

5-2-2006

Self-deception and Other-deception in Personality Assessment: Detection and Implications

Mary Lynn Starke

University of Missouri-St. Louis, marystarke@yahoo.com

Follow this and additional works at: <https://irl.umsl.edu/dissertation>



Part of the [Psychology Commons](#)

Recommended Citation

Starke, Mary Lynn, "Self-deception and Other-deception in Personality Assessment: Detection and Implications" (2006).
Dissertations. 609.

<https://irl.umsl.edu/dissertation/609>

This Dissertation is brought to you for free and open access by the UMSL Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Dissertations by an authorized administrator of IRL @ UMSL. For more information, please contact marvinh@umsl.edu.

Self-deception and Other-deception in Personality Assessment: Detection and Implications

by

Mary Lynn Starke
M.A.

A DISSERTATION

Submitted to the Graduate School of the

UNIVERSITY OF MISSOURI- ST. LOUIS
In partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY
in

Industrial/Organizational Psychology

May, 2006

Advisory Committee

Therese Macan, Ph.D.

Jim Breaugh, Ph.D.

Alice Hall, Ph.D.

Mark Tubbs

Abstract

The present study utilized multiple methods of detecting self-deception and other-deception and explored potential implications for organizations hiring individuals exhibiting these tendencies. Participants were 242 undergraduate business students who completed self-ratings of extraversion and agreeableness under both “answer honestly” instructions and “answer as if you are applying for a job” instructions. Additionally, they completed the impression management and self-deceptive enhancement scales of the BIDR, the fake good scale and the good impression scale of the CPI, and took part in a role play with a trained observer. Individuals who knew the participants well provided ratings of participants’ adjustment, integrity, interpersonal skills, satisfaction with life, extraversion and agreeableness. Results suggested that participants who had views of themselves that were closely aligned with the views of those who knew them well were rated as having higher levels of adjustment, integrity, interpersonal skills, and satisfaction with life. Measures of other-deception were positively related to levels of adjustment, integrity, and interpersonal skills while self-deception was associated with higher levels of interpersonal skills and lower levels of satisfaction with life. Future research is warranted to determine whether these results are observed in the applicant population.

Self-Deception and Other-deception in Personality Assessment: Detection and Implications

Personality and job performance studies from the 1900s through the mid-1980s consisted primarily of researchers investigating the relationships of individual scales from a myriad of personality inventories to various aspects of job performance. Based on these studies, researchers concluded that there were no meaningful relationships between personality and job performance: “There is not generalizable evidence that personality measures can be recommended as good or practical tools for employee selection” (Guion & Gottier, 1965 p. 159).

However, the application of meta-analysis techniques and an emerging consensus regarding the Big Five structure of personality has increased the respect and usage of personality measures in both the research and the applied world. As a comparison, consider that in 1983 a study by the American Society of Personnel Administrators (ASPA, 1983) found that only 9% of organizations were using personality for external selection and only 4% were using this type of testing for internal selection/promotion. In a more recent study, Ryan, McFarland, Baron, & Page (1999) found that 21-50% of organizations were utilizing personality testing.

One meta-analysis was particularly instrumental in the growth of personality testing research and application (Barrick & Mount, 1991). The authors examined the relationship between the Big Five personality dimensions and various job performance criteria (i.e., job proficiency, training proficiency and personnel data) across five occupational groups (i.e., professionals, police, managers, sales, and skilled/semi-skilled). They found that conscientiousness predicted performance ratings across jobs and settings ($r = .20$) and that

other Big Five characteristics were found to be predictive within certain occupational groups. For example, extraversion was found to be predictive of performance in sales jobs while extraversion and openness to experience were found to be valid predictors of training proficiency across occupations. Another meta-analysis (Tett, Jackson, & Rothenstein, 1991) found that when researchers choose personality tests on the basis of job analysis, conducted confirmatory analyses, and studied incumbents with reasonable job tenure, validity coefficients were even larger than those reported by Barrick and Mount (1991). Additionally, personality tests have found favor in both academia and industry as they exhibit little adverse impact (Ones & Viswesvaran, 1998; Hough, 1998) and add incremental validity above and beyond ability measures (Salgado, 1998; Day & Silverman, 1989).

However, personality tests are still not without critics. One frequently cited criticism is that applicants may be able to inflate their scores and may do so in order to increase their likelihood of obtaining a job (Hogan & Hogan, 1992). Numerous studies lend credence to this concern as research suggests that individuals can distort their scores when instructed to do so and that some distortion is actually occurring in applicant settings (e.g., Alliger & Dwight, 2000; Becker & Colquitt, 1992; Dalen, Stanton, & Roberts, 2001, Furnham, 1990; Furnham, 1997; Kluger & Collela, 1993; LoBello & Sims, 1993; Martin, Bowen, & Hunt, 2002; Scandell & Wlazelek, 1996; Viswesvaran & Ones, 1999).

To address concerns with response distortion, researchers have taken steps to assess the impact on validity. Most notably, researchers have focused on the impact of distortion on criterion-related validity and on the rank ordering of job candidates. However, the results of these studies have been mixed and offer no definitive conclusions about the impact of

distortion (e.g., Christiansen, Goffin, Johnston, & Rothstein, 1994; Ones & Viswesvaran, 1998; Topping & O’Gorman, 1997; Borkenuau & Amelang, 1985; Ruch & Ruch, 1967).

Examination of the research reveals that there are several areas of concern regarding the design of previous studies that may shed light on the conflicting results. One glaring issue is the lack of consistency in the terminology used across research methods. For example, over time, the tendency to provide an inaccurate portrayal of the self came to be known as “faking”. However, the method of detecting this inaccuracy has varied widely. For example, some researchers use social desirability scales to detect an inaccurate portrayal while others use a directed faking study, and still others may compare applicant scores to incumbent scores. Although all of these methods likely have some merit, there is little evidence to suggest exactly what phenomena they are capturing. Furthermore, there are no research studies that are known to the author that examine whether these methodologies are capturing the same experience. As a result, it appears that researchers have put the proverbial “cart before the horse”, making it difficult (if not impossible) to examine the collective research data and draw reasonable conclusions.

In an effort to provide a clearer context for the interpretation of existing research and in an attempt to build a better foundation for future research, is necessary to take a closer look at what phenomenon traditionally-used methodologies are capturing. Some researchers suggest that one promising framework for interpretation is the extent to which studies capture *why* distortion is occurring (e.g., Paulhus, 1984). Two commonly cited reasons for the distortion are self-deception and other-deception. In the case of self-deception, individuals provide an overly positive view of themselves because they lack the self-sight necessary to provide a realistic self- description. On the other hand, individuals engaging in other-

deception may be consciously distorting their scores in order to obtain a favorable outcome (i.e., in the selection process this would be the job). The evidence suggests that self-deception and other-deception are indeed distinct constructs (e.g., Robins & Paulhus, 2001) that have differing implications for organizations (e.g., Paulhus, 1986). For example, self-deception has been linked to a variety of positive outcomes including adjustment:

There is a large and replicated literature showing that well-adjusted people have positively biased self-images; consequently, well adjusted people tend to ignore minor criticisms, discount their failures, and avoid negative thoughts, and expect to succeed in most of their undertakings (Hogan, 1991 p. 905).

This may be important to organizations as individuals having positive views of the self may be more likely to work harder and longer on tasks (e.g., Felson, 1984). In turn, this perseverance may lead to more effective performance and greater likelihood of achieving goals (e.g. Bandura, 1977; Baumesiter, Hamilton, & Trice, 1985).

On the other hand, other-deception has been linked to the ability to effectively engage in self-presentation behaviors. That is, there is evidence to suggest that in some cases, self-presentation behaviors may lead to deception (e.g., Feldman, Forrest, & Happ, 2002). The ability to engage in self-presentation behaviors is frequently cited as being important in those jobs that involve much interpersonal interaction (e.g., salesperson, customer service agent).

Recognizing that these two types of deception are important considerations, the present study will use these constructs as a framework for examining what some traditionally-used research methods are capturing. This research will include a simultaneous examination of a directed faking study, social desirability scales, and external criteria measures in an attempt to discern whether they capture self-deception and/or other-deception. This research makes significant practical and theoretical contributions to the literature. From a theoretical standpoint, it will be a good first step in examining what these multiple methods

are capturing, rather than, as previous research has done, making an assumption that all methods are measuring the same construct. As such, this study will provide a better context for interpretation of past research and a firmer foundation for future work. From a practical standpoint, the results of this study will give organizations a better understanding of what traditionally used methods are capturing. As a result, decision-makers can have a better context for interpretation of results and an understanding of which research methodologies will allow them to capture self-deception and/or other-deception.

A closer look at self-deception and other-deception: Why is the differentiation important?

In a selection context, the primary concern is that individuals will manipulate their scores in order to cast themselves in a more favorable light. Typically, one attempts to cast oneself in a positive light by answering affirmatively to socially sanctioned behaviors (e.g., I help out my coworkers when they need me) and negatively to behaviors that are not socially sanctioned (e.g., I manipulate others). This tendency is frequently termed socially desirable responding. More specifically, the former is often called attribution while the latter is termed denial (e.g., Paulhus, 1984). However, evidence suggests these are equivalent measures of the same construct (e.g., Ramanaih & Martin, 1980).

One key area that is frequently overlooked in the discussion of this type of response distortion is that individuals may make these attributions and/or engage in this denial either at a conscious or unconscious level. That is, individuals may inaccurately perceive they possess the favorable traits or behaviors they report (or conversely they don't possess the unfavorable traits or behaviors they deny). As such, they believe their self-reports to be accurate when they are actually inflated. In contrast, individuals who engage in other-deception realize they do not possess the favorable behavior or traits that they report but consciously distort their

responses in order to cast themselves in a more favorable light and thus, increase the likelihood of a favorable outcome.

It is important to make these distinctions for several reasons. First, research suggests that self-deception is consistent across situations, but other-deception is contextually bound (i.e., its presence and level are dependent on the situation under which the assessment occurs as well as the purpose of the assessment) (Paulhus, 1984; Gudjonsson, 1990). That is, depending upon the circumstances, individuals may choose to consciously distort responses to a self-report questionnaire in order to increase the likelihood of a positive outcome. However, across circumstances, an individual's level of self-deception is expected to remain constant. Second, there may be differing outcomes for organizations according to the type of deception that is occurring.

Self-deception

The first factor, self-deception, has been defined as the “process of giving positively biased responses to questionnaire items when the responses are not true but the respondent believes they are” (Paulhus, 1984, 1986) and has been found to be prevalent in personality testing (e.g., Lewinshon, Mischel, Chaplin, & Barton, 1980). This finding is not surprising as research suggests that most people hold a very positive view of themselves (see Greenwald, 1980 for review). The majority of individuals rate themselves as better than the average person in terms of traits (Brown, 1966) and abilities (e.g., Larwood & Whittaker, 1977; Campbell, 1986) and rate themselves more favorably on personality attributes than do outside observers (Lewinshohn, Mischel, Chaplin, & Barton, 1980). Furthermore, research suggests that most individuals process positive information about the self differently than negative information. For example, most people recall positive personality information more easily than negative personality information (e.g., Kuiper & Derry, 1982; Kuiper &

MacDonald, 1982; Kuiper, MacDonald, & Shaw, 1985) and information about success more easily than information regarding failure (Silverman, 1964).

Traditional conceptualizations of mental health have been that psychologically healthy people were those who maintained a close contact with reality, including a realistic view of the self:

An impartial and objective attitude toward oneself is held to be a primary virtue, basic to the development of all others. There is but a weak case for chronic self-deception with its crippling self-justifications and rationalizations that prevent adaptation and growth. And so may it be said that if any trait of personality is intrinsically desirable, it is the disposition and ability to see oneself in perspective (Allport, 1937; p. 422).

This view has been shared by numerous researchers throughout the years (e.g., Vaillant, 1977; Jourard & Landsman, 1980; Haan, 1977) but has been challenged in the recent literature. This challenge has come as a result of several studies that link the tendency to self-enhance to a variety of positive outcomes including adjustment (Taylor & Brown, 1988), optimism (Scheier & Carver, 1985), and a sense of general capability (Holden & Fekken, 1989). Moreover, significant correlations have been reported between self-deception and good mental health (Linden, Paulhus, & Dobsen, 1986) and self-esteem (Paulhus & Reid, 1991).

Additionally, research suggests that self-deception plays a role not only in one's perception of the present self, but also in the perception of the past and future selves (e.g., Greenwald, 1980; Ross, 1989; Dunning, Meyerowitz & Holzberg; 1989). One reason this finding is important is because there is evidence to suggest that those who expect positive outcomes are more likely to set higher goals for themselves. Furthermore, they are more likely to pursue those goals more vigorously even in the face of setbacks (Bandura, 1989). Similarly, Taylor and Brown (1988), in an integration of the literature, conclude that self-

deception promotes psychological adjustment as well as “higher motivation, greater persistence, more effective performance, and ultimately, greater success” (p. 199). In an organizational setting, it is likely these considerations translate into higher levels of job performance.

On the flipside, lack of self-deceptive positivity about one’s future has been linked to negative outcomes such as depression (e.g., Abramson, Seligman, & Teasdale, 1978). Evidence also suggests that those individuals who have a more balanced view of themselves have low self-esteem and/or are moderately depressed (e.g., Coyne & Gotlib, 1983). These individuals offer self-appraisals that are more consistent with appraisals by objective observers (e.g., Lewinsohn, Mischel, & Barton, 1980) and with evaluations by known others (e.g., Brown, 1986). In turn, low self-esteem and depression have been linked to decreased levels of job performance (e.g., Bono & Judge, 2003; Druss, Schlesinger, & Allen, 2001).

Despite the positive aspects of self-deception, there is also evidence that not all outcomes of self-deception are beneficial. For example, an enhanced version of the self may lead individuals to be closed to outside information that threatens their positive self-image. In turn, this may impede their development and learning. Research indicates that this is indeed occurring. For example, one study found that inflated self-perceptions could not be changed even when confronted with videotaped performance of a task (Robins & John, 1997). Furthermore, individuals engaging in self-deception have been found to make slow progress in the early stages of a training program (Lee & Klein, 2002).

An additional concern is that self-enhancers may have problems building and maintaining interpersonal relationships. For example, receiving feedback inconsistent with one’s self-image may make these individuals hostile toward their evaluators. Within an

organizational context, this might translate into poorer working relationships between supervisors and subordinates. A longitudinal study found that self-enhancement had long-term negative implications for interpersonal relationships outside of work (Colvin, Block, & Funder, 1995). That is, men who self-enhanced were described by their friends as being “condescending in relations with others” and “having hostilities toward others”. In contrast, men who did not self-enhance were described as “sympathetic and considerate”, “having a clear-cut and consistent personality”, and as “having a giving way with others”. Female self-enhancers were described as having “hostility toward others”, “self-defeating”, and having a “brittle ego defense system”. Women who did not self-enhance were described by their friends as being “liked and accepted by people”, “cheerful”, and “having social poise and presence”.

Other-deception

For the purpose of the present study, other-deception will refer to consciously portraying oneself in a manner that is inconsistent with one’s perception of “true self”. Typically, individuals engage in this other-deception in order to gain a desirable outcome (e.g., in the case of selection the desirable outcome would be a job). Other-deception is one type of self-focused impression management (also called self-presentation) tactic.

Impression management has been referred to as “a creative process in which actors mold their outward behavior while taking into careful consideration the specific context and the target audience” (Goffman, 1959). Research suggests that individuals engage in a variety of impression management tactics when going through a selection process including both self-focused and other-focused behaviors. *This terminology is not to be confused with self-deception and other-deception.* As the term suggests, self-focused impression management tactics include those behaviors that draw attention to the participant’s positive attributes

and/or accomplishments (e.g., self-promotion, entitlements) while other-focused tactics refer to those tactics that focus on the person they are trying to impress (e.g., ingratiation, flattery). The present study will focus on personality testing (not an area where the participant interacts with another individual), a context in which the concentration will be solely on self-focused impression management tactics.

