Summer 2012

Deception Detection: Using Eye-Tracking Technology to Measure Faking in a Simulated Applicant Setting

Luke Simmering

Deception Detection: Using Eye-Tracking Technology to Measure Faking in a Simulated Applicant Setting

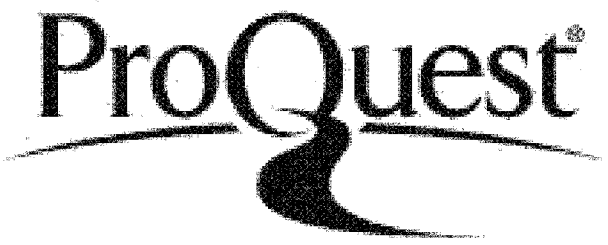
by

Luke Simmering

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

COLLEGE OF EDUCATION
LOUISIANA TECH UNIVERSITY

August 2012



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

LOUISIANA TECH UNIVERSITY

THE GRADUATE SCHOOL

06/08/2012

Date

We hereby recommend that the dissertation prepared under our supervision by Luke Simmering

entitled

Deception Detection: Using Eye Tracking Technology to Measure Faking in a Simulated Setting

be accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Supervisor of Dissertation Research

Head of Department

Department

Recommendation concurred in:

Advisory Committee

Approved:

Director of Graduate Studies

Approved:

Dean of the Graduate School

Dean of the College

GS Form 13a
(6/07)

ABSTRACT

The impact of faking on selection assessments and the need for methods to detect such behavior has drawn increased attention of researchers in the selection field over the last quarter century. The overarching purpose of this study was to assess the validity of utilizing eye-tracking technology in the detection of applicant faking on personality measures. Specifically, this study examined the physiological cues of response latency, eye fixation, and pupil dilation and their association with deception in the context of personality assessment in a job seeking scenario. The results indicated that individuals engaged in faking behavior had significantly more eye fixations and recorded significantly higher scores on the paper and pencil measure of cognitive load. In addition, results suggest that the experimental conditions likely accounted for the alterations in cognitive load regardless of the level of social desirability of items.

APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Dissertation. It is understood that "proper request" consists of the agreement, on the part of the requesting party, that said reproduction is for his personal use and that subsequent reproduction will not occur without written approval of the author of this Dissertation. Further, any portions of the Dissertation used in books, papers, and other works must be appropriately referenced to this Dissertation.

Finally, the author of this Dissertation reserves the right to publish freely, in the literature, at any time, any or all portions of this Dissertation.

Author 

Date 07-14-2012

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

The writing of the dissertation has been one of the greatest challenges I have ever faced. Without the guidance of Dr. Tilman Sheets, Dr. Frank Igou, Dr. Tony Young, and Dr. Mitzi Desselles this project would have never been possible. I would like to especially thank Dr. Sheets for giving me the opportunity to be a doctoral student and the mentorship he provided. Further, the support of my fellow classmates provided the opportunity to brainstorm ideas and a sense of camaraderie during the dissertation process. In addition, I am grateful to my parents and sisters for instilling in me a desire to do great things. Most importantly, I would like to thank my wife Morgan for her support and patience. Especially for the late night writing sessions that surely kept her awake. I am also grateful for the birth of my son, William, who in the past year has provided motivation, a work-ethic, and a desire to be a better person. Lastly, I would like to thank the Sundown Tavern of Ruston which offered a first-class establishment to decompress after a long day of dissertation writing.

CHAPTER ONE

INTRODUCTION

The impact of faking on selection and the need for methods to detect such behavior has drawn increased attention from researchers in the selection field over the last quarter century. In addition, Griffith and Peterson (2006) pointed out that response distortion or faking on personality and integrity measures within the employee selection sector is a major area of concern for many practitioners. Donovan et al. (2003) found that more than 30% of recent job applicants admitted that they had engaged in faking behaviors by altering their responses on a personality-based assessment to fit a more socially desirable applicant profile. Douglas, McDaniel, and Snell (1996) found that applicant faking decreases the criterion-related validity of a personality measure. The decrease in utility of personality measures caused by applicant faking reduces the overall performance of an organization's employee selection system by creating disarray in the linkage between an organization's metrics and employee performance.

MacKinnon (1944) referred to personality as the factors inside people that explain their behavior. Allport (1960) suggests that personality is the dynamic arrangement of psychophysical systems that influence an individual's adjustment to the environment. The use of trait-based personality assessment within the employee selection domain has dramatically increased since the early 1990s (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006). This increase is primarily due to the development of the five-factor model

of personality (Digman, 1990) and research supporting the utility of personality measures in the prediction of job performance across a wide array of criteria (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991). Although personality assessment has established its utility as a tool in employee selection, it has undergone much scrutiny regarding its use.

The definition of personality most applicable to employee selection is that of Cattell (1950), who described personality as "that which permits a prediction of what a person will do in a given situation" (p. 2). One consideration that underlies personality is that individuals may portray different personality characteristics in different environments or situations (Buss, 1989). The employee selection environment invokes an interesting display of personality characteristics. In particular, the employee selection setting creates an environment that increases the potential for applicants to distort their actual personality to improve their chances of being hired (Barrick & Mount, 1996).

Handler and Hunt (2003) evaluated the financial impact of making a bad hire. They investigated the relative cost of having a single false positive (i.e., an individual who matched the selection assessment requirement, but whose performance was below an acceptable criterion). To identify the costs, Handler and Hunt (2003) accounted for the cost of advertising for a position, the human resource staffing costs, interview time, salary, and even any relocation and training expenses. Based on estimated U.S. cost averages, they calculated that the cost of a bad hire or false positive is roughly $50,000. Donovant et al. (2003) pointed out that such costs can be reduced through the use of deception detection techniques to screen out applicants who are faking on employee selection assessments.

Applicant faking has been described as both an individual difference and a situational influence (McFarland & Ryan, 2006). In other words, some degree of faking may be due to a trait inherent within an individual, while the environment also acts as a catalyst or motivator that promotes faking behavior. Although not within a real-life context of employee selection, several studies have reported higher mean personality scale scores among job applicants when compared with job incumbents (Griffin, 2002; Hough, 1998; Robe, Zickar, & Schmit, 2001; Rosse, Stecher, Miller, & Levin, 1998).

While evidence has accumulated to suggest that applicant faking is ubiquitous, some studies have asserted that faking does not necessarily impact the validity of these measures (Barrick & Mount, 1996; Hough, 1998; Smith & Ellingson, 2002). However, these studies operationally defined and measured faking through the use of impression management and socially desirable responding scales. This is problematic because self-report faking detection scales are themselves sensitive to faking (Zickar & Robie, 1999) and have disputed validity (Griffith & Peterson, 2008). As a result, inferring conclusions about the effects of faking from studies that utilize self-report scales is challenging. Due to the preponderance of evidence illustrating the issues associated with the use of self-report detection techniques (Viswesvaran & Ones, 1999), alternatives to self-report measures may provide supplemental information to detect applicant faking.

Van Hooft and Born (2011) used eye-tracking technology to detect applicant faking. The authors presented participants with a big-five personality instrument and an integrity measure in a within-subjects design where each participant was instructed to respond honestly in one condition and fake in another condition. As supported in previous research (Viswesvaran & Ones, 1999), participants were able to successfully fake and make themselves appear more socially desirable on the measures when

instructed to do so. Mean scores on the measures within the faking condition were significantly higher than the mean scores in the honest condition, especially for the conscientiousness and emotional stability factors. Participants completed the measures in the faking condition significantly faster with an average of eye fixation time and an average number of eye fixations that was less than in the honest condition.

These findings support van Hooft and Born's (2011) premise that faking involves a more primitive process with lower cognitive load than responding honestly. Van Hooft and Born go on to theorize that the lower cognitive load in the faking condition results from the directions to respond in a socially desirable manner, while the honest condition provokes the participants to make self reflections or interpretations of the item content. This is in agreement with Holtgraves (2004) who suggested that faking is a less complex process because the goal is to respond based on faking instruction combined with the desirability of the item. However, this finding is in contrast to the majority of research on lying and deception (Depaulo et al., 2003; Holtgraves, 2004; McDaniel & Timm, 1990) which supports the theory that faking behaviors are more cognitively complex than telling the truth. Van Hooft and Born (2011) acknowledged that their results may have been an artifact of their methodology. Specifically, their methodology rendered little motivation for subjects who were instructed to fake, which significantly differs from a real-world context. They suggested that future research should utilize experimental manipulations that better mimic actual faking conditions.

The Present Study

The present study attempted to extend the findings of van Hooft and Born's (2011) research by utilizing an experimental methodology that attempts to invoke a more realistic motivational component in the faking scenario. Van Hooft and Born (2011) point

out that researchers should utilize experimental manipulations that better mimic actual faking conditions. The methodology used in the present study involved providing an incentive for participants who fake the personality measurement in an attempt to match an ideal profile. Specifically, the purpose of this study was to explore and investigate the utility of using the physiological variables of response latency, eye fixation, and pupil dilation to detect faking on personality-based measures within a simulated applicant setting. The items used on the personality-based measure were determined by a preliminary survey to identify a diverse mix of socially desirable items. Van Hooft and Born (2011) pointed out that researchers need to study the relationship between eye-tracking variables and specific items that differentiate strongly between fakers and honest responders. Rather than strictly focus on the mean differences between big-five personality scales as done by van Hooft and Born (2011), this study assessed the response processes of the participants across a diverse array of socially desirable personality-based items based on a preliminary survey to differentiate socially desirable and non-socially desirable items. As a non-invasive technology that does not rely on self-report data, eye-tracking technology may provide assessment at the physiological level to explore response processes elicited when an individual is engaged in faking on personality assessments.

Review of the Literature

Within the research literature, faking on personality assessments is referred to as response distortion, impression management, social desirability, displaying unlikely virtues, and self-enhancement (Viswevaran & Ones, 1999). One topic of particular interest within this body of research is the concern that applicants in an employee selection environment may be motivated to fake or distort their responses on a

personality-based measure in order to increase their probability of being hired (Griffith & Peterson, 2006). Assuming the motivation to fake exists, the need for better understanding the underlying complex cognitive processes involved in faking is clear (Vrij, 2008).

Faking and Employee Selection

Various definitions of applicant faking behavior on personality-based measures exist. Zeigler et al. (2011) have described faking as, "...a response set aimed at providing a portrayal of the self that helps a person to achieve personal goals" (p.8). This definition posits the notion that faking within the employee selection context is a deliberate act to improve one's chances of being hired. Various studies have supported the theory that applicant faking results in problems in employment decision making (Rosee et al., 1998; Weiner & Gibson, 2000; Barrick & Mount, 1996; Dunnet et al., 1981). In one study, Griffith, Chmielowski, and Yoshita (2007) found that between 22% and 49% of applicants fake personality measures. Although personality measures have provided incremental validity evidence in the prediction of job performance (Tett et al., 1991; Barrick & Mount, 1991) when added to other employee selection measures (Goffin et al., 1996; Schmidt & Hunter, 1998), personality measures have tended to rely on self-report methods that can easily be faked (Hough & Oswald, 2000; McFarland & Ryan, 2006). While personality measures have demonstrated sufficient utility as an employee selection measure, the ease at which they can be faked may dramatically alter the utility within the employee selection domain. Surprisingly, several self-help books have been published to teach applicants how to respond on personality assessment in a way that reveals a positive personality profile (e.g., Hoffman, 2001). Organizational leaders and employee

selection experts are mutually concerned that dishonesty on a selection test may transfer to dishonesty in the workplace (Peterson & Griffith, 2006).

Gammon and Griffith (2011) performed a study to evaluate individual differences between fakers and non-fakers. Their research was prompted by previous research that examined academic cheating by looking at individual differences in integrity, locus of control, and self-efficacy between cheaters and non-cheaters. Gammon and Griffith took a similar approach in looking at the differences between fakers and non-fakers within a sample of real applicants. Their study found that fakers revealed significantly lower integrity and self-efficacy while scoring higher in external locus of control and counterproductive work behaviors compared to the non-fakers. This study illustrated the inherent differences between individuals who fake on personality assessments as an applicant and the negative behaviors fakers may display once employed.

Research suggests that applicants who fake on personality or non-cognitive measures tend to fake more on items that require less deceptiveness (Donovan et al., 2003). While there are differences in the types of behaviors involved in faking (Donovan et al., 2003), it is important to understand the various methods of faking which applicants may use to gain an advantage. Snell, Sydell, and Lueke (1999) posited that successful faking requires the motivation or intent to fake along with the ability to fake. Further, they suggested that individual differences (e.g., cognitive ability), experiential factors (e.g., job experience, job knowledge) as well as test characteristics (e.g., item format, item transparency) influence the applicant's overall ability to successfully fake on a non-cognitive measure. The authors stated that demographics (e.g., age, gender), dispositional factors (e.g., integrity, Machiavellianism, locus of control), and perceptual factors (e.g., others' attitudes and behaviors, fairness) influenced the applicants

motivation to fake on the measure. Their model accounted for both individual difference characteristics and situational factors in an explanation of applicant faking behavior. The authors suggested that applicant faking should not be regarded as a unitary process that is comparable for every test taker. Their view has been supported by other research which indicates at least two different forms of faking exist: slight and extreme faking (Zickar, Gibby, & Robie, 2004; Zickar & Robie, 1999).

McFarland and Ryan (2006) extended the theory advanced by Snell et al. (1999) to describe applicant faking behaviors on non-cognitive measures. The authors proposed that faking is cued by the intention to fake which is, in turn, moderated by the applicants' ability to fake (e.g., self monitoring, perception of item transparency) and the opportunity for the applicant to fake. Their model posited that the intention for the applicant to fake will be strongest when both the ability to fake and opportunity to fake is high. Further, the intention to fake is influenced by the applicant's internal beliefs towards faking. The extent to which an applicant is morally opposed to faking on the measure will influence the applicants faking intentions. The relationship between attitudes or beliefs towards faking and the intention to fake is moderated by situational factors such as desire for the job and warnings about faking on the measure. Their model proposed that attitudes toward faking, subjective norms, and perceived behavioral control predict the intention to fake. Specifically, the researchers found that attitudes toward faking, subjective norms, and perceived behavioral control predicted the intention to fake and actual faking behavior.

In a similar vein, Mueller-Hanson et al. (2006) developed a model to aid in understanding an applicant's intention to fake on personality or any other non-cognitive measure. Their model proposed that sensitivity of the situation, readiness to fake, ability

to fake, conscientiousness, and emotional stability are the key factors driving an applicant's intention to fake. Similar to the McFarland and Ryan (2006) model, the intention to fake tends to provoke actual faking behavior. In testing their model, Mueller-Hanson et al. (2006) found that perceptions of the situation, conscientiousness, and emotional stability were all significantly related to intention to fake and predicts actual faking behaviors. Their model lends support to Snell et al. (1999) in that personality traits predicted applicant faking behavior.

Although models of faking behavior (Mueller-Hanson et al., 2006; Snell et al., 1999) provide an intricate and informative method of understanding applicant faking and its antecedents, they do not explore the cognitive processes involved in the actual act of faking. Krosnick (1999) reviewed various models of applicant faking behavior and proposed a four-step process of responding that includes comprehension, retrieval, judgment, and mapping. The process is described as follows: the test taker must first comprehend and encode the item to create a mental representation of its content. The test taker then retrieves the most valuable item content. Retrieved information is compared with the test taker's mental representation of the item. Finally, the test taker maps this information onto the rating scale. Krosnick referred to this as a model of optimal process, occurring when people are motivated to answer the items of the questionnaire in a sincere manner. Factors such as cognitive ability, fatigue, and motivation may deter the applicant from following this optimal process and resort to adjusting their effort to meet their expectation of the importance of the test. Krosnick referred to this strategy as satisficing. While Tourangeau and Rasinksi (1988) hypothesized a similar model of the applicant faking process, they included an additional step during the mapping phase called editing. Editing is essentially an individual's inclination toward socially desirable responding.

Zeigler (2011) conducted a qualitative study using a cognitive interview technique to reveal the cognitive process involved in faked personality questionnaire responses. Participants were initially instructed to assume they are applying for a position as a student in undergraduate psychology. They were informed that to gain placement in the program, they had to take a personality measure. Participants were instructed to think out load and voice their thoughts as they completed the personality measure. The aggregation and coding of participants' verbalizations while completing the personality measure revealed interesting faking strategies. For instance, many of the participants would determine what their actual choice would be under normal circumstances and adjust slightly in the favorable direction. Extreme fakers endorsed the highest possible category on the scale to portray the most positive or socially desirable personality profile. While this study is similar to a previous study by Zeigler (1999), some extensions to the model provided additional insight into the faking process. In particular, the interaction between the applicant and the situation were observed in different stages of the study. Specifically, the study revealed that applicants first evaluate the importance of an item in terms of the situational demands. If the applicant judges the item as unimportant, no faking occurs. Ziegler's (2011) model also included general mental ability to account for the cognitive effort it takes to maintain faking across the entirety of the assessment process. Although faking research tends to involve real applicant samples, Zeigler (2011) asserted the importance of using laboratory settings in the exploration of basic cognitive processes. Overall, Zeigler's (2011) study provides the opportunity to glimpse into the cognitive process that a test taker undergoes when faking on a personality measure.

