that IM items will demonstrate a stronger relationship with Conscientiousness. Collectively, the tests of these hypotheses will serve to clarify the true nature of SD as being a construct that reflects both "substance" and "style;" and that the focus of the item content determines which aspect is tapped.

CHAPTER 4

METHOD

Participants

Two hundred ninety undergraduate students from a large southeastern university participated in the current study (147 males, 138 females, 5 unspecified). The mean participant age was 19.90 (sd = 2.85). Participants volunteered for the study and received research credit used towards completing a course requirement. Participants were told that the experiment was entitled: "Personality Focused advertising." Power analysis was conducted to determine the necessary sample size for the current study (Cohen, 1988; Green, 1991). The *a priori* alpha level for all analyses was set at $\alpha = .05$. Due to the sparse amount of research regarding the effect of motivational context on the relationship between SD and Conscientiousness, a conservative effect size, $f^2 = .05$ was chosen (Cohen, 1988). Power was set at a value of .90 for all N size calculations. This sample size was determined to be sufficient to enable the correct rejection of false null hypotheses without being so high as to detect practically meaningless effects. Subsequent power analyses using the effect sizes calculated from the observed R^2s revealed that power was at least .95 for all analyses.

Inducing IM Pressure

One hundred sixty-five participants (87 males, 73 females, 5 unspecified) were in the Low IM Pressure condition. To induce Low IM pressure, respondents were led to believe that their responses were anonymous and that they would be used for research purposes only. They were specifically told to respond honestly and assured that their responses would not affect them

in any way. We believe that elevations in SDR occurring in this Low IM pressure condition represent the extent to which individuals engage in SDR in the absence of motivation to distort their responses. Furthermore we believe that the SDR occurring in this Low IM pressure condition will be primarily of the SDE variety, and that the level of SDE will be positively related to Conscientiousness.

One hundred twenty five (60 males and 65 females) were in the high IM pressure condition. In developing a stimulus suitable for inducing High IM pressure, several factors were taken into account. Unlike previous studies which asked participants to role play and pretend they were job applicants applying for a job, we created a selection environment germane to this particular sample. Deception was used to make participants believe that they were in a selection process for what, to them, would seem to be a highly desirable position.

A scenario describing a highly desirable research opportunity was printed on a graphic flyer and distributed to participants. The flyer stated that a pseudo-research organization was looking for 30 individuals who were dependable, hard working, and achievement striving (see Appendix A). All of the descriptors of the ideal candidate were the facet level descriptors for the conscientiousness personality dimension. This was done to give respondents a target personality dimension for them to fake "towards." It must be noted, however, that none of these descriptor words (e.g. hard-working, dependable, achievement striving) appeared in the actual survey items.

Individuals filled out a certificate that included their contact information and were led to believe that this certificate would be redeemable for cash and extra research credit if they were accepted into the pseudo-experiment. Individuals in the High IM pressure condition were also presented with a brief slide show providing additional details of the pseudo-experiment. The pseudo-selection instrument was identical for participants in both "IM pressure" conditions. The

instructions, however, differed slightly between the high (Appendix B) and low (Appendix C) "IM pressure" conditions.

Experimental sessions were scheduled and the assigned level of "IM pressure" for each session was determined randomly. Participants were tested in the experimental sessions in groups of 40-60. Participants were seated next to each other in an auditorium in a setting similar to that in which large numbers of applicants are tested simultaneously. Prior research has indicated that the number of people present when completing a non-cognitive measure may influence the degree of SDR in which individuals engage (Richman, Kiesler, Weisband, & Drasgow, 1999). However, no research has identified the effect of group size on individual SDR specifically in the presence of "IM pressure." Therefore, the choice of our group size stems from an attempt to provide ecological validity for the study by creating an environment that approximates a "real world" selection context as might be encountered in a large group testing situation such as testing for Military occupations.

Measures

A 130 item measure (see Appendix D) was developed using scales obtained from the International Personality Item Pool (IPIP); (Goldberg, 1999). The IPIP is a database of nonproprietary psychological instruments with published reliability and validity indices. Conscientiousness was assessed using a 20 item scale with internal consistency reliabilities of $\alpha = .90$ in the Low IM pressure condition and $\alpha = .93$ in the High IM pressure condition. (See Tables 1 and 2 for Descriptive statistics, intercorrelations among variables measured in the study, and obtained reliabilities for the Low and High IM pressure conditions). Respondents were instructed to respond to items on an optical scan sheet using a 5-point scale to indicate the extent to which statements describe them (A = *very inaccurate,* B = *inaccurate,* C = *neither accurate*

nor inaccurate, D = *accurate*, E = *very accurate*). Sample items included "*Love order and regularity"* and "*Shirk my duties."* (International Personality Item Pool, 2001).

Self-Deceptive Enhancement was assessed using a 10 item scale with internal consistency reliability of $\alpha = .79$ in the Low IM pressure condition and $\alpha = .80$ in the High IM pressure condition. Respondents responded using a five point, *very inaccurate-very accurate* scale indicating the extent to which statements describe them. Again, SDE items include statements that reflect cognitions that have been determined to be psychologically threatening but statistically common. Sample items include *"Feel comfortable with myself"* and *"Worry about what people think of me."* (International Personality Item Pool, 2001).

Impression Management was assessed using a 20 item scale with internal consistency reliability of $\alpha = .80$ in the Low IM pressure condition and $\alpha = .83$ in the High IM pressure condition. Respondents responded using a five-point, *very inaccurate-very accurate* scale indicating the extent to which statements describe them. Again, IM items include statements that reference the extent to which individuals claim to engage in socially desirable overt behaviors. Sample items include *"Would never take things that aren't mine"* and *"Use flattery to get ahead."* (International Personality Item Pool, 2001).

Procedure

Participants in the "Low IM Pressure" condition were seated in the experimental room and given the inventory. For the "Low IM Pressure" condition, our intention was to ensure that respondents believed their responses had no bearing on any future decision that will affect them in anyway. By creating a condition in which we could reasonably expect low incidence of deliberate positive self-presentation to occur, we hoped to establish an IM baseline. Participants were advised that the results of their inventory would remain confidential; be used for research

purposes only; and not affect them in any way. Since participants in the Low IM pressure condition were given little reason to distort their responses, we expect that the distortion that does occur will primarily be of the SDE variety. Once respondents completed the inventory, they were debriefed (Appendix F) and given research credit.

Participants in the "High IM Pressure" condition were given the flyer describing a pseudo-research opportunity (Appendix A). They were told that the inventory was a selection device being used to determine who would be chosen to participate in the pseudo-experiment. To insure that individuals found the pseudo-experiment desirable, participants were given the opportunity to self-select out of the selection process. Individuals who were not interested in "applying" for the pseudo-experiment were asked to sit on the other side of the room, and told that they would be given a different inventory to be used for research purposes only. However, all participants in the High IM pressure condition elected to take what they were told was the "selection" instrument. It was assumed that individuals who chose to remain as "applicants" believed that a desirable outcome was predicated on the results of the inventory, and were therefore motivated to present themselves in a positive light to a greater degree than individuals in the non-evaluative, Low IM pressure condition. To make the manipulation more salient for respondents, they filled out their contact information on the bottom of the pseudo selection form. They were advised that after careful review of their application survey, an experimenters' signature would be required to acknowledge their successful screening thus allowing them to redeem their reward.