Although the present study examines deliberate, other-deception, it is still helpful to have a general understanding of research surrounding impression management as it provides a context in which deliberate, other-deception may occur. Research suggests that impression management techniques occur in a variety of human resource activities including interviews, performance appraisals, and training. For example, in the interview, it has been found that applicant's impression management techniques influence interviewer decisions, independent of applicant qualifications (Gilmore & Ferris, 1989). Another study found that job applicants employing self-focused impression management tactics received higher ratings, more recommendations for a job offer, and fewer rejections than applicants using other-focused impression management tactics (Delery & Kacmer, 1998). In a longitudinal study, it was found that individuals who engaged in impression management behaviors (both self-focused and other-focused) influenced supervisory ratings of performance through perceived similarity and liking (Wayne & Liden, 1995). Furthermore, it has been found that individuals set higher goals on a new task when it is observed by someone than when it is not (Ferris & Porac, 1984). The authors interpret this behavior as a strategic form of self-focused impression management.

Some researchers suggest that these impression management behaviors should not be factored out of the human resource decisions and activities as they represent trait-related

variance that is applicable to job performance. Specifically, it has been suggested that the ability to engage in these self-presentation behaviors or impression manage may be important in those jobs that involve much interpersonal interaction. That is, awareness of social norms and expectations and the ability to present oneself in a manner appropriate to the demands of the situation may be quite functional (suggested by Rosse, Stecher, Miller & Levin, 1998). However, the studies examining the link between deception and job performance have failed to adequately distinguish between self-deception and other-deception (e.g., Viswesvaran, Ones & Hough, 2001).

Similarly to self-deception, not all consequences of engaging in this type of behavior are positive. For example, impression management behaviors are viewed as negative when the actions are seen as incongruent with an observer's view of situationally appropriate behavior. When this occurs, undesirable attributions and affective reactions are likely (e.g., Ferris, Judge, et al.; 1994 Gardner & Martinko, 1988). Furthermore, although self-promotion may lead others to give positive evaluations of the actor's ability, it may create the impression that the actor is manipulative (e.g., Gurevitch, 1984).

Moreover, research supports the supposition that self-presentation strategies do, in fact, in some situations lead to deception or the presentation of clearly false information. For example, one laboratory study examined dyadic interactions in which the participant was told to either appear likeable, competent, or told to simply get to know his/her partner (Feldman, Forrest, & Happ, 2002). After a ten-minute conversation, participants were asked to review a videotape of the process and identify the instances in which they deceived their partners. The authors then coded the lies into lie content (i.e., feelings, achievement, actions & plans, explanations, and facts), rationale for the lies (i.e., self-oriented or other-oriented) and type of

lie (outright, exaggerations, or subtle). The authors examined the differences observed. Overall, it was found that those individuals assigned to either the competent or likable conditions told more lies than those in the control condition (i.e., get to know your partner). As hypothesized, self-presentation goals were related to the content, type of lie, and rationale for the lie being told. Specifically, people in the competent condition told more self-oriented lies and more lies about plans and achievements. People in the likeable condition told more outright lies and exaggeration and more lies about feelings. Similarly, another study (Kashy & DePaulo, 1996) found that people who told more lies were more concerned with self-presentation than those who told fewer lies. Another study utilized the Good Impression (GI) scale of the GPI as a measure of deliberate other-deception. No evidence was found for a relationship between other-deception and job performance for either executives or managers (Viswesvaran, Ones, & Hough, 2001). Clearly, these findings are applicable to selection studies where applicants are presumably quite motivated to provide a positive self-presentation.

People also report being overtly dishonest during the selection process. For example, in one study (McDaniel, Douglas, & Snell, 1997), a group of applicants who had recently posted their resumes on the World Wide Web were asked to report what faking behaviors they engaged in during the application process. Of the applicants, 42% reported giving false opinions, one-third stated they had exaggerated their work experience, and one quarter admitted to such behaviors as inflating their pay rates, denying being fired from a previous job, and presenting themselves as more agreeable than they actually are. People also report using non-existent equipment. For example, one study of electrician applicants found that 35% reported using a piece of equipment that did not exist (Pannone, 1984). Another study

(Anderson & Warner, 1984) found that 45% of applicants for state jobs indicated that in the past they had observed or performed a nonexistent task.

These results suggest that self-deceptors, other-deceptors, and those who do not engage in deception may possess different underlying characteristics that have important implications for job performance. According to the research, there are both positive and negative implications for organizations hiring people possessing these tendencies. As such, it is important that organizations recognize what types of deception that occur in selection settings and potential implications for job performance.

However, complicating matters further, as the following literature review reveals, there are no definitive answers regarding what type of deception any single methodology captures. Therefore, it becomes necessary to take a more holistic approach to examining applicant deception. The present study will take such an approach by simultaneously using several methodologies (i.e., within-subjects, two social desirability scales, and observations from others) in an attempt to understand how the use of multiple tactics may allow us to identify patterns in the data that will enable us to more accurately identify self-deceptors, other-deceptors, and non-deceptors. Based on the findings of previous research, a “self-deception profile”, “other-deception profile”, “no deception profile”, and “mixed deception profile” will be created. That is, profiles will be designed that reflect how individuals who are primarily “self-deceptors”, “other-deceptors”, “non-deceptors”, and “mixed deceptors” would be expected to perform in directed faking studies, on social desirability scales, and how their self-ratings would compare to outside raters. Following participant assignment to these categories, outcome data will be examined in order to determine whether the expected

pattern of results emerges in terms of each participant's adjustment, well-being, integrity, and interpersonal skills.

Methods for Examining Deception

Directed Faking Studies

One method frequently used to examine deception (and a method that will be used in the current study) is a "directed faking" study. This methodology typically involves participants in either a laboratory or applied setting who are given a personality test under differing instructional sets. A common within-subjects methodology includes two conditions: one condition in which applicants are told to 'answer honestly' and another condition in which the same individuals are given some variant of 'fake good' instructions (e.g., fake good, answer if you are applying for a job, answer as if you are applying for a specific job, etc.). The difference scores obtained for each individual are then computed. These difference scores are sometimes interpreted as conscious score inflation or other-deception.

A between-subjects version of this methodology is also quite common. That is, control group participants are given answer honestly instructions and experimental group participants are given some variant of 'fake good' instructions. Next, the responses between the groups are compared and differences are interpreted as score inflation. These between-subject and within-subject studies frequently do not explicitly make the differentiation between self and other-deception. Although the author is not aware of any studies explicitly examining this assumption, given that there is random assignment to groups, it could be argued that self-deception would be constant across groups and therefore other-deception would be the measurement captured.

The Big Five personality traits (i.e., openness to experience, conscientiousness, agreeableness, extraversion and neuroticism) are often the focus of between-subjects studies. The majority of studies involving the Big Five using this methodology have demonstrated that individuals are able to inflate their scores when instructed to do so (e.g., Furnham, 1997; Scandell & Wlazelek, 1996). However, there have been some discrepancies as to which of the Big Five variables are most susceptible to impression management. For example, one study found that individuals were only able to significantly alter their scores to appear to be more conscientious, more agreeable, and less neurotic (Furnham, 1997). However, a recent meta-analysis (Viswesvaran & Ones, 1999) combined the results of 51 studies that examined Big Five score inflation and found that all Big Five characteristics were equally susceptible to score inflation. Additionally, the authors found that if instructed to “fake good” respondents in within-subjects designs were able to raise their scores by approximately .75 standard deviations and in between-subjects designs by approximately .50 standard deviations. The authors also found that within-subject designs produced larger effect sizes and greater variability across the Big Five factors than between-subjects designs.

Similarly, research utilizing this methodology suggests that individuals can significantly alter their scores on both overt and covert integrity tests (It should be noted that some measures of the Big Five may be included in the covert integrity tests—especially measures of conscientiousness). A meta-analysis of 14 studies (thirteen student samples and one prisoner sample) found that overt integrity tests (i.e., those tests that ‘judge the likelihood of counterproductive behavior based on responses to questions designed to measure thoughts, feelings, and expected behaviors involving honesty and punishment and deviance as well as admissions of past misbehavior’) found that individuals could raise their scores by

approximately one standard deviation when instructed to “fake good” and by one and a half standard deviations when coached about how to respond to the inventory (Alliger & Dwight, 2000). It was also found that individuals who were asked to “fake good” on a personality-based integrity test could raise their scores by as much as one-half of a deviation.

Personality tests and integrity tests are not the only selection tools that have been examined using this research design. Numerous studies have found that individuals are able to inflate their scores on biodata instruments when instructed to do so (e.g., Becker & Colquitt, 1992; McFarland & Ryan, 2000; Kluger, Reilly, & Russell, 1991). One study found that the items that were more objective and verifiable were less amenable to score inflation (Becker and Colquitt, 1992). As mentioned previously, several interesting studies have been conducted in which applicants were asked for very concrete and verifiable background information and yet they flagrantly misrepresented themselves (e.g. Pannone, 1984; Anderson, Wagner & Spencer, 1984).

To the author’s knowledge, the assumption that directed faking studies capture deliberate, other-deception has not been empirically tested. However, given that the design of the study generally consists of a comparison of scores obtained in two different scenarios, one in which individuals presumably have motivation to portray a specific image (i.e., answer as if applying for a job) and the other situation in which they do not (i.e., answer honestly), it is a reasonable hypothesis that the design captures other-deception. As such, the present study will hypothesize and test whether directed faking studies are able to capture other-deception.

Social Desirability Scales

Another method that will be used to detect deception in the present study is the use of social desirability scales. This method is both common and highly controversial. Social desirability scales are included in most commercially available measures of personality (Viswesvaran, Ones, & Hough, 2001) and typically contain items that are either socially desirable but unlikely to be true (e.g., “Before voting I thoroughly investigate the qualifications of all candidates”) or socially undesirable items that are likely to be true (e.g., I like to gossip at times”) (Crowne & Marlowe, 1960). The scores on these scales are commonly used to correct the inventory for the effects of social desirability or to disqualify applicants from the selection process.

Social desirability scales are popular as they are typically inexpensive, easy to administer, and add little time to the test administration. However, evidence suggests these scales do not differentiate self-deception and other-deception and bear little relationship to one another, further suggesting they are measuring different constructs (e.g., Paulhus, 1991). This is problematic as most social desirability scales are purported to measure “other-deception” and are typically interpreted within that context. As such, within some practical settings, individuals who surpass a cut-off score on a social desirability scale are deemed “fakers” and disqualified from the selection process, when in fact, they may actually believe their responses to be honest portrayals of the self.

Another concern is that social desirability scales frequently exhibit relationships with the observed personality scales of interest (e.g., Furnham, 1986; Messick, 1960) and there is little agreement regarding the interpretation of this covariation. Some researchers interpret this relationship as an indication of contaminated scales scores (e.g., Holden & Fekken,

1989) while others claim the overlap reflects true overlap between the personality traits of interest and the social desirability scales (e.g., Nicholson & Hogan, 1990; McCrae & Costa, 1983, Smith & Ellingson, 2002). The results of one meta-analysis (Ones, Viswesvaran, & Reiss, 1996) support the latter by suggesting estimated population correlations of .18 and .13 for the relationship between social desirability and emotional stability and conscientiousness, respectively. Studies utilizing a clear distinction between self-deception and other-deception may shed some light on this area of disagreement.

There has been some progress in creating instruments that differentiate self-deception from other-deception. One such instrument that has shown promise is Paulhus' Balanced Inventory of Desirable Responding (BIDR). Based on earlier work by Sackeim and Gur (1979), this instrument was specifically designed to capture self-deception and other-deception as independent constructs by using the Impression Management scale (IM) and Self-deceptive Enhancement scale (SDE). When completing the IM scale, respondents are asked to rate the degree to which they typically perform socially desirable, but uncommon behaviors (e.g., I always obey the laws even if I am unlikely to get caught). In contrast, when completing the SDE scale, respondents are asked to rate their level of agreement to items that reflect a lack of self-insight (e.g., I am not a safe driver when I exceed the speed limit).

Numerous studies have demonstrated that the BIDR does differentiate between these two independent constructs. One manner in which validity was examined was through the examination of scale correlates to personality variables. In one study, Paulhus, Reid, and Delognis (unpublished) found that SDE, but not IM, correlated with optimism. Furthermore, another study found that SDE and IM exhibited different relationships with the Big Five. Specifically, the primary correlates of SDE were high extraversion and low neuroticism

while IM was associated with agreeableness and conscientiousness (Paulhus, 1998). Those high in SDE were also found to have high self-esteem, low neuroticism, depression, empathic distress, and social anxiety (Paulhus & Reid, 1991).

Researchers have used the BIDR to further explore the relationship between commonly used social desirability scales, self-deception and other-deception and have found varied relationships. For example, one study used the BIDR to determine the relationships between response distortion, the NEO's Positive Presentation Management (PPRM) scale, and self-deception and other-deception (Reid-Seiser & Fritzsche, 2001). Initial validation work on the PPRM suggested that the authors intended it to capture deliberate other-deception. However, the findings of the study using the BIDR indicated that Paulhus' measure of self-deceptive enhancement was most closely related to the PPRM. Another study (Paulhus, 1991) found that the Marlowe-Crowne Social Desirability Scale, loaded on both self-deception and other-deception (although a little more so on other-deception) whereas the Edwards Social Desirability scale loaded primarily on the self-deception factor. In contrast, the Wiggins SD scale, and the CPI GI scale and Eysenck's lie scale load primarily on other-deception. No research has directly compared these social desirability scale scores to other methods of assessing self-deception and other-deception. Clearly, many social desirability scales are failing to differentiate between self-deception and other-deception. However, two scales that can be reasonably deemed to be capturing other-deception are the GI scale and the fake good scale of the CPI. In order to understand why it is reasonable to make this assumption, it is necessary to understand the manner in which they were developed. The GI scale was constructed using an approach in which participants completed an experimental booklet of items under normal circumstances and then under

instructions to rate themselves so that they presented a very favorable self-portrait (i.e., such as if they were applying for a desirable job or in any situation in which they wanted to be judged as admirable or praiseworthy). The authors then compared the results of these two protocols and identified those items that changed significantly in terms of the reported degree of endorsement. Forty items that changed significantly were identified and retained for inclusion in the scale. The fake good scale was derived similarly to the GI. That is, the authors administered protocols under both “fake good” and “answer honestly” instructions. Then, an equation was derived as it best separated the experimentally produced protocols from the normative protocols. Given the method in which these two equations were developed, it is reasonable to hypothesize that they capture other-deception.

The present study will include four scales: the good impression scale, fake good scale, and the SDE and IM scales of the BIDR. The BIDR scales of IM and SDE are hypothesized to capture other-deception and self-deception, respectively. The GI scale and fake good scale are hypothesized to capture other-deception. By using all of these scales, it is expected that it can be ascertained which scales are capable of capturing which type of deception.

External Criteria Measures

Another method used for detecting invalid protocols is external criteria measures. Two types of criteria are typically used: an objective operational criterion or a social judgment criterion. Examples of operational criteria are: talkativeness defined by counting the number of words one speaks in a conversation or sales ability defined by number of units sold (Paulhus & Robins, 2001). The most obvious advantage to this approach is its objectivity. However, on the downside, it is not always possible to find an objective criterion

that will cover all domains of the construct of interest. Furthermore, in some cases, there are simply no objective criteria available (Paulhus & Robins, 2001). Unfortunately, this is generally the case in the measurement of personality. As a result, it's often necessary to use some sort of social judgment criterion in order to have a measure on which to compare the self-report measure of personality. Several types of social judgment criteria are frequently used including (but not limited to): (a) informed observers and (b) trained raters. In an effort to balance the strengths and weaknesses of each approach, both of these methods will be used in the present study.