Rosse et al. (1998) compared job applicant personality scores to job incumbent personality scores and found that applicants had levels of faking greater than three

standard deviations above the mean for the incumbent sample. The researchers asserted that this type of response distortion can dramatically alter the rank order of the applicants and decreases the utility of personality assessment as an employee selection tool. Within the study, the differences in personality scale scores were especially large for the Neuroticism, Agreeableness, and Consciousness scales, due to the socially desirable characteristics of these items within the employee selection context.

Douglas, McDaniel, and Snell (1996) investigated the effects of situational induced response bias on the validity, reliability, and various psychometric characteristics of non-cognitive assessments. Specifically, the researchers found that faking increases the internal consistency of the scales due to a lack of homogeneity found in honestly responding participants. Honest responding is much less consistent than more homogeneous or consistent endorsement of positive behaviors by applicants engaged in faking. Although faking increased the internal consistency of non-cognitive measures, the authors maintained that faking decreased the construct validity of the personality scales. This is due to the authors' finding of a different factor structure between faking and honest applicants. Additionally, the results of their study indicated that faking decreased the criterion-related validity (e.g., the correlation between the personality scales and measures of job performance). In general, scales with adequate validity when administered to honest responders lacked validity when administered to faking responders. This research illustrated how faking in the applicant setting dramatically alters the validity of personality scales.

A similar study conducted by Mueller-Hanson, Heggestad, and Thornton (2003) examined the effects of faking on the criterion-related validity of job selection measures utilizing an honest versus faking condition design. Their results demonstrated that

applicant faking had a negative impact on the criterion-related validity of personality assessment scores and called into question the inferences that can be made in selection decisions. The impact of faking on the criterion-related validity of their predictor was most noticeable at the high end of the distribution. The authors found less error in prediction at the bottom third of the distribution of fakers than the top third. The study illustrated how applicants who fake can move toward the top of the score distribution, dramatically altering the utility of personality tests used in the employee selection. The results of the study supported the theory that faking may alter the validity of personality measures leading to selection of employees whose job performance may be lower than expected.

Self Presentation

Paulhus and Trapnell (2010) suggested that personality can be viewed as self-presentation. They asserted that every action by an individual communicates information about the actor and carries with it a degree of inauthenticity. That is, some actions by individuals are designed to portray a desired image rather than an accurate representation of one's actual personality. Their concept of self-presentation, in large part, included a bias in which there is an inaccurate portrayal of oneself. This bias is perhaps even more apparent in employee selection settings where applicants present themselves in a more positive fashion during interviews than they do after they have been hired (Rosse et al., 1998). Several studies have examined the extent to which individuals make themselves appear more socially desirable than participants who are instructed to respond honestly (Viswesvaran & Ones, 1999). In general, previous studies indicated that people are proficient at distorting their responses on personality measures in a socially desirable direction (Birkeland et al., 2006).

Additionally, Birkeland et al. (2006) found that studies utilizing an induced faking paradigm suggest that participants can tailor their "fake-good" responses on personality measures to fit their perception of the personality traits of employees in the target job. For example, individuals applying for an accountant position are relatively effective at projecting a personality profile that fits the positive personality profile for an accountant. Other studies have found that participants' "fake-good" personality profiles are very similar to personality profiles the participants produced when describing an ideal employee for a specific job (Martin, Bowen, & Hunt, 2002). People tend to self-promote themselves with potential dating partners more than they do in basic interactions with their friends (Rosee, Stecher, Miller, & Levin, 1998). Repeated public self-presentations can lead to automatization where biased self-descriptions can become integrated into habitual self-presentations (Hogan, 1983; Johnson & Hogan, 1981). Self-presentation bias was also illustrated in a study by Richman, Weisband, Kiesler, & Drasgow (1999) where participants in proctored questionnaire administrations engaged in more socially desirable responding than did participants in an anonymous internet administration. According to Baumeister (1982) self-presentation bias involves the discrepancy in how individuals present themselves more favorably in a public setting or to an audience when compared to being alone.

Paulhus and Trapnell (2010) conceptualized the individual differences in self-presentation as pervasive aspects of an individual's personality. Specifically, they separated self-presentation bias by an audience (public vs. private) and content distinction (agentic vs. communal). The authors describe an agentic image as one that involves "getting ahead" while a communal image involves "getting along". The four subtypes of self-presentation (public-agentic/communal; private-agentic/communal) must be treated

separately to gain a holistic understanding of self-presentation. Their model emphasized the understanding of individual differences in self-presentation rather than the psychological processes that maintain these differences. The authors went on to express the need for future research on self-presentation which focuses on whether one's behavior is public and identifying the corresponding appropriate images that are presented to the audience.

Buss (1980) emphasized that previous research supports the theory that impression management plays an important role in self-presentation in that the awareness of an audience alters an individual's behavior. The author also pointed out that if the context is private, there is no need to manage one's impression. Other research has suggested that individuals default to one of a limited number of basic self-presentation identities or roles that requires a social intelligence which is much more elaborate than notions of general intelligence (Jones & Pittman, 1982).

Some people confine themselves to only one role or self-presentation style while others have more flexibility in their self-presentation (Paulhus & Martin, 1988). However, people tend to revert to their predominant self-presentation style when experiencing stress (Leary, 1990). The most comprehensive taxonomy of self-presentation styles has been proposed by Jones and Pittman (1982). These authors suggested that people present themselves to exhibit or convey intimidation, supplication, ingratiation, self-promotion, or exemplification. Further research has supported the idea that these five primary styles are prevalent in everyday interactions (Bolino & Turnley, 1999). Although various researchers have identified several self-presentation styles (Holden & Evoy, 2005), the two themes of agentic versus communal self-portrayals are of particular interest because they are connected to several of the self presentation styles

(Carey & Paulhus, 2008). The two factors, agentic (competent, strong, clever) and communal (cooperative, dutiful), provide a succinct framework for understanding self-presentation styles.

Paulhus and Trapnell (2010) outlined three primary individual differences in self-presentation. These include the attunement allotted to self-presentation, the motivation to engage in self-presentation, and the amount of distortion expressed in self-presentation. Attunement involves the concept of self monitoring, which entails how an individual attends to the social demands of their current situation and adjusts their behavior to act appropriately. According to Snyder (1974), self-monitoring is an individual's ability to monitor his social environment and, using the information obtained, alter self-presentation to fit the social environment. Closely related to self-monitoring is Buss' (1980) concept of public self-consciousness. Buss investigated public and private self-awareness and described how some individuals are especially vigilant and reactive to public attention of their behavior. Paulhus and Trapnell (2010) presented a view of self-presentation which takes into account the influence of personality and motivational individual differences. For example, subclinical psychopathology leads to the planning of self-presentation stemming from an ego-centric personality.

Additionally, a motivational basis for self-presentation can stem from chronic insecurity and the need for approval (Crowne & Marlowe, 1964). Chronic insecurity is related to factors such as the fear of negative evaluation (Watson & Friend, 1969), subclinical narcissism (Morf & Rhodewalt, 2001), and a perfectionist self-presentation (Sherry, Hewitt, Flett, Lee-Baggley, & Hall, 2007). Swann (1990) provided a framework to aid in understanding the possible motivations of self-presentation that involve the

concepts of self-enhancement and self-verification. An individual's primary motivations are to either to enhance oneself or gain verification from the environment. Robbins and John (1997) provided a broader taxonomy that outlines four styles which take into account why an individual's self-perceptions may depart from reality. These four styles include the egoist who is motive by self-enhancement, the politician who is motivated by popularity, the consistent seeker who is motivated by consistency, and the scientist who is motivated by accuracy.

Socially Desirable Responding

The most pervasive and relevant construct related to faking, within the measurement of personality dimensions, is the concept of socially desirable responding. Socially desirable responding primarily refers to self-presentation on self-report questionnaires (Paulhus, 1991). For example, when individuals are asked to rate their own personality, they tend to rate their personality in a favorable way (Edwards, 1970). When socially desirable responding is measured as a stable individual difference, the tendency is considered a social desirability response style (Jackson & Messick, 1962), versus a response set which is driven by the context of the situation (Zerbe & Paulhus, 1987). Socially desirable responding is assessed and measured to gain insight into the validity of a self-report questionnaire and serves as a gauge to monitor dishonest or fabricated responses. Socially desirable responding not only encompasses positivity bias, but may also include responses aimed at the portrayal of unfavorable images, like being mentally ill (Baer, Rinaldo & Berry, 2003) or incompetent (Furnham & Henderson, 1982). Socially desirable responding has been conceptualized as a behavior pertaining to questionnaires, a personality construct that generalizes to other-self presentation contexts (i.e., interviews), and as an accurate portrayal of a desirable personality (Paulhus &

Trapnell, 2010). Diversity in the conceptualization of socially desirable responding may impede the singular convergence among socially desirable responding measures (Paulhus, 1984).

One way to assess socially desirable responding is to collect socially desirable ratings of a large variety of items and then assemble the items with the most extreme desirable ratings as an assessment of socially desirable responding (Saucier, 1994). This method assumes that an individual who claims the high desirability items and disclaims the low desirability items is responding on the basis of the items' desirability rather than accuracy. This method has endured much criticism due to the possibility that some people may have an abundance of desirable qualities. If this is the case, people may be responding accurately or providing a realistic profile of an individual with desirable qualities (Paulhus & Trapnell, 2010).

Another method of researching socially desirable responding is to utilize a role playing scenario where one group is instructed to fake good on a questionnaire and another group is instructed to respond honestly while completing the same questionnaire. Items that discriminate between the two groups are then used to build a new scale measuring the tendency to respond in a socially desirable manner (Wiggins, 1959). This type of "role playing" scenario was the methodology used to create the Minnesota Multiphasic Personality Inventory (MMPI) Malingering Scale (Gough, 1950).

In a study examining social desirability response differences between two separate samples, Rosee et al. (1998) found that job applicants had higher scores on social desirability and impression management scales compared to non-job applicants. The combination of simulated applicant settings and scenarios in applicant faking research, illustrated the propensity for applicants to answer in a socially desirable manner, and

lends support to the theory that job applicants distort their responses to appear more desirable for a specific occupation. A more elaborate way researchers have measured socially desirable responding is through the use of a rational method of test construction based on specific hypotheses regarding the underlying construct (Eysenck & Eysenck, 1979). This methodology incorporates items designed to trigger different responses in honest respondents and respondents who are likely motivated to appear socially desirable. Content validation models incorporating subject matter experts has been the most common way to identify the items that tap the underlying construct of interest (Nunnally & Bernstein, 1994).

The most wide-ranging and comprehensive program of construct validity pertaining to socially desirable responding was carried out by Crowne and Marlowe (1964) in the development of their socially desirable responding measure. The authors stated that the concept of need for approval underlies the socially desirable responding measure and that high scores on their measure are not just positive, but are improbably positive. Further Crowne and Marlowe asserted that their assessment measures a distinct personality construct. McCrae and Costa (1983) suggested that although various methods of measuring socially desirable responding exist, socially desirable measures serve as an index of one's personality traits rather than true desirable responding and that high scorers actually have a socially desirable character.

McCrae and Costa (1983) found that high scoring respondents on the Marlowe-Crowne Socially Desirable Responding Scale had responses similar to their spouses' ratings. According to the authors, this suggests high scores on the socially desirable responding scales reflect good social adjustment rather than desirable responding. Milham and Jacobson (1978) suggested an integrative perspective that

reconciles between the response distortion and accuracy positions is needed and hypothesized that both can be explained by the construct of need for approval. They proposed that high scorers on socially desirable responding scales realize that portraying conventional behavior is the best way to gain approval and that deceit works in situations where detection is unlikely. In a similar vein, Schlenker and Weigold (1990) suggested that some people will attempt to ensure others view them as well-adjusted by denying certain misleading facts or exaggerating evidence. Paulhus and Trapnell (2010) proposed that one-dimensional measures of socially desirable responding tap an unclear combination of distortion and reality which is sometimes difficult to understand.

Although socially desirable responding has been commonly viewed as a singular construct, factor analyses have consistently revealed two distinct and independent clusters of items within socially desirable responding measures labeled Alpha and Gamma (Wiggins, 1964). The Alpha cluster showed a strong relationship with the MMPI K scale (Hathaway & McKinley, 1951), while the Gamma factor was related to Hathaway and McKinley's (1951) MMPI Lie scale and Marlowe and Crowne's (1964) socially desirable responding scale. Paulhus (1991) examined these two distinct factors and developed a measure consisting of two scales: Self-Deceptive Enhancement (SDE) and Impression Management (IM). The combination of these two measures laid the foundation for the Balanced Inventory of Desirable Responding (Paulhus, 1991).

The Impression Management scale within the Balanced Inventory of Desirable responding was sensitive to faking (Paulhus , 1995) and has been shown to moderate the validity of personality scales (Holden, 2007). This effect on the utility of personality scales prompted research into the NEO Personality Inventory (NEO-PI-R). The NEO-PI-R is considered the most comprehensive and well-validated assessment of the

five-factor model (Costa & McCrae, 1992). Although the NEO-PI-R is considered a premiere measure of the big-five personality constructs, the inventory has its critics. Ben-Porath and Waller (1992) pointed out that the NEO-PI-R does not employ the use of validity scales that offer detection of response bias, inconsistent responding, or response distortion. Instead, the NEO-PI-R provides validity checks, structured with direct questions to the respondent about answering the items accurately and honestly (Marshall, 2005).

Some authors have defended the absence of validity scales within the NEO-PI-R, as the use of such scales may actually decrease the validity of personality measures (McCrae & Costa, 1983; McCrae et al., 1989). These authors believe that the use of social desirability measures as moderator variables may reduce the correlations between self-report and observer ratings on the dimensions. However, studies have shown the susceptibility of the NEO-PI-R to faking (Ballenger, Caldwell-Andrews, & Baer, 2001) and have illustrated the potential need for validity scales within the measure. Although the argument remains unsettled, Schinka, Kinder, & Kremer (1997) developed a validity scales for the NEO-PI-R. The objective of the NEO-Validity Scales is to detect random responding (Inconsistency Scale; INC), attempts to present oneself in a positive manner (Positive Presentation Management; PPM), and attempts to present oneself in an overly negative light (Negative Presentation Management; NPM). In addition to random responding and faking, Paulus et al. (1984) suggested that socially desirable responding is common among personality measures and is comprised of impression management, a deliberate effort to manipulate one's image in a positive manner, and self-deception, the unintentional process associated with self-enhancement. It is difficult for studies to pinpoint which form of socially desirable responding a participant is exhibiting, due to

intentions possibly occurring at the subconscious level; therefore, the specific constructs of impression management and self-deception are not specified within the NEO-Validity Scales.

Schnika et al. (1997) were able to identify items from their existing NEO Scales that could be used to infer validity, rather than creating a set of new items to measure response validity. Studies that have utilized the NEO-Validity Scales with the NEO-PI-R have illustrated the ability of the scales to effectively identify participants instructed to answer in a positive manner (Ballenger et al., 2001; Caldwell-Andrews, Baer, & Berry, 2000) or in a negative manner (Berry, Bagby, Smerz, Rinaldo, Caldwell-Andrews, & Baer, 2001; Caldwell-Andrews et al., 2000). While there are conflicting studies evaluating the utility of the NEO-Validity Scales, subsequent studies evaluated the scales ability to differentiate honest and faking respondents. Young and Schinka (2001) assessed the NEO-Validity Scales in treatment seeking individuals and their results indicated the scales predicted socially desirable responding. Additionally, a study by Marshall et al. (2005) compared the factor structure of the NEO-PI-R within sample of respondents who were motivated in different ways to respond in a socially desirable manner and found a stable factor structure across the groups. These findings illustrated the NEO-PI-R's ability to maintain reliable and valid psychometric properties in samples with socially desirable responding. While there were mean differences between differently motivated groups in answering (fake good, fake bad, etc.), the factor structure remained intact.

Bagby et al. (2003) completed a study in which participants completed the NEO-PI-R under normal conditions and fake-good instructions as job applicants. Results indicated that the "fake good" participants had lower scores on Neuroticism and higher

scores on Extraversion than the normal group. In a similar study, Berry et al. (2001) found the NEO-Validity Scales to be significantly different in the expected direction for samples that faked bad and for psychiatric outpatients. The study supported the construct validity for the use of the NEO-Validity Scales in distinguishing a faking bad response set. The faking bad sample differed from the psychiatric outpatient sample, demonstrating that the NEO-Validity Scales can differentiate between faking and true personality disorders. Morasco, Gfeller, and Elder (2007) examined the relationship between the the NEO-Validity Scales and the validity scales of the Minnesota Multiphasic Personality Inventory (MMPI). The authors found significant correlations between the validity scales of both measures. Participants within the study whose scores were elevated on the validity scales (invalid responses), produced significantly different clinical profiles on the NEO-PI-R and MMPI than participants with valid responses.