Once these individuals completed the inventory, they completed a biographical information form (Appendix E). The Biographical form contained three distracter items and five items that measured the extent to which respondents found the pseudo research opportunity

attractive. The internal consistency reliability for this five item "attractiveness composite" was high ($\alpha = .88$). Respondents responded using a five point, *very inaccurate-very accurate* scale indicating the extent to which statements describe them. Sample items include "*The [Pseudo Experiment] seems like a great program to participate in."* and *"I would like to be selected to participate in the [Pseudo Experiment]."* Although no specific hypotheses were made regarding these data, they served as an additional manipulation check and as the focus of post hoc analyses. It seems plausible that within the High IM pressure condition there may be variability among individuals regarding the extent to which they viewed the "pseudo experiment" and proposed compensation as being desirable. Respondents were then debriefed and advised of the true nature of the experiment (Appendix G). They were asked not to divulge the nature of the deception to their classmates.

Participants from both conditions received a final debrief form (Appendix H) in which they were given a link to an on-line personality inventory as well as the contact information for the principal investigator. Those interested were able to take the on-line inventory and receive an interpretive report.

CHAPTER 5

RESULTS

Tests of Hypotheses

Means, Standard Deviations, and intercorrelations among the Big Five dimensions, Impression Management, and Self-Deceptive Enhancement for the Low and High IM pressure conditions are presented in Tables 1 and 2 respectively. The intercorrelations among the "Big-Five" personality dimensions are nearly identical to established mean inter-factor correlations across multiple measures of the "Big Five" (John & Srivastava, 1999) for those data from the Low IM pressure condition. This is encouraging, as it supports the notion that the IPIP version of the "Big Five" inventory demonstrated psychometric properties nearly identical to other widely used measures. It was interesting to note that the intercorrelations among the dimensions from the High IM pressure condition were consistently stronger. This observation may indicate changes in the underlying factor structure of the Big Five when respondents fake (Schmit & Ryan, 1993).

The "IM pressure" variable was dummy coded such that individuals in the "Low IM pressure" condition were coded as "0," and individuals in the "High IM pressure" condition were coded as "1." Additive subscale composite scores were formed for IM, SDE, and each of the "Big Five" personality dimensions. Product terms were computed by multiplying the IM and SDE subscale scores by the respective values contained in the dichotomized "IM pressure" vector.

All hypotheses were tested using regression analysis and the *a priori* alpha level was set at $\alpha = .05$ (two-tailed) for rejection of the null hypotheses (see Table 3 for a summary of the regression analyses for tests of *a priori* hypotheses). Because of the coding scheme employed, the unstandardized regression coefficient associated with the "IM Pressure" variable is equivalent to the mean difference in observed scores across High and Low IM pressure conditions for the personality dimensions examined in the tests of hypotheses 1a, 1b, and 1c. Hypothesis 1a was tested by regressing the Conscientiousness variable on the IM Pressure variable. The resultant t-value from the test of the regression coefficient associated with IM pressure was statistically significant, thus providing support for hypothesis 1a ($b = 11.68$, $t(289) = 8.22$, $p = .000$).

Hypothesis 1b was tested by regressing the Self-Deception variable on the IM pressure variable. The resultant t-value from the test of the regression coefficient associated with IM pressure was statistically significant, thus providing support for hypothesis 1b ($b = 4.957$, $t(287) = 6.744$, $p = .000$).

Hypothesis 1c was tested by regressing the Impression Management variable on the IM pressure condition variable. The resultant t-value from the test of the regression coefficient associated with IM Pressure was statistically significant, thus providing support for hypothesis 1c ($b = 10.194$, $t(286) = 8.333$, $p = .000$).

Correlations among the variables examined in the tests of hypotheses 2a and 2b appear in tables 4 and 5 respectively. Hypothesis 2a was tested by regressing Conscientiousness on SDE, IM Pressure, and the SDE x IM pressure product term. The resultant t-value from the test of the regression coefficient associated with the product term was not statistically significant ($b = .352$,

t(285) = 1.863, p = .064); suggesting that the relationship between SDE and Conscientiousness is unaffected (not moderated) by IM pressure.

Additionally, we tested the difference between the observed bivariate correlations between Conscientiousness and SDE across the high and low IM pressure conditions. This was accomplished by using the Fisher *z*' transformation and testing the null hypothesis that the difference between correlations is zero (Cohen, Cohen, West, & Aiken, 2003). The resultant z-value from the test of the difference between correlations was not statistically significant (z = 1.68, p = .093; 95% confidence interval of the difference between correlations = -.06<.13<.32). The results of both of these tests support the notion that the relationship between Conscientiousness and SDE remains relatively stable irrespective of whether respondents are distorting their responses.

Hypothesis 2b was tested by regressing Conscientiousness on IM, "IM pressure," and the IM x IM pressure product term. The resultant t value from the test of the regression coefficient associated with the product term was statistically significant (b = .270, t(284) = 2.322, p = .021) providing support for the hypothesis that IM pressure moderates the relationship between IM scale scores and conscientiousness scale scores.

Again, we tested the difference between the observed bivariate correlations between Conscientiousness and SDE across the high and low IM pressure conditions using the Fisher *z*' transformation. The resultant z-value from the test of the difference between correlations was statistically significant (z = 2.667, p = .008; 95% confidence interval of the difference between correlations = .03<.22<.41). The results of these tests support the notion that the relationship between Conscientiousness and IM becomes somewhat stronger when respondents are distorting their responses.

Because SDE was measured using a 10 item scale, and IM measured using a 20 item scale, standardized regression coefficients were used to evaluate hypotheses 3 and 4. Hypothesis 3 was tested by regressing the Conscientiousness variable on the SDE and IM variables simultaneously using only scores obtained in the Low IM pressure condition. The standardized regression coefficient associated with SDE was both statistically significant and higher than the coefficient associated with IM (β_{SDE} = .447, t(159) = 6.229, p = .000; β_{IM} = .329, t(159) = 5.112, p = .000). While the results of this test provided apparent support for hypothesis 3, we conducted a test of whether the observed difference in standardized regression coefficients was statistically significant. This was accomplished by first inverting the correlation matrix between IM and SDE. The values contained in the inverted matrix were then entered into the following equation to obtain the standard error of the difference between the standardized regression coefficients (Cohen et al., 2003):

$$SE\beta_1 - \beta_2 = \sqrt{\frac{1 - R_Y^2}{n - k - 1}(r^{ii} + r^{jj} + 2r^{ij})} \qquad [5\text{-}1]$$

Where r^{ii} corresponds to the value appearing in ith row and ith column of the inverse matrix. The test of the difference between the two standardized regression coefficients was performed by dividing the observed difference between these coefficients by the standard error of the difference obtained in equation 2.

$$t = \frac{\beta i - \beta_j}{SE\beta_i - \beta_j} \qquad [5\text{-}2]$$

The resultant t value, t(159) = -1.487 was not statistically significant at the $p < .05$ level.

Hypothesis 4 was tested by regressing the Conscientiousness variable on the Self-Deception and Impression management variables simultaneously using only scores obtained in the High IM pressure condition. The standardized regression coefficient associated with

Impression Management was both statistically significant and slightly higher than the coefficient associated with Self-Deception (β_{IM} = .433, t(122) = 6.229, p = .000; β_{SDE} = .432, t(122) = 6.210, p = .000) providing some apparent support for hypothesis 4. However, when the above test was performed, the obtained t value associated with the test of this difference, t(122) = .0167, was not statistically significant.