Preferably, informed observers are those individuals who have had frequent and prolonged contact with the target individual (e.g., friends, significant others, parents, peers, children) and therefore have numerous observations of behavior across a variety of naturalistic settings. Research suggests that informed observer ratings can add incremental validity to self-report measures, perhaps because outside observers can offer a more realistic view of the target individual. For example, Mount, Barrick, and Strauss (1994) found that outside ratings of personality by supervisors, coworkers, and customers added incremental validity to self-report measures in the prediction of performance ratings.

Despite the advantages of using informed raters, one frequently cited criticism is that these observers are typically poorly equipped to rate others and may be more likely to fall prey to common rater biases such as leniency bias and halo effect. Furthermore, they may typically only see the target individual within a small domain of their full behavior (e.g., a supervisor only seeing the behavior exhibited at work). In an effort to circumvent these biases, these ratings will be used in conjunction with ratings provided by trained observers.

In the present study, trained raters will engage participants in a scripted interaction. Throughout the interaction, the rater will systematically collect and record information and then make ratings based on this information. There are two main advantages to using this approach. The first advantage is uniformity. All target individuals are observed in the same controlled environment. The second advantage is that the same trained raters are used to make ratings. As such, there is consistency across ratings and less likelihood of the occurrence of common biases. On the downside, these are not ratings obtained in a naturalistic setting. However, as previously mentioned, these ratings, combined with information obtained from informed observers should provide insight into some aspect of social reality.

Developing Categories

Based on these research findings, it is possible to hypothesize a data pattern for the different types of deception. The following describes the expected pattern of outcomes for those expected to be self-deceptors (i.e., those engaging primarily in self-deception), other-deceptors (i.e., those engaging primarily in other-deception), non-deceptors (i.e., those who are not engaging in any type of deception), and a mixed category (i.e., those who score high in both self-deception and other-deception):

Category 1: Self-deceptor. Self-deceptors will be identified through the use of social desirability scales, a directed faking study, and comparisons of self-report data to ratings made by trained raters and outside observers. One social desirability scale that will be used to identify self-deceptors will be the SDE scale of the BIDR (i.e., those individuals who answer affirmatively to 7 of the items on the SDE scale). The other social desirability scales that will be used are the GI scale and the fake-good scale of the CPI. As previously discussed,

these scales are purported to capture only other-deception (given that they were developed using a “directed faking” approach), self-deceptors will not surpass the threshold scores on each of these scales. Based on previous research, it is also expected that self-deception will remain constant across situations while other-deception is contextually bound. Therefore, it is expected that in a directed faking study, personality scores of self-deceptors will remain constant from Time 1 (i.e., answer honestly) to Time 2 (i.e., answer as if you are applying for a job). In this study, the personality inventory that will be used in the two different administrations is the Goldberg Inventory. It is also expected that those individuals who are engaging in self-deception perceive themselves as more favorable than others perceive them. As a result, it is expected that their self-ratings would be higher than the ratings provided by others.

It is also hypothesized that individuals who engage primarily in self-deception will have a pattern of results consistent with previous research on self-deceptive enhancement (e.g., Taylor & Brown 1988; Schier & Carver, 1985; Holden & Fekken, 1989; Linden, Paulhus, & Dobsen, 1986; Paulhus & Reid, 1991; Bandura, 1989). That is, it is expected that self-deceptors will be rated by others as higher in self-esteem, adjustment, general well-being, and integrity than other-deceptors. Furthermore, it is expected that self-deceptors will be rated by others as higher in self-esteem, adjustment, and general well-being than non-deceptors and that those individuals high in self-deception will be rated as lower in interpersonal skills than other-deceptors.

As a summary:

Category 1: Self-deceptors will be identified using the following decision rules:

- a) Pass the threshold level on the SDE scale of the BIDR only (i.e., those who answer affirmatively to seven or more of the twenty items on the SDE scale of the BIDR and do not exceed a score of fourteen on the IM scale of the BIDR).
- b) The good impression scale and “fake good” scales will not surpass the threshold level (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) Personality data received at Time 1 (i.e., answer honestly) and Time 2 (i.e., answer as if applying for a job) are not different.
- d) Self-reports on the Goldberg inventory will be higher than those reports by informed observers and trained raters in the role play.

The hypothesized outcome data that is expected for self-deceptors is as follows:

- a) Self-deceptors will be rated significantly higher in adjustment than other-deceptors and non-deceptors.
- b) Self-deceptors will be rated significantly higher in general well-being than other-deceptors and non-deceptors.
- c) Self-deceptors will be rated significantly higher in integrity than those identified as other-deceptors and mixed-deceptors.
- d) Self-deceptors will be rated significantly lower in interpersonal skills than those individuals rated as other-deceptors and mixed-deceptors.

Category 2: Other-Deceptors. Other-deceptors will be identified using the same methods used to identify self-deceptors. That is, other-deceptors will be identified through the use of social desirability scales, a directed faking study, and comparisons of self-report data to others' ratings. Other-deceptors are expected to be those individuals who surpass the threshold scores of the IM scale (of the BIDR) and the fake good scale and the GI scale (of the CPI). It is also expected that these individuals will have a change in personality scores from T1 (i.e., answer honestly) to T2 (i.e., answer as if you are applying for a job). That is, they will present a higher socially desirable profile at T2 than at T1. Finally, at T2, as individuals are consciously trying to present a favorable image of themselves, the other-deceptors should present self-report personality profiles that are more favorable than those reported by others.

Outcome data for other-deceptors is expected to be consistent with the outcomes expected for those individuals who engage in impression management. That is, based on previous research (e.g., Gilmore & Ferris, 1989; Kacmer et al., 1992; Wayne & Liden, 1995; Rosse, Stecher, Miller & Levin, 1998; Feldman, Forrest, & Happ, 20023; Kasher & DePaulo, 1996), it is expected that other-deceptors will be reported as being higher on interpersonal skills than those categorized as self-deceptors. It is also expected that they will be rated lower in integrity, adjustment, self-esteem, and general well-being than self-deceptors.

As a summary:

Category 2: Other-deceptors will be identified using the following decision rules:

- a) Those who surpass the threshold level on the IM scale of the BIDR only (i.e., respond affirmatively to at least fourteen out of twenty of the items designed to measure other-deception).
- b) The social desirability scale and “fake good” scale will surpass the threshold level (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) A change in personality scores will be observed between T1 (answer honestly) and T2 (answer as if you are applying for a job).
- d) At T2, the self-reports on the Goldberg inventory will be higher than the ratings provided by others. At T1, the self-reports will be consistent with the ratings provided by others.

Outcome data patterns expected are as follows:

- a) Other-deceptors will be rated significantly lower in adjustment, general well being, and integrity than self-deceptors.
- b) Other-deceptors will be rated significantly higher in social skills than self-deceptors and non-deceptors.
- c) Other-deceptors will be have significantly lower levels of integrity than non-deceptors.

Category 3: Non-deceptors. Non-deceptors will be identified as those individuals who one would expect to have good self-insight and do not engage in conscious other-deception. That is, they are aware of their own personalities and do not intentionally take action to distort their personality data. As such, it is expected that they would not surpass the

threshold score on the SDE scale or the IM scale of the BIDR. Furthermore, it is expected that the social desirability scale and the fake good scales should not be triggered. It is also expected that they would present themselves the same if they were answering honestly and if they were applying for a job. Finally, self-reports should not be different than others' reports of them.

Non-deceptors are those individuals who have a realistic view of themselves and do not consciously try to manage the impression they make. As a result, the outcomes expected are consistent with those expected who have similar tendencies in terms of seeing themselves realistically. Based on previous research (e.g., Abramson, Seligman & Teasdale, 1978; Coyne & Gotlieb, 1983; Lewinsohn, Mischel, & Barton, 1980; Brown, 1986), it is expected that non-deceptors will be lower in self-esteem and well-being than self-deceptors. They will have higher integrity and lower levels of interpersonal skills in comparison to other-deceptors.

Category 3: Non-deceptors will be identified using the following decision rules:

- a) Those who do not trigger either the IM scale or SDE scale on the BIDR.
- b) The good impression scale and "fake good" scale will not surpass the threshold levels (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) The scores on the Goldberg Inventory will not be different between Time 1 (i.e., answer honestly) and Time 2 administration (i.e., answer as if you are applying for a job).
- d) Self-reports on the Goldberg Inventory will be consistent with those reported by the outside observers.

Outcome data expected:

- a) Non-deceptors will have significantly lower levels of adjustment and general well-being than self-deceptors and mixed-deceptors.
- b) Non-deceptors will have significantly higher integrity than other-deceptors and mixed-deceptors.
- c) Non-deceptors will have significantly lower levels of interpersonal skills than other-deceptors and mixed-deceptors.

Category 4: Mixed. A fourth and final category may emerge that for the purpose of the present study, will be called the “mixed category.” It will contain those individuals who display both self-deception and other-deception tendencies. Individuals in the mixed category will be those who trigger the SDE scale of the BIDR (due to the self-deception component), the GI scale and the fake good scale will surpass the thresholds and T1 data will be less than T2 data (due to the other-deception component), and Goldberg inventory self-reports will be higher than other reports (due to both self-deception and other-deception).

Mixed deceptors will be identified using the following decision rules:

- a) Trigger both the SDE scale and the IM scale on the BIDR.
- b) The GI scale and the fake good scale will both surpass the threshold level (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) At the T2 administration, self-report profiles will be higher than at the T1 administration.
- d) Goldberg inventory self-report data will be higher than reports from outside observers and trained raters at both T1 and T2.

The following outcomes are expected:

- a) Those in the mixed category will have significantly higher levels of adjustment and well-being in comparison to those in the no-deception and other-deception only categories.
- b) Those in the mixed category will have significantly lower integrity than those in the self-deception category and non-deception category.
- c) Those in the mixed category will have significantly higher levels of interpersonal skills in the self-deception and no-deception categories.

Method

Participants

A total of 320 participants were recruited through business school courses at a Midwestern university. Of these participants, 242 (76%) completed all four aspects of the study (i.e., they responded to the Goldberg Inventory under “answer honestly” and “answer

as if you are applying for a job” conditions, took part in the role play, and submitted information from an outside observer). Given that the main goal of this study was to examine the role of multiple indicators of deception simultaneously, subsequent analyses included only those 242 participants with complete data.

Of these 242 participants, 124 (51%) were men and 107 (44%) were women. The remaining 11 (5%) did not report gender. The sample consisted of 174 (72%) Whites, 25 (10%) African-Americans, 14 (6%) Asians, and 16 (7%) reported other ethnic backgrounds. The remaining 13 (5%) did not indicate race. Participants’ ages ranged from 19 to 56 with a mean age of 24.33 years and a standard deviation of 5.95 years. Of the group, 3 (1%) reported their class status as freshmen, 5 (2%) as sophomores, 120 (50%) as juniors, and 97 (40%) as seniors. Seventeen (7%) did not report year in school.

Measures

Goldberg Inventory

The personality inventory developed by Goldberg (1992) was used to collect self-report data from participants (see Appendix A). Additionally, informed observers made ratings of the participant using this questionnaire. This inventory was created in order to replace the Big Five factor markers created by Norman (1963). The inventory consists of 100 adjectives designed to measure the Big Five personality traits. However, in the interest of time, only those 40 adjectives used to assess extraversion and agreeableness were included in the present study. These scales were retained as they were deemed applicable to the job of customer service agent as well as being observable personality characteristics.

Participants responded on a scale of 1 to 7 to indicate the extent to which each adjective described them (or if they were an informed observer, rated how descriptive each

adjective was of the participant). Goldberg has found correlations among his inventory and other measures of the Big 5 that ranged from .46-.69 (Goldberg, 1992). This scale has also been used in previous research for the purpose of obtaining informant-rated personality. When used for collecting informant ratings, it has been found to add incremental validity in predicting job performance above and beyond self-report data (e.g., Mount, Barrick, & Strauss, 1994).

For the present study, four combined mean scores were computed for each participant. These mean scores were computed for each participant on the self-report ratings of personality in both the “answer honestly” and “answer as if you are applying for a job” conditions, and the ratings that were made by informed observers and the trained role players. These scores were calculated by computing the mean of all responses to the 40 items on the inventory (after re-keying so that each was phrased affirmatively). The scale demonstrated adequate reliability across all four scoring conditions (alphas ranged from .91-.97).

Social Desirability Scales

Good impression scale and fake good scale. The good impression scale and fake good scale of the CPI was given to participants. The good impression scale consists of 40 items to which individuals responded either true or false to indicate whether the item described them. Gough developed these items by using Ruch’s (1942) strategy of testing under normal instructions and then with instructions to fake good. Following the administration, an item analysis showed significant differences in endorsement between the two conditions on 40 of the items. According to the technical manual (Gough, 1996), scores of thirty and above are used to indicate impression management.

The fake good scale is composed of a combination of five of the CPI's folk scales (i.e., dominance, empathy, responsibility, well-being, and flexibility) along with the good impression scale (see Appendix B). According to the technical manual, scores above 60.60 indicate that the respondent is faking the inventory (Gough, 1996). The cut-off score was set in accordance with the findings of research conducted by Lanning (1989). To determine cut-scores for the CPI, the author used 5 archival data sets from the CPI archives and 8 computer-generated data sets to develop regression equations that could detect invalid protocols (i.e., regression equations that could detect fake-good, fake-bad, and random response patterns). Using the framework of signal detection theory, he determined cut-scores and tested these scores using 5 additional data sets. Results suggested that using the cut-scores was an effective means of identifying participants who were faking-good, faking-bad, and responding randomly.

The items on the good impression scale and the fake good scales of the CPI were originally designed to be answered dichotomously (i.e., either true or false). However, for the purpose of consistency with other scales in this study, participants were asked to respond on a 1-7 scale and the items were later dichotomized. It was necessary to return these items to a dichotomy as the equations and cutoff scores created by Gough (1996) were based on dichotomous responses. Therefore, all items for which participants responded 5, 6, or 7 were recoded to 1 and all items rated as 1, 2, 3, or 4 were recoded to 0. The decision was made to adopt this coding strategy as items that were scored as 1, 2, 3, or 4 were those items that participants deemed to either be neutral or not true. A decision was made to include the rating of 4 in the "O" coded group as a rating of "neutral" indicated that the participants did not have a strong feeling as if the item was indicative of them. Those items that were coded

5, 6, or 7 indicated that participants had rated those items as true of themselves. In accordance with the guidelines set forth by Gough, a sum was then computed using the dichotomized items. The scale demonstrated adequate reliability ($\alpha = .87$).

Recall that the fake good scale is composed of five scales from the CPI (i.e., dominance, empathy, good impression, well-being, and flexibility). It was necessary to compute sums on each of these subscales as they were then entered into the fake good scale equation as unit weights (see Appendix B for the equation). A reliability coefficient was not computed for the FG scale as it is derived using a combination of various scale scores.

Balanced Inventory of Desirable Responding. The BIDR was given to participants (see Appendix C). The instrument consists of 40 total statements, 20 of which refer to impression management (IM) and the remaining 20 refer to self-deceptive enhancement (SDE). Participants responded on a seven-point Likert scale (1 = not true of them, 7 = very true of them) to indicate how true the items were for them.

In previous studies, reliabilities for the instrument have ranged from .83-.86 (Paulhus, 1998). In the present study the impression management scale demonstrated adequate reliability ($\alpha = .85$) but a lower reliability coefficient was observed for the SDE scale ($\alpha = .68$).

In previous research, evidence of validity has been found for both scales. For example, the IM scale has been found to be responsive to the demands for impression management. That is, previous research has found that IM scores increase from private to public conditions (Paulhus, 1984; Booth-Kewley, Rosenfield, Edwards, and Alderton, 1992) and that anonymous testing decreases scores on the IM (Paulhus, 1998). The SDE scale has

been linked to a lack of insight into how one is seen by others and over-claiming (Paulhus, 1998).