An additional scale of socially desirable responding is The Self Deceptive Enhancement scale (Paulhus, 1998). Paulhus found that the scale was related to narcissism as well as a disconnection with reality, as indicated by a discrepancy in self-ratings compared to peer-ratings. Further, the scale has been found to be positively related to self-perceptions of mental health (Bonanno, Field, Kovacevic, & Kaltman, 2002) and task performance (Johnson, 1995). However, Paulhus demonstrated that individuals who score high on the Self Deceptive Enhancement scale were also perceived negatively over time by others.

Self Enhancement

Another conceptualization of self-presentation that is very similar to the construct of socially desirable responding is self-enhancement. Generally, self-enhancement can be considered as an overly positive self-evaluation (Taylor & Brown, 1988).

Self-enhancement is primarily measured through social comparison or self-criterion discrepancies to gain a normative comparison of one's self-beliefs compared to the general population. Further, self-enhancement focuses on self-illusions or distortions in private self-beliefs rather than purposeful dissimulation (impression management). However, scores on self-enhancement measures should vary similarly to socially desirable responding measures (Carey & Paulhus, 2008). That is, although socially desirable responding and self-enhancement are theoretically distinct constructs, the measurement of both has been very similar.

The most popular method of conceptualizing self-enhancement has through social comparison (Kwan, John, Kenny, Bond, & Robins, 2004). Social comparison research has illustrated that people tend to rate themselves as above average on lists of various traits (Alikcke, 1985). Since it is highly unlikely that a large majority of individuals are actually better than average, it is more likely that self-enhancement is taking place. The operational definition of social comparison makes it difficult to differentiate self-enhancement from true discrepancies in positive traits (Klar & Giladi, 1999). Some individuals are above average across a wide array of traits (Block & Colvin, 1994). However, the social comparison conceptualization does not incorporate an authentic criterion to check the validity of the self-evaluation.

Self-enhancement has also been operationally defined as the overestimation of one's positivity in relation to a criterion, creating a criterion discrepancy (Paulhus & Trapnell, 2010). The intent of criterion discrepancy measures has been to identify difference scores and residual scores to indicate the extent to which respondents' self-ratings surpass their criterion scores. Criterion discrepancy assessment of self-enhancement has a negative association with stable adjustment (Bonanno et al.,

2002; Kwan et al., 2004; Paulhus, 1998). Paulhus et al. (2003) presented one of the more interesting operational definitions and measurements of self-enhancement in their over-claiming technique. Their over-claiming technique focused on measuring each individual's departure from reality by asking respondents to rate their familiarity with a group of people, places, items, or events, some of which do not exist. Responses are scored via signal detection method to illustrate an accuracy and bias score for each respondent based on their accuracy in claiming real items and disclaiming fake items. The authors reported that the accuracy index within the over-claiming technique was correlated highly with general mental ability, while the bias index was correlated highly with self-enhancement measures such as narcissism.

Self-enhancement has also been assessed through Krueger's (1998) method, which is sometimes referred to as the idiosyncratic weighting method. This method gathers participants' self-ratings as well as their desirability ratings for each item to create weights for each rating of desirability as deemed by the rater. Krueger stated that this is in contrast to the assumption or social consensus that socially desirability is shared among all respondents and that this method provides a more individualistic profile of self-enhancement. Finally, Kwan et al. (2004) provided a method of operationally defining self-enhancement using an existing social relations model to decompose self-perception into perceiver effect, target effect, and unique self-perception. The authors believed that this method provides a holistic view of one's perception relative to others by having all participants rate each other. These various methods of measuring self-enhancement are dependent on self-awareness. Self-awareness can be examined by first understanding that the manner in which people present themselves within an employee selection context is likely divergent from their actual personalities. This gap

between their actual personality and what they present can be a gauge of their self-awareness.

Accurate assessment of oneself is also dependent on the degree to which one is self-aware. Fletcher and Bailey (2003) defined self-awareness broadly as the amount of agreement between how individuals view themselves and how others view them. In other words, they proposed that the assessment of the self-awareness construct should, to some extent, include the perceptions of how others see the person in question. According to these authors, the concept relates more to the measurement of character rather than traits.

It is only recently that researchers have begun to address self-awareness within the framework of differentiating individuals on some external construct. Self-awareness appears to have a positive relationship with supervisory leadership effectiveness (Fletcher, 1997). In addition, self-awareness has been positively related to performance across various settings (Bass & Yammarino, 1991). Nasby (1989) also found that highly self-aware individuals were better at incorporating comparisons of behavior into their self-perception, and their perception of themselves was both more reliable and valid. On the other hand, people with low self-awareness were more likely to negate feedback about themselves, experience career derailment, and have poor attitudes towards their job (Ashford, 1989).

Within the world of work, individuals with high levels of self-awareness have more positive relationships in the workplace (Wexley, Alexander, Greenwalt, & Couch, 1980) and are more receptive to different kinds of feedback (Wohlers & London, 1989). While the construct of self-awareness is difficult to measure (Fletcher & Bailey, 2003), an index of self-other ratings score may provide an avenue to pursue further analysis. Warr & Bourne (1999) used a "gap analysis" as a way to conceptualize a measure of

self-awareness. The gap analysis consists of deriving a congruence score, as is commonly utilized in 360-degree studies. Church (1997) provided another way self-awareness can be assessed through the use of difference scores. Difference scores are created by computing the square root of the sum of squared differences between the self-report and average other-rating score for each self-other item contrast, and then the rating score is divided by the total number of items for the sample.

Fletcher & Baldry (2000) reported that many factors such as the influence of friendship, quantity of contact between self and the raters, and the various performance evaluation methods held by various rating groups are related to bias in assessment of self-awareness. These confounds are hard to control and suggest that it is important to interpret self-awareness scores with some apprehension. An additional point of confusion arises from Fletcher and Baldry's (2003) assertion that self-awareness should be considered a personality trait and an ability. Similarly, Fletcher, Taylor, and Glanfield (1996) studied self-awareness within the framework of personality and found that participants' accuracy in self-awareness was related to their scores on factor O of the Cattel 16PF assessment. The authors stated that factor O of the 16PF identifies those that are less confident, more likely to be anxious and less likely to feel accepted within groups. Roush and Atwater (1992) also looked at the relationship between self-awareness and personality. The authors used the Myers-Briggs Type Indicator (MBTI) and found that introverts and sensing types scored higher on self-awareness than individuals identified as other MBTI types. Finally, Mabe and West (1982) found that a high need for achievement is an additional indicator of accurate self-assessment. Variation within the assessment of self-enhancement may be due to many factors, one of which may be accuracy of an individual's perception of self.

Although self-enhancement would be considered a negative characteristic by many, some researchers have claimed that self-enhancement is adaptive (Taylor & Brown, 1988). Several studies have indicated that individuals who claim to be above average on a variety of traits also scored high on a measure of self-esteem (Brown, 1986) as well as other adaptive outcomes (Cambell et al., 2002). This seems to contradict traditional conceptions of mental stability that promote the importance of perceiving oneself accurately (Allport, 1960). Colvin and Block (1994) acknowledged this contradiction and suggested that positive illusions may be supportive in mood regulation and may provide momentary relief from negative affect.

While a temporary positive affect may occur due to such perceptions, the authors pointed out that over time positive illusions of oneself may damage self-efficacy or overall mental health. These illusions were investigated by Colvin et al. (1995) utilizing the criterion discrepancy operational definition of self-enhancement. The researchers found that when an external criterion was used to evaluate outcomes, discrepancy showed long-term negative maladaptive outcomes. In addition, an earlier study by John and Robbins (1994) compared self-rated performance, others' ratings of performance, and the amount of money earned in a group exercise. They found that higher scores on the discrepancy between self-rated performance and the performance indicators were negatively related with ratings of adjustment by trained psychologists.

Colvin et al. (1995) conducted a longitudinal study demonstrating that self-enhancement was related to poor social skills and psychological adjustment five years before and five years after the measurement of self-enhancement. Further, their study revealed that in a confrontational situation, self-enhancers were rated negatively by both expert raters and peers. Robins and Beer (2001) found short term affective benefits

of self-enhancement. However, long term damage to self-esteem and academic engagement occurred as disconfirming evidence came to light. Taylor and Armor (1996) suggested that self-enhancement should be viewed as an adaptive strategy rather than a trait. The authors acknowledged that while there may be utility of self-enhancement for certain environments or situations in gaining the short-term benefits, individuals must return to a more accurate self-perception to maintain mental health.

Arousal Based Deception Detection

The study of deception detection has often involved the application of arousal-based protocols used by law enforcement (Frank et al., 2008). Arousal-based procedures originate from two primary sources: physiological cues and behavioral cues. An underlying assumption in using these protocols is that there is fear that the deceit will be discovered. The fear then leads to an individual reacting differently while engaged in deception than when being honest. For instance, law enforcement agencies commonly use the polygraph examination as an arousal-based method to examine physiological responses while an individual is being questioned. The National Research Council (2003) examined the utility of the polygraph and found that the technique discriminates lying from honest responses at rates above random chance. However, the National Research Council cautioned against the use of the polygraph due to the ability of individuals to manipulate or alter the results of the examination. They recommended that other measures or methods of deception detection should be investigated.

In addition to the polygraph, several technologically based methods of detecting the somatic responses involved in deception have been described including; the polygraph, electrogastrogram, radar vital signs monitor, facial expressions, eye blinks, eye saccades, eye fixations, voice stress analysis, thermal imaging, and truth serums

(sometimes referred to as narcoanalysis) (Heckman & Happel, 2005). In addition, more elaborate systems such as the electroencephalography (EEG), magentocenephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and the near infrared spectroscopy (fNIRS) have been utilized to detect deception. Heckman and Happel (2005) provided a thorough review of the most common forms of the physiological measurement tools used to detect deception.

The electrogastrogram (EGG) assesses the nerves that control the stomach muscles to linkage between the gastrointestinal tract and the central nervous system (Hutson, 2005). However, the EGG procedure takes approximately three hours to complete in a clinical setting and is subject to the measurement of individual artifacts such as respiration and cardiac signals. Due to the time, costs, and lack of research involving the EGG, it may not be the best avenue to detect deception.

Geisheimer and Greneker (2001) reported that the first Radar Vital Signs Monitor (RVSM) was developed by the Department of Defense in the 1980s to measure physiological signals at a distance using electromagnetic waves. The Georgia Tech Research Institute further developed the RVSM to measure an individual's heartbeat, respiration, and eye blinks using electromagnetic waves. The RVSM's ability to detect deception has been reported as relatively similar to that of the polygraph. As a result, the RVSM is subject to many of the same limitations as the polygraph and lacks an appropriate amount of scientific validation (Heckman & Happel, 2005).

The observation and categorization of facial expressions illustrated in the Facial Action Coding System (FACS; Ekman & O'Sullivan, 1991) may have some utility in identifying deception. The FACS requires extensive training of interpreters and yields

results that may not be any better than the polygraph or any other deception detection technique (Heckman & Happel, 2005).

Voice stress analysis (VSA) uses technology to detect laryngeal micro-tremors in the voice that are caused by the amount of blood in the vocal chords to drop as a result of stress (Rice, 1978). VSA was initially developed for the U.S. Army as a lie detector that could be utilized remotely and to serve as an alternative to the polygraph. Research using the VSA has not yielded a consistent relationship between micro-tremors and deception, and it has demonstrated deception detection rates barely above chance (National Research Council, 2003).

Thermal imaging has been used to assess changes in an individual's skin surface temperature. Although research has indicated that thermal imagining is better than chance at detecting deception through analysis of changes in facial blood flow, additional research is needed to further assess its validity in an applied setting (Warmelink et al. 2011).

The use of truth serums or narcoanalysis is an invasive procedure used to lower individuals' inhibitions under the premise that they will be more forthright and honest in their responses. Research has not shown a consistent relationship between the use of narcoanalysis and detecting deception (Heckman & Happel, 2005).

In addition to these non-intrusive measures of deception, several newer technologies emerged in the search for valid methodologies aimed at detecting deception. Heckman & Happel (2005) reviewed a number of these types of technologies and found limited validity associated with their detection of deception in the literature. The authors described electroencephalography (EEG) as a measure of the changes in the electrical field potential within the neural systems of the brain through the use of electrodes place

on an individual's head. They pointed out that although EEG has some utility in identifying deception, it is very expensive, requires a trained interpreter, and is vulnerable to certain artifacts that can confound the measurement. In addition to the EEG, the authors described the positron emission tomography (PET) as a tool which monitors the blood flow in the brain. They reported that PET has limited utility in the detection of deception and that there are similar technologies which monitor blood flow of the brain which appear to show more promise.

In particular, the authors discussed the use of functional magnetic resonance imaging (fMRI) which uses a static magnetic field to measure oxygenated blood to identify which areas of the brain are undergoing the most neural activity. They pointed out that although attempts have been made to commercialize the use of the fMRI for deception detection, further studies are needed to validate its use in applied settings. Finally, the authors also discussed functional near-infrared spectroscopy (fNIRS), which is similar to the fMRI as it measures oxygenated and deoxygenated blood during brain activation. The fNIRS is much more portable technology compared to the PET, EEG, and fMRI. As is the case with fMRI, further research is needed to assess its utility in detecting deception.

The review of deception detection technologies provided by Heckmen et al. (2005) illustrates the need for additional research to measure the physiological components of deception. Further, there is a need for technologies that are non-intrusive, cost effective, portable, and provide utility based on a solid foundation of scientific research. Eye-tracking technology provides the potential to detect deception through a non-intrusive methodology that is cost effective. The use of eye-tracking technology has recently gathered increased attention due to substantial developments in using low cost

computers with sophisticated software in place of expensive proprietary hardware (Hansesn & Pece, 2005).

In contrast to technologically-based systems of deception detection, Inbau et al. (2001) described the Behavioral Analysis Interview (BAI), a method of deception detection used by law enforcement. According to the authors, the BAI attempts to identify behavioral cues to deception elicited through specific interview protocols. However, research has challenged the notion that deception is always accompanied by behavioral indicators.

Overall, several researchers have questioned the utility of arousal-based techniques of deception detection used by law enforcement agencies (DePaulo et al., 2003; Ekman, O'Sullivan, & Frank, 1999; Vrij & Mann, 2001), and many agree that a collection of cognitive, behavioral, and autonomic reactions occur when an individual is lying (Burgoon & Buller, 1994; DePaulo et al., 2003; Vrij, 2000; Zuckerman, DePaulo, & Rosenthal, 1981). Various behavioral indicators have been observed in individuals engaged in deception, such as errors in speech, hesitations in speech, and changes in voice (Ekman et al., 1999; Vrij & Mann, 2001). However, these behaviors may be a result of fear or guilt, rather than deception (Ekman & O'Sullivan, 1991; Vrij, 2000). Further, when people are engaged in making statements with serious outcomes tied to their responses, their behavior can be expected to change from their baseline behavior (Depaulo et al., 2003; Vrij, 1999).

The attempt to manage one's deceptive behavior, conceptualized as impression management, requires an individual to control indicators of nervousness while maintaining a response pattern that is consistent (Buller & Burgoon, 1996). Research has indicated that impression management can manifest itself in various physiological cues

similar to those found in faking research, such as body movement and facial expressions (DePaulo et al., 2003; Vrij, 1999). Therefore, discriminating between true physiological cues of deception and impression management can be difficult when observed through the lens of arousal-based protocols. However, as suggested by the National Research Council (2003), further research should be aimed at identifying new ways or methods of detecting deception at the physiological level.

Faking and Cognitive Load

Cognitive load can be conceptualized as the extent to which cognitive resources are utilized by activities that facilitate learning or the processing of information that occurs in working memory (Chandler, 1991). Researchers have suggested that deception leads to an increase in cognitive load due to increased effort at self-monitoring (Buller & Burgoon, 1996). When individuals engage in deception, they must monitor an array of physical and emotional changes in order for the deception or lie to be effective. Supplying a deceptive answer requires that the individual generate a response, avoid providing an inconstant response, and commit cognitive functioning to the details within the response for future responses (Granhag & Vrij, 2005).

Vrij, Fisher, Mann, and Leal (2008) examined two primary approaches to detecting deception in relation to cognitive load. The 'mere cognitive load' approach has asserted that the act of lying requires more attention and working memory demands than telling the truth. If this assertion is correct, then lying should be detectable by measuring response times, pupil dilation, and eye movement (Walczyk, 2005). Specifically, deception should be detected via longer response times, greater pupil dilation, and fewer eye movements (fixations). Vrij et al. (2008) proposed an additional approach to

detecting lying entitled the 'increase cognitive load' approach which aims to invoke cognitive load through the manipulation of experimental conditions.

These conditions may have participants recount events in reverse chronological order (Vrij et al., 2008) or interrogating participants while they are multi-tasking. Several hypotheses have been made regarding how and why an increase in cognitive load typically accompanies deception (Vrij et al., 2007). One proposition is that the creation of a lie is a cognitively demanding task. Another proposal is that liars do not take their credibility for granted, while individuals who tell the truth presume that they will be believed (Vrij et al., 2008). A third proposition is that liars are more likely than honest responders to monitor the reaction of others to assess the believability or credibility of their lie. A fourth proposition is that lying requires role-play activities to sustain the lie. A fifth proposition for why cognitive load accompanies deception is that suppressing the truth is a cognitively demanding task. Overall, the research suggests that the act of lying, deceiving, or faking requires additional cognitive resources compared to responding or acting honestly.