Post Hoc Analyses

Additional analyses were conducted to determine the extent to which IM, SDE, and Conscientiousness scores obtained in the high IM condition could be predicted from the "attractiveness composite" (ATTR) scores. It seems reasonable that there would be individual differences in the extent to which respondents viewed both the "Pseudo-Experiment" and the compensation they believed was predicated on their being selected into said experiment, as attractive prospects. Furthermore, it seems plausible that individuals viewing the outcome on which their responses were predicated as highly attractive, would be apt to engage in increased levels of response distortion and that this distortion would be detected in scores on IM scales. The analyses of the ATTR scale proceeded by first regressing the Conscientiousness variable on the ATTR composite (see Table 6). The resultant t-value from the test of the regression coefficient associated with the ATTR composite was statistically significant (b = .750, t(121) = 3.070, p = .003) thus supporting the notion that individuals viewing the outcome that their responses would determine as being desirable, tended to respond in a manner that suggested they possessed high levels of Conscientiousness.

Second, the IM variable was regressed on the ATTR composite. The resultant t-value from the test of the regression coefficient associated with the ATTR composite was statistically significant (b = .825, t(121) = 4.068, p = .000) thus suggesting that individuals viewing the

outcome that their responses would determine as being desirable, tended to engage in higher levels of Impression Management.

Third, the SDE variable was regressed on the ATTR composite. The resultant t-value from the test of the regression coefficient associated with the ATTR composite was statistically significant (b = .250 t(121) = 2.051, p = .042) thus suggesting that individuals viewing the outcome that their responses would determine as being desirable, tended to engage in higher levels of Self-Deceptive Enhancement.

Finally, the ATTR composite was regressed on both the IM and SDE variables simultaneously. In this multivariate analysis, the t-value from the test of the regression coefficient associated with SDE was not statistically significant (b_{SDE} = .005, t(120) = .074, p = .941) while the t-value from the test of the regression coefficient associated with IM was statistically significant (b_{IM} = .144, t(120) = 3.440, p = .001). This finding is particularly interesting as it provides support for the notion that individuals who found the outcome attractive engaged in deliberate response distortion that was detected in their responses to the IM scale. While the perceived attractiveness of the test outcome was also related to SDR in the bivariate analysis, the effect diminished when IM was controlled. Ostensibly, individuals who are likely distorting their responses exhibited higher IM scores, but not higher SDE scores when controlling for IM. To the extent that individuals with a vested interest in the test outcome did, in fact, engage in increased levels of response distortion, this finding is consistent with the notion that the IM scale detected deliberate response distortion while the SDE scale did not detect response distortion above and beyond IM.

Table 1

Correlations Between Sex, "Big Five" Personality Dimensions, Impression Management, and Self-Deception for Individuals in the ***Non-Evaluative*** *"Low IM Pressure" Condition*

Variable	*n*	*M*	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1. Sex	152	.47[a]	.50	--							
2. Impression Management	152	63.09	10.24	.105	(.80)						
3. Self-Deception	152	33.67	6.40	-.285*	**.237***	(.79)					
4. Conscientiousness	152	68.60	11.80	-.015	**.429***	**.526***	(.90)				
5. Extraversion	152	73.20	14.22	.140	-.076	.352*	.204*	(.93)			
6. Agreeableness	152	70.00	9.10	.186*	.627*	.214*	.287*	.140	(.80)		
7. Openness to Experience	152	74.36	10.72	.186*	.008	.116	.054	.200*	.109	(.84)	
8. Neuroticism	152	52.70	12.94	.250*	-.275*	-.733*	-.368*	-.337*	-.351*	-.084	(.90)

Note. [a] The 'Sex' variable was dummy coded as 0=Male and 1=Female. *$p < .05$ two tailed. Values contained in parentheses refer to the internal consistency (α) reliabilities of each scale. Values relevant for hypothesized relations appear in bold.

Table 2

Correlations Between Sex, "Big Five" Personality Dimensions, Impression Management, and Self-Deception for Individuals in the ***Evaluative*** *"High IM Pressure" Condition*

Variable	*n*	*M*	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1. Sex	121	.54[a]	.50	--							
2. Impression Management	121	73.26	10.53	.211*	(.83)						
3. Self-Deception	121	38.56	6.02	-.126	**.527***	(.80)					
4. Conscientiousness	121	80.28	12.35	.085	**.662***	**.669***	(.93)				
5. Extraversion	121	78.33	13.77	.070	.373*	.545*	.530*	(.93)			
6. Agreeableness	121	76.25	10.74	.362*	.685*	.181*	.439*	.388*	(.86)		
7. Openness to Experience	121	78.80	11.25	.049	.231*	.301*	.302*	.290*	.194*	(.87)	
8. Neuroticism	121	44.44	12.53	.039	-.494*	-.758*	-.659*	-.576*	-.419*	-.306*	(.91)

Note. [a] The 'Sex' variable was dummy coded as 0=Male and 1=Female. *$p < .05$ two tailed. Values contained in parentheses refer to the internal consistency (α) reliabilities of each scale. Values relevant for hypothesized relations appear in bold.

Table 3

Summary of Regression Analyses for Tests of Hypothesized Relations

Hypothesis	Criterion	Predictor(s)	*b*	SE_b	*β*	*p*	R^2
1a)	Conscientiousness	IM Pressure[a]	**11.680***	1.421	.436*	.000	.190
1b)	Self-Deception	IM Pressure	**4.957***	.735	.370*	.000	.137
1c)	Impression Management	IM Pressure	**10.194***	1.223	.442*	.000	.195
2a)[b]	Conscientiousness	1. Self-Deception	.994*	.119	.497*	.000	
		2. IM Pressure	-6.850	7.043	-.255	.332	
		3. Self-Deception x IM Pressure	**.352**	.189	.518	.064	.474
2b)	Conscientiousness	1.Impression Management	.502*	.077	.433*	.000	
		2.IM Pressure	-13.133	8.10	-.491	.106	
		3.Impression Management x IM Pressure	.**270***	.116	.752*	.021	.435
3)[c]	Conscientiousness	1. Impression Management	.379*	.074	**.329***	.000	
		2. Self-Deception	.830*	.120	**.447***	.000	.379
4)[d]	Conscientiousness	1. Impression Management	.509*	.082	**.433***	.000	
		2. Self-Deception	.886*	.143	**.432***	.000	.568

Note. [a] The 'IM Pressure' variable was dummy coded as 0=Low IM Pressure/Non-Evaluative Testing Context and 1=High IM Pressure/Evaluative Testing Context.
[b] Predictors were entered simultaneously in all multiple regression analyses for tests of hypotheses 2a-4. [c] Analysis of responses from the 'Low IM Pressure' condition only.
[d] Analysis of responses from the 'High IM Pressure' condition only. *p<.05 two tailed. Values relevant for hypothesized relations appear in bold.