In the present study, the scales were computed in accordance with the guidelines set by Paulhus (1998). That is, only extreme responses were scored (i.e., 6 or 7 after re-keying). The rationale behind this scoring is that the extreme responses are more likely to indicate distortion (Paulhus, 1984). Therefore, each time a participant indicated an extreme response (i.e., 6 or 7 after re-keying) to an item, they were assigned one point. As such, an individual could receive a score of 0 to 20 for both the IM scale and the SDE scale with 0 being the lowest possible score and 20 being the highest possible score. According to the guidelines set forth by the author, scores in excess of 7 on the SDE scale are indicative of self-deceptive enhancement while scores in excess of 14 on the IM scale are suggestive of impression management.

Outcome Measures

Paulhus scale of adjustment. Adjustment was measured using a four-item scale developed by Paulhus (1998). Informed observers used a seven point scale to (1 = not descriptive of participant, 7 = very descriptive of participant), rate whether the participant is happy, likes himself/herself, is well-adjusted, and is mentally healthy (see Appendix D). Reported reliabilities for this scale range from .68-.80 (Paulhus, 1998) and adequate reliability was found in the present study ($\alpha = .84$). For each participant an average score was computed with each item receiving equal weight.

Satisfaction with Life Scale. Well-being of the participants was assessed by asking informants to complete the Satisfaction with Life Scale (SWLS) about the participant (see Appendix E). The SLWS is a five-item scale designed to measure one's basic satisfaction

with aspects of his/her life. The five items are rated on a 1-7 scale (1= strongly disagree, 7= strongly agree). Due to the original scale being self-report, the items were modified only so that they could be used by informed observers to make ratings of participants.

The SWLS has shown adequate test-retest correlations of .80 and above in the short-term (Diener, Emmons, & Larsen, 1985) and respectable correlations in the longer-term (e.g., .50 in ten weeks and .54 in four years) (Pavot & Diener, 1993) and adequate reliability was found in the present study ($\alpha = .79$). Furthermore, the SLWS has been shown to demonstrate convergent validity with a variety of constructs of interest, including other self-report measures of satisfaction with life satisfaction and well-being (e.g., Pavo, Diener, Colvin, & Sandik, 1981). An average score on satisfaction with life was computed for each participant.

Interpersonal Competence Scale. Informants were asked to rate the interpersonal skills of the participants using Baird and Holland's (1968) Interpersonal Competence Scale (see Appendix F). The scale is based upon previous work by Foote and Cattrell (1955) and consists of 20 items designed to assess basic interpersonal skills. An example item is "I have a reputation for being able to cope with difficult people." The items were re-written to be reflective of the participant, rather than the informant. Furthermore, five items were removed that address physical condition (e.g., I have good coordination). Holland and Baird reported an internal consistency (KR-20 because the items were originally designed to be answered either true or false) of .67 for women and .69 for men. The authors do not report any suggested hypotheses for why the differences in reliability across genders was observed. Evidence for construct validity has been found as the scale positively correlates with a variety of constructs including self-confidence, speaking ability, cheerfulness, sensitivity to

others' needs, social competency, and leadership (Holland & Baird, 1968). In the present study, the instrument demonstrated adequate reliability ($\alpha = .84$). An average score was computed on interpersonal skills for each participant with each item receiving equal unit weight.

Informant-rated integrity. Informed observers were asked to rate participants on five items designed by the researcher to capture integrity (see Appendix G). Informed observers were asked to respond to the items on a 1-7 Likert scale (1 = not descriptive of the participant, 7 = very descriptive of the participant). The scale was found to demonstrate adequate reliability ($\alpha = .82$). For each participant, an average score was computed on integrity.

Design and Procedure

This procedure section consists of two parts. The first part of this section is an in-depth description of how each traditionally-used “faking” study (i.e., directed faking study, social desirability scales, and external criteria measures) was conducted in the current research design. The second part of this section is a design overview and is provided to give the reader a broad overview of the data collection processes. It also consists of a sequential summary of the steps involved. Providing these two sections separately allows the reader to gain a better understanding of how each type of study was conducted and an overview of the steps involved.

Types of studies

Directed Faking Study

The directed faking study was completed in two administrations, taking place approximately two weeks apart. The administration instructions were

counterbalanced in an effort to eliminate any possible order effects. It should be noted that order effects were observed for the difference scores between participant responses in the “answer honestly” administration and the “answer as if you are apply for a job” administration.

In the “answer honestly” administration, participants were given the following instruction set before taking the extraversion and agreeableness scales of the Goldberg inventory:

“As you are completing these materials, imagine that you are being assessed for your own self-knowledge and that your results will be returned to you in an anonymous manner. Therefore, it is in your own best interest to answer these questions honestly and gain the most self-insight possible from the experience.”

In the “answer as if you are applying for a job” administration, participants completed the scales of the Goldberg inventory as well as the CPI’s fake good and good impression scales and the BIDR under the following instruction set:

“I’d like you to imagine you are applying for a job. As part of the selection process you were asked to respond to this personality questionnaire. Keep in mind that this is part of the selection process.”

Social Desirability Scales

The fake good scale, the good impression scale, and the Balanced Inventory of Desirable Responding were administered in the “answer as if you are applying for a job” condition. It was decided to administer these scales under this instruction set rather than the “answer honestly” condition because under this administration, participants presumably have a motivation to consciously distort their responses, therefore, giving the scales the opportunity to detect other-deception.

External Criteria Ratings

Informed observers. Informed observers were identified by all participants in the first administration (regardless of whether it was the “answer honestly” condition or “answer as if you are applying for a job” condition). Participants were asked to choose someone who met the following criteria: someone whom they felt “knows them well”, has known them for at least one year, and someone with whom they interact at least once per week. Participants were given a packet to give their observer. These observers were asked to complete the following items to describe the participant: extraversion and agreeableness scales of the Goldberg inventory, Paulhus’ scale of adjustment, Satisfaction with Life Survey, Interpersonal Competence Scale, and the integrity scale designed by the researcher. They were asked to return the data in an envelope addressed to the researcher.

Trained raters. Participants took part in a 10-minute scripted interaction with one of two trained role players (see Appendix H). Prior to beginning the study, the role players discussed and practiced the script to ensure that the roles would be played similarly. In order to ensure consistency, the role players also calibrated on scoring and role playing several times throughout the data collection process. Following the completion of the study, an independent sample t-test indicated that the scores reported by Trained Rater #1 ($M = 4.50$, $SD = .92$) did not differ significantly from the scores reported by Trained Rater #2 ($M = 4.55$, $SD = 1.01$), $t(233) = .45$, $p > .05$.

In the role play, the participant played the role of a customer service agent while the trained role player adopted the role of an angry customer. Following the completion of the role play, the role player made ratings of the participant’s extraversion and agreeableness using the Goldberg Inventory scales of extraversion and agreeableness.

Design overview

Undergraduate students were invited to take part in this study through recruitment efforts in undergraduate business courses. Participants were told that researchers were interested in studying the personality characteristics of undergraduate students. In all classes, instructors offered extra credit points in return for student participation. The study was conducted using a two-part data collection process, with the administration sessions spaced approximately two weeks apart. In addition, the manipulations were counterbalanced in an effort to avoid potential order effects with 117 participants receiving the “answer honestly” instructions in the first administration and 124 participants receiving the “answer as if you are applying for a job” instructions in the first administration. A one-way ANOVA revealed that order effects were present. Participants receiving “answer honestly” instructions first had larger difference scores ($M = .13$, $SD = .48$) than those who received the “answer as if you are applying for a job” instructions first ($M = .003$, $SD .40$) $F(1, 239) = 5.33$, $p < .05$, $\eta^2 = .02$.

In the first administration, all participants were asked to complete informed consent forms, report demographic information, and were given a packet of information to distribute to their informed raters. One-half of the participants (hereafter Group 1) were asked to complete the Goldberg inventory scales of agreeableness and extraversion under “answer honestly” instructions. The other half of the participants (hereafter Group 2) were given the scales on the Goldberg inventory and asked to respond to under “answer as if you are applying for a job” instructions. Group 2 participants were also asked to complete the BIDR, fake good scale, and good impression scale at this time.

In the second administration, all participants were given a verbal reminder to ensure that their informed observers completed the materials in the packet and returned them to the researcher. Group 1 participants were asked to complete the Goldberg inventory scales under “answer as if you are applying for a job” instructions. They were also asked to complete the BIDR, fake good, and good impression scales. Group 2 participants completed the Goldberg inventory under “answer honestly” instructions. Additionally, all participants took part in a 10-minute role-play with a trained rater.

The following Table provides a summary of the multiple steps that were involved in the data collection process:

Administration	Group 1	Group 2
Time 1	<ul style="list-style-type: none"> • Informed consent • Demographic information • “Answer Honestly” instruction • Goldberg inventory • Distribution and instruction on informant packages 	<ul style="list-style-type: none"> • Informed consent • Demographic information • “Answer as if applying for job” instruction • Goldberg inventory • BIDR, fake good, and good impression inventories • Distribution and instruction on informant packages
Time 2	<ul style="list-style-type: none"> • “Answer as if applying for job” instructions • Goldberg inventory • BIDR, Fake good, and good impression inventories • Reminder of informant ratings • Role play • Debrief form 	<ul style="list-style-type: none"> • “Answer honestly” instruction • Goldberg inventory • Reminder of informant ratings • Role play • Debrief form

Results

Descriptive Analyses

Overall means, standard deviations, internal consistency coefficients, and correlations were computed for all scales, and are shown in Table 1. As can be seen in Table 1, although all 242 participants took part in all aspects of the study, they did not always fully complete the surveys. As a result, several of the scale scores have sample sizes that are less than 242 (i.e., good impression, fake good scale, adjustment, integrity, interpersonal skills, and satisfaction with life). When any data were missing from the scales, participants were excluded from any analyses involving those instruments. Sample sizes for the analyses ranged from 226-242.

Computation of Difference Scores

Five difference scores were computed for each participant on the Goldberg Inventory across the four different administrations of the scales (i.e., “answer honestly”, “answer as if you are applying for a job”, informed observer and role play). One difference score was computed for each participant by subtracting individual scores in the “answer honestly” condition from individual scores in the “answer as if applying for a job” condition. Difference scores were also computed for the following combinations: apply for a job-role play, apply for a job-informed other, answer honestly -role play, and answer honestly-informed other (see Table 1).

Manipulation Check

In an effort to understand whether the instruction manipulation was successful, it was necessary to compare the group in the “answer honestly” condition to the group scores in the “answer as if you are applying for a job” condition. According to t-test results, scores were

significantly higher in the “answer as if you are applying for a job” condition ($M = 5.46$, $SD = .63$) than they were in the “answer honestly” condition ($M = 5.39$, $SD = .61$) $t(241) = 2.32$, $p < .05$, suggesting that the manipulation was successful.

Categorization

Data analysis. The following constraints were used to place participants into the category of self-deceptor, other-deceptor, non-deceptor, or mixed-deceptor. If participants did not fit the following constraints, they would be placed into an “other” category.

Category 1: Those who fit the following criteria were coded as self-deceptors:

- a) Pass the threshold level on the SDE scale of the BIDR only. (i.e., those who answer affirmatively to seven or more of the twenty items on the SDE scale of the BIDR).
- b) The good impression scale and “fake good” scales will not surpass the threshold level (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) Personality data received at Time 1 (i.e., answer honestly) and Time 2 (i.e., answer as if applying for a job) are not different.
- d) Self-reports on the Goldberg inventory will be higher than those reports by informed observers and trained raters in the role play.

Category 2: Those who fit the following criteria were coded as other-deceptors.

- a) Those who surpass the threshold level on the IM scale of the BIDR only (i.e., respond affirmatively to at least fourteen out of twenty of the items designed to measure other-deception).
- b) The social desirability scale and “fake good” scale will surpass the threshold level (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60). In the T2 condition (i.e., answer honestly condition).
- c) A change in mean personality scores will be observed between T1 (answer honestly) and T2 (answer as if you are applying for a job).
- d) At T2, the self-reports on the Goldberg inventory will be higher than the ratings provided by others. At T1, the self-reports will be consistent with the ratings provided by others.

Category 3: Those who fit the following criteria were coded as non-deceptors:

- a) Those who do not trigger either IM scale or SDE scale on the BIDR.
- b) The good impression scale and “fake good” scale will not surpass the threshold levels (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) The mean scores on the Goldberg Inventory will not differ between Time 1 (i.e., answer honestly) and Time 2 administration (i.e., answer as if you are applying for a job).
- d) Self-reports on the Goldberg inventory will be consistent with those reported by the outside observers.

Category 4: Those who fit the following criteria were coded as mixed-deceptors:

- a) Trigger both the SDE scale and the IM scale on the BIDR.
- b) The GI scale and the fake good scale will both surpass the threshold level (i.e., the GI scale being above 30 and the fake good scale exceeding 60.60).
- c) At the T2 administration, self-reports on the Goldberg inventory will be higher than at the T1 administration.
- d) Goldberg inventory self-report data will be higher than reports from outside observers and trained raters at both T1 and T2.

Category 5: Those participants that did not fit into any of the above-mentioned groups will be placed into an “other” category.

Using the above-mentioned criteria, no participants were included in the self-deceptor, other-deceptor, and mixed deceptor groups ($n = 0$). However, 93 participants met the criteria for inclusion in the non-deceptor group. All remaining participants were then placed into Category 5 and were denoted as “some deception” group ($n = 146$). By definition, those in the non-deceptor group were found to be free from engaging in any type of deception while those in the “some deception” group were identified by one or more criteria as having engaged in some type of deception. As such, an examination of the differences between these two groups of participants was conducted.

Means, standard deviations, reliability coefficients, and correlations were computed for the two separate groups (see Table 3, 4, 5, and 6). Four one-way ANOVAs were computed to determine whether these two groups demonstrated any significant differences on

the outcome variables (see Table 7). No significant differences were found on scores of adjustment $F(1, 237) = .01, p > .05, \eta^2 = .00$, integrity $F(1, 237) = .38, p > .05, \eta^2 = .00$, interpersonal competency $F(1, 237) = .42, p > .05, \eta^2 = .00$, or satisfaction with life $F(1, 237) = .24, p > .05, \eta^2 = .00$.

Alternative 2: Loosening the Original Constraints

Given that the objective of this study was to compare different methods of assessing deception and better understand potential outcomes of self-deception and other-deception, it was necessary to conduct additional analyses. One such analysis was an attempt to relax the original constraints (e.g., lowering the IM score requirement) to determine whether there were any other reasonable methods for identifying self-deceptors, other-deceptors, mixed deceptors, and non-deceptors. However, this method was also ineffective in identifying different types of deceptors.

Alternative 3: Analysis by Type of Study

Given that the aforementioned strategies did not result in categorizing participants across the different study types, analyses were conducted separately by the four individual study types as a means of understanding the data. However, it should be noted that in many of the studies below, the groups consistently contained unequal sample sizes. This unequal distribution among the groups led to lower power, lessening the likelihood of finding significant differences among the groups even if they did exist. In some cases, the distribution was so unequal that it was deemed inappropriate to conduct further analyses. However, where appropriate, two separate analyses were conducted for each different type of study. In the first analysis, the original constraints were used to form groups for comparisons purposes. In the second analysis, a mean cut-score approach was used. That is, those who

scored above the mean were included in Group 1 and those who scored below the mean were included in Group 2.

All difference scores were analyzed using the two aforementioned approaches as well as a third method. The third approach was conducted in order to gain an understanding of whether the direction of difference scores was related to the outcome variables. That is, it was important to understand whether there were any outcome differences for the participants who rated themselves higher than they were rated by others compared to those who rated themselves lower than they were rated by others. In the directed faking study, analyses were conducted to determine whether there were any differences in outcomes dependent upon whether participants rated themselves higher in the “answer honestly” condition compared to the “answer as if you are applying for a job” condition. Additionally, for all participants, comparisons were made between the scores obtained in the “apply for a job” condition and scores from the role play, the “apply for a job condition” and scores from informed others, “answer honestly” condition and scores from the role play and finally the differences between the “answer honestly” condition and the informed others. For each of these comparisons, one-way ANOVAs were computed using the cut-score as 0 to separate participants into two different groups: those that had a difference score above 0 and those that had a difference score that was less than 0. Participants who had a difference score of 0 (i.e., their ratings matched the ratings given by others perfectly) were dropped from subsequent analyses. The decision was made to drop these participants as there were not enough individuals to comprise a group that would be large enough to effectively compare against the other groups.