Many deception-based studies have suggested that several behaviors act as indicators of an increase in cognitive load. Ekman and Friesen (1969) found that individuals told to respond to a cognitively complex task exhibited a decrease in excess body movement while responding. Research has also revealed that an increase in cognitive load is indicated by a decrease in eye blinking (Harrigan & O'Connell, 1996; Wallbott & Scherer, 1991).

Cognitive Capacity

Cognitive Capacity Theory maintains that individuals have limited cognitive resources at any given time and describes cognitive processing ability as the fixed

amount of capacity accessible for single or multiple tasks (Kahneman, 1973). In other words, at any one time, various tasks must compete for a limited amount of cognitive capacity. When a lone task is being performed, all accessible resources can be used for that task. If an individual is engaged in multiple tasks, cognitive resources are divided among the tasks. As a result, when a limited amount of available capacity has to be distributed among many tasks, fewer resources are available for future tasks and performance dramatically declines. Kahneman (1973) referred to this as the divided attention effect, where performance on any task will be impaired if the person is performing a second task in parallel. The complexity of a task is a primary determinant of the amount of cognitive resources or capacity required to perform the task. Simply stated, complex tasks require more cognitive resources than simple tasks.

Researchers have examined cognitive processing using capacity models that look at performance on a secondary task as an outcome of cognitive load (Eysenck & Eysenck, 1979; Johnston & Heinz, 1978). The theory of divided attention, in alignment with capacity models, has been extensively researched (Eysenck & Eysenck, 1979; Johnston et al., 1972, Johnston & Heinz, 1979; Shulman & Fisher, 1972). Both capacity models and cognitive load theories provide a foundation to explore the physiological processes individuals undergo when engaged in deception. In the present study, cognitive resources and capacity should be diminished for individuals engaged in applicant faking or deception. Therefore, the measurement of behavioral cues found within the eye-tracking technology should differentiate between participants engaged in deception and those responding honestly.

Cognitive Ability and Faking

Since deception is a cognitive process, individuals with a larger cognitive capacity or cognitive ability may be able to fake better on personality-based questionnaires. Research suggests that individuals who score higher on a general mental ability tests are able to fake more effectively in a motivated setting (Griffith et al., 2006; Pauls & Crost, 2005). Connecting the situational requirements with the specific items on the assessment may be a form of intelligence that enables an individual to fake effectively (Pauls & Crost, 2005). Intelligence may aid an individual in understanding the underlying meaning of the item in relation to social desirability within a specific context or occupation. Additionally, higher scores on a cognitive ability assessment appear to be related to faking on extraversion, conscientiousness, and emotional stability personality dimensions (Mersman and Shultz, 1998). Individuals higher in cognitive ability may be more acutely aware of the advantage to faking a socially desirable profile on a personality measure for the sake of gaining employment. That is, individuals with a higher cognitive ability may better understand the benefits and consequences of presenting themselves in a positive light.

Since research has indicated a positive relationship between intelligence and specific job knowledge (Schmidt, Hunter, & Outerbridge, 1986), individuals with an "adopted applicant schema" appear to know how to match the job characteristics with specific items within the personality measure to present the most positive personality profile (Vasilopoulos et al., 2006). Additionally, Christiansen et al. (2005) found that more intelligent applicants formulate very accurate implicit job theories that help them match their responses to the stereotypical profile of the job. The authors went on to state that having an accurate implicit job theory will increase the ability to inflate specific

personality scales in a way that meshes with the job. Based on this premise, individuals with a higher cognitive ability and job aptitude may be better at accurately identifying the most pertinent or desirable items within a personality assessment as it relates to a particular job.

Mersman and Shultz (1998) found that cognitive ability was positively related to ability to fake the personality scales of extroversion, consciousness, and emotional stability. Subjects with a higher general mental ability tended to provide profiles on non-cognitive selection instruments that coincide with an ideal candidate for a position. Biderman and Nguyen (2004) assessed the relationship between cognitive ability and faking ability on non-cognitive measures using structural equation modeling (SEM) techniques and also found that cognitive ability was related to faking ability. However, after controlling for cognitive ability, none of the scales on the personality measure was related to their ability to fake. Cognitive ability was related to the ability to fake effectively. In addition, individuals with higher levels of cognitive ability may experience less cognitive load when faking due to an increased cognitive capacity.

Within the present study, the participants' self-reported grade point average was used to quantify their general mental ability. The grade point average of the participants was used due to research pinpointing college grade point average as an indicator of future job performance (Roth et al., 1996). Further, it has been suggested that grade point average is a useful predictor of job performance because it reflects intelligence, motivation, and other abilities pertinent to a job (Baird, 1985). It is theorized that the participants' grade point averages will have a significant negative correlation with the participants' cognitive load scores on the NASA-TLX. That is, individuals with a higher cognitive load will have a lower score on the intelligence measure. It is presumed that

individuals with a higher cognitive ability have an increased cognitive capacity that enables them to fake utilizing minimal cognitive resources.

Physiology of Faking

In their examination of automatic versus controlled self-presentation, Paulhus, Graf, and Van Selst (1989) reported that self-descriptions were more positive while an individual was experiencing high cognitive load (e.g., engaged in a distracter task). Physiological measurement of faking indicated that an individual's depletion in energy and performance directly after self-presentation episodes could be accurately assessed (Vohs, Baumeister, & Ciarocco, 2005). Galliot et al. (2007) found that depletion of energy after a self-presentation could be renewed with a boost in glucose. Brain imaging techniques have been used to reduce self-enhancement by activating the medial prefrontal cortex with transcranial magnetic stimulation (Kwan et al., 2007). Zuckerman et al. (1981) stated that lying is cognitively more complex than telling the truth, resulting in a higher cognitive load. Studies such as these are instrumental in understanding the physiological implications of applicant faking.

The Autonomic Nervous System (ANS) plays a role in the regulation of cognitive load and the associated physiological responses (Iani, Gopher, & Lavie, 2004). Specifically, the ANS involves the sensory impulses derived from the blood vessels' reaction to nerve signals from the brain (Hagemann, Waldstein, & Thayer, 2003). The ANS is a subdivision of the peripheral nervous system which regulates bodily activities that are not controlled consciously. Therefore, individuals are not consciously aware of these signals, but they may elicit automatic responses or functions throughout the body. These unconscious bodily adjustments altered through the autonomic nerves reaction to

stimuli can alter the size of the pupil, the digestive functions of the stomach, the rate of respiration, and dilation of the blood vessels (LeDoux, 2003).

The ANS is divided into the parasympathetic nervous system (PNS) and the sympathetic nervous system (SNS) (Grings & Dawson, 1978). The PNS involves the functionality of basic bodily functions such as breathing while the SNS involves more active functions involved with the fight, flight, or fright response. The particular physiological responses to SNS activation of the most interest in detecting deception are increased sweat, elevated blood pressure, dilated pupils, and increased heart and respiratory rates (Fried, 1980). Engaging in deception often results in increased levels of anxiety, which in turn, is associated with the arousal of the ANS (Horowitz, Horowitz, & Cope, 1986)

LeDoux (2003) proposed that both physiological arousal and emotional experience are produced simultaneously through the same nerve fibers. Therefore, emotions are tied to physiological reactions and the arousal associated with emotions is mediated through activity generated in the ANS. Along these lines, Levenson and Ekman (2002) discussed differences in ANS activity among specific emotions. Specifically, their study investigated facial movements related to the emotions of anger, disgust, fear, happiness, and sadness and found emotion specific autonomic nervous system activity which was related to facial expressions. An understanding of the ANS is vital to comprehending the physiological responses of individuals engaged in deception.

Response Time. Response time or response latency is the duration of time between the delivery of a stimuli and the response. Response latency has commonly been used as an indication of lying or deception. In general, it has been commonly assumed that liars or fakers will take longer to respond. The theory is that lying is more

cognitively complex and an increase in response times usually indicates more complex cognitive processing (Callegaro, Yang, Bhola, & Dillman, 2005). Generally, a respondent enacts four cognitive steps when answering a close-ended question: (a) understanding the question; (b) remembering the necessary information to form an answer; (c) evaluating the retrieved information; and (d) choosing an answer from the response options (Tourangeau, Rips, & Rasinkski, 2000). When an individual is motivated to respond a certain way, such as in an applicant setting, this cognitive process takes even longer (Callegaro et al.).

Previous research has suggested that individuals with a well-developed schema will provide faster responses (Walczyk et al., 2005). A schema is the organized knowledge that can be easily accessed and used to increase comprehension of new material (Alvarez & Risko, 1989). Research has demonstrated that individuals who are well-rehearsed or practiced in deception take longer lying than telling the truth (Vendemia, Buzan, and Green, 2005). In addition, Holden et al. (1992) pointed out that when job applicants compare their answers to their schemata they respond more slowly when lying. In sum, researchers suggested that engaging in deception requires complex cognitive processes that require more time than responding honestly. Thus, latency of response by an individual engaged in a certain task may serve as an indication of lying due to increased cognitive load.

Pupil Dilation. Measurement of pupil size of an individual to gain insight regarding the pupils' response to light, drugs, or an arousal state is referred to as pupillometry or pupillography (Fried, 1980). Dilation of the pupil is controlled by the parasympathetic and sympathetic nervous system (Wilhelm et al., 2001). The parasympathetic nervous system (PNS) controls the constriction of the pupil and is

responsible for the reflex of light and near response, while pupil dilation is altered by the sympathetic nervous system (SNS) (Wilhelm et al., 2001). Physiological arousal has been found to alter or affect the sympathetic nervous system (SNS activation) which can alter the pupil size (Grings & Dawson, 1978). Specifically, intellectual processes and physical effort may cause pupil dilation (Hess, 1972). Further, Hess (1972) found that the pupil dilation can increase 10-20% due to increased cognitive activity, such as solving multiplication problems. Additionally, pupil dilation has been found to be proportional to the difficulty of a task (Kahneman & Beatty, 1966), and dilation tends to occur within two to seven seconds from the onset of a stimulus (Hess, 1972).

Deception appears to impact pupil dilation due to an increase in cognitive load (Beatty & Lucero-Wagoner, 2000). Research has indicated a fairly consistent relationship between lying and an increase in pupil size (DePaulo et al., 2003). The linkage between cognitive load and pupil size has been examined by utilizing cognitively demanding tasks such as adding numbers (Granholm et al., 1996) or retrieving information from one's memory, thinking about a difficult topic, and pausing during a complex speech (Beatty & Lucero-Wagoner, 2000). Increased in pupil dilation has been demonstrated in individuals performing a mental arithmetic problem (Ahern & Beatty, 1979), as well as in individuals engaged in verbal comprehension tasks (Wright & Kahneman, 1971). These findings illustrate how general cognitive processing may alter the pupil size of an individual. The current study utilized the measurement of pupil dilation to examine its relationship to faking on certain items within the personality assessment.

Eye Fixations. Eye movements also appear to be related to deception. Rayner (1998) suggested that eye movements (i.e., eyes moving too quickly during saccades), or lack of eye movement (i.e., fixations), may provide clues to cognitive processes. Along

these lines, Rayner found that increased cognitive load is related to an increase in eye fixations. Further, Rayner's research indicated that eye saccades and fixations may serve as an indication of cognitive load. In a related study, Doherty-Sneddon et al. (2002) found that individuals engaged in cognitively demanding activities tended to break eye contact or look away. That is, individuals tended to avoid visual stimulation when confronted with cognitive demands in order to minimize distractions. Although there may be a corresponding increase in eye movements and breaking eye contact when people lie, individuals who engage in deception may be aware of this stereotype and may make a conscious effort to increase eye gaze to bolster perceptions of honesty (Eckman & Friesen, 1969).

In addition to eye fixations, DePaulo et al. (2003) found that individuals have larger pupil sizes when lying than when telling the truth. In addition to larger pupil sizes, Depaulo et al. found that the liars were less cooperative in the experiment and were less likely to admit not remembering something. The authors suggested that pupil size may play a valuable role in serving as a cue to deception and provide a direct indication of information-processing activities of the individuals engaged in deceptive behavior.

Issues with Within/Subjects Design

The study of applicant faking or socially desirable responding has used various experimental techniques. Many studies utilized between-subjects designs that explore the score differences between job incumbents and job applicants (Birkeland et al., 2006), score differences on personality scales between experimentally manipulated conditions of faking good and responding honestly (Mueller-Hanson, Heggestad, & Thornton, 2003), or score differences on socially desirable responding scales across various conditions (Veswesvaran & Ones, 1999). As an alternative to between-subjects designs, some

researchers have utilized within-subjects designs using the aforementioned conditions of experimentally manipulated conditions of faking good, faking bad, and responding honestly (McFarland & Ryan, 2000).

These within-subjects designs sometimes compare across research and applicant conditions (Ellingson, Sackett, & Connelly, 2007; Griffith et al., 2007). Within-subjects designs are prevalent in the literature and have several benefits. These benefits are based upon the assumption that individuals will respond differently to items on a personality measure as job applicants where the situation exerts motivational influences on respondent behavior that are absent in situations unrelated to employment outcomes (Peterson, Griffith, Converse, & Gammon, 2011). Further, the within-subjects design provides a measurement of an individual's score change across conditions. Mesmer-Magnus and Viswesvaran (2006) reviewed the research designs utilized in faking research and identified several benefits to within-subjects designs. For example, the authors found that within-subjects designs provided greater statistical power and typically resulted in larger effect sizes compared to between-subjects designs.

Although within-subjects designs provide several benefits, Mesmer-Magnus and Viswesvaran (2006) noted that within-subject designs are subject to many threats to internal validity. These threats include testing, history, and maturation (Cook & Campbell, 1979). One threat to the internal validity for within-subjects designs is the effect of the order of administration (Ellingson et al., 2007; McFarland, Ryan, & Ellis, 2002; McFarland & Ryan, 2000). This occurs in within-subjects designs because participants may become bored, tired, fatigued, or less motivated across conditions resulting in inaccurate responding. Within-subject research designs often utilize a counterbalancing technique to hopefully cancel out the error of these order effects (Cook

& Campbell, 1979); however, this methodology has flaws. For example, McFarland and Ryan (2000) found order effects when the honest condition precedes the condition in which the participants are instructed to respond in a positive fashion or "fake-good".

McFarland et al. (2002) performed a similar study and found a significant order effect for the conscientiousness personality scale. That is, individuals who provided honest responses prior to "fake-good" responses reported greater score increases in the applicant condition. Further, a study by Ellingson, Sackett, and Hough (1999) found that participants that responded honestly prior to responding in the faking condition resulted in higher "fake-good" scores on the personality measure. In addition, they found that participants that responded to the "fake-good" instructions prior to responding honestly had higher honest condition scores. The authors referred to this as a "repeated measures effect". Pauls and Crost (2005) illustrated order effects by having participants complete a personality questionnaire in an honest response condition before providing responses under various faking conditions. Some of the participants provided responses within both the "fake-good" and "fake-bad" conditions. The researchers found that the agreeableness scores on the personality measure for the "fake-bad" condition were higher when the participants were engaged in the "fake-good" condition first. In addition, they found that impression management scores for participants in the "fake-good" condition were higher when participants provided "fake-bad" responses.

A method used to counteract the order effect for within-subjects experimental designs in applicant faking research is the applicant-applicant pretest-posttest design utilized by Hogan et al. (2007). This design has applicants complete the personality assessment several times in each condition and any improvement across test administrations serves as an indication of applicant faking. Kelley, Jacobs, and Farr

(1994) utilized this method (applicant-applicant retest) and found MMPI profile changes of an individual profile across multiple administrations. In addition, using this methodology, Landers, Sacket, and Tuzinski (2011) found increases in extreme responding of participants re-taking a personality assessment.

The changes or order effects for the honest-applicant administration order could be due to various factors. Peterson et al. (2011) suggested that exposure to items a second time inflates scores in the applicant context due to prior knowledge of the test content and the additive influence of faking. Further, the authors suggested that providing honest responses first may motivate consistency in responses resulting in a decreased motivation to fake. Their research investigated these order effects, and consistent with previous research, found that participants' applicant condition agreeableness and amount of faking indices were significantly higher when the honest condition was administered prior to the applicant condition.

Overall, it seems that order effects are prevalent for applicant faking research that utilizes within-subjects designs. The present study utilizes a between-subjects experimental design to eliminate order effects and gain more credible and valid data pertaining to detection of faking using the eye-tracking technology. Additionally, a within-subjects component of levels of cognitive load is used. Although power and effect sizes may be reduced due to the between-subjects design (Mesmer-Magnus & Viswesvaran, 2006), this method should provide a more realistic view of the physiological differences between faking and honest conditions in a controlled experiment (McFarland et al., 2002).