Table 4

Correlations Among Variables in Test of Hypothesis 2a

Variable	*n*	*M*	*SD*	*1*	*2*	*3*	*4*
1. Conscientiousness	289	73.59	13.31	(.93)			
2. Self-Deception	289	35.85	6.65	.651*	(.82)		
3. IM Pressure	289	.43[a]	.50	.436*	.370*	--	
4. Self-Deception x IM Pressure	289	16.72	19.58	.506*	.480*	.980*	--

Note. [a] The 'IM Pressure' variable was dummy coded as 0=Low IM Pressure and 1=High IM Pressure. * $p < .05$ two tailed. Values contained in parentheses refer to the internal consistency (α) reliabilities of each scale.

Table 5

Correlations Among Variables in Test of Hypothesis 2b

Variable	*n*	*M*	*SD*	*1*	*2*	*3*	*4*
1. Conscientiousness	288	73.58	13.28	(.93)			
2. Impression Management	288	67.51	11.45	.625*	(.85)		
3. IM Pressure	288	.43[a]	.50	.439*	.442*	--	
4. Impression Management x IM Pressure	288	31.80	37.02	.505*	.545*	.983*	--

Note. [a] The 'IM Pressure' variable was dummy coded as 0=Low IM Pressure and 1=High IM Pressure. * $p < .05$ two tailed. Values contained in parentheses refer to the internal consistency (α) reliabilities of each scale.

Table 6

Summary of Regression Analyses from Post Hoc Examinations.

Criterion	Predictor(s)	*b*	SE_b	β	*p*	R^2
1) Conscientiousness[a]	ATTR Composite	.750	.244	.269	.003	.072
2) Impression Management	ATTR Composite	.825	.203	.347	.000	.120
3) Self-Deception	ATTR Composite	.250	.122	.183	.042	.034
4) ATTR Composite	1. Impression Management	.144	.042	.343	.001	
	2. Self-Deception	.005	.073	.007	.941	.120

Note. [a] These analyses all include only responses form Hi IM pressure condition. *p<.05 two tailed.

CHAPTER 6

DISSCUSSION

Review of Hypotheses Tests

The significant observed mean differences in Conscientiousness, Self-Deceptive Enhancement, and Impression Management between the Low and High IM pressure groups provided strong support for hypotheses 1a, 1b, 1c. We feel comfortable making the assertion that respondents in the High IM pressure condition were motivated to inflate their responses, and in fact, did so to a considerable extent (e.g., approximately one standard deviation increase in mean conscientiousness scores between Low and High IM pressure conditions).

However, these observed mean differences are not of a magnitude that would suggest respondents simply identified the "right" answer and responded to the items in the most socially desirable manner possible (as is likely to occur in an experiment employing a "fake good" instruction set). Therefore, we feel our data reflect the responses of individuals engaging in response distortion in a strategic and naturalistic manner. The fact that the standard deviations are similar for each of the personality variables across Low and High IM pressure conditions lends further support to the notion that individuals in the High IM condition were not simply responding in the most socially desirable manner possible, as if they were, it would be reasonable to expect a degree of range restriction, particularly in responses to the "target" personality dimension; in this case, Conscientiousness.

Hypothesis 2a was supported in that the functional relationship between SDE and Conscientiousness remained constant across conditions in which respondents were likely honest

(Low IM pressure) versus when they were likely distorting their responses (High IM pressure). This finding supports the notion that SDE may be the facet of SDR substantively related to Conscientiousness and that controlling for SDE when making performance predictions would not yield corrected scores closer to "true" Conscientiousness scores. Future research examining whether SDE serves as a predictor of performance would provide further clarity as to the true nature of SDE as a construct. It very well may be related substantively to Conscientiousness, but it remains to be seen whether it is related to the aspects of Conscientiousness that are useful for prediction of performance.

Hypothesis 2b was supported in that the relationship between IM and Conscientiousness changed and became higher in the high IM pressure condition. The fact that this functional relationship was not stable, and became stronger in the condition in which respondents were likely faking supports the notion that IM scales detect deliberate distortion. Therefore, it seems reasonable that controlling for IM would yield corrected Conscientiousness scores that more closely reflect "honest" Conscientiousness scores. Future research examining the effect of controlling for IM in predictive validity studies across a variety of occupations is needed to determine the extent to which, if any, controlling for IM improves the criterion related validity of Conscientiousness. Regardless of whether it does, the results of this study have demonstrated that SDR, or specifically, IM is not simply a redundant measure of personality.

Interestingly, while hypotheses 2a and 2b were supported indicating that the functional relationship between Conscientiousness and SDE remained constant while the relationship between Conscientiousness and IM changed across IM pressure conditions, this difference in Conscientiousness and IM across Low and High IM pressure conditions was more pronounced among males. The correlations between Conscientiousness and IM for males in the Low and

High IM Pressure conditions were r = .343, p = .001 and r = .643, p = .000 respectively. For females, the correlations between Conscientiousness and IM in the Low and High IM Pressure conditions were r = .555, p = .000 and r = .671, p = .000 respectively. Conversely, the change in the relationship between Conscientiousness and SDE across Low and High IM pressure conditions was more pronounced among females. The correlations between Conscientiousness and SDE for males in the Low and High IM pressure conditions were r = .606, p = .000 and r = .653, p = .000 respectively. For females, the correlations between Conscientiousness and SDE in the Low and High IM Pressure conditions were r = .515, p = .000 and r = .706, p = .000 respectively (see Table 7).

This finding is quite interesting as it appears that there may be gender differences, not only in the "quantity" of response distortion individuals engage in, but in the "type" of distortion respondents engage in. It appears that females may be more apt to distort their responses on items referencing cognitions, whereas males may be more likely to distort their responses to items referencing explicit behaviors. In Table 2 we see that females scored significantly higher than males on the IM scale in the High IM pressure condition. While it may be tempting to interpret this as evidence of females engaging in greater levels of response distortion; the *change* in the relationship between IM and the substantive personality dimension Conscientiousness in the High IM pressure condition may be a more accurate reflection of response distortion. Taking this view, one would infer that males engaged in response distortion to a greater extent.

We had no hypotheses related to gender, but our findings suggest there may be gender differences in how respondents view items referencing cognitions versus those referencing behaviors. Future research examining gender differences in naturally occurring response distortion may shed light on this matter. If there are, in fact, gender differences in the types of

items detecting response distortion researchers should consider the sample characteristics of studies declaring that "faking doesn't matter." A disproportionate representation of a particular gender combined with the use of a SD scale with particular item content (e.g. referencing cognitions vs. behaviors) could potentially influence the results and subsequent inferences. Furthermore, meta-analyses combining "faking" studies with varying sample characteristics and SD operationalizations may serve to cancel out important relationships among variables.

Hypotheses 3 and 4 were partially supported. The relationship between SDE and Conscientiousness was stronger than the relationship between Conscientiousness and IM in the Low IM pressure condition, when respondents were presumably responding honestly. This provides support for the notion that SDE items reflect information related to the substantive personality dimension. Conversely, the relationship between Conscientiousness and IM was stronger in the High IM pressure condition, when respondents were presumably faking. This provides support for the notion that IM items detect deliberate faking. However, these hypotheses did not receive full support, as the test of the difference between the standardized regression coefficients was not statistically significant. Additionally, we wish to make it clear that the interpretability of standardized regression coefficients (β's) is a debated issue. While some experts suggest that standardized coefficients are useful for determining the relative importance of each variable to others in a given sample (Schumacker & Lomax, 1996), others caution that β's are affected by the variances and covariances of variables under study (Pedhazur, 1997) and should therefore be interpreted cautiously, if at all. Referring to Tables 1 and 2 and inspecting the SDs associated with the variables involved in these hypotheses will reveal that they all demonstrate a similar degree of dispersion about their respective means. In the present analyses, examination of β's provides a method for examining the interrelationship between

three variables, and the change of this interrelationship across IM Pressure conditions. Again two points seem particularly noteworthy and worthy of repeating.