Directed faking study. The original hypothesis of this study was that the difference scores between the “answer honestly” and the “answer as if you are applying for a job” conditions were suggestive of other-deception. The original constraints were that difference scores in excess of 2 were suggestive of other-deception (i.e., participants who raised their scores by 2 or more points in the “answer as if you are applying for a job” condition). However, in the present study, there were no participants who raised their score by 2 or more points. As a result, no analysis were computed using the original constraints.

In order to use the mean cut-score approach, it was necessary to compute a mean for the difference scores ($\underline{M} = .06$, $\underline{SD} = .43$). Using the mean as a cut-score approach, 95 participants scored above the mean while 145 scored below the mean. No participants had a difference of exactly .06. A one-way ANOVA was computed using each of the outcome variables as dependent variables and whether participants scored above or below the mean as the grouping variable (see Table 8). Using this approach, no significant results were observed on reported levels of adjustment, integrity or satisfaction with life (see Table 8). Significant differences were observed for the outcome variable of interpersonal skills $F(1, 238) = 4.41$, $p < .05$, $\eta^2 = .02$, with participants with mean difference score greater than .06 receiving higher interpersonal competency ratings ($\underline{M} = 5.27$, $\underline{SD} = .77$) than those with difference scores less than .06 ($\underline{M} = 5.05$, $\underline{SD} = .81$).

A frequency analysis revealed that 53% of participants actually rated themselves lower in the “answer as if you are applying for a job” condition than they did in the “answer honestly” condition. Three participants (1%) rated themselves the same in the “answer honestly” condition as they did in the “answer as if you are applying for a job” condition (i.e., they had a difference score of 0). The remaining 46% rated themselves higher in the

“answer as if you are applying for a job” condition than they did in the “answer honestly” condition. The three participants who rated themselves the same across the two conditions were dropped from the analysis and four one-way ANOVAs were computed using whether difference scores were greater than or less than 0 as the grouping variable. No significant differences were observed for adjustment, integrity, interpersonal skills, or satisfaction with life.

Social Desirability Scales

GI scale. Of the 242 participants, only 9 surpassed the cutoff of 30 that was suggested by Gough (1996) as indicative of impression management. Given that using this cut-off score would lead to a very unequal distribution across the two groups and that one group would have only 9 participants, it was deemed inappropriate to conduct group comparisons.

The mean score for the GI scale was computed ($\underline{M} = 23.81$, $\underline{SD} = 6.61$). Using the mean cut-score approach, 121 participants were found to have rated themselves higher than the mean while 118 participants scored themselves lower than the mean on the GI. No participants scored exactly 23.81. No significant results were observed for satisfaction with life and interpersonal skills (see Table 9). However, significant results were obtained for adjustment $F(1,237) = 5.67$, $p < .05$, $\eta^2 = .02$, and integrity $F(1,237) = 4.05$, $p < .05$, $\eta^2 = .02$. Participants who scored higher than the mean on the GI were rated as having higher levels of adjustment ($\underline{M} = 6.00$, $\underline{SD} = .90$) than those participants who scored themselves lower than the mean ($\underline{M} = 5.73$, $\underline{SD} = .88$). Similarly, participants who scored higher on the GI were rated as having higher levels of integrity ($\underline{M} = 6.40$, $\underline{SD} = .84$) than those who scored lower than the mean on the GI ($\underline{M} = 6.17$, $\underline{SD} = .92$).

Fake good scale. In the present study, only 3 of the 242 participants surpassed the cutoff level of 60.60 that was deemed by Gough (1996) as indicative that the participant is inflating scores on the CPI. Given the unequal distribution across the two groups and the fact that one group would consist of only 3 participants, it was deemed inappropriate to conduct analyses using the original constraints.

Using the mean cut-score approach ($\underline{M} = 56.14$, $\underline{SD} = 4.05$), 111 participants were placed into the above the mean group while 115 participants were placed into the below the mean group (see Table 10). No participants scored directly at the mean. Significant results were not obtained for integrity or satisfaction with life. Significant results were obtained for the outcome variables of adjustment $F(1, 224) = 4.26$, $p < .05$, $\eta^2 = .02$, and interpersonal skills $F(1, 224) = 6.65$, $p < .05$, $\eta^2 = .03$. Participants who scored higher than the mean on the FG scale were rated as having higher levels of adjustment ($\underline{M} = 6.00$, $\underline{SD} = .87$) than those who scored below the mean ($\underline{M} = 5.76$, $\underline{SD} = .91$). Participants who scored higher than the mean were rated as having higher levels of interpersonal skills ($\underline{M} = 5.28$, $\underline{SD} = .76$) than those who scored lower than the mean ($\underline{M} = 5.01$, $\underline{SD} = .81$).

Self-deceptive enhancement scale. Of the group of participants, 81 reported SDE scores that were in excess of 7, the score set by Paulhus as being indicative of self-deceptive enhancement (see Table 11). The rest of the group ($n = 159$) reported results on the SDE scale that were less than or equal to 7. Participants were placed into one of two groups: those who had an SDE score greater than 7 and those who had an SDE score that was less than or equal to 7. ANOVAs were conducted to determine whether significant differences existed between the two groups. Significant results were not obtained for adjustment, integrity, or interpersonal skills. Significant results were obtained for satisfaction with life $F(1, 238) =$

7.15, $p < .05$, $\eta^2 = .03$, with those participants who scored themselves lower than 7 on the SDE scale ($M = 4.93$, $SD = 1.01$) being rated as having higher levels of satisfaction with life than those who scored themselves higher than the mean ($M = 4.55$, $SD = 1.10$).

Using the mean cut-score approach ($M = 4.56$, $SD = .60$), participants were sorted into two groups, those who scored themselves above the mean on the SDE ($n = 114$) and those who scored themselves below the mean ($n = 126$). No participants scored exactly the same as the mean. A one-way ANOVA revealed no significant results on adjustment, integrity, or interpersonal skills. Significant results were obtained for satisfaction with life $F(1, 238) = 4.02$, $p > .05$, $\eta^2 = .02$, with participants reporting higher levels of SDE being rated as having lower levels of satisfaction with life ($M = 4.65$, $SD = 1.09$) than those participants reporting levels of SDE below the mean ($M = 4.93$, $SD = 1.00$).

Impression management scale. Using the scores set forth by Paulhus as indicators of other-deception (i.e., IM scores of more than 14), only 14 participants were identified as impression managers. Given the small sample size that would be present in the group identified as impression managers and the unequal distribution across the two groups, it was deemed to be inappropriate to conduct comparisons using the original constraints.

Using the mean cut-score approach ($M = 4.13$, $SD = .90$) participants were sorted into two groups: those with scores above the mean ($n = 123$) and those with scores below the mean ($n = 117$). No participants scored directly at the mean. No significant results were obtained for adjustment, integrity, satisfaction with life or interpersonal skills (see Table 12).

Apply for a job and role play score. The first difference scores analyzed were the differences in self-report in the “apply for a job” condition and how participants were perceived by the trained raters in the role play (see Table 13). For each participant, the score

obtained in the role play condition was subtracted from the self-report score in the “answer as if you are applying for a job condition”. Using the original constraints (i.e., that participants much score themselves at least two points higher on the Goldberg Inventory than they were rated in the role play), participants were sorted into two categories: those who rated themselves two or more points higher in the “answer as if you are apply for a job” condition ($n = 37$) being placed into one category while those who did not surpass this threshold ($n = 203$) were placed into the second category. No participant had a difference score of exactly two. Four one-way ANOVAs revealed no differences for adjustment, integrity, interpersonal skills, or satisfaction with life.

A mean cut –score approach was also taken ($M = .93$, $SD = 1.01$). Of the group, 116 participants had difference scores that were above the mean while 124 had difference scores that were below the mean. No participant had a difference score of exactly .93. No significant results were obtained for adjustment, integrity, interpersonal skills, or satisfaction with life.

Of the group, 201 (84%) scored themselves higher in the “answer as if you are applying for a job” condition than they were scored by the role players, while 38 (16%) scored themselves lower in the “answer as if you are applying for a job” condition than they were rated by role players. It should be noted that one participant was dropped from this analysis as the score reported by the role players was the exact same score reported by the participant. Three one-way ANOVAs revealed no significant differences in adjustment, integrity, or satisfaction with life (see Table 13). However, significant differences were observed in rating of interpersonal skills $F(1, 237) = 4.05$, $p < .05$, $\eta^2 = .02$ with those participants who scored themselves higher than they were rated in the role play receiving

higher ratings on interpersonal skills ($\underline{M} = 5.18$, $\underline{SD} = .79$) than those individuals who rated themselves lower than they were rated in the role play ($\underline{M} = 4.90$, $\underline{SD} = .81$).

Apply for a job and informed other score. The second difference scores computed were the differences between participants' self reports in the "apply for a job" condition versus how they were rated by their informed observers (see Table 14). Participants were divided into two groups: those who scored themselves at least 2 points higher than they scored themselves in the "answer as if you are applying for a job condition" than they were rated by their informed observer. However, only three participants rated themselves two or more points higher than they were rated by their informed observers. As a result, no further analysis were computed using the original constraints.

The mean difference between the two conditions was computed to be $-.11$. That is, on average, participants scored themselves $.11$ lower than they were scored by their informed observers. Using this mean, participants were sorted into two groups: those who had a difference score greater than the mean ($\underline{n} = 113$) and those who had a difference score less than the mean ($\underline{n} = 127$). Four one-way ANOVAs were computed comparing the two groups on the four outcome variables. Significant results were obtained for adjustment $F(1, 238) = 34.78$, $p < .0001$, $\eta^2 = .13$, integrity $F(1, 238) = 5.77$, $p < .05$, $\eta^2 = .02$, interpersonal skills $F(1, 238) = 45.38$, $p < .001$, $\eta^2 = .16$, and satisfaction with life $F(1, 238) = 24.22$, $p < .001$, $\eta^2 = .09$. For all outcome variables, those participants who rated themselves higher than the mean difference were rated lower on all outcome variables.

Of the group, 43% rated themselves higher than they were rated by their informed observers. The remaining 57% rated themselves lower than they were rated by their informed observers. No participants rated themselves exactly the same as they had been rated by their

informed observer. Four one-way ANOVAs were computed using whether participants scored themselves higher or lower than their informed observers as the grouping variable and the four outcome variables as dependent variables. Significant results were obtained for adjustment $F(1, 238) = 47.66, p < .0001, \eta^2 = .17$, integrity $F(1, 238) = 4.08, p < .05, \eta^2 = .02$, interpersonal skills $F(1, 238) = 44.72, p < .0001, \eta^2 = .16$, and satisfaction with life $F(1, 238) = 30.84, p < .001, \eta^2 = .12$. Those participants who rated themselves higher on the Goldberg Inventory than they were rated by informed observers were rated as having lower levels of adjustment, integrity, interpersonal skills, and satisfaction with life than those who saw themselves less favorably than their informed observers.

Answer honestly scores and role play scores. Difference scores were obtained by subtracting the average role play scores from the self-report scores that were reported in the “answer honestly condition”. Out of the sample, 32 participants had a two or more point difference in scores while the remaining 207 participants did not have a two point difference in scores. Four one way ANOVAs were computed. Significant results were not observed for adjustment, integrity, and satisfaction with life. However significant results were obtained for interpersonal skills $F(1, 237) = 6.72, p < .05, \eta^2 = .03$ with those participants who had larger difference scores ($M = 5.47, SD = .68$) being seen as having higher levels of interpersonal skills than those who had lower reported difference scores ($M = 5.08, SD = .81$).

Using the mean cut-score approach ($M = .87, SD = 1.02$), 121 participants of the sample had difference scores that were below the mean difference score while 119 participants had scores that were above the mean group difference score. Four one-way ANOVAs were computed and no significant differences were found for adjustment, integrity, interpersonal skills or satisfaction with life (see Table 15).

Participants were then divided into two groups, those who scored themselves higher than they were rated by the role players ($n = 194$) and those who rated themselves lower than they were rated by the trained role players ($n = 45$). Four one-way ANOVAs revealed no significant differences among the groups on adjustment, integrity, interpersonal skills, or satisfaction with life (see Table 15).

Answer honestly and informed others score. Difference scores were computed by subtracting the scores obtained in the informed other condition from the self-report scores in the answer honestly condition (see Table 16). Using the original constraints (i.e., the difference score must be two or greater) only 3 participants were included in the group that met the original constraints. As a result, no analysis was conducted using the difference score of 2 as the grouping factor.

Using the mean cut-score approach ($M = -.18$, $SD = .87$), participants were divided into two groups: those who had a difference score greater than the mean ($n = 110$) and those who had a difference score that was less than the mean ($n = 130$). No participants had a difference score of exactly $-.18$. Four one-way ANOVAs revealed significant differences for adjustment $F(1, 238) = 45.09$, $p < .001$, $\eta^2 = .16$, interpersonal skills $F(1, 238) = 44.43$, $p < .001$, $\eta^2 = .16$, and satisfaction with life $F(1, 238) = 18.89$, $p < .01$, $\eta^2 = .07$. Participants who had difference scores that were greater than $-.18$ were rated as having lower levels of adjustment ($M = 5.48$, $SD = .93$) than those who had difference scores that were less than $-.18$ ($M = 6.19$, $SD = .72$). Similarly, those participants who had a difference scores that were greater than $-.18$ were rated as having lower levels of satisfaction with life ($M = 4.49$, $SD = 1.10$) than those who had difference scores that were less than $-.18$ ($M = 5.06$, $SD = .94$). Significant results were not obtained for integrity.

Next, participants were grouped into those who scored themselves higher than their informed observers ($n = 89$) and those who score themselves lower than they were scored by their informed observers ($n = 147$). Four people were dropped from this analysis as they had scores that were exactly that same as they were rated by their informed observers. Four one-way ANOVAs revealed significant results for adjustment $F(1, 234) = 44.20, p < .001, \eta^2 = .16$, integrity $F(1, 234) = 4.90, p < .01, \eta^2 = .02$, interpersonal skills $F(1, 234) = 42.69, p < .001, \eta^2 = .15$ and satisfaction with life $F(1, 234) = 24.61, p < .001, \eta^2 = .10$.

Alternative 4: Using Regression

Using ANOVAs dictates that participants must be divided into different groups for comparison purposes. While valuable, this approach leads to a loss of variability in the indicators of deception (e.g., a participant who scores 1 point below the mean on a social desirability scale is treated the same as a participant who scores 1.5 points below the mean). In an effort to take advantage of the full range of scores of the predictor variables, two regression analyses were computed: one to assess whether hypothesized indicators of self-deception would predict any of the four outcome variables and a second to assess whether hypothesized indicators of other-deception would predict outcomes. As stated in the original hypotheses of this study, the indicators of self-deception were scores on the SDE and the difference scores for the following combinations: answer honestly-informed other, answer honestly-role play, and apply for a job-answer honestly. The hypothesized indicators of other-deception are scores on the fake good scale, impression management scale, and the following difference scores: apply for a job-informed other, apply for a job-role play, and apply for a job-answer honestly.

Indicators of self-deception. It was hypothesized that SDE scores and several difference scores (i.e., difference scores between the “answer honestly” condition and informed other ratings, difference scores between the “answer honestly” condition and role play ratings, and difference scores between the “answer honestly” and “answer as if you are applying for a job) would be predictors of the outcome variables. As such, four separate regression equations were computed in which these variables served as the independent variables and each of the four outcome variables served as dependent variables (i.e., adjustment, integrity, interpersonal competency skills, and satisfaction with life). As there were no a priori hypotheses about the strength of the predictors, all four variables were entered into the equation simultaneously (see Table 17).