Construct Validity

The process of construct validation is appropriate any time a test score is interpreted as an assessment of some characteristic or variable that is not currently operationally defined (Cronbach & Meehl, 1955). Construct validation investigates whether a particular scale is related with the theorized psychological construct of interest. The current study examined the relationship between pupil dilation, eye fixation, response latency and a validated measure of cognitive load (i.e., NASA-TLX) to assess convergent construct validity. The need to assess construct validity stems from the lack of previous research to identify the link between these physiologically-based variables and cognitive load in a simulated applicant setting.

Construct validity is sometimes more easily understood through the Multitrait-Multimethod Matrix (MTMM) developed by Campbell and Fisk (1959). The MTMM examines correlations of traits across different methods of assessment. For instance, measuring cognitive load via eye-tracking technology and through the use of the NASA-TLX cognitive load measure is an example of a multi-method, same trait analysis. Correlations that fit the hypothesized pattern provide evidence for the construct validity of the measurements provided by the eye-tracking technology. Further, this analysis of validity can be referred to as convergent construct validity, as it provides evidence that the measurements provided by the eye-tracking technology are related to the theoretical construct (Pennington & Donald, 2003). The use of the NASA-TLX also provides a manipulation check to ensure that the experimental methodology in the study actually induced cognitive load.

In addition to exploring the relationship between the measures provided by the eye-tracking technology and cognitive load, the use of socially desirable responding

items from the Marlowe-Crowne Socially Desirable Responding Scale should provide additional evidence of construct validity. The Marlowe-Crowne Socially Desirable responding scale was used because it is one the oldest and most widely used social desirability scales (Crowne & Marlowe, 1964).

Binning and Barrett (1989) pointed out that the terms construct validation and theory development imply the same process as both refer to the action of finding constructs through the development of construct assessments and examining the relationships among the various measures. Based on this description, the authors suggested that assessments are nothing more than methods for sampling certain behaviors within a specific construct domain.

Participant Motivation

Van Hooft and Born (2011) provided instructions to researchers for the creation of experimental manipulations that mimic actual faking conditions. Specifically, the authors suggested informing participants in the faking condition that they should attempt to deceive or fake the questionnaire without getting caught. A financial incentive, such as the one used in the present study, may increase external validity (Bassi, Morton, and Williams, 2011). The labor theory of cognition (Smith & Walker, 1993) suggests that financial incentives can effectively induce motivation in subjects. Further, Smith and Walker posited that incentives will cause subjects to devote cognitive resources in making choices in the same way they would outside of the experimental setting.

In addition to financial incentives within an experiment, verbal praise or feedback may increase the intrinsic motivation of the participants (Cameron & Pierce, 1994). Along these same lines, a study by Deci, Koesnter, and Ryan (1999) found that positive feedback improved free choice behavior and self-reported interest in the study. In the

present study, the participants were told that immediate feedback would be provided after completion of the personality assessments to gauge their level of performance or 'faking ability.' This, in conjunction with the financial incentive, was used to increase participants' motivation to fake well and more closely mirror a real applicant mindset or scenario.

Summary

Current methods of detecting faking on personality-based measures have utilized validity scales, socially desirable responding scales, and even measures of integrity. However, these methods have not always accurately identified individuals who are faking to gain employment on a personality measure. In addition, these techniques have relied on self-report data that have questionable validity. Due to the prevalence of evidence illustrating the problems with using self-report detection techniques (Viswesvaran & Ones, 1999), techniques that do not utilize self-report measures may be a better choice in the detection applicant faking.

Physiologically-based assessments such as eye gaze, response time, and pupil dilation are objectively measurable and are difficult for examinees to modify, change, or fake. These methods overcome deliberate coaching to provide fake symptoms and provide investigation of the cognitive process of faking. The physiological processes of deception or faking may more accurately detect faking without the underlying issues related to self-report measures. The methods utilized in this study investigate cognitive cues such as response time, eye focus, and pupil dilation to investigate the cognitive process of faking on personality measures within a simulated applicant setting.

Hypotheses

Hypothesis One. Participants' pupil dilation will increase significantly as items on the personality scale increase in social valence from low to moderate to high.

Hypothesis Two. Participants will have a higher number of eye fixations as items on the personality scale increase in social valence from low to moderate to high.

Hypothesis Three. Participants will show significant increases in response latency (milliseconds of response per stimulus items presented) as items on the personality scale increase in social valence from low to high.

Hypothesis Four. Participants assigned to the faking condition will have a significantly higher cognitive load as measured by pupil dilation than participants in the honest condition across the three item valence categories

Hypothesis Five. Participants assigned to the faking condition will display a significantly higher ratio of eye fixations per second than participants assigned to the honest condition.

Hypothesis Six. Participants assigned to the faking condition will demonstrate a significantly longer latency in completing responses to stimulus questions (i.e., time of completion) than participants assigned to the honest condition.

Hypothesis Seven. Participants assigned to the faking condition will have higher scores on the NASA-TLX, a cognitive load measure, than participants assigned to the honest condition.

Hypothesis Eight. Participants' self-reported grade point average will be inversely related to their scores on the NASA-TLX cognitive load measure.

Hypothesis Nine. There will be positive relationships between the eye-tracking data, including the number of total number of eye fixations, the total number of dilations,

the total for latency (the sum of the participant's average across the three valence categories) and the total NASA-TLX for participants in the faking condition.

CHAPTER TWO

METHOD

The purpose of this study was to determine if there are differences in certain physiological and behavioral measures, such as gaze fixations (i.e., saccades), pupil dilation, and response time can be used to detect faking on personality items answered by participants in a simulated applicant setting. A questionnaire was developed based on a panel of judges' ratings of item social desirability. Specifically, six doctoral students comprised the panel of judges. The panel rated the degree to which they believed participants, with the intent of obtaining the job like the one used in this study, would be inclined to fake each particular item. After signing an Informed Consent form (see Appendix A), participants were presented a generic job description to provide them with basic knowledge, skills, and abilities needed for the job to which they were applying.

A between-subjects design was utilized to assess physiological differences between the experimental condition (asked to fake positively to the items on the questionnaire) and a control condition (asked to respond to the questionnaire honestly).

Participants

The sample for this experiment consisted of adult college students ranging in age from 18 to 47 years who were enrolled in undergraduate/graduate college courses at a midsized southern university. A total of 80 participants completed the study. However,

due to initial technical problems, eye-tracking data (gaze fixations and pupil dilation) was successfully captured for only 78 participants.

A brief description of the study was presented in several classrooms to recruit volunteers for the study and to inform students of the basic premise of the study and the length of time to which potential participants were committing. A signup sheet was circulated and participants volunteered for one-hour blocks of time. For their participation, participants received extra credit. In addition, participants were notified that they would be entered into a raffle for a $20 gift card for their participation. The raffle was used as an attempt to increase participant motivation and engagement in the study.

Instruments

Demographic Questionnaire. Demographic information was gathered from each participant which included their sex, age, ethnicity, student classification (year in school), and college major (see Appendix B). In addition to the basic demographic information, participants were asked to report their cumulative college grade point average.

International Personality Item Pool. Items from the International Personality Item Pool Big Five Personality (IPIP) scale were selected based on the preliminary social desirability ratings survey completed by the panel of judges. Specifically, the IPIP sub-scales (Goldberg et al., 2006) for Neuroticism ($\alpha = .86$), Extraversion ($\alpha = .86$), Openness to Experience ($\alpha = .82$), Agreeableness ($\alpha = .77$), and Conscientiousness ($\alpha = .81$) were used within the initial item pool. That is, all were included in the survey administered to the panel that rated the social desirability of each item. Based on the panel's evaluations, certain items were selected for use in the primary study (see Appendix C). The five items, deemed most socially desirable by the panel of judges, were utilized as the high valence items. These five items were selected by general mean

score calculations. In addition to the five highly socially desirable items, eight moderate to low socially desirable items were also used as a stimulus within the study. A total of 13 stimulus items were used in the study, representing various levels of social desirability to gauge response differences between the levels of social desirability.

Marlowe-Crowne Social Desirability Scale (MCSDS). Items from the Marlowe-Crowne Social Desirability Scales (MCDS) (Crowne & Marlowe, 1960) were used to assess the level of socially desirable responding occurring in each participant. The MCDS was designed for administration to non-pathological individuals with higher scores indicating a need for approval (Crowne & Marlowe, 1960). The shortened 11-item MCDS (Reynolds, 1982) was utilized to determine the degree to which participants respond in a socially desirable manner. The MCDS utilizes two types of items to measure social desirability. Some of the items ask about common behaviors that may be described as desirable, for example, admitting mistakes (Crowne & Marlowe, 1964). There are also items on the MCDS that ask about behaviors that are not considered socially desirable, such as gossiping (Crowne & Marlowe, 1960).

Crowne & Marlowe (1964) reported the internal consistency reliability of the MCSDS has a range from .73 to .88. The authors also note that a convergent validity study of the MCDS using undergraduate students found a significant relationship with the construct of need for approval. In addition, the authors found that individuals who score high on the scale tend to respond to social reinforcement, inhibit aggression, and are more likely to be influenced by external pressure.

Cognitive Load NASA-TLX. Hart and Staveland's (1988) NASA Task Load Index (TLX) cognitive load measure was used to assess work load (see Appendix E). Within the NASA Task Load Index (TLX), perceived subjective cognitive load or work

load is viewed as a multidimensional construct involving the participants' appraisal of their: (a) mental demand, (b) physical demand, (c) temporal demand, (d) performance, (e) effort level, and (f) frustration level. These dimensions were presented as individual items in a paper-pencil questionnaire format. Participants used a 21-point scale to indicate their level of work load on each item.

Work Experience Questionnaire. All participants were asked to complete a work experience questionnaire prior to using the eye-tracking equipment. This questionnaire asked participants about their most recent two jobs, their job responsibilities and educational level. Participants were also asked specific job knowledge questions relating to the IPIP items administered during the eye-tracking phase. Specifically, participants' knowledge, skills, and abilities related to items on the IPIP were included to evaluate participants' "adopted applicant schema" (Vasilopoulos et al., 2006). Individuals with specific experience and knowledge for areas identified within the generic job description may be better at faking or presenting a positive personality profile.

Eye-tracking Software. ITU Gaze Tracker (Agustin, Skovsgaard, Hansen, & Hansen, 2009) and Ogama (Voßkühler, 2009), two open source programs, were used to capture the eye-tracking data of participants in the study. Ogama uses a slide show format to present stimuli while recording data through the ITU Gaze Tracker software. Ogama also provides pre-processing and filtering of ITU Gaze Tracker data. When used with a camera, infrared lights, and the ITU Gaze Tracker program, Ogama gathers point-of-regard data. Point-of-regard is the position on the display being viewed by the subject. This may be superimposed in real time over the researcher's scene image display. Two computers were used to administer the stimulus and record the eye-tracking data.

Eye-tracking Hardware

A Thorlabs DC high resolution camera was used in conjunction with the Ogama eye-tracking software. A Fujinon lens, which allows the infrared spectrum of light to enter the camera, was attached to the camera to enhance focus on eye movements and changes in pupil size using infrared imaging. The camera and lens were mounted to a tripod stand (see Figure 1). In addition, two lights, each containing 24 infrared LEDs, were mounted approximately five inches below and five inches to each side of the camera to increase the clarity and visibility of eye movements. A chin rest was mounted 26 inches from the presentation screen (Hansen & Hansen, 2006), and could be adjusted up and down a range of 10 inches to alter the height appropriately for each participant. The chinrest was designed to decrease the incidental movement during data collection, reducing error in the measurement of eye movements and pupil dilation.

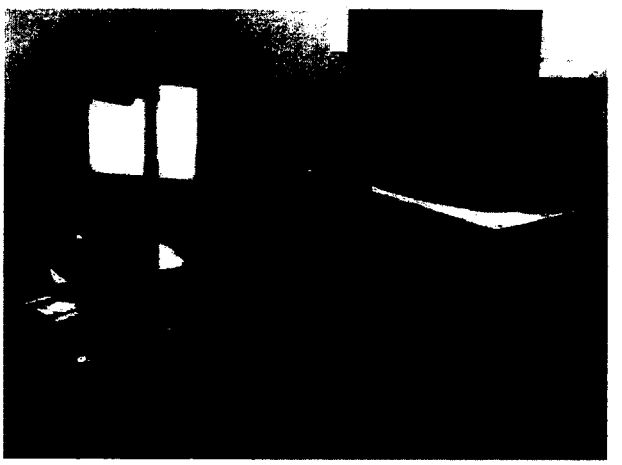


Figure 1. Eye-tracking system.

Generic Job Description

A job description (see Appendix F) was provided for each participant to review before they responded to the IPIP and MCSD items. Some of the content for the job

description was derived from the personality items that were selected for use in the study. The O*NET Occupational Database was used as the primary source for content for the job description's knowledge, skills, and abilities. An 'ideal profile' of responses to the personality items was created using ratings of panel judges to allow the experimenter to assess each participant's ability to fake effectively. Specifically, items were categorized as high, moderate, or low in social desirability (i.e., valence).

Procedure

Participants were assigned to the faking or honest condition. Each participant completed the short demographic questionnaire prior to administering the initial personality measures. Participants were informed of the general purpose of the study and the basic protocol. In addition, participants were given a job description to review that provided details about the knowledge, skills, and abilities of a hypothetical job. Participants were then provided with a brief orientation to the eye-tracking apparatus and the system was calibrated.

Participants in the faking condition were told to respond to the questions presented on the screen as if they were applying for a job which required the knowledge, skills, and abilities outlined in the generic job description. Participants in the faking condition were also instructed to respond to the items as they would expect an ideal applicant might respond to the items. In addition, these participants were informed that if they faked the responses that typify the ideal applicant's response profile with minimal detection, they would be entered into a drawing for a gift card. The ideal profile used the mean of the panel judges, plus or minus five points (i.e., the range of panel). Participants in the faking condition received immediate feedback regarding their ability to match the ideal profile and whether their performance was good enough to be entered into the drawing. All other

participants, those in the faking condition who did match the ideal applicant profile and those in the honest condition, were entered into a separate drawing for another gift card of the same value.

Participants assigned to the honest condition were also given the brief demographic survey and the generic job description for review. They were instructed to complete the measure honestly and accurately.

After the initial instructions, the personality questionnaire items were presented one by one using the Ogama presentation software. Non-intrusive items (i.e., items with low social valence) were presented first in order to collect a baseline for each participant. Upon completion of the IPIP and Marlowe-Crowne items selected from the preliminary assessment, the participants were disconnected from the eye-tracking technology and directed to a table to complete the NASA-TLX cognitive load measure. Upon completion of the cognitive load measure, participants were asked to complete a brief questionnaire asking them to reflect how well they thought they had faked and what cognitive process they used while faking. Participants also reported on the extent to which it was difficult to fake their responses. Finally, participants were given feedback on how well their profile matched the ideal profile for the job.

CHAPTER THREE

RESULTS

Because of the large amount of data generated by the eye-tracking system, "data condensing" was performed. The average pupil dilation and eye fixations for each stimulus slide (e.g., item) was used as the unit of analysis. For each participant the dilation data was reduced to the average pupil dilatation, per stimulus slide. In addition, the number of eye fixations per stimulus slide was divided by the total number of seconds taken to complete the slide to create eye fixations variable. Latency was defined as the total amount of time, in milliseconds, participants took to complete their response to each stimulus slide containing one of the personality items that were deemed low, moderate or high valence.

Hypotheses one through three were tested using a 2 x 3 (faking/non-faking condition by low, moderate and high valence) mixed repeated measures analysis of variance (ANOVA). Hypotheses four through seven were tested using a multivariate analysis of variance (MANOVA) in order to control for potential collinearity between fixation and dilation data. Analysis of variance procedures were used to compare and investigate within and between group differences. The within-subjects factor was question type with three levels: high, low, and moderate item valence. The between-subjects factor was the faking condition with two levels: respondents were asked to fake

on the personality items to portray an ideal employee or to respond honestly. Effect sizes were reported as eta squared ($\acute{\eta}^2$) coefficients.

Results for Within-Subjects Hypotheses

Hypothesis One. Participants' pupil dilation will increase significantly as items on the personality scale increase in social valence from low to moderate to high.

Means and standard deviations for dilation are reported in Table 1 by item valence category. No significant main effect for social valence level was found, $F(1,77) = .161$, $p = .290$, $\acute{\eta}^2 = .004$. Pupil dilation did not appear to increase with social valence level. The findings of the present research do not support previous research which indicated that increasing cognitive load increases pupil dilation (DePaulo et al., 2003; Beatty & Lucero-Wagoner, 2000).

Table 1

Item Valence (Means and Standard Deviations)

	Item Valence		
Variable	Low	Moderate	High
Dilation	75.91 (16.25)	76.05 (15.84)	75.49 (16.55)
Latency (ms)	3977.91 (1241.15)	3882.52 (1109.24)	3963.25 (1450.93)
Fixations/second	1.57 (.72)	1.57 (.85)	1.52 (.75)

Note. N = 78.

Hypothesis Two. Participants will have a higher number of eye fixations as items on the personality scale increase in social valence from low to moderate to high.