First, SDE exhibits a stronger relationship with Conscientiousness in the Low IM pressure condition than IM, while IM exhibits a stronger relationship with Conscientiousness than SDE in the High IM pressure condition. This observation provides additional support for the notion that when individuals are responding honestly (Low IM pressure condition) the facet of SD most highly related to Conscientiousness (e.g., SDE) is due to the fact that these items are measuring a similar underlying substantive personality construct. However, when individuals are engaging in response distortion (High IM pressure condition) the SD dimension demonstrating a strong relationship with Conscientiousness (e.g., IM) may reflect the fact that both IM and Conscientiousness are measuring the same construct that was detected by the IM scale and contained in the Conscientiousness scores in the form of a systematically favorable self-portrayal.

Second, examination of the standardized regression coefficients from hypotheses 3 and 4 combined with an examination of the interrelationships among "Big-Five" dimensions in Tables 1 and 2 provides strong support for the independence of SDE and IM scales as measuring distinct constructs. In fact, the Bivariate correlation between SDE and IM in the presumably honest Low IM pressure condition, while significant, is less than the correlation between SDE and Conscientiousness, Extraversion, and Neuroticism suggesting again that Socially Desirable responding manifests as two related but conceptually distinct facets, SDE and IM, and that this distinction is predicated on whether items focus on cognitions or behaviors.

Finally, regression diagnostics were performed for all previous analyses to verify that the underlying assumptions for employing regression analyses were, in fact, met. The plot of the

standardized residuals on the standardized predicted values from the regression of Conscientiousness on IM from the High IM pressure group revealed several points of interest and concern. As can be seen in Figure 1, the residual plot reveals an apparent violation of the assumption of homoscedasticity of residuals. This apparent violation was tested by a technique that was a slight variation of the modified Levene's test as presented in Cohen et al., (2003). We calculated the absolute values of the unstandardized residuals from the regression of Conscientiousness on IM from the High IM pressure condition. The bivariate correlation between the absolute deviation of the residuals across predicted Conscientiousness scores from the high IM pressure condition, and the IM scores was statistically significant ($r = -.205$, $p = .022$). This same residual analysis was performed for the regression of Conscientiousness on IM from the Low IM pressure condition ($r = -.083$, $p = .294$) as well as for the regression of Conscientiousness on SDE from the High ($r = -.165$, $p = .066$) and Low ($r = -.078$, $p = .323$) IM pressure conditions. Only the residual analysis from the regression of Conscientiousness on IM from the high IM pressure condition revealed statistically significant heteroscedasticity. Two important points follow from this finding. First, the previous regression analyses involving Conscientiousness and IM (e.g., hypotheses 1a and 2b) should be interpreted cautiously, as the homoscedasticity of residual assumption underlying the OLS regression model was violated. That said, inferences are routinely made in the extant literature, perhaps inappropriately, without a passing reference to residual analysis. Therefore, we feel that the discovery of this violation presents a limitation rather than a fatal flaw in the current study.

Furthermore, this finding provides some directions for future research. In brief, the pattern of the residuals in Figure 1 reveals that the error of prediction between Conscientiousness and IM decreases as IM scores increase. This pattern of response is consistent with one in which

individuals in the top of the conscientiousness distribution are scoring at a similarly high level in the IM distribution. In the most extreme case, individuals are scoring as high as possible on both scales, in which case there would be nearly a perfect correspondence between an individual's standing on both conscientiousness and IM in a cumulative score distribution. The apparent increase in the precision of predictions as IM scores increase may reflect the fact that Conscientiousness scores are inflated by response distortion elicited in the High IM pressure condition. Similarly, this finding may indicate instability of factor structure across levels of Conscientiousness in that IM and Conscientiousness become ostensibly identical constructs at the high end of the score distribution. That this occurred only in the High IM pressure condition when respondents are presumably "faking" to some extent suggests that these highest scorers on Conscientiousness are distorting their responses and that IM is not simply capturing redundant substantive variance related to Conscientiousness, but rather it is capturing response distortion. Tests of measurement equivalence/invariance of personality and SDR scales obtained across evaluative and non-evaluative contexts may shed further light on this issue.

General Discussion

While difficult to define, deliberate distortion or "Faking" can simply be thought of as that which occurs when it seems plausible for it to occur. Our findings suggest that deliberate distortion; or the "style" of the "substance vs. style" debate; is evidenced by the variance shared by IM and Conscientiousness when there is reason to fake. It is quite possible that SDR, as a whole, is a construct that can be characterized as having both a "substantive" and a "style" component. The support for our hypotheses indicate that Self-Deceptive Enhancement may be the form of socially desirable responding that is related to an individual's personality (substance), and should not be viewed as a contaminate in personality measures. Elevations in

Impression Management, however, may represent response distortion and its removal may improve the predictive accuracy of Conscientiousness scales when used in a selection context.

The results of this study are of particular interest to HR professionals intent on developing correction procedures that improve the predictive validity of personality measures. It seems plausible that a correction technique could be employed that partials out variance in Conscientiousness accounted for by the effects of IM. To the extent that IM scales capture deliberate distortion and "faking," partialling IM from Conscientiousness should result in a "purified" Conscientiousness score free from "style." However, Conscientiousness scores should not be "overcorrected" in the sense that variance shared exclusively with SDE should not be removed, as the variance shared with SDE may represent that portion of SD which reflects real individual differences (substance). Perhaps the most important point for future research is the implication of the heteroscedasticity of residuals in the functional relationship between Conscientiousness and Impression Management when respondents are distorting their responses to varying degrees. Making performance predictions based on observed Conscientiousness scores while control for IM using the OLS regression model, may yield biased predicted performance estimates. A weighted least squares model may be more appropriate when attempting to control for the effects of IM in Conscientiousness scores.

Our utilization of a student sample will likely be viewed by some as a limitation to the external validity of this study. In some respects, we believe that student samples may be preferable to applicant and incumbent samples when the goal is to obtain equivalent experimental and control groups. Recent investigations of faking related issues have successfully employed student samples (Mueller-Hanson et al., 2003). Our sample represented respondents who will soon be job applicants applying for a broad array of jobs. Our method of inducing IM

pressure was specifically catered to a student sample such that we feel comfortable using students as a proxy for applicants in studying the faking phenomena. We believe that a far more important issue is to study "Faking" that is elicited from respondents rather than forced out with "fake good" instruction sets, and we believe that we succeeded in doing this. Nevertheless, future research examining the differential effect of items referencing cognitions versus behaviors for detecting faking should be carried out using applicant samples from various occupations.