Examination of the F statistics revealed that all four regression equations were significant ($p < .05$). In predicting adjustment, examination of the beta weights revealed that out of the four predictors, two of the difference scores were predictive: answer honestly-informed other ($\beta = -.51$) and apply for a job-answer honestly ($\beta = .18$) ($p < .01$). Overall, the predictors accounted for 25% of the variance in adjustment. For integrity, 6% of the variance was accounted for by the predictors. Only one of the four predictors was significant which was difference score of answer honestly-informed other negatively predicted ratings on integrity ($\beta = -.23$). In predicting interpersonal skills, all four variables were significant. SDE score ($\beta = -.13$), answer honestly – informed other ($\beta = -.58$), answer honestly – role play ($\beta = .11$) and apply for job – answer honestly ($\beta = .20$) and 31% of the variance was accounted for by these variables. Only one predictor was significant in predicting satisfaction with life which was answer honestly-informed other ($\beta = -.40$). Overall, the variables accounted for 15% of the variance in satisfaction with life.

Indicators of Other-deception.

In order to examine the relationship between the hypothesized indicators of other-deception, four additional regression equations were computed using the outcome variables as the dependent variables. However, in these regression equations, the GI scale of the CPI, the fake good scale of the CPI, IM scale of the BIDR, the difference scores between the “answer as if applying for a job” condition and role play condition, and “answer as if you are apply for a job” and informed other conditions. For all four outcome variables, examination of the F statistic revealed that all regression equations were significant ($p < .01$).

Overall, the variables accounted for 25% of the variance in adjustment scores. The difference score between answer honestly and informed other negatively predicted adjustment ($\beta = -.52$). For the outcome variable of integrity, only 8% of the variance was accounted for by the predictors with the only significant predictor being apply for a job-informed other ($\beta = -.23$). For interpersonal skills, the predictors explained 30% of the variance with the difference score of “apply for a job” and informed other was significant ($\beta = -.54$). For satisfaction with life, 15% of the variance was explained by the predictors and the only significant variable was the difference score between the scores in the “apply for a job” condition and the informed other condition ($\beta = -.40$).

Discussion

Multiple studies have been conducted examining the role of deception in personality testing. The present study was intended to add to this body of literature by using multiple methods of detecting deception simultaneously and identifying potential implications of both self-deception and other-deception. Although it was unsuccessful in testing the original hypotheses, it still made multiple contributions to the literature. First, it utilized multiple

methods of detecting deception simultaneously and provided insight into potential implications. Second, it examined deception as a two-dimensional construct, using self-deception and other-deception as the organizing framework. Although several significant results were found, it is necessary to preface these findings with a word of caution. First, given the large number of analyses, the opportunity for Type I error is also increased. Second, according to the effect size conventions set forth by Cohen (1988), the effect sizes for these results were small (i.e., effect sizes ranged from .02-.17). As a result, it should be noted that the theoretical and practical implications of the results may be quite limited as only small differences were found between the groups on outcome variables. However, the results are still important to discuss as they provide some insight into the relationships between these methods of detecting deception and important outcomes. Additionally, they provide direction for future research. Replicating this study in an applicant sample or with another sample that may have a hypothesized higher level of motivation to perform well than participants in the present study may reveal larger effect sizes. In turn, a finding of larger effect sizes would suggest greater practical and theoretical implications.

One of the most consistent themes throughout the findings was that there were significant relationships between how people saw themselves in both Goldberg Inventory self-report conditions and how they were rated by others who knew them well. Specifically, the greater the discrepancy between how people rated themselves in the “answer honestly” condition and how they were rated by someone who knew them well, the lower they were rated on adjustment, integrity, interpersonal skills, and satisfaction with life. The same pattern held true for the discrepancy between how people rated themselves in the “answer as if you are applying for a job” condition and how they were rated by someone who knew them

well. Those individuals who had views of themselves that were closely aligned with how others saw them were rated as having higher levels of adjustment, integrity, interpersonal skills, and satisfaction with life. Another finding was that having self-ratings higher than ratings received from informed others was related to receiving lower ratings on the outcome variables. Taken together, these results suggest that individuals who have views of themselves that are consistent with how others perceive them are better adjusted, possess greater interpersonal skills, have higher levels of integrity, and higher satisfaction with life. While inconsistent with some of the more recent thinking on psychological well-being (e.g., Taylor & Brown, 1988; Scheier & Carver, 1985), these findings are consistent with more traditional conceptualizations of good mental health (e.g., Allport, 1987; Vaillant, 1977; Jourard & Landsman, 1980; Haan, 1977), in which it was believed that having a realistic view of oneself was a necessary condition for good psychological health.

However, some words of caution are important in interpretation of the data. Examination of the correlation coefficients suggests some evidence for a methods effect. Specifically, moderate to large correlations were observed between all four outcome variables and the responses to the ratings on the Goldberg Inventory (r 's range from .27-.74) and some moderate to large correlations among the outcome variables themselves (r 's range from .21-.62). Taken together, these results suggest that informed raters may have rated participants high or low across all variables. As the survey including the informed-other ratings of the Goldberg Inventory and the outcome variables were given in a single administration, informed observers may have responded positively or negatively across the entire survey, not effectively discriminating their scores on the differing scales.

Another key finding was that when participants were grouped as either non-deceptors (i.e., participants who were not identified by any methods as engaging in other-deception or self-deception) or some-deceptors (participants who were identified by one or more of the methods as engaging in self-deception and/or other-deception), no significant differences were observed between the two groups, suggesting that deception, when defined as a one-dimensional construct, is not related to differences in adjustment, integrity, interpersonal skills, and satisfaction with life.

While no significant findings were observed when deception was defined as a one-dimensional construct, several significant differences were observed when defining deception in terms of self-deception and other-deception. Specifically, higher ratings on adjustment were related to several indicators of other-deception. That is, those individuals who were hypothesized to be engaging in other deception were seen as being better adjusted. These findings are consistent with some researchers' views that the ability to present oneself in a manner that is appropriate to the demands of the situation is quite functional (as suggested by Rosse, Stecher, Miller & Levin, 1998). It may be that those individuals who can adapt their personality to meet the demands of a variety of situations, come across to those around them as being better adjusted.

Differences in scores on the GI scale were found to be significantly related to the outcome variable of integrity. Specifically, those individuals who had higher ratings on the good impression scale were seen as having higher levels of integrity. These findings are consistent with previous research by Gough (1996) that suggest that high scorers on the GI are often seen in a positive light, as appreciative, cooperative, conscientious, and thorough.

For interpersonal skills, there was some indication that both indicators of self-deception and other-deception were important predictors of how one's level of interpersonal skills were perceived. It may be that those individuals who engage in other-deception have a flexible personality and are able to change their behaviors to align with the needs of a social situation, thus leading others to have the perception that they have higher levels of interpersonal skills. Those high in self-deception may overlook their own flaws and therefore come across as more confident and in turn, as having higher levels of interpersonal skills.

Several results suggest that individuals who are high in self-deception may be less satisfied with their lives than those individuals who are low in self-deception. These findings are somewhat consistent with some of the work by Paulhus (1988) which indicates that early in relationships with others, individuals who are high in self-deception are initially seen as interesting, confident, and well-adjusted. However, with more exposure they are seen as arrogant, hostile, and maladjusted, all variables that one would hypothesize to be negatively correlated with satisfaction with life.

Understanding why specific hypotheses can not be tested

Although it is certainly important to have an understanding of the significant results, it is also critical to gain insight into why the participants could not be placed into the hypothesized groups. One contributing factor was that many hypothesized indicators of other-deception identified very few participants as engaging in this type of deception. For example, one hypothesized indicator of other-deception was that participants would raise their score two points in the "apply for a job" condition in comparison to the "answer honestly" condition. However, in the present study, no participants met this criteria and overall, the mean change score for the entire group was quite small ($M = .06$). Similar results

were found for all three social desirability scales hypothesized to be indicators of other-deception (i.e., GI, IM, FGCPI). That is, given the original constraints, very few participants were identified as other-deceptors (i.e., sample sizes ranging from 3-14). Likewise, few participants were identified by the difference scores that were hypothesized to be indicators of other-deception (“apply for a job” – informed other difference score yielding only 3 participants and the “apply for a job”-role play difference score yielding only 37 participants).

Although limited research has been done in the field to examine the base rates of deception, several comparisons can be made between the results of this study and the norm groups reported by the authors of the scales. Across the three scales, the rate of participants being flagged as engaging in other-deception is quite similar to the norm groups reported by the authors. Based on 78 samples of participants, Gough (1996) estimated the range of participants surpassing the FGCPI cut-score to be 0% to 8.5%. Based on his normative database ($n = 3000$), only approximately 2% of people would be expected to surpass the cut-score of the GI. For the IM scale of the BIDR, 30% of participants would be expected to surpass the threshold level. While the majority of the results in the present study were consistent with findings from the authors’ normative databases, it was anticipated that the rate of other-deception would be higher in the present study as the students had some motivation to present themselves in a favorable light in the “answer as if you are applying for a job” condition. In creating the normative samples, much of the research had focused on individuals who may have had limited motivation to portray themselves in a favorable light (e.g., members of an inventor’s club). Additionally, these percentages are much lower than those found in previous research. For example, one study reports that 42% of job applicants

reported giving false opinions (McDaniel, Douglas, & Snell, 1997) and 45% of job applicants indicated that they had observed or performed nonexistent tasks (Anderson & Warner, 1984). These study results would suggest a much higher number of other-deceptors in the population than was found in this study. However, an alternative explanation is that the manipulation in the present study (i.e., the “answer as if you are applying for a job”) instructions may have been too weak to elicit other-deception. A final explanation may be that the rate of other-deception for job applicants may indeed, be quite low and therefore, not a pervasive problem.

Sample sizes varied for identifying self-deceptors. For example, the SDE scale identified 81 participants as engaging in self-deception and was consistent with the Paulhus’ normative sample (i.e., 30% of his norm group was identified as self-deceptors). However, fewer participants were identified using the difference score approach. For the difference score of “answer honestly” – role play only 32 participants were flagged as self-deceptors and for “answer honestly” –informed other, only 3 participants were identified.

Limitations and future research

Several limitations were observed in the research. Although the sample size was relatively large (i.e., $n = 242$), it may have been too small to include an adequate number of self- deceptors and other-deceptors. Future research should focus on gaining a better understanding of the base rates of self-deceptors and other-deceptors in both the general population and applicant population. A better understanding of frequency of occurrence would allow for a deeper understanding of whether examining deception is of practical and theoretical importance forwarding future research. That is, if the rate of deception is quite small, it may be that future studies may not be warranted.

However, other possible causes of the small sample sizes should be investigated. An alternative explanation is that participants in this study did not have an adequate level of motivation to engage in other-deception. Although participants were given instructions to respond as if they were applying for a job, in reality, they may have had very little motivation to consciously portray themselves in a positive light. That is, they were aware that there were no tangible rewards for presenting a profile that was socially desirable and as a result, may have had little motivation to present a positive personality profile. Future researchers could replicate this study in the lab but take actions that may raise the level of motivation (e.g., offer a cash prize). This study could also be replicated in the field, where motivation is likely much higher, given that often whether an individual does or does not receive a job is at least partially contingent upon their responses to a personality profile.

Another possibility in this study is that the manipulation (i.e., the instructions) may not have been too weak to elicit other-deception in the present study. In an effort to avoid demand effects, the instructions in this research stayed at a very general level, informing participants that this personality survey was just one part of the selection process. As a result, they may have put an inadequate level of emphasis on the importance of this step in the process or felt motivated to portray themselves positively in order to receive a job offer. Future research should consider creating stronger instructions, those that emphasize the importance of the personality test in the selection process and/or emphasize that the participant is very interested in the job in question.

The present study was an improvement over prior research as it included a role play as an additional method of detecting deception. However, the role play used in this study had several limitations. Although the simulation was designed to engage individuals in

approximately 30 minutes of dialogue, the interactions actually lasted only approximately 10 minutes. If researchers could design role plays that lasted for longer periods of time, more behavioral observations could be made and perhaps, more valid ratings of personality. In an ideal setting, participants would be observed by multiple trained raters in multiple role plays.

This research added to the literature base by providing some insight into several different implications of both self-deception and other-deception. However, there was some evidence to suggest that the results may have been obscured by a methods effect. In future research, informed observers should be given adequate education regarding how to provide effective and valid ratings of others. Additionally, it may be important to gather data from multiple informed observers and perhaps not have the same individuals who rate outcome variables also rate personality.

This study also focused heavily on the use of difference scores. Several criticisms of difference scores have been made in the literature and may not be an effective means of measuring congruence. In a summary of the literature, Edwards (1993) provides his interpretations of the some of the most worrisome problems when using difference scores including: reduced reliability, ambiguous interpretation, confounded effects, and untested constraints. In the future, researchers should consider using different methods of measuring congruence.

In an effort to understand the roles of social desirability scales in detecting deception, the present study simultaneously examined multiple scales. However, one of the most pressing areas for future research lies in gaining an understanding of many more of the most commonly-used scales. Currently, many organizations use this tool as a method for screening “fakers” and as such, may drop people from an applicant pool solely based upon their

responses to this scale. Practitioners need to have a better understanding of what the scales are truly capturing before using this as a decision-making tool. While the current research relies on the BIDR, fake good scale, and good impression scales, future research should examine what other commonly used scales are capturing. It may be that some of these scales are not able to adequately differentiate between the different types of deception or perhaps some are capturing self-deception while others may be capturing other-deception. A deeper understanding of what the scales are measuring may also shed some light on the meaning of the much-disputed covariation between personality scales and social desirability scales (e.g., Holden & Fekken, 1989; Nicholson & Hogan, 1990; McCrae & Costa, 1983, Smith & Ellingson, 2002).

Another area for future research lies in examining the outcomes associated with each type of deceptor. Although the present study examined four potential outcomes, there are multiple other areas to examine. One of the most critical areas, is an understanding of the relationship between type of deceptor and overall job performance. As this link is researched and hopefully becomes clear, it is a good first start in understanding the role that this information should play in the selection process.

Another limitation is that the current study does not take into account levels of deception. That is, once individuals reach a specific threshold, they will be treated equally to others who have surpassed the said threshold. That is, there are no provisions to measure level of faking. Research suggests that this may be an important consideration. For example, one researcher suggests that the amount of self-deception one is engaging is an important consideration in determining outcomes (Baumeister, 1989). He suggests that there is an optimal level of illusion (in his definition of illusion he includes one's perception of the self

as well as one's perception of the world), a level of illusion at which people are happiest and function most effectively. However, deviating from this optimal level of illusion is likely to cause problems. Future research should examine the role that differing levels of deception have for those individuals that are identified as self-deceptors, other-deceptors, or mixed-deceptors and ascertain the impact that it may have on outcomes.

Future research should also consider additional methods of assessing personality (e.g., continue studying assessment centers, how to use interviews to assess personality) as people will continually be attempting to find ways to "outsmart" paper and pencil tests. As Ehrenreich (2001) suggests, some approaches are quite simplistic:

My approach to pre-employment personality tests has been zero tolerance vis-à-vis the obvious "crimes"--drug use and theft--but to leave a little wriggle room elsewhere, just so it doesn't look like I'm faking out the test. My approach was wrong. When presenting yourself as a potential employee, you can never be too much of a suck-up (p. 124).

Other approaches are more sophisticated. That is, there are currently books on the market written by individuals with a technical understanding of personality inventories and use their educational and professional backgrounds to "coach" others into "beating" the test. As more individuals become involved in this "trade", it will become necessary to produce more sophisticated tests as well as use other avenues as a means of collecting personality information about others.