Means and standard deviations for eye fixations are reported in Table 1 by item valence. No significant effect for eye fixations was found, $F(1,77) = 1.47$, $p = .232$, $\acute{\eta}^2 =$

.019. Eye fixations did not appear to increase with social valence level. This finding is not consistent previous research which has indicated that increasing cognitive load should increase the number of eye fixations (Rayner, 1998). In the present research, differences in item valence were not related to differences in eye fixations.

Hypothesis Three. Participants will show significant increases in response latency (milliseconds of response per stimulus items presented) as items on the personality scale increase in social valence from low to high.

Means and standard deviations for the amount of time in milliseconds the participants took to complete each stimulus slide are reported in Table 1 by item valence. No significant effect was found for, $F(1,77) = .381, p = .684, \acute{\eta}^2 = .005$. No significant differences were found in the amount of time it took participants took to complete their responses to the stimulus slides for the three levels of valence (i.e., high, low, and moderate). This finding is not consistent with previous research which suggested that items or tasks which increase cognitive load will also increase the amount of time taken to complete a specific task (Tourangeau, Rips, & Rasinkski, 2000; Callegaro et al., 2005).

Results for Between-Subjects Hypotheses

Hypothesis Four. Participants assigned to the faking condition will have a significantly higher cognitive load as measured by pupil dilation than participants in the honest condition across the three item valence categories

Means and standard deviations for the participants' dilation for each condition are reported in Table 2. No significant difference in pupil dilation were found between the faking and honest conditions, $F(1,77) = 1.953, p = .166, \acute{\eta}^2 = .025$. This finding is not

consistent with previous research which has indicated a relationship between cognitive load and increased pupil dilation (DePaulo et al., 2003; Kahneman & Beatty, 1966).

Table 2

Participant Condition (Means and Standard Deviations)

	Condition	
Variable	Faking	Honest
Dilation	78.31 (15.14)	73.27 (16.65)
Latency (ms)	4198.84 (1357.50)	3701.67 (831.83)
Fixations/second	1.78 (.80)	1.33 (.60)
NASA-TLX	33.45 (12.45)	18.75 (9.44)

Note. N = 78.

Hypothesis Five. Participants assigned to the faking condition will display a significantly higher ratio of eye fixations per second than participants assigned to the honest condition.

Means and standard deviations for the participants' ratio of eye fixations per second within each condition are reported in Table 2. Participants assigned to the faking condition displayed a significantly higher ratio of eye fixations per second than participants assigned to the honest condition, $F(1,77) = 7.625, p = .007, \eta^2 = .091$. Participants assigned to the faking condition engaged in more fixations per second ($M = 1.78$, $SD = 0.80$) than participants in the honest condition ($M = 1.34$, $SD = 0.60$). This finding is consistent with previous research which has indicated that eye fixations increase as cognitive load increases (Rayner, 1998).

Hypothesis Six. Participants assigned to the faking condition will demonstrate a significantly longer latency in completing responses to stimulus questions (i.e., time of completion) than participants assigned to the honest condition.

Means and standard deviations for response latency are reported in Table 2. No significant difference in response latency was found among the three social valence of item $F(1,79) = 3.901, p = .052, \hat{\eta}^2 = .044$. These findings are not consistent with previous research findings which has indicated that when an individual is motivated to respond a certain way, such as an applicant who is motivated to get a job, the cognitive load may increase and response time may take longer (Callegaro et al., 2005).

Hypothesis Seven. Participants assigned to the faking condition will have higher scores on the NASA-TLX, a cognitive load measure, than participants assigned to the honest condition.

Means and standard deviations for participants' score on the NASA-TLX within each condition are reported in Table 2. A significant difference in mean score was found between participants assigned to the faking condition and participants assigned to the honest differences between conditions on the NASA-TLX, $F(1,80) = 35.397, p = .000, \hat{\eta}^2 = .311$. Specifically, participants in the faking condition scored higher on the NASA-TLX ($M = 33.45, SD = 12.45$) than participants in the honest condition ($M = 18.75, SD = 9.44$). These results supported previous research which has indicated deception leads to an increase in cognitive load (Buller & Burgoon, 1996; Vrij et al., 2008).

Results for Additional Analyses

Hypothesis Eight. Participants' self-reported grade point average will be inversely related to their scores on the NASA-TLX cognitive load measure.

A Pearson's Product Moment Correlation was calculated to examine the relationship between the participants' self-reported grade point average and participants' total score on the NASA-TLX (see Table 3). This score is a composite of the six scales on the NASA-TLX. The correlation between participants self-reported grade point average and total score on the NASA-TLX was not statistically significant, r (79) = -.028, p = .809.

Table 3

Correlation Matrix

	Fixations/sec	Dilation	Latency (ms)	Shipley-2
NASA-TLX	.131	-.105	.321*	-.028
Fixations/sec		-.083	.140	-.153
Dilation			-.137	-.154
Latency (ms)				-.198

*p < 0.01

Hypothesis Nine. There will be positive relationships between the eye-tracking data, including the number of total number of eye fixations, the total number of dilations, the total for latency (the sum of the participant's average across the three valence categories) and the total NASA-TLX for participants in the faking condition.

A zero order correlation matrix was used to examine the relationship among the variables. The analysis revealed a statistically significant positive relationship between the NASA-TLX and latency (r = .321, n=80, p =.004) but no significant relationships were found between participants' total score on the NASA-TLX and eye fixations (r = .131, n = 78, p = .254) and between participants' total score on the NASA-TLX and their pupil dilation (r = -.105, n = 78, p = .359). While the latency and cognitive load

relationship provides some evidence that latency may be an indication of increased cognitive load, the lack of a relationship between the NASA-TLX and both dilation and eye fixations is not consistent with the findings of previous research (Grings & Dawson, 1978; Hess, 1972).

CHAPTER FOUR

DISCUSSION

The primary purpose of this study was to assess the utility of eye-tracking technology in the detection of applicant faking on personality measures. Specifically, this study aimed to identify physiological cues associated with deception or faking in a job applicant scenario.

Eye-tracking technology has been utilized in previous research to measure deception (DePaulo et al., 2003; Callegaro et al., 2005) and it has demonstrated a strong and consistent relationship with cognitive load (Walczyk, 2005). The assumption has been that individuals engaged in deception on personality measures experience increased cognitive load due to the cognitively complex task of searching one's memory for a response that matches their perceived ideal profile while maintaining consistency in their responses during the duration of the examination.

This study provided partial evidence to support the utility of employing the use of eye-tracking technology to detect faking on personality measures. Differences in physiological data were recorded through the eye-tracking technology, revealing a significant increase in the number of eye fixations per second for participants within the faking experimental condition. Further, standardized paper-pencil assessment of cognitive load revealed similar differences between the experimental conditions. That is, individuals engaged in faking behavior had significantly more eye fixations per second

and recorded significantly higher scores on the paper-pencil measure of cognitive load. Although these results lend support to the use of eye-tracking technology as a form of detecting applicant faking on personality-based measures, other physiological measures did not. There were no statistically significant differences in levels of valence across any of the eye-tracking variables.

Conclusions

Interpretation of Hypotheses One through Three. The analyses failed to support previous research findings of a positive relationship between cognitive load and deception, pupil dilation (DePaulo et al., 2003), eye fixations (Rayner, 1998), and response latency (Tourangeau, Rips, & Rasinkski, 2000), the valence of the item did not alter dilation, fixation, or latency in this study. This indicates that item valence did not influence the dependent variables. These results do not support Hypotheses One, Two, and Three and failed to support previous research involving deception and its influence on pupil dilation (DePaulo et al., 2003), eye fixations (Rayner, 1998), as well as latency (Callegaro et al., 2005).

One explanation could be cognitive load was sufficiently increased through the experimental condition such that item valence yielded no increase in cognitive load. That is, the experimental condition may have increased the participant's cognitive load to an extent that there was a restriction in the amount it could increase due to the influence of item valence. Although certain items are commonly used to indicate faking on a personality measure, they may significantly alter a person's physiological response. The simulated applicant scenario used in this experiment seems to have altered the participants' cognitive load more than item valence. Further, the panel may not have been the most effective method for identifying differences in social desirability in personality

items. A larger panel of raters with a greater number of personality items may have provided more differentiation between valence categories.

An actual applicant scenario could possibly elicit individual differences in cognitive load due to item valence in the absence of direct instructions to fake. Specifically, an actual applicant scenario may include individuals who don't intend on faking on the personality measure but end up engaging in socially desirable responding (i.e., faking) when presented items deemed highly related to the job of interest.

Interpretation of Hypotheses Four through Seven. The analyses yielded mixed results as some of the proposed physiological indicators of cognitive load were statistically different for participants in the honest condition while other indicators failed to indicate statistical differences between conditions. Specifically, the number of eye fixations per second revealed a significant difference between the experimental conditions, with participants in the faking condition revealing significantly more eye fixations per second per stimulus slide than participants in the honest condition. This supported the notion that an increase in eye fixations is an indicator of an increase in cognitive load for individuals engaged in deception and supports previous research on deception and eye fixations (Rayner, 1998). However, the analysis also revealed no significant differences between experimental conditions for the variables of pupil dilation and response latency. These findings are counter to previous research linking deception and pupil dilation (DePaulo et al., 2003) as well as the cognitive load incurred through deception and response latency (Callegaro et al., 2005).

The NASA-TLX cognitive load measure revealed significant differences between the faking and honest experimental conditions. Specifically, participants within the faking condition had significantly higher scores on the measure indicating participants'

perception of a more cognitively complex task. This finding is aligned with previous research on cognitive load (Vrij et al., 2008) which suggested that the act of deception provokes an increase in cognitive load. This finding illustrates the fidelity of the experimental manipulation of condition type.

Although the between-subjects analyses revealed mixed results, the findings illustrated a general increase in cognitive load for participants within the faking experimental condition. Not only does this provide interesting information pertaining to applicant faking and physiological behaviors, the results help to validate the experimental manipulation. The use of a student population in a contrived scenario as used in this study resulted in the anticipated increase in cognitive load in the experimental condition. This indicates that subjects followed the instructions to fake good in this study. This lends support to the use of students as subjects in scenario-based research.

Interpretation of Hypothesis Eight. A correlation analysis was used to investigate the relationship between cognitive ability and cognitive load. Specifically, the relationship between the participants' self-reported grade point average and the NASA-TLX measure of cognitive load was assessed to investigate how participants' cognitive ability alters their self-reported cognitive load. This relationship is of interest due to the preponderance of previous research suggesting that individuals with a higher cognitive ability have additional cognitive capacity to endure cognitive load during activities that possibly provoke cognitive load (i.e., deception). No relationship was found, indicating no link between participants' cognitive ability and their self-3reported cognitive load while engaged in deception. This does not support Hypotheses Eight and deviates from previous research which has indicated that individuals with higher cognitive functioning

have an increased ability to engage in deception or faking while enduring minimal cognitive load (Griffith et al., 2006).

Interpretation of Hypothesis Nine. The purpose of this analysis was to investigate whether the dependent measures of fixations, dilation, and latency are related to cognitive load. The analysis revealed a significant positive relationship between the NASA-TLX and the amount of time taken to complete the personality assessment. This is consistent with previous research on response latency and cognitive load (Tourangeau, Rips, & Rasinkski, 2000) and lends evidence of convergent construct validity. However, the number of eye fixations per second and pupil dilation did not have a significant relationship with the NASA-TLX. This is counter to the hypotheses and does not provide support for the construct validity of using these variables in the measurement of cognitive load.

Implications

Eye-tracking technology has been used to investigate reading, information processing, visual search (Rayner, 2008), and the relationship between eye fixations and cognitive load (Chandler, 1991). Eye-tracking technology has also been used to study individual differences in studies on optimism (Isaacowitz, 2005) and trait anxiety (Calvo & Avero, 2002). This study explored the use of eye-tracking technology in detecting applicant faking on a personality assessment. Specifically, this study attempted to provide an additional avenue of deception detection that does not rely on self-reported impression management or socially desirable responding scales (Zickar & Robie, 1999). The exploration of alternatives to the measurement of faking is needed due to issues with the validity of self-report methodologies of faking detection (Griffith & Peterson, 2008). The use of response latency, eye fixation, and pupil dilation have been used in previous

research and appear to show promise as ways to measure deception and cognitive load (Beatty & Lucerno-Wagoner, 2000; Walczyk et al., 2009,; Zuckerman et al., 1981).

This study evaluated the response processes of individuals engaged in deception on a personality assessment and contrasted these responses with individuals responding honestly. Further, this study assessed the physiological reactions at the item level to discern if item content influenced the cognitive load of an individual beyond the initial influence of the experimental condition. The method utilized was suggested by van Hooft and Born (2011) as a way to gauge how items with a high social desirability alter the physiological reactions compared to items with a low social desirability. The results of this study suggest that item content does not influence eye fixation, pupil dilation, and response latency as they relate to deception beyond the influence of the experimental condition. That is, the assignment or designation of a participant into the faking condition provoked cognitive load for all participants regardless of item valence.

In addition to assessing the valence of the personality items and their influence on the participants' cognitive load, analyses were performed to investigate the differences between subjects in the faking and honest conditions. The NASA-TLX measure of cognitive load revealed significant differences between the conditions, such that subjects within the faking condition reported significantly higher scores than those in the honest condition. This is consistent with previous research suggesting that the act of deception is cognitively more complex than telling the truth and results in higher cognitive load (Zuckerman et al., 1981). This study also found that subjects within the faking experimental condition revealed more eye fixations per second than subjects in the honest condition. This supports previous eye-tracking research describing an increase in cognitive load related to an increase in the number of eye fixations (Rayner, 1998).

Although the analysis revealed significant differences in cognitive load via the NASA-TLX and the average number of eye fixations, no significant differences were found for the pupil dilation or response latency. The failure to find any difference in pupil diameter or dilation between in the faking and honest conditions is contrary to previous research (DePaulo et al., 2003). In addition, the results suggest that there was no difference between conditions on response latency. This is also in contrast to previous research that suggests an increase in response time corresponds with increased cognitive load (Callegaro et al., 2005).

This study attempted to explore the relationship between participants' perception of cognitive load when faking and cognitive ability. The hypothesis that cognitive ability would be positively related to perceptions of cognitive load was based on previous research which suggested that cognitive ability is associated with an increased ability to fake "good" on personality assessments (Griffith et al., 2006).

Convergent construct validity evidence was assessed in this study to identify the relationship between a validated measure of cognitive load (NASA-TLX) and eye fixation, pupil dilation, and response latency. This analysis was performed to assess the validity of using eye-tracking variables as measures of cognitive load. The results indicated limited construct validity evidence, through the significant positive relationship of response latency and cognitive load. However, no significant relationships existed between the NASA-TLX and pupil dilation or eye fixations.

The lack of statistically significant differences in physiological responses from the eye-tracking technology and response latency between categories of item valence was counter to what was hypothesized. However, the data indicated that the level of cognitive load varied between experimental conditions. In short, the experimental conditions likely

account for the alterations in cognitive load regardless of the item content. The greater number of eye fixations in the faking condition was consistent with previous research and provided support for using this variable in studying the extent to which an applicant is engaged in faking behavior on a personality assessment.

Limitations

The current study had several limitations that may have influenced the results. One of the limitations was lack of diversity in the participant sample. The participant sample was a relatively homogenous group of college students with little variation in age and educational background. Participants' educational level is above the average education of the general population and any generalization of the results should be done with caution. Perhaps one of the biggest limitations regarding the sample is the lack of experience the subjects had in the workforce. Because the sample was comprised of students with little experience in the world of work, they may have had difficulty relating to a real world applicant setting.

Another limitation involved the use of a hypothetical scenario. While the study provided separate experimental conditions, the faking condition required participants to pretend or imagine that they were completing a personality assessment as part of an employee selection process and encouraged them to put their best foot forward. The simulation may have lacked fidelity or experimental realism compared to an actual applicant setting. Although an incentive was used to increase the motivation of the participants to fake, a real applicant setting may have revealed results more consistent with the literature. In addition, the study did not include a baseline measure of cognitive load. A baseline for each subject would have allowed a comparison with participants' final cognitive load measure. This would have allowed a within-subjects comparison of

the impact the experimental condition. A baseline of cognitive load may have provided insight into the lack of differences in cognitive load between different levels of item valence.

Additionally, participants were not randomly assigned to experimental conditions. Therefore, possible order effects were not eliminated through the random assignment methodology. Random assignment was not integrated due to issues with the eye-tracking software. To ensure proper data collection, it was important to first collect data for individuals in the faking condition to consistently administer the protocols. However, post-hoc analyses were performed to ensure no significant differences between the honest and faking conditions on any of the demographic variables.

Future Research

Future research should pursue the examination of physiological and behavioral measures related to deception or faking on personality measures within an actual applicant setting. Utilizing an applicant scenario with a college student sample may not have captured the motivation and cognitive load that actual job candidates' experience. Future research should also include participants across various occupations and demographic backgrounds. Diversity of participants may provide more generalizability to the entire workforce. Future research should incorporate a baseline measurement of all physiological data derived from the eye-tracking technology to monitor the effect of the experimental manipulation. This baseline would aid in identifying the true differences in cognitive load rather than individual differences in the physiological variables of interest.