One of the purposes for conducting this study was to clarify the nature of the relationship between SDR and personality. While we can not say that we have provided a resolution to the substance versus style debate, we have made several useful contributions to this area of research. First, we have demonstrated that the two factor conceptualization of SDR based on Paulhus' model relates to Conscientiousness differentially across evaluative and non-evaluative contexts. Second, confirmation of our hypotheses provides support for the notion that SD can be conceived of as both a redundant personality dimension, as well as a contaminate; but that the role it plays seems to depend on whether the SD items involve claiming or denying cognitions versus overt behaviors. Furthermore, our findings indicate that there may be gender differences in patterns of faking with respect to item content that should be examined further. Third, we have revealed that the underlying assumptions for using the OLS regression model to make predictions with faking corrections may be violated and should be explicitly tested in future research and practice. Perhaps most importantly, we have provided evidence that supports the notion that the item content of a Social Desirability scale influences the extent to which the scale actually detects faking. This means that controlling for faking, as measured by unidimensional SD scales, may be a poor practice which overcorrects personality scores by removing both contamination as well as information related substantively to the personality trait.

Table 7

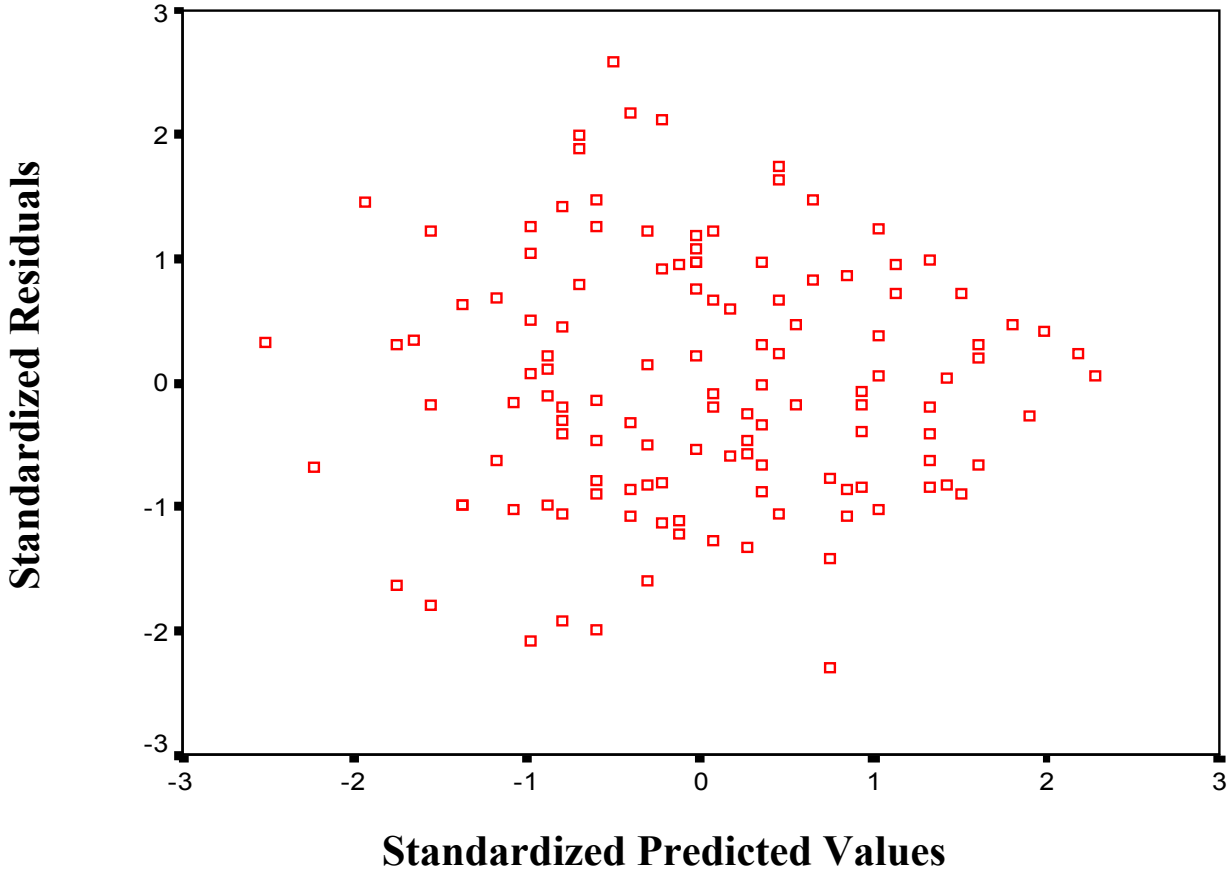
Correlations Between Conscientiousness, Impression Management, and Self-Deception for Males and Females in both the Non-Evaluative "Low IM Pressure," and Evaluative "High IM Pressure" Condition.

Variable	*n*	*M*	*SD*	*1*	*2*	*3*	*n*	*M*	*SD*	*1*	*2*	*3*
	Low IM Pressure						High IM Pressure					
Males												
1. Conscientiousness	85	68.89	10.71	--			60	79.13	11.12	--		
2. Impression Management	85	62.28	10.60	.343*	--		60	71.07	10.35	.643*	--	
3. Self-Deception	85	35.35	5.86	.606*	.306*	--	60	39.53	6.09	.653*	.674*	--
Females												
1. Conscientiousness	72	68.28	13.20	--			65	81.25	13.15	--		
2. Impression Management	72	64.25	10.02	.555*	--		65	75.32	10.10	.671*	--	
3. Self-Deception	72	31.72	6.44	.515*	.269*	--	65	37.86	5.76	.706*	.455*	--

Note. *$p < .05$ two tailed.

Figure 1

Residual Plot from the Regression of Conscientiousness on IM from Evaluative, "High IM Pressure" Condition.



REFERENCES

Alliger, G. M., & Dwight, S. A. (2000). A meta-analytic investigation of the susceptibility of integrity tests to faking and coaching. *Educational and Psychological Measurement, 60(1),* 59-73.

Angleitner, A., & Wiggins, J. S. (Eds.). (1986). *Personality Assessment via Questionnaires: Current Issues in Theory and Measurement.* Berlin: Springer-Verlag.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51(6),* 1173-1182.

Barrick, M. R., & Mount, M. K. (1991). The big-five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44(1),* 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(3),* 261-272.

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

Behling, O. (1998). Employee selection: Will intelligence and conscientiousness do the job? *Academy of Management Executive, 12,* 99-106.

Cascio, W. F., (1998). *Applied psychology in human resource management.* New Jersey: Prentice Hall.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum Associates.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum Associates.

Costa, P. T., & McCrae, R. R. (1997). Stability and change in personality assessment: The revised NEO personality inventory in the year 2000. *Journal of Personality Assessment, 68(1),* 86-94.

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

Davies, M. F., & French, C. C. (1998). Self-deceptive enhancement and impression management correlates of EPQ-R dimensions. *Journal of Psychology, 132(4),* 401-406.

Dwight, S. A., & Feigelson, M. E. (2000). A quantitative review of the effect of computerized testing on the measurement of social desirability. *Educational and Psychological Measurement, 60(3),* 340-360.

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(2),* 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(1),* 122-133.

Finch, D. W. (2002). *Rater personality and judgments of sexual harassment*. Unpublished master's thesis, University of Georgia, Athens, Georgia.

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

Furnham, A. F. (1997). Knowing and faking one's five-factor personality score. *Journal of Personality Assessment, 69(1),* 229-243.

Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social Desirability response effects: Three alternative models. *Academy of Management Journal, 26(2),* 321-331.