Despite the limitations of the current study, it is an initial start in gaining some insight into the reasons why individuals present (or do not present) accurate portrayals of the self. Furthermore, it examines how some traditionally used methods of examining "faking" can capture these different types of deception. The present research also identifies potential implications for organizations hiring these individuals. It is hoped that the present study will

make other researchers and the applied world take a more critical eye toward their approach to measuring and assessing deception.

References

- Abramson, L. Y, Seligman, M. E. , & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.
- 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*, 59-72.
- Allport, G.W. (1937). *Personality: A Psychological Interpretation*. New York: H. Holt and Company.
- Anderson, C. D., Warner, J. L., & Spencer, C. C. (1984). Inflation bias in self-assessment examinations: Implications for valid employee selection. *Journal of Applied Psychology, 69*, 574-580.
- Ballesteros, R.F., & Zamarron, M.D. (1996). New findings on social desirability and faking. *Psychological Reports, 79*, 612-614.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology, 25*, 729-735.
- Barrick M.R., & Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.
- Barrick M.R., Mount, M.K., & Perkins, J. P. (1994). Validity of observer ratings of the big five personality factors. *Journal of Applied Psychology, 79*, 272-280.

- Barrick, M.R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection & Assessment, 9*, 9-30.
- Bass, B.M. (1957). Faking by Sales Applicants of a Forced Choice Personality Inventory. *Journal of Applied Psychology, 41*, 403-404.
- Baumeister, R.F.; Hamilton, J.C.; & Tice, D.M. (1985). Public versus private expectancy of success: Confidence booster or performance pressure? *Journal of Personality and Social Psychology, 48*, 1447-1457.
- Baumeister, R.F. (1989). The optimal margin of illusion. *Journal of Social & Clinical Psychology, 8*, 176-189.
- Becker, T.E., & Colquitt, A. L. (1992). Potential versus actual faking of a biodata form: An analysis along several dimensions of item type. *Personnel Psychology, 45*, 389-406.
- Bono, J.E., & Judge, T. A. (2003). Core self-evaluations: A review of the trait and its role in job satisfaction and job performance. *European Journal of Personality, 17*, 5-18.
- Booth-Kewley, S., Edwards, J.E., & Rosenfeld, P. (1992). Impression Management, social desirability, and computer administration of attitude questionnaires: Does the computer make a difference? *Journal of Applied Psychology, 77*, 562-566.
- Borkenuau, P. & Amelang M. (1985). The control of social desirability in personality inventories: A study using the principal-factor deletion technique. *Journal of Research in Personality, 19*, 44-53.
- Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition, 4*, 353-376.

- Campbell, J.D. (1986). Similarity and uniqueness: The effects of attribute type, relevance, and individual differences in self-esteem and depression. *Journal of Personality and Social Psychology*, 50, 281-294.
- Christiansen, N.D., Goffin, R., Johnston, N. & Rothstein, M.G. (1994). Correcting for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology*, 47, 847-860.
- Colvin, C.R., Block, J., & Funder, D.C. (1995). Overly positive self-evaluation and personality: Negative implications for mental health. *Journal of Personality and Social Psychology*, 68, 1152-1162.
- Coyne, J. C. & Gotlib, I. H. (1983). The role of cognition in depression: A critical appraisal. *Psychological Bulletin*, 94, 472-505.
- Crowne, D.P. & Marlowe, D.A. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 355-360.
- Dalen, L.H., Stanton, N.A., & Roberts, A.D. (2001). Faking personality questionnaires in personnel selection. *Journal of Management Development*, 20, 729-741.
- Day D.V. & Silverman S.B. (1989). Personality and job performance: Evidence of incremental validity. *Personnel Psychology*, 42, 25-36.
- Delery, J.E., & Kacmar, K.M. (1998). The influence of applicant and interviewer characteristics on the use of impression management. *Journal of Applied Social Psychology*, 28, 1649-1669.
- Diener, E., Emmons, R., Larsen, J. & Griffen, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49, 71-75.

- Druss, B.G. & Schlesinger, M.; Allen, H.M. (2001). Depressive symptoms, satisfaction with health care, and 2-year work outcomes in an employed population. *American Journal of Psychiatry*, 158, 731-734.
- Dunning, D., Meyerowitz, J.A. & Holzberg, A.D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality & Social Psychology*, 57, 1082-1090.
- Edwards, J.R. (2002) Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *Measuring and Analyzing Behavior in Organizations: Advances in measurement and data analysis* (pp. 350-400), San Francisco, CA: Jossey-Bass.
- Ehrenreich, B. (2001). Nickel and Dimed: On (not) getting by in America. New York: Henry Holt and Company.
- Ellingson, J.E., Smith, D., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86, 122-133.
- Feather, N. T. (1962). The study of persistence. *Psychological Bulletin*, 59, 94-115.
- Feather, N. T. The relationship of persistence at a task to expectation of success and achievement related motives. *Journal of Abnormal & Social Psychology*, 63, 552-561.
- Felson, R.B. (1981). Ambiguity and bias in the self-concept. *Social Psychology Quarterly*, 44, 64-69.
- Felson, R. B. (1984). The effect of self-appraisals of ability on academic performance. *Journal of Personality & Social Psychology*, 47, 944-952.

- Feldman, R.S., Forrest, J.A., & Happ B.R. (2002). Self-presentation and verbal deception: Do self-presenters lie more? *Basic & Applied Social Psychology, 24*, 163-170.
- Fernandez-Ballesteros, R. & Zamarron, M.D. New findings on social desirability and faking. *Psychological Reports, 79*, 612-614.
- Ferris, G. R., & Porac, J. F. (1984). Goal setting as impression management. *Journal of Psychology, 117*, 33-36.
- Ferris, G. R., Judge, T. A, Rowland, K.M., Fitzgibbons, D.E. (1994). Subordinate influence and the performance evaluation process: Test of a model. *Organizational Behavior & Human Decision Process, 58*, 101-135.
- Furnham, A. (1986). Response bias, social desirability, and dissimulation. *Personality and Individual Differences, 7*, 385-400.
- Furnham (1990). The fakeability of the 16PF, Myers-Briggs, and FIRO-B personality measures. *Journal of Personality and Individual Differences, 7*, 711-716.
- Furnham, A. (1997). Knowing and faking one's five-factor personality score. *Journal of Personality Assessment, 69*, 229-243.
- Gardner, W. L. & Martinko, M. J. (1988). Impression management in organizations. *Journal of Management, 14*, 321-338.
- Gilmore, D. C., & Ferris, G. R. (1989). The effects of applicant impression management tactics on interviewer judgments. *Journal of Management, 15*, 557-564.
- Goffman, Erving (1959). *The presentation of self in everyday life*. New York: Doubleday.
- Gough, H.G. (1996). *California Psychological Inventory Manual*. Palo Alto, CA: Consulting Psychologists Press.

- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35, 603-618.
- Griffith, R. (1998). Faking of noncognitive selection devices: Red herring is hard to swallow. (Doctoral dissertation, University of Akron, 1998) Dissertation Abstracts International, 58, 5681.
- Gudjonsson, G. H. (1990). The relationship of intellectual skills to suggestibility, compliance and acquiescence. *Personality & Individual Differences*, 11, 227-231.
- Guion, R.M; Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135-164.
- Gurevitch, Z. D. (1984). Impression formation during tactical self-presentation. *Social Psychology Quarterly*, 47, 262-270.
- Haan, N. (1977). Coping and defending. New York: Academic Press.
- Holden, R.R. & Fekken, G.C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers? *Journal of Research in Personality*, 23, 180-191.
- Hogan, R. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 873-919). Palo Alto, CA: Consulting Psychologists Press.
- Holden, R. R. (1998). Detecting fakers on a personnel test: Response latencies versus a standard validity scale. *Journal of Social Behavior and Personality*, 13, 387-398.
- Holden, R. R. (1995). Incremental validity of response latencies for detecting fakers on a personality test. *Journal of Research in Personality*, 29, 362-372.
- Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, 11, 209-244.

- Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D. (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. & Ones, D. S. (2002). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N.Anderson, D. Ones, H.K. Sinangil, K. Handan, & C. Viswesvaran, *Handbook of industrial, work and organizational psychology* (233-277). Thousand Oaks, CA: Sage Publications Ltd.
- Imada, A.S., & Hakel, M.D. (1977). Influence of nonverbal communication and rater proximity on impression and decisions in simulated employment interviews. *Journal of Applied Psychology, 82* , 295-300.
- Jourard, S.M. & Landsman, T. (1980). *Healthy personality: An approach from the viewpoint of humanistic psychology*. New York: Macmillan.
- Johnson, J. A. & Hogan R. (1981). Moral judgments and self-presentations. *Journal of Research in Personality, 15*, 57-63.
- Kacmar, K.M., Delery, J.E. & Ferris, G.R.(1992). Differential effectiveness of applicant impression management tactics on employment interview decisions. *Journal of Applied Social Psychology, 22*, 1250-1272.
- Kashy, D.A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality & Social Psychology, 70*, 1037-1051.
- Kluger, A.N., & Collella, A. (1993). Beyond the mean bias: The effect of warning against faking on biodata item variances. *Personnel Psychology, 46*, 763-780.

- Kluger, A.N., Reilly, R.R., & Russell, C.J. (1991). Faking Biodata Tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology, 76*, 889-896.
- Kristof-Brown, A. Barrick, M.R., & Franke, M. (2002). Applicant impression management: dispositional influences and consequences for recruiter perceptions of fit and similarity, *Journal of Management, 28*, 27-46.
- Kuiper, N. A., & Derry, P. A. (1982). Depressed and nondepressed content self-reference in mild depressives. *Journal of Personality, 50*, 67-80.
- Kuiper, N. A.; MacDonald, M. R. (1982). Self and other perception in mild depressives. *Social Cognition, 1*, 223-239.
- Kuiper, N. A., Olinger, L. J., MacDonald, M. R., & Shaw, B. F. (1985). Self-schema processing of depressed and nondepressed content: The effects of vulnerability to depression. *Social Cognition, 3*, 77-93.
- Larwood, L. & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology, 62*, 194-198.
- Lautenschlager, G. J. (1996). Within-subject measures for the assessment of individual differences in faking. *Educational and Psychological Measurement, 46*, 309-316.
- Lee, S., & Klein H.J. (2002). Relationships between conscientiousness, self-efficacy, self-deception, and learning over time. *Journal of Applied Psychology, 87*, 1175-1182.
- Levin, R.A. & Zickar, M.J. (2002). Investigating self-presentation, lies, and bullshit: Understanding faking and its effects on selection decisions using theory, field research and simulation. In J.M. Brett & F. Drasgow (Eds.), *Psychology of Work: Theoretically Based Empirical Research*. Mahwah, N.J. : Lawrence Erlbaum, 2002.

- Lewinsohn, L. M, Mischel, W., Chaplin, W. , Barton, R (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology*, 89, 203-212.
- Linden, W., Paulhus, D.L., & Dobson, K.S. (1986). Effects of response styles on the report of psychological and somatic distress. *Journal of Consulting and Clinical Psychology*, 54, 309-313.
- LoBello, S.G., & Sims, B.N. (1993). Fakeability of a commercially produced pre-employment integrity test. *Journal of Business and Psychology*, 8, 265-273.
- Mahar, D., Cologon, J., & Duck, J. (1995). Response strategies when faking personality questionnaires in a vocational selection setting. *Personality and Individual Differences*, 18, 605-609.
- Martin, B.A., Soven, C.C. & Hunt, S.T. (2002). How effective are people at faking on personality questionnaires? *Personality and Individual Differences*, 32, 247-256.
- McFarland L.A. & Ryan A.M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, 85, 812-821.
- McCrae, R. R & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51, 882-888.
- McDaniel, M.A., Douglas, E.F. & Snell, A.F. (1997). *A survey of Deception Among Job Seekers*. Presented at 12 annual SIOP Conference, St. Louis.
- Mersman, J.L., and Shultz K.S. (1998). Individual differences in the ability to fake on personality measures. *Personality and Individual Differences*, 24 , 217-227.

- Merydith, S.P. & Wallbrown, F.H. (1991). Reconsidering response sets, test-taking attitudes, dissimulation, self-deception and social desirability. *Psychological Reports, 69*, 891-905.
- Messick, S. (1960). Dimensions of social desirability. *Journal of Consulting Psychology, 24*, 1960.
- Mount, M.K., Barrick, M.R., & Strauss, J. P. (1994). Validity of observer ratings of the Big 5 personality factors. *Journal of Applied Psychology, 79*, 272-280.
- Nicholson, R. A. & Hogan, R. (1990). The construct validity of social desirability. *American Psychologist, 45*, 290-292.
- 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*, 600-679.
- Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11*, 245-269.
- Ones, D. S., Viswesvaran, C., & Schmidt, F.L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*, 679-703.
- Pannone, R.D. (1984). Predicting test performance: A content valid approach to screening applicants. *Personnel Psychology, 37*, 507-514
- Paulhus, D.L. (1984). Two components of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598-609.
- Paulhus, D.L. (1986). Self-deception and impression management in test responses. In A.

- Angleitner & J.S. Wiggins (Eds.), *Personality assessment via questionnaire* (pp. 17-59). New York: Springer Verlag.
- Paulhus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Paulhus, D.L. & Reid, D.B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, *60*, 307-317.
- Paulhus, D.L. (1998). *Paulhus Deception Scales*. North Tonawanda, NY: Multi-Health Systems.
- Paulhus, D.L. (1998). Interpersonal and interapsyhic adaptiveness of trait self-enhancement: A Mixed blessing? *Journal of Personality and Social Psychology*, *74*, 1197-1208.
- Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment*, *5*, 164-172.
- Ramanaiah, N. V., & Martin, H. J. (1980). On the two-dimensional nature of the Marlowe-Crowne Social Desirability Scale. *Journal of Personality Assessment*, *44*, 507-514.
- Reid-Seiser, H. L., & Fritzsche, B. A. (2001). The usefulness of the NEO PI-R Positive Presentation Management Scale for detecting response distortion in employment contexts. *Personality & Individual Differences*, *31*, 639-650.
- Robie, C. (no date available). Faking and personality measurement. PDI Whitepaper.
- Robins, R.W. & Paulhus, D.L. (2001). The character of self-enhancers: Implications for organizations. In B.W. Roberts & R. Hogan. (Eds.), *Personality psychology in the workplace. Decade of behavior*. (pp. 193-219). Washington, DC: American Psychological Association.