In general, future research should focus on examining the utility of eye-tracking technology and other physiological tools of deception detection. This accumulation of research could help drive changes in employee selection practices and the detection of

faking on personality measures. Further, studies should assess the extent to which the presence of physiologically based deception detection techniques decreases the likelihood of applicants engaging in faking due to the possibility of getting caught. Perhaps the presence of the technology alone may deter applicants from engaging in deception and increase the validity of the personality measure.

Research uncovering cues to deception on personality measures may help to increase the validity of the assessments which, in turn, should lead to less biased forms of employee selection and an overall increase in the performance of the workforce at large.

REFERENCES

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science, 205*, 1289-1292.

Agustin, J., Skovsgaard, H., Hansen, J. P., & Hansen, D. W. (2009). Low-cost gaze interaction: Ready to deliver the promises. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4453-4458). New York: Association for Computing Machinery.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50* , 179-211.

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49*, 1621-1630.

Allport, G. W. (1960). *Personality and social encounter.* Boston: Beacon.

Alvarez, M.C., & Risko, V.J. (1989). Schema activation, construction, and application. *ERIC Clearinghouse on Reading and Communication Skills.* Retrieved from http://www.ericdigests.org/pre-9213/schema.htm

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.

American Psychological Association. (2007). *Publication manual of the American Psychological Association* (5th ed.). Washington DC.

Baer, R. A., Rinaldo, J. C., & Berry, D. T. R. (2003). Response distortions in self-report assessment. In R. Fernandez-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (pp. 861-866). London, England: Sage Publications.

Baird, L. L. (1985). Do grades and tests predict adult accomplishment? *Research in Higher Education, 23*, 3-85.

Bagby, R. M., & Marshall, M. B. (2003). Positive impression management and its influence on the Revised NEO Personality Inventory: A comparison of analog and differential prevalence group designs. *Psychological Assessment, 2002*, 333-339.

Ballenger, J. F., Caldwell-Andrews, A., & Baer, R. A. (2001). Effects of positive impression management on the NEO Personality Inventory Revised in a clinical population. *Psychological Assessment, 13*, 254-260.

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

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personal constructs. *Journal of Applied Psychology, 81*, 265-72.

Bassi, A., Morton, A., & Williams, K. (2011). The effects of identities, incentives, and information on voting. *Journal of Politics, 73*, 558-571.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*, 3-26.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, G. Berntson, & L. G. Tassinary (Eds.). *Handbook of psychophysiology* (pp. 142-162). Cambridge: Cambridge University Press.

Ben-Porath, Y. S., & Waller, N. G. (1992). Five big issues in clinical personality assessment: A rejoinder to Costa and McCrae. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 4*, 23-25.

Berry, D. T. R., Bagby, R. M., Smerz, J., Rinaldo, J. C., Caldwell-Andrews, A., & Baer, R. A. (2001). Effectiveness of NEO-PI-R research validity scales for discriminating analog malingering and genuine psychopathology. *Journal of Personality Assessment, 76*, 496-516.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology, 74*, 478-494.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M.A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14*, 317-333.

Block, J., & Colvin, R.D. (1994). Do positive illusions foster mental health: Separating fiction from fact. *Psychological Bulletin, 116*, 21-27.

Bolino, M. C., & Turnley, W. H. (1999). Measuring impression management in organizations: A scale development based on the Jones and Pittman taxonomy. *Organizational Research Methods, 2*, 187-206.

Bonanno, G.A., Field, N.P., Kovacevic, A., & Kaltman, S. (2002). Self-enhancement as a buffer against extreme adversity: Civil war in Bosnia and traumatic loss in the United States. *Personality and Social Psychology Bulletin, 28*, 184-196.

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition, 4*, 353-376.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*, 203-242.

Burgoon, J. K., & Buller, D.B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal dynamics. *Journal of Nonverbal Behavior, 18*, 155-184.

Buss, A. H. (1980). *Self-consciousness and social anxiety.* San Francisco: Freeman.

Buss, A.H. (1989). Personality as traits. *American Psychologist, 44*, 1378-1388.

Caldwell-Andrews, A., Baer, R. A., & Berry, D. T. R. (2000). Effects of response sets on NEO PI-R scores and their relations to external criteria. *Journal of Personality Assessment, 74*, 472-488.

Callegaro, M., Yang, Y., Bhola, D. S., & Dillman, D. A. (2005). Response latency as an indicator of optimizing. A study comparing job applicants and job incumbents' response latency on a web survey. Retrieved April 23, 2011, from http://mariocallegaro.info/pdf/papers/Callegaro_Yang_Bhola_Dillman_ResponseLatencyAsAnIndicatorOfOptimizingRC33.pdf

Calvo, M. G., & Avero, P. (2002). Eye movement assessment of emotional processing in anxiety. *Emotion, 2*, 105-117.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*, 363-423.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and divergent validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Carey, J., & Paulhus, D.L. (2008). *The structure of self-presentation scales.* Presented at the meeting of the American Psychological Association, Boston, MA.

Cattell, R. B. (1950). *Personality.* New York: McGraw-Hill.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.

Christensen, N. D., Burns, G. N., & Montogmery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance, 18*, 267-307.

Christie, R., & Geis, F.L. (1970). *Studies in Machiavellianism.* New York: Academic Press.

Colvin, C. R., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin, 116*, 3-20.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68*, 1152-1162.

Cook, T. D., & Campbell, D. T. (1979*). Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.

Costa, P. T. Jr., & McCrae, R. R. (1992). *The Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J. (1955). Processes affecting scores on "understanding others" and "assumed similarity." *Psychological Bulletin, 52*, 177-193.

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

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627-668.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-118.

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

Dillman, D. A., & Redline, C. D. (2004). Testing paper self-administered questionnaires: Cognitive interview and field test comparisons. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin & E. E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 299-318). Hoboken, NJ: Wiley.

Doherty-Sneddon, G., Bruce, V., Bonner., L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental Psychology, 38*, 438-445.

Donovan, J. J., Dwight, S. A., & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance, 16*, 81-106.

Douglas, E. F., McDaniel, M.A., & Snell, A.F. (1996, August). The validity of non-cognitive measures decays when applicants fake. Paper presented at the annual meeting of the Academy of Management, Cincinnati, OH.

Dunnet, S., Koun, S., & Barber, P. (1981). Social desirability in the Eysenck Personality Inventory. *British Journal of Psychology, 72*, 19-26.

Edwards, A. L. (1970). *The measurement of personality traits by scales and inventories.* Oxford, England: Holt, Rinehart & Winston.

Ekman, P., & Friesen, W. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*, 49-98.

Ekman, P., O'Sullivan, M. & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10,* 263-266.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46,* 913-920.

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

Ellingson, J. E., Sackett, P. R., & Connelly, B. S. (2007). Personality assessment across selection and development contexts: Insights into response distortion. *Journal of Applied Psychology, 92,* 386-395.

Eysenck, M. W., & Eysenck, M. C. (1979). Processing depth, elaboration of encoding, memory store and expended processing capacity. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 472-484.

Fletcher, C., & Bailey, C. (2003) Assessing self awareness: Some issues and methods. *Journal of Managerial Psychology, 18,* 395-404.

Frank, M. G., Menasco, M. A., and O'Sullivan, M. (2008). Human behavior and deception detection. In J. G. Voeller (Ed.), Wiley Handbook of Science and Technology for Homeland Security. New York: Wiley & Sons.

Fried, R. (1980). Pupillometry: The Psychology of the Pupillary Response. *Journal of Personality Assessment, 44,* 441-444.

Furnham, A., & Henderson, M. (1982). The good, the bad, and the mad: Response bias in self-report measures. *Personality and Individual Differences, 3,* 311-320.

Gailliot, M. T., Schmeichel, B. J., & Maner, J. K. (2007). Differentiating the effects of selfcontrol and self-esteem on reactions to mortality salience. *Journal of Experimental Social Psychology, 43*, 894-901.

Gammon, A. R., Griffith, R. L, & Kung, M. (2011). *How do real applicants who are fakers compare to nonfakers?* Poster session presented at the 26th annual meeting of the Society for Industrial and Organizational Psychology, Chicago, IL.

Godfrey, D. K., Jones, E. E., & Lord, C. G. (1986). Self-promotion is not ingratiating. *Journal of Personality and Social Psychology, 50*, 106-115.

Goffin, R. D., Rothstein, M. G., & Johnston, N. G. (1996). Personality testing and the assessment center: Incremental validity for managerial selection. *Journal of Applied Psychology, 81*, 746-756.

Gough, H.G. (1950). The F minus K dissimulation index for the MMPI. *Journal of Consulting Psychology, 14*, 408-413.

Granhag, P. A., & Vrij, A. (2005). Detecting deception. In N. Brewer, & K. Williams (Eds.). *Psychology & Law: An empirical perspective* (pp. 43-92). New York: Guilford Press.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*, 457-461.

Griffith, R. L., & Peterson, M. H. (2006). *A closer examination of applicant faking behavior.* Greenwich, CT: Information Age.

Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review, 36*, 341-357.

Griffith, R. L., & Peterson, M. H. (2008). The failure of social desirability measures to capture applicant faking behavior. *Industrial and Organization Psychology: Perspectives on Science and Practice, 1*, 308-311.

Grings, W. W., & Dawson , M. E. (1978). *Emotions and bodily responses: A psycho physiological Approach.* New York: Academic Press.

Handler, C., & Hunt, S. (2006). *Bad hire calculator.* Retrieved January 2012, from http://www.adpselect-info.com/badHireCalculator.html

Hagemann, D., Waldstein, S., & Thayer, J. (2003). Central and autonomic nervous system, integration of emotions. *Brain and Cognition, 52*, 79-87.

Hathaway, S. R., & McKinley, J. C. (1951). *MMPI manual.* New York: Psychological Corporation.

Harrigan, J. A., & O'Connell, D. M. (1996). Facial movements during anxiety states. *Personality and Individual Differences, 21*, 205-212.

Hart, S. G., & Staveland, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.). *Human mental workload* (pp. 239-250). Amsterdam: North Holland Press.

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

Heckman, K. E., & M. D. Happel. (2006). Mechanical detection of deception: A short review. *Educing information interrogation: science and art-foundations for the future* (63-93). Washington, DC: National Defense Intelligence College Press.

Hess, E. H. (1972). Pupillometrics: A method of studying mental, emotional, and sensory processes. In N. S. Greenfield, & R. A. Sternbach (Eds.), *Handbook of psychophysiology* (pp. 491-531). New York: Holt, Rinehart & Winston.

Hoffman, E. (2001). Ace the corporate personality test. New York: McGraw-Hill.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *Nebraska symposium on motivation* (Vol. 30, pp. 55-89). Lincoln, NE: University of Nebraska Press.

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 91*, 9-24.

Holden, R. R., Kroner, D. G., Fekken, G. C., & Popham, S. M. (1992). A model of personality test item response dissimulation. *Journal of Personality and Social Psychology 63*, 272-279.

Holden, R. R., & Evoy, R. A. (2005). Personality inventory faking: A four-dimensional simulation of dissimulation. *Personality and Individual Differences,* 39, 1307-1318.

Holden, R. R. (2007). Socially desirable responding does moderate personality scale validity both in experimental and non-experimental contexts. *Canadian Journal of Behavioral Science, 39*, 184 201.

Holtgraves, T. M. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin. 30*, 161-172.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70*, 125-132.

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

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

Iani, C., Gopher, D., and Lavie, P. (2004) Effects of task difficulty and invested mental effort on peripheral vasoconstriction. *Psychophysiology, 41*, 789-798.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th ed.). Gaithersburg, Maryland: Aspen Publishers.

Isaacowitz, D. M. (2005). The gaze of the optimist. *Personality and Social Psychology Bulletin, 31*, 407-415.

Jackson, D. N., & Messick, S. (1962). Response styles and the measurement of psychopathology. *Psychological Bulletin, 55*, 243-252.

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology, 66*, 206-219.

Johnson, J. A., & Hogan, R. (1981). Moral judgments and self-presentations. *Journal of Research in Personality, 15*, 57-63.

Johnson, E. A. (1995). Self-deceptive responses to threat: Adaptive only in ambiguous circumstances. *Journal of Personality, 63*, 759-791.

Johnston, W. A., Griffith, D., & Wagstaff, R. H. (1972). Speed, accuracy, and ease of recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 512-520.

Johnston, W. A., & Heinz, S. P. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology, 107*, 420-435.

Johnston, W. A., & Heinz, S. P. (1979). Depth of non-target processing in an attention task. *Journal of Experimental Psychology, 5*, 168-175.

Jones, E. E., & Pittman, T. S. (1982). Toward a theory of strategic self-presentation. In J. Suls (Ed.), *Psychological perspectives on the self* (pp. 231-262). Hillsdale, NJ: Lawrence Erlbaum.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583-1585.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kelley, P. L., Jacobs, R. R., & Farr, J. L. (1994). Effects of multiple administrations of the MMPI for employee screening. *Personnel Psychology, 47*, 575-591.

Klar, Y., & Giladi, E. E. (1999). Are most people happier than their peers, or are they just happy? *Personality and Social Psychology Bulletin, 25*, 586-595.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*, 537-567.

Krueger, J. (1998). Enhancement bias in descriptions of self and others. *Personality and Social Psychology Bulletin, 24*, 505-516.

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review, 111*, 94-110.

Kwan, V. S. Y., Barrios, V., Ganis, G., Gorman, J., Lange, C., Kumar, M., Shepard, A., & Keenan, J. P. (2007). Assessing the neural correlates of self-enhancement bias: A transcranial magnetic stimulation study. *Experimental Brain Research, 182*, 379-385.

Landers, R. N., Sackett, P. R., & Tuzinski, K. A. (2011). Retesting after initial failure, coaching rumors, and warnings against faking in online personality measures for selection. *Journal of Applied Psychology, 96*, 202-210.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin, 107*, 34-47.

Leary, M. R. (1995). Self-presentation: Impression management and interpersonal behavior. Madison, WI: Brown & Benchmark.

LeDoux J. (2003). The emotional brain, fear, and the amygdala. *Cellular & Molecular Neurobiology, 23*, 727-738.

Levenson, R. W., & Ekman, P. (2002). Difficulty does not account for emotion-specific heart rate changes in the directed facial action task. *Psychophysiology, 39*, 397-405.

MacKinnon, D. W. (1944). The structure of personality. In J. McV. Hunt (Ed.), *Personality and the Behavior Disorders* (pp. 3-48). New York: Ronald Press.

Marshall, M. B., De Fruyt, F., Rolland, J. P., & Bagby, R. M. (2005). Socially desirable responding and the factorial stability of the NEO-PI-R. *Psychological Assessment, 17*, 379-384.

Martin, B. A., Bowen, C. C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality and Individual Differences, 32*, 247-256.

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

McCrae, R. R. (1989). Why I advocate the five-factor model: Joint analyses of the NEO-PI and other instruments. In D. Buss and N. Cantor (Eds.), *Personality psychology: recent trends and emerging directions* (pp. 237-245). New York: Springer.

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

McFarland, L. A., Ryan, A. M., & Ellis, A. (2002). Item placement on a personality measure: Effects on faking behavior and test measurement properties. *Journal of Personality Assessment, 78*, 348-369.

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*, 979-1016.

McDaniel, M. A., & Timm, T. (1990). *Lying takes time: Predicting deception in biodata using response latency.* In H. Wing (Chair), Alternative predictors. Symposium conducted at the 98th Annual conference of the American Psychological Association, Boston, MA.

Mersman, J. L., & Shultz, K. S. (1998). Individual differences in the ability to fake on personality measures. *Personality and Individual Differences, 24,* 217-228.

Mesmer-Magnus, J., & Viswesvaran, C. (2006). Assessing response distortion in personality tests: A review of research designs and analytic strategies: In R. L. Griffith and M. H. Peterson (Eds.). *A closer examination of applicant faking behavior.* Greenwich, CT: Information Age.

Millham, J., & Jacobson, L. I. (1978). The need for approval. In H. London & J. E. Exner (Eds.), *Dimensions of Personality* (365-390). New York: Wiley.

Morasco, B. J., Gfeller, J. D., & Elder, K. A. (2007). The utility of the NEO-PI-R validity scales to detect response distortion: A comparison with the MMPI-2. *Journal of Personality Assessment, 88,* 276-283.

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry, 12,* 177-1961.

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

Mueller-Hanson, R., Heggestad, E. D., & Thornton, G. C. (2006). Individual differences in impression management: an exploration of the psychological processes underlying faking. *Psychology Science, 48,* 288-312.

National Research Council (2003). *The polygraph and lie detection.* Washington, DC: The National Academic Press.

Nunnally, J. C., & Bernstein, I. H. (1995). *Psychometric Theory*, New York: McGraw-Hill.

Pauls, C. A., & Crost, N. W. (2005). Cognitive ability and self-reported efficacy of selfpresentation predict faking on personality measures. *Journal of Individual Differences, 26*, 194-206.