Goldberg, L. R. (1995). What the hell took so long? Donald W. Fiske and the big-five factor structure. In P. E. Shrout & S. T. Fiske (Eds.), *Personality Research, Methods, and Theory* (pp. 29-43). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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: Tillburg University Press.

Goldberg, L. R. (in press). The comparative validity of adult personality inventories: Applications of a consumer-testing framework. In S. R. Briggs, J. M. Cheek, & E. M. Donahue (Eds.), *Handbook of Adult Personality Inventories.*

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26(3)*, 499-510.

Guion, R. M. (1993). The need for change: six persistent themes. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 481-496). San Francisco, CA: Jossey-Bass.

Hays, W. L. (1994). *Statistics* (5th ed.). Fort Worth: Harcourt.

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

Holden, R. R. (1998). Detecting fakers on a personnel test: Response latencies versus a standard validity scale. *Journal of Social Behavior and Personality, 13(2),* 387-399.

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

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11(2/3),* 209-244.

International Personality Item Pool (2001). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individuals Differences (http://ipip.ori.org/).

John, O. P., & Srivastava, S. (1999). The big five taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102-138). New York: Guilford Publications, Inc.

Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology, 85(5),* 751-765.

Keppel, G., & Zedeck, S. (1998). *Data Analysis for Research Designs.* New York: Freeman.

Klimoski, R. J. (1993). Predictor constructs and their measurement. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 99-134). San Francisco, CA: Jossey-Bass.

Kline, P. (1993). *Personality: The Psychometric View.* New York, NY: Routledge.

Kurtz, J. E., & Parrish, C. L. (2001). Semantic response consistency and protocol validity in structured personality assessment: The case of the NEO-PIR. *Journal of Personality Assessment, 76,* 315-332.

Kluger, A. N., & Colella, A. (1993). Beyond the mean bias: The effect of warning against faking on biodata item variances. *Personnel Psychology, 46(4),* 763-781.

Lautenschlager, G. J., & Flaherty, V. L. (1990). Computer administration of questions: More desirable or more social desirability? *Journal of Applied Psychology, 75,* 310-314.

Lautenschlager, G.J. (1994). Accuracy and faking of background data. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook* (pp. 391-419). Palo Alto, CA: Consulting Psychologists Press.

Moorman, R. H., & Podsakoff, P. M. (1992). A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behavior research. *Journal of Occupational & Organizational Psychology, 65(2),* 131-140.

Muchinsky, P. M. (1996). The correction for attenuation. *Educational and Psychological Measurement, 56(1),* 63-75.

Mueller-Hanson, R., Heggestad, E. D., & Thorton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology, 88(2),* 348-355.

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

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(6),* 660-679.

Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11(2/3),* 245-269.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46(3),* 598-609.

Pedhazur, E J. (1997). *Multiple Regression in Behavioral Research: Explanation and Prediction* (3rd ed.). Orlando, FL: Harcourt Brace & Company.

Piedmont, R. L., McRae, R. R., Riemann, R. & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology, 78,* 582-593.

Popp, E. C. (2002). *A model for reducing positive bias on personal employment reference forms*. Unpublished master's thesis, University of Georgia, Athens, Georgia.

Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84(5),* 754-775.

Rosse, J. G., Stecher, M. D., & Miller, J. L. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology, 83(4),* 634-644.

Sackett, P. R., & Arvey, R. D. (1993). Selection in small N settings. In N. Schmitt, & W. C. Borman (Eds.), *Personnel Selection in Organizations* (pp. 418-447). San-Francisco: Jossey-Bass.

Schmidt, F. L. (1993). Personnel Psychology at the Cutting Edge. In N. Schmitt, & W. C. Borman (Eds.), *Personnel Selection in Organizations* (pp. 497-515). San Francisco: Jossey-Bass.

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(6),* 966-974.

Schmitt, N., & Landy F. J. (1993). The concept of validity. In N. Schmitt, & W. C. Borman (Eds.), *Personnel Selection in Organizations* (pp. 275-309). San-Francisco: Jossey-Bass.

Schrader, A. D., & Osburn, H. G. (1977). Biodata faking: Effects of induced subtlety and position specificity. *Personnel Psychology, 30,* 395-404.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling.* Mahwah, NJ: Lawrence Erlbaum Associates,

Shrout, P. E., & Fiske S. T. (Eds.). (1995). *Personality Research, Methods, and Theory: A Festschrift Honoring Donald W. Fiske.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Smith, D. B., & Ellingson, J. E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology, 87(2),* 211-219.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15,* 72-101.

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 3,* 271-295.

Tupes, E. C. & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings.* USAF ASD Technical Report, No. 61-97.

Wiggins, J. S. (1988). *Personality and Prediction: Principles of Personality Assessment.* Malabar, Florida: Robert E. Krieger Publishing.

Wilkerson, J. M., Nagao, D. H., & Martin, C. L. (2002). Socially desirable responding in computerized questionnaires: When questionnaire purpose matters more than the mode. *Journal of Applied Social Psychology, 32(3),* 544-559.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review, 12(2)*, 250-264.

APPENDICES

APPENDIX A

TEXT FROM PSEUDO-RESEARCH FLYER

Athens Consumer Research Program

We are currently recruiting 30 participants for a series of consumer research studies in which we will be asking tomorrow's leaders and professionals to evaluate advertisements prior to their release in the marketplace.

In exchange for participating in three 1 hour sessions, qualified applicants will receive $150 and 5 hours of research credit that can be used to meet research requirements this semester, or will carry over to next semester.

Potential participants must pass the evaluation given in the "Personality Focused Advertising" experiment at the University of Georgia.

Remember, we're interested in hearing from tomorrow's professionals. We are seeking hardworking, achievement striving, dependable individuals.

We look forward to meeting you!!

$150

Five Research Credit Hours

12/2004

Expiration Date:

University of Georgia. Department of Psychology

Experimenter signature required to certify that the above individual successfully passed the selection process: _____

APPENDIX B

HIGH "IM PRESSURE" INSTRUCTIONS

Athens Consumer Research Program

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. Please read each statement carefully, and then fill in the bubble that corresponds to the letter on the scale.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

APPENDIX C

LOW "IM PRESSURE" INSTRUCTIONS

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. Since there will be nothing to be lost or gained based on your responses, please 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. All responses will be used for research purposes only, and will be kept strictly confidential. Only the primary researcher will have access to your responses. Please read each statement carefully, and then fill in the bubble that corresponds to the letter on the scale.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

APPENDIX D

PERSONALITY INVENTORY

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

1. Panic easily.
2. Waste my time.
3. Am out for my own personal gain.
4. Have a good word for everyone.
5. Cut others to pieces.
6. Get stressed out easily.
7. Am not easily frustrated.
8. Get back at others.
9. Treat all people equally.
10. Follow through with my plans.
11. Am often down in the dumps.
12. Keep in the background.
13. Would never cheat on my taxes.
14. Talk to a lot of different people at parties.
15. Return extra change when a cashier makes a mistake.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

16. Use swear words.
17. Get excited by new ideas.
18. Have a rich vocabulary.
19. Am very pleased with myself.
20. Have frequent mood swings.
21. Worry about what people think of me.
22. Enjoy thinking about things.
23. Feel comfortable around people.
24. Don't see things through.
25. Do things according to a plan.
26. Tend to vote for liberal political candidates.
27. Easily resist temptations.
28. Am exacting in my work.
29. Hold a grudge.
30. Have a sharp tongue.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