- Robins, R.W. & John, O. P. (1997). The quest for self-insight: Theory and research on accuracy and bias in self-perception. In R. Hogan, J.A. Johnson & S. Briggs (Eds.), *Handbook of personality psychology*. (pp. 649-679).
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96, 341-357.
- Rosse, J.G., Stecher, M.D., Miller, J.L. & Levin, R.A. (1998). The impact of response distortion on pre-employment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634-644.
- Ruch F.L. & Ruch L.W. (1967). The K factor as a (validity) suppressor variable in predicting success in selling. *Journal of Applied Psychology*, 51, 201-204.
- Ryan, A.M., McFarland, L., Baron, H., & Page, R.(1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52, 359-391.
- Sackeim, H.A., & Gur, R. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology*, 47, 213-215.
- Salgado, Jesus F.(1998). Big Five personality dimensions and job performance in Army and civil occupations: A European perspective. *Human Performance*, 11, 271-288.
- Scandell, D.L. & Wlazelek, B.G. (1996). Self-presentation strategies on the NEO-Five factor inventory: Implications for detecting faking. *Psychological Reports*, 79, 1115-1121.
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219-247.
- Silverman, I. (1964). Self-esteem and differential responses to success and failure. *Journal*

- of Abnormal and Social Psychology, 69, 115-119.*
- Smith, B. D. & Ellingson, J.E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology, 87, 211-219.*
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 4, 526-537.*
- Taylor, S.E., & Brown, J.D. (1988). Illusion and well-being: A social-psychological perspective on mental health. *Psychological Bulletin, 103, 193-210.*
- Tett, R. P., Jackson, D.N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44, 703-742.*
- Thumin, F.J. & Barclay A.G. (1993). Faking behavior and gender differences on a new personality research instrument. *Consulting Psychology Journal, 45, 11-19.*
- Topping, G.D. & O’Gorman, J.G. (1997). Effects of faking on validity of NEO-FFI. *Journal of Personality and Individual Differences, 23, 117-124.*
- Viswesvaran, C., & Ones, D.S. (1999). Meta-analyses of fakeability estimates: Implications for Personality Measurement. *Educational and Psychological Measurement, 59, 197-210.*
- Viswesvaran, C.; Ones, D. S. & Hough, Leetta M. (2001). Do impression management scales in personality inventories predict managerial job performance ratings? *International Journal of Selection & Assessment, 9, 277-289.*
- Wayne, S.J., & Liden, R. C. (1995). Effects of impression management on performance ratings: A longitudinal study. *Academy of Management Journal, 38, 232-260.*

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

Appendix A

Goldberg Scales

Please rate the (participant/yourself) on each item using the following scale:

1-----2-----3-----4-----5-----6-----7
Not Neutral Very
true of participant true of participant
of participant

01. ___ Extraverted
02. ___ Talkative
03. ___ Assertive
04. ___ Verbal
05. ___ Energetic
06. ___ Bold
07. ___ Active
08. ___ Daring
09. ___ Vigorous
10. ___ Unrestrained
11. ___ Introverted
12. ___ Shy
13. ___ Quiet
14. ___ Reserved
15. ___ Untalkative
16. ___ Inhibited
17. ___ Withdrawn
18. ___ Timid
19. ___ Bashful
20. ___ Unadventurous
21. ___ Kind
22. ___ Cooperative
23. ___ Sympathetic
24. ___ Warm
25. ___ Trustful
26. ___ Considerate
27. ___ Pleasant
28. ___ Agreeable
29. ___ Helpful
30. ___ Generous
31. ___ Cold
32. ___ Unkind
33. ___ Unsympathetic
34. ___ Distrustful

- 35. _____ Harsh
- 36. _____ Demanding
- 37. _____ Rude
- 38. _____ Selfish
- 39. _____ Uncooperative
- 40. _____ Uncharitable

*Items 1-20 assess extraversion. Items 21-40 assess agreeableness.

Appendix B

Fake Good Scale

Formula for fake good scale = 41.255 + .273 (Dominance) + .198 (Empathy) +.538
(Good Impression)- .255 (Well-being) -.168 (Flexibility)

Appendix C

Balanced Inventory of Desirable Responding-Version 6

Use the scale below to indicate how true each item is of you.

1-----2-----3-----4-----5-----6-----7
Not true Neutral Very true

- 01. ___ My first impressions of people usually turn out to be right.
- 02.* ___ It would be hard for me to break any of my bad habits.
- 03. ___ I don't care to know what other people really think of me.
- 04.* ___ I have not always been honest with myself.
- 05. ___ I always know why I like things.
- 06.* ___ When my emotions are aroused, it biases my thinking.
- 07. ___ Once I've made up my mind, other people can seldom change my opinion.
- 08.* ___ I am not a safe driver when I exceed the speed limit.
- 09. ___ I am fully in control of my own fate.
- 10.* ___ It's hard for me to shut off a disturbing thought.
- 11. ___ I never regret my decisions.
- 12.* ___ I sometimes lose out on things because I can't make up my mind soon enough.

13. ___ The reason I vote is because my vote can make a difference.
- 14.* ___ My parents were not always fair when they punished me.
15. ___ I am a completely rational person.
- 16.* ___ I rarely appreciate criticism.
17. ___ I am very confident of my judgments.
- 18.* ___ I have sometimes doubted my ability as a lover.
19. ___ It's alright with me if some people happen to dislike me.
- 20.* ___ I don't always know the reasons why I do the things I do.
- 21.* ___ I sometimes tell lies if I have to.
22. ___ I never cover up my mistakes.
- 23.* ___ There have been occasions when I have taken advantage of someone.
24. ___ I never swear.
- 25.* ___ I sometimes try to get even rather than forgive and forget.
26. ___ I always obey laws, even if I'm unlikely to get caught.
- 27.* ___ I have said something bad about a friend behind his or her back.
28. ___ When I hear people talking privately, I avoid listening.
- 29.* ___ I have received too much change from a salesperson without telling him or her.
30. ___ I always declare everything at customs.
- 31.* ___ When I was young, I sometimes stole things.

32. ___ I have never dropped litter on the street.
- 33.* ___ I sometimes drive faster than the speed limit.
34. ___ I never read sexy books or magazines.
- 35.* ___ I have done things that I don't tell other people about.
36. ___ I never take things that don't belong to me.
- 37.* ___ I have taken sick-leave from work or school even though I wasn't really sick.
38. ___ I have never damaged a library book or store merchandise without reporting it.
- 39.* ___ I have some pretty awful habits.
40. ___ I don't gossip about other people's business.

* Indicates items that will be reverse-scored.

** Items 1-20 assess SDE. Items 21-40 assess IM.

Appendix D

Adjustment scale

Please rate the participant on each of the items using the following scale:

1-----2-----3-----4-----5-----6-----7
Not Neutral Very
true of participant true of participant

1. ___ The participant is well-adjusted.
2. ___ The participant is happy.
3. ___ The participant likes himself/herself.
4. ___ The participant is mentally healthy.

Appendix E

Satisfaction with Life Scale (SWLS)

Please rate the participant on each of the items using the following scale:

1-----2-----3-----4-----5-----6-----7
Not Neutral Very
true of participant true of participant

1. ___ In most ways the participants sees life as close to ideal
2. ___ The participant perceives that conditions of life are excellent.
3. ___ The participant is satisfied with life.
4. ___ So far, the participant has gotten the important things wanted in life.
5. ___ If the participant could live life over he/she would change almost nothing.

Appendix F

Interpersonal Competency Scale

Please rate the participant on each of the items using the following scale:

1-----2-----3-----4-----5-----6-----7
 Not Neutral Very
 true true of participant
 of participant

1. ___ The participant has a reputation for being able to cope with difficult people.
2. ___ The participant finds it easy to talk with all kinds of people.
3. ___ The participant finds it easy to play many roles (e.g., student, leader, follower, church-goer, athlete, traveler, etc.).
4. ___ The participant would be good at playing charades.
5. ___ People seek the participant out to tell him/her their troubles.
6. ___ The participant has unusual skills for making groups, clubs, or organizations function effectively.
7. ___ The participant has unusual skills for assessing the motivation of other people.
8. ___ If the participant wants to, he/she can be a very persuasive person.
9. ___ The participant has a clear picture of what he/she is like as a person.
10. ___ The participant knows what he/she wants to do with life.
11. ___ The participant is shrewd and insightful about other people.
12. ___ The participant would enjoy being an actor/actress.
13. ___ Most of the time, the participant has an optimistic outlook.
14. ___ The participant has good practical judgment.
15. ___ Others believe the participant has good practical judgment.

Appendix G

Integrity Scale

Please rate the participant on each of the items using the following scale:

1-----2-----3-----4-----5-----6-----7
Not Neutral Very
true of participant true of participant

1. ___ The participant is an honest individual.
2. ___ The participant does not lie.
3. ___ The participant does not take things that do not belong to him/her.
4. ___ The participant is trustworthy.

Appendix H

Role play information

Scenario

To be read to participants:

During this simulation you will assume that you work for a fictitious company that provides cell phone products and services. The background information provides you with information on the organization, your role in the organization, and the products and services you represent.

As you take on the role outlined in this simulations, you will be most effective if you handle the situation the way YOU think is best.

Background information on Reliable Communications (to be distributed to participants):

Founded in 1957 by a group of Marketing experts, “Reliable Communications, Inc. is a rapidly growing organization in the telecommunications industry. Reliable Communications success can be traced to a strong commitment to providing quality service at a low cost. Reliable Communications has communicated internally and externally its commitment to unparalleled customer service and satisfaction. It has invested heavily in internal customer service training and other initiatives related to providing the best possible customer service.

Your role:

- You were recently hired as a Customer Service Specialist in the Reliable Communications’ Webster Grove store
- Your boss is the store manager, Bob Smith. He has let you know that customer service is key and he gives you a lot of leeway in determining the best way to satisfy customers. He has let you know that offering “perks” to disgruntled customers is okay. He lets you decide what those are but says you must have solid rationale in defense of your decisions.
- You are responsible for selling cell phone products and services to customers who visit the Webster Grove store. Additionally, you are responsible for handling a variety of customer complaints on matters ranging from billing problems to poor service connections.
- Your only other peer is Jason Wilson and he is currently gone on vacation in Tahiti and can not be reached for the next 3 weeks.
- Your boss is out sick for the day and you are the only person in the store.

Information on Reliable Communications Cell phone products and Services

Reliable Communications provides 1 and 2 year service contracts. All customers who receive new service contracts receive a free phone (currently, Reliable Communications carries only one type of cell phone).

The table below describes the monthly plans available and the cost of additional services.

Type of plan (all plans include free standard phone)	Monthly Cost	Anytime Minutes (all minutes are anytime)	Cost per month to add Voicemail	Cost per month add 3-way calling	Cost per month add call waiting	# of Free months with 1 year service (# of Free months with 2 year service
Standard	39.99	80	2.50	1.00	1.50	1	2
Enhanced	49.99	120	1.50	1.00	1.50	1	2
Premium	69.99	500	1.00	.50	.50	1	2
Extra Premium	109.99	3000	Included	Included	Included	2	3

Products offered:

Product	Price
Hands free headset	39.99
Phone cover	15.00
Car Charger	42.00
Spare Battery	25.00

Scenario (to be given to the trained rater)

You will play the role of Chris Hoff, a physician at the local hospital. As a physician, your cell phone is very important to you. You must be able to stay in constant contact with the hospital in case of an emergency. Your boss has insisted that you use Reliable Communications for this business cell phone as they are good providers.

Characteristics that describe Chris Hoff

- Mildly impatient and insistent on good customer service
- Extremely busy and overwhelmed with stressful work
- Generally reasonable if treated in a respectful manner
- Very skeptical about Reliable Communications. You didn't want to use them in the first place, but your boss told them you had to use them for you work cell phone purchase. Now that you have experienced problems with them, you are even more convinced that there is no value in using the products.

Reason for your visit to the store:

- For the last two months in a row, you were invoiced for three way calling which you did not sign up for
- You had signed up for voicemail but it is not yet set up
- You also did not receive the one month free service the salesperson offered you when you signed your contract

Tone for the simulation:

- Irritated
- Looking for answers and solutions to the invoice problem
- Looking for assurance that future problems won't come up in the future
- Interaction should start out mildly tense but become less/more so dependent on the participant's interactions with you

Interaction guide

If Lee Says:	Then Chris responds:
Lee opens by asking Chris to share his/her problem.	The problem? The problem is that your company has been ripping me off. I'm getting charged for a Voice Mail feature that I don't have. When the sales person talked to me, I signed a contract stating that the feature would be added. Also, I'm having problems with getting my free month of service that I was offered.
Lee opens the conversation by attempting to ask for more information (e.g, copy of the invoice, name, etc.)	I can give you this information, but I really want to know what you are going to do to help me.
Lee asks who your salesperson was	I think it was Jason. However, I want to make sure that you can help me today.
Lee asks you to calm down without offering any	Why should I? You would be as upset as I am if

kind of resolution to your problems.	you have all of these problems. You are already charging me an arm and a leg for this service and you are still trying to take me for a ride. I get pushed around all the time by your people to buy more products.
Lee says it must have been a mistake.	This isn't the first time I've had problems with this company. . I used this company for my personal cell phone and ended up getting charged 3 times in one month! Plain and simple, I need some answers and help or I'll be forced to cancel my contract.
Lee says he/she doesn't know why you were incorrectly invoiced, but that he/she can take care of the situation.	I need this taken care of today. What can you do for me?
Lee tries to defer the problem to someone else	Isn't there anything you can do for me now? This has been a big hassle for me.
If Lee can't make any progress:	If you can't help me, please find someone who can.
Lee claims he/she can't verify your information because he/she does not have computer access:	That sounds like an excuse to me. Here's my copy of my invoice.
Lee apologizes, takes ownership of the problem, acknowledges you are correct, agrees the mistakes have been unacceptable, expresses empathy for your situation, and outlines clear action steps	If Chris feels that it is genuine, he she starts to settle down and become less skeptical. (e.g., I appreciate that you are trying to help me)
Lee offers credit.	Well yes, I would expect your to but how can you assure me that this won't happen again.
Lee asks Chris what solution he/she is looking for.	Chris provides a solution that is relatively easy for Lee (e.g., free service for a month, free products, credit on her account)
Lee offers solutions without asking Chris what he/she wants	Chris becomes irritated and resistant to Lee's suggestions
Lee offers to talk to the appropriate people to make sure it won't happen again	That's good, but I'd like some compensation for the trouble you've cause us. What are you going to offer me
Lee recaps the action items and the timeframes (without being prompted)	Chris becomes docile and the tension is all but gone.

Table 1

Descriptive Statistics for all Participants

Measures	N	Mean	SD	Reliability
Goldberg Inventory				
Self-Answer Honestly	242	5.39	0.61	0.91
Self-Apply For Job	242	5.46	0.63	0.92
Roleplay	242	4.52	0.97	0.97
Informed Other	242	5.57	0.64	0.91
CPI Scales				
Good Impression	239	23.81	6.61	0.87
Fake Good Scale	226	56.14	4.05	n/a
BIDR				
Self-deceptive Enhancement	242	4.56	0.60	0.68
Impression Management	242	4.13	0.89	0.85
Outcome Measures				
Adjustment	240	5.87	0.89	0.84
Integrity	240	6.28	0.88	0.82
Interpersonal skills	240	5.13	0.80	0.84
Satisfaction w/ Life	240	4.78	1.05	0.79
Goldberg Inventory-difference scores				
Apply for job-answer honestly	242	0.06	0.43	n/a
Apply for job-role play	242	0.93	1.01	n/a
Apply for job-informed other	242	-0.11	0.90	n/a
Answer honestly-role play	242	0.87	1.02	n/a
Answer honestly-inf. Other	242	-0.18	0.87	n/a

Table 2

Pearson Correlations for all Participants

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Goldberg–self answer honestly																
2. Goldberg-self-apply for job	.75**															
3. Goldberg-role play	.22**	.27**														
4. Goldberg-informed other	.04	.01	-.15*													
5. Good Impression scale	-.11	-.01	-.08	.07												
6. Fake Good scale	-.04	.04	-.13	.14*	.89**											
7. SDE scale	.25**	.35**	-.06	-.11	.10	.15*										
8. Impression management scale	.07	.22**	-.06	-.07	.07	.12	.54**									
9. Adjustment	-.08	-.13*	-.04	.56**	.10	.06	-.12	-.05								
10. Integrity	.00	-.07	-.06	.27**	.15*	.09	-.09	-.02	.25**							
11. Interpersonal Skills	.07	.00	-.05	.74**	.13*	.18**	-.03	.00	.58**	.25**						
12. Satisfaction With Life	-.06	-.10	-.12	.42**	.00	.02	-.10	-.04	.62**	.21**	.45**					
13. Apply for Job-Answer Honestly	-.30**	.41**	.08	-.05	.14*	.12	.17**	.22**	-.08	-.09	-.09	-.06				
14. Apply for Job-Role Play	.26**	.38**	-.80**	.14*	.07	.15*	.28**	.20**	-.04	.01	.05	.05	.18*			
15. Apply for Job-Informed Other	.50**	.70**	.29**	-.71**	-.06	-.07	.33**	.21**	-.49**	-.24**	-.53**	-.37**	.32	.16*		
16. Answer Honestly-Role Play	.39**	.20**	-.82**	.16*	.01	.01	.20**	.10	-.01	.05	.09					