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

Paulhus, D. L. (1995). Bypassing the will: The automatization of affirmations. In D. W. Wegner & J. W. Pennebaker (Eds.) *Handbook of mental control.* (pp. 573-587). Hillsdale, NJ: Erlbaum.

Paulhus, D. L. (1998). Interpersonal and intra-psychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74*, 1197-1208.

Paulhus, D. L., & Trapnell, P. D. (2010). Self-presentation: An agency-communion framework. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality psychology* (pp. 492-517). New York: Guilford.

Paulhus, D. L., & Martin, C. L. (1988). Functional flexibility: A new conception of interpersonal flexibility. *Journal of Personality and Social Psychology, 55*, 88-101.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556-563.

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* (17-60). San Diego: Academic Press.

Paulhus, D. L., Graf, P., & Van Selst, M. (1989). Attentional load increases the positivity of self-presentation. *Social Cognition, 7*, 389-400.

Pennington, D. (2003). *Essential Personality*. Clearwater, FL: Arnold Publishing.

Peterson, M. H., Griffith, R. L., Converse, P. D., & Gammon, A. R. (2011). Using within-subjects designs to detect applicant faking. In P. D. Converse (Chair), *Detecting deception: Techniques for assessing applicant faking on personality measures*. Symposium conducted at the 26th Annual Conference for the Society for Industrial and Organizational Psychology, Chicago, IL.

Peterson, M. H., & Griffith, R. L. (2006). Faking and job performance: A multifaceted issue. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 233-261). Greenwich, CT: Information Age Publishing.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research, *Psychological Bulletin, 124*, 372-422.

Rayner, K., Miller, B., & Rotello, C. M. (2008). Eye movements when looking at print advertisements: The goal of the viewer matters. *Applied Cognitive Psychology, 22*, 697-707.

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

Rice, B. (1978). The new truth machines. *Psychology Today, 12*(1), 61-78.

Robins, R. W., & John, O. P. (1997). Effects of visual perspective and narcissism on self-perception: Is seeing believing? *Psychological Science, 8*, 37-42.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychology, 80*, 340-352.

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

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

Saucier, G. (1994). Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology, 66*, 141-154.

Schinka, J., Kinder, B., & Kremer, T. (1997). Research validity scales for the NEO-PI-R: Development and initial validation. *Journal of Personality Assessment, 68*, 127-138.

Schlenker, B. R., & Weigold, M. F. (1990). Self-consciousness and self-presentation: Being autonomous versus appearing autonomous. *Journal of Personality and Social Psychology, 59*, 820-828.

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

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

Sherry, S. B., Hewitt, P. L., Flett, G. L., Lee-Baggley, D. L., & Hall, P. A. (2007). Trait perfectionism and perfectionist self-presentation in personality pathology. *Personality and Individual Differences, 42*, 477-490.

Shulman, H. G. & Fisher, R. P. (1972). Expected value as a determinant of the distribution of attention. *Journal of Experimental Psychology, 93*, 343-348.

Smith, V. L., & Walker, J. (1993), Monetary rewards and decision cost in experimental economics, *Economic Inquiry 31*, 245-261.

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

Snell, A. F., Sydell, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review, 9*, 219-242.

Snyder, M. (1974). Self - monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*, 526-537.

Swann, W. B. J. (1990). To be adored or to be known? The interplay of self-enhancement and self-verification. In R. M. Sorrentino, E. T. Higgins (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior*, (pp. 408-448). New York: Guilford Press.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Taylor, S. E., & Armor, D. A. (1996). Positive illusions and coping with adversity. *Journal of Personality, 64*, 873-898.

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.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive-processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*, 299-314.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

Van Hooft, Edwin A. J., & Born, M. P. (2011). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology, 97*, 301-316.

Vasilopoulos, N. L., Cucina, J. M., Dyomina, N. V., Morewitz, C. L., & Reilly, R. R. (2006). Forced-choice personality tests: A measure of personality and cognitive ability? *Human Performance, 19*, 175-199.

Vendemia, J. M., Buzan, R. F., Green, E. P., & Schillaci, M. J., (2006). Effects of preparedness to deceive on ERP waveforms in a two-stimulus paradigm, *Journal of Neurotherapy, 9*, 45-70.

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

Voßkühler, A., (2009). *OGAMA description (for Version 2.5): A software to record, analyze and visualize gaze and mouse movements in screen based environments.* Retrieved from http://www.ogama.net/sites/default/files/pdf/OGAMA-DescriptionV25.pdf

Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology, 88*, 632-657.

Vrij, A. (1999). Interviewing to detect deception. In A. Memon and R. Bull (Eds.), *Handbook of the psychology of interviewing* (pp. 317-326). Chichester: Wiley & Sons.

Vrij, A. (2000). Detecting lies and deceit: The psychology of lying and its Implications for professional practice. Chichester: Wiley & Sons.

Vrij, A. & Mann, S. (2001) Telling and detecting lies in a high stake situation: the case of a convicted murderer. *Applied Cognitive Psychology. 15*, 187-203.

Vrij, A., Edward, K., & Bull, R. (2001). Police officers' ability to detect deceit: The benefit of indirect deception detection measures. *Legal and Criminological Psychology, 6*, 185-196.

Vrij, A., Mann, S., Kristen, S., & Fisher, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*, 499-518.

Vrij, A., Fisher, R. P., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling, 5*, 39-43.

Walczyk, J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., Zha, P. (2005). Lying person to person about life events: A cognitive framework for lie detection. *Personnel Psychology, 58*, 141-170.

Wallbott, H. G. & Scherer, K. R. (1991). Stress specifics: Differential effects of coping style, gender, and type of stressor on automatic arousal, facial expression, and subjective feeling. *Journal of Personality and Social Psychology, 61*, 147-156.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*, 448-457.

Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., & Fisher, R. (2011). Thermal imaging as a lie detection tool at airports. *Law and Human Behavior, 35*, 40-48.

Weiner, J. A., & Gibson, W. M. (2000*). Practical effects of faking on job applicant attitude test scores.* Paper presented at 15th Annual Conference of the Society for Industrial and Organizational Psychology, New Orleans.

Wexley, K. N., Alexander, R., Greenwalt, J. & Couch, M. (1980). Attitudinal congruence and similarity as related to interpersonal evaluation in manager-subordinate dyads. *Academy of Management Journal, 23*, 320-330.

Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology, 23*, 419-427.

Wiggins, J. S. (1964). Convergence among stylistic response measures from objective personality tests. *Educational and Psychological Measurement, 24*, 551-562.

Wilhelm, B., Giedke H., Ludtke H., Bittner E., Hofmann A., & Wilhelm H., (2001). Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. *Journal of Sleep Research. 10*, 1-7.

Wright, P. & Kahneman, D. (1971). Evidence of alternative strategies of sentence retention. *Quarterly Journal of Experimental Psychology 23*, 197-213.

Young, M. S. & Schinka, J. (2001). Research Validity scales for the NEO PI-R: Additional evidence for Reliability and validity. *Journal of Personality Assessment, 76*, 412-420.

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

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

Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods, 7*, 168-190.

Ziegler, M., (2011). Applicant faking: A look into the black box. *The Industrial Organizational Psychologist, 49*, 29-36.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology,* (pp. 1-59). New York: Academic Press.

APPENDIX A

HUMAN SUBJECTS CONSENT FORM

HUMAN SUBJECTS CONSENT FORM

The following is a brief summary of the project in which you are asked to participate. Please read this information before signing the statement below.

TITLE OF PROJECT: "The Use of Eye-tracking Technology to Detect Faking on Personality Measures in a Simulated Applicant Setting"

PURPOSE OF STUDY/PROJECT: This project aims at using eye-tracking technology, in conjunction with the Marlowe-Crowne Social Desirability Scales (MCSDS), to detect faking on the IPIP-NEO Conscientiousness and Openness personality scales within a simulated applicant setting. The project will assess the relationship between the eye-tracking technology and the Marlowe-Crowne Scale to evaluate construct validity evidence. The use of the NASA-TLX cognitive load questionnaire will be used to assess level of cognitive load between the faking and non-faking conditions. In addition, various simulated job types will be used among the simulated applicants to assess differences in faking, cognitive load, and eye movements (saccades and fixations).

PROCEDURE: After volunteering for the study, you will be assigned to one of three experimental job application scenarios. You will be informed of the purpose of the study and will be directed to fill out an informed consent form. After filling out the informed consent form, you will be directed to the eye-tracking laboratory where you will take the personality measures for this study while attached to the eye-tracking technology. An alternative extra credit activity will be offered by your instructor should you choose not to participate in this study. Your participation in this project is voluntary and you may discontinue participation at any time without penalty. Completing the questionnaires is expected to require approximately 50-60 minutes. If you meet certain criteria, you will be eligible to enter a raffle for a $20 gift card.

INSTRUMENTS: The NEO IPIP short form Consciousness and Openness Personality Scales, the Marlowe-Crowne Social Desirability Scales (MCSDS), the Positive Presentation Management (PPM) Scale, a demographic questionnaire requesting the participant's sex, age, major, and GPA, the NASA-TLX cognitive load measure, and the eye-tracking technology.

RISKS/ALTERNATIVE TREATMENTS: There is little risk in this research, but it is possible that some questions might be disturbing to participants. If a question is disturbing to participants, Dr. Young, at 257-2449, is available to process the concern, as is the Louisiana Tech Counseling Center.

BENEFITS/COMPENSATION: Students will be given credit for participation and are offered alternative ways of acquiring credit that require a similar investment of time and energy. This work should lead to more balanced and accurate diagnostic procedures for personality issues and strengths in normal and clinical populations. A $20 gift-card will be used as an incentive to increase the motivation for participants to fake well when in the faking condition. In addition, they will learn the nature of psychological experimentation u sing high tech equipment (the infrared eye-tracking system).

The following disclosure applies to all participants using online survey tools: This server may collect information and your IP address indirectly and automatically via "cookies".

I attest by entering my personal information below that I have read and understood the following description of the study, "The Use of Eye-tracking Technology to Detect Faking on Personality Measures in a Simulated Applicant Setting" and its purposes and methods. I understand that my participation in this research is strictly voluntary and my participation or refusal to participate in this study will not affect my relationship with Louisiana Tech University or my grades in any way. Further, I understand that I may withdraw at any time or refuse to answer any questions without penalty. Upon completion of the study, I understand that the results will be freely available to me upon request. I understand that the results of my survey will be confidential, accessible only to the principal investigators, myself, or a legally appointed representative. I have not been requested to waive nor do I waive any of my rights related to participating in this study.

CONTACT INFORMATION: The principal experimenters listed below may be reached to answer questions about the research, subjects' rights, or related matters.

PROJECT DIRECTOR: Dr. Tony Young, Dr. Jeff Walczyk, Luke Simmering
EMAIL: las041@latech.edu, tyoung@latech.edu, walczyk@latech.edu
PHONE: 318-257-3413, 318-257-2449

Members of the Human Use Committee of Louisiana Tech University may also be contacted if a problem cannot be discussed with the experimenters:
Dr. Les Guice (257-3056)
Dr. Mary M. Livingston (257-2292 or 257-4315)

Tech email address:
Psych instructor's name or class meeting time:

APPENDIX B

DEMOGRAPHIC QUESTIONNAIRE

Demographic Questionnaire:

What is your gender? _____Male _____Female
How old are you (years)? _____
What is your student classification?
____ a. Freshmen
____ a. Sophomore
____ a. Junior
____ a. Senior
What is your college major? _____
What is your current Grade Point Average (best estimate)? _____
Please list your two most recent jobs along with a brief description of the tasks you performed.
Job 1 Title: _____
Job 1 Responsibilities: _____
Job 2 Title: _____
Job Responsibilities: _____
Have you taken a personality assessment as part of the screening process for a job you applied for (please circle)? (Yes/No)

APPENDIX C

SOCIALLY DESIRABLE SURVEY ITEM POOL

IPIP Social Valence Ratings

This survey is aimed at exploring the social valence of items on various personality scales. Specifically, the level to which the item portrays behaviors or attitudes that would be deemed important. That is, which items describe overt behaviors or attitudes that can easily be modified to portray a more positive personality profile and put your best foot forward as part of a job application.

Personality Items

The following items below are all from the International Personality Item Pool. Please rate the social valence of each item. Keep in mind the 'fake ability' of the item. That is, which items portray overt behaviors or attitudes that would be deemed important for an individual applying for a job. For instance, the item 'I am always on time' would have a very high social valence because it is overtly positive. Compared to the item 'I enjoy it when it rains' having a low social valence because no clear linkage can be made between answering the item a specific way and perceptions of positive or negative behaviors or attitudes.

Please rate the social valence or 'fake ability' of the following items. The answer options of the statements below are typically 'please rate your level of agreement'. Use this to help interpret the meaning of the item but please rate the valence of the item. *

1. Am always prepared.
2. Pay attention to details.
3. Get chores done right away.
4. Carry out my plans.
5. Make plans and stick to them.
6. Waste my time.
7. Find it difficult to get down to work.
8. Do just enough work to get by.
9. Don't see things through.
10. Shirk my duties.
11. Have a good word for everyone.
12. Believe that others have good intentions.
13. Respect others.
14. Accept people as they are.
15. Make people feel at ease.
16. Have a sharp tongue.
17. Cut others to pieces.
18. Suspect hidden motives in others.
19. Get back at others.
20. Insult people.
40. Don't talk a lot.
41. Believe in the importance of art
42. Have a vivid imagination.
43. Tend to vote for liberal political candidates.
44. Carry the conversation to a higher level.
45. Enjoy hearing new ideas.
46. Am not interested in abstract ideas.
47. Do not like art.
48. Avoid philosophical discussions.
49. Do not enjoy going to art museums.
50. Tend to vote for conservative political candidates.
51. I sometimes feel resentful when I don't get my way.
52. On a few occasions, I have given up doing something because I thought too little of my ability
53. There have been times when I felt like rebelling against people in authority even though I knew they were right
54. No matter who I'm talking to, I'm always a good listener.
55. I can remember "playing sick" to get out

21. Often feel blue.
22. Dislike myself.
23. Am often down in the dumps.
24. Have frequent mood swings.
25. Panic easily.
26. Rarely get irritated.
27. Seldom feel blue.
28. Feel comfortable with myself.
29. Am not easily bothered by things.
30. Am very pleased with myself.
31. Feel comfortable around people
32. Make friends easily.
33. Am skilled in handling social situations.
34. Am the life of the party.
35. Know how to captivate people.
36. Have little to say.
37. Keep in the background.
38. Would describe my experiences as somewhat dull.
39. Don't like to draw attention to myself.

of something.
56. There have been occasions when I took advantage of someone.
57. I'm always willing to admit it when I make a mistake.
58. I sometimes try to get even rather than forgive and forget.
59. When I don't know something I don't at all mind admitting it.
60. I am sometimes irritated by people who ask favors of me.
61. I have never deliberately said something that hurt someone's feelings.
62. Am trusted to keep secrets.
63. Keep my promises.
64. Believe that honesty is the basis for trust.
65. Can be trusted to keep my promises.
66. Am true to my own values.
67. Lie to get myself out of trouble.
68. Am hard to understand.
69. Feel like an imposter.
70. Like to exaggerate my troubles.

APPENDIX D

SOCIALLY DESIRABLE SURVEY ITEMS

Item	Valence
I shirk my duties	High
I pay attention to details.	High
I waste my time	High
I can remember playing sick to get out of something	High
I'm always willing to admit it when I make a mistake	High
I believe in the importance of art.	Low
I tend to vote for liberal political candidates	Low
I believe others have good intentions	Low
I do not like art.	Low
I rarely get irritated	Low
I make people feel at ease	Moderate
I make friends easily	Moderate
I have a sharp tongue	Moderate

APPENDIX E

NASA TASK LOAD INDEX

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

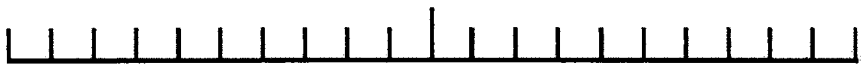
Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?



Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

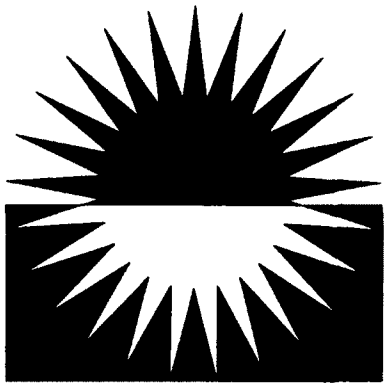
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low — Very High

APPENDIX F

JOB DESCRIPTION



Position: Analytics Supervisor

Position Overview

The Analytic Supervisor is in charge of managing the day to day details of the analytics staff. Specifically, the Analytics Supervisor spends abundant time organizing plans, schedules, while monitoring possible operational mistakes.

Requirements

A proficient detail orientation
Managing an extensive daily task list
Problem solving ability
Working well with others
Assuring project deadlines

Other Skills/Abilities

Previous experience managing others
At least a High School Diploma or Equivalent

NOTE: This job description is not intended to be all-inclusive. Employee may perform other related duties as negotiated to meet the ongoing needs of the organization.