31. Respect others.
32. Am not always what I appear to be.
33. Am the life of the party.
34. Like to take responsibility for making decisions.
35. Seldom feel blue.
36. Believe there is never an excuse for lying.
37. Carry out my plans.
38. Avoid contacts with others.
39. Am hard to get to know.
40. Enjoy hearing new ideas.
41. Seldom get mad.
42. Complete tasks successfully.
43. Know how to captivate people.
44. Sympathize with others' feelings.
45. Have a low opinion of myself.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

46. Carry the conversation to a higher level.
47. Have sometimes had to tell a lie.
48. Feel threatened easily.
49. Don't always practice what I preach.
50. Am not interested in abstract ideas.
51. Believe in the importance of art.
52. Am easy to satisfy.
53. Do not enjoy going to art museums.
54. Need a push to get started.
55. Trust what people say.
56. Am not always honest with myself.
57. Use flattery to get ahead.
58. Feel comfortable with myself.
59. Like myself.
60. Can say things beautifully.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

61. Would describe my experiences as somewhat dull.
62. Rarely talk about sex.
63. Make a mess of things.
64. Have a vivid imagination.
65. Find it difficult to get down to work.
66. Find it difficult to approach others.
67. Have little to say.
68. Keep others at a distance.
69. Dislike myself.
70. Don't mind being the center of attention.
71. Worry about things.
72. Sometimes have trouble making up my mind.
73. Am always prepared.
74. Remain calm under pressure.
75. Make friends easily.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

76. Leave things unfinished.
77. Finish what I start.
78. Pay attention to details.
79. Would never take things that aren't mine.
80. Always know why I do things.
81. Believe that I am better than others.
82. Insult people.
83. Retreat from others.
84. Try to follow the rules.
85. Warm up quickly to others.
86. Do not like poetry.
87. Make demands on others.
88. Cheer people up.
89. Tell the truth.
90. Suspect hidden motives in others.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

91. Don't like to draw attention to myself.

92. Just know that I will be a success.

93. Am not interested in theoretical discussions.

94. Am concerned about others.

95. Am filled with doubts about things.

96. Misuse power.

97. Tend to vote for conservative political candidates.

98. Rarely lose my composure.

99. Believe that others have good intentions.

100. Always admit it when I make a mistake.

101. Have difficulty understanding abstract ideas.

102. Contradict others.

103. Enjoy wild flights of fantasy.

104. Feel comfortable with myself.

105. Rarely look for a deeper meaning in things.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

106. Start conversations.

107. Accept people as they are.

108. Rarely get irritated.

109. Get chores done right away.

110. Often feel blue.

111. Break rules.

112. Shirk my duties.

113. Am relaxed most of the time.

114. Am skilled in handling social situations.

115. Get back at others.

116. Rarely overindulge.

117. Do just enough work to get by.

118. Don't talk a lot.

119. Know that my decisions are correct.

120. Avoid philosophical discussions.

Please use the rating scale below to describe how accurately each statement describes *you*.

Response Options

A: Very Inaccurate
B: Moderately Inaccurate
C: Neither Inaccurate nor Accurate
D: Moderately Accurate
E: Very Accurate

121. Am not easily bothered by things.

122. Fear for the worst.

123. Make plans and stick to them.

124. Mess things up.

125. Don't put my mind on the task at hand.

126. Cheat to get ahead.

127. Am likely to show off if I get the chance.

128. Believe that too much tax money goes to support artists.

129. Make people feel at ease.

130. Do not like art.

APPENDIX E

HIGH "IM PRESSURE" BIOGRAPHICAL FORM

We would like to know more about you. The following items focus on your personal and demographic characteristics. Please respond honestly so that we can insure a good "fit" with students we decide to accept.

131. Have you participated in any consumer research in the past?
a) yes
b) no

132. How many hours a week, on average, do you play video games?
a) do not play video games
b) 1-2 hours
c) 3-5 hours
d) 6 or more hours

133. How many hours a day, on average, do you watch television?
a) do not watch television
b) 1 hour
c) 2 hours
d) 3 or more hours

Please respond to the following items using this format:

a=*strongly disagree,* **b**=*disagree,* **c**=*neutral,* **d**=*agree,* **e**=*strongly agree*

134. Overall the "Athens Consumer Research Program" seems like a great program to participate in.

135. The "Athens Consumer Research Program" offers generous compensation for students who participate in their research program.

136. If more students found out about this program, many of them would like to participate in it.

137. I like the idea of being able to have my research credits count towards fulfilling the research requirements in a future semester.

138. I would like to get selected to participate in the "Athens Consumer Research Program".

APPENDIX F

LOW "IM PRESSURE" DEBRIEFING FORM

Thank You for your participation. The inventory you just completed will not be used for establishing a database for an advertising agency. The true purpose of the study was to examine individuals' responses to personality items in the absence of pressure to distort their responses. All responses will be kept confidential and will be used for research purposes only. The primary researcher will stare the raw data in a secure location, and will be the only individual with access to it. Feel free to contact the experimenter if you have additional questions:

Lawrence Michels

Michels@arches.uga.edu

APPENDIX G

HIGH "IM PRESSURE" DEBRIEFING FORM

Thank You for your participation. The inventory you just completed is not, in fact, a selection instrument. The "Athens Consumer Research Program" is a fictitious entity, and there is not an additional experiment. Additionally, there is not a $150 reward nor a five hour research credit reward. You will receive the 1.5 hours of research credit for this portion of the experiment you were guaranteed. The true nature of this study was to examine individual responses to personality items in a context which put respondents under pressure to present themselves favorably. In order to achieve this degree of pressure, it was necessary to lead you to believe that you would be evaluated on the basis of your responses. In actuality, all responses will be used for research purposes only, and will be kept strictly confidential. The principle researcher will maintain all surveys and answer sheets in a secure location, and will be the only individual with access to these data. All identifying materials not related to the survey (such as the "advertisement" certificate) will be destroyed within one week.

It is absolutely critical that the true nature of this experiment remains confidential and that other potential participants are not given any information about it. If information is shared with other students, the integrity of the experiment could be destroyed. Please keep the details of this experiment confidential for the remainder of the semester.

If you have any additional questions, please contact the experimenter.

Thank You

Lawrence Michels

Michels@arches.uga.edu

APPENDIX H

FINAL DEBRIEF

If you are interested in learning more about personality inventories, or would like to take an on-line version of an inventory very similar to the one you just took, go to the following address and follow the links:

http://ipip.ori.org/ipip/new_home.htm

If you take the full length inventory (300 items), a descriptive report will be generated for you in which you will learn where you stand on each of 30 individual facets across five broad personality dimensions including: Extraversion, Neuroticism, Openness to experience, Agreeableness, and Conscientiousness.

Keep in mind that this particular personality inventory is used to measure personality within the "normal" range. This means that while it may be informative, very high or very low scores are not necessarily indicative of psychopathology. Therefore, the results should not be used to self-diagnose.

Have Fun, and if you have any additional questions please contact the principal researcher:

Lawrence Michels
Michels@arches.uga.edu
706-542-3035