5-2023

THE EFFECT OF RESPONSE FORMAT ON FAKING IN PERSONALITY MEASUREMENTS USED FOR PERSONNEL SELECTION

Gilberto Sanchez

THE EFFECT OF RESPONSE FORMAT ON FAKING IN PERSONALITY MEASUREMENTS USED FOR PERSONNEL SELECTION

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Industrial and Organizational Psychology

by

Gilberto Sanchez

May 2023

THE EFFECT OF RESPONSE FORMAT ON FAKING IN PERSONALITY MEASUREMENTS USED FOR PERSONNEL SELECTION

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

by

Gilberto Sanchez

May 2023

Approved by:

Kenneth Shultz, Committee Chair, Psychology

Ismael Diaz, Committee Member

Janet Kottke, Committee Member

ABSTRACT

The primary purpose of this study was to explore intentions of faking during personality measurements and determine the best option to address intentions of faking during high stakes situations when comparing Likert-Scale items with Multidimensional Forced Choice (MFC) measures. Participants (N=618) participated in the study which consisted of answering items on impression management and self-deceptive enhancement. Afterward, participants were placed randomly in an “answer honestly group” or “answer as if you are applying to your dream job” to fill out a Big 5 personality measurement. Findings from the study indicated that the personality traits of conscientiousness and neuroticism were positively correlated with intentions of faking. When comparing which personality trait was exaggerated between both the focus and control groups, the findings indicated that conscientiousness and agreeableness were scored higher among the “answer as if you are applying to your dream job” group. There were differences between MFC measures and Likert-scale items based on Cohen’s d, but there wasn’t substantial differences. Overall, the findings from this study indicated that their personality traits such as conscientiousness and neuroticism are associated with intentions of faking and that individuals may exaggerate in certain personality traits when placed on high stake situations.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank professor Frieze, a faculty member from CSUDH who motivated me to go to grad school. I appreciate your mentorship and guidance. Without you, I wouldn't be here. I also want to thank my thesis advisor, Dr. Shultz. Thank you for your support and guidance through this process and for the words of motivation you gave me which made me realize that is goal was attainable. Special shout out to my cohort and coworkers from the City of LA. I had a pleasure meeting you all. Thank you for the memories and motivation.

Lastly, I want to thank my mom and dad who taught me the value of hard work and discipline at such a young age. Thank you for believing in me and always supporting my dreams. Also, my brother who served as my #1 support system during this process.

Thank you all!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

LITERATURE REVIEW

Introduction

Personality measurements are used by organizations to identify what type of personality an applicant has to predict an applicant's likelihood of success on the job. The primary reason being organizations face a high cost of making inaccurate hiring decisions (Henle et al., 2005). Therefore, organizations typically identify ways to improve their screening process to have more useful information about applicants (Stabile, 2002). It has been shown that organizations that implement personality measures in their selection process for promotional positions typically outperform their competitors (Gatewood et al., 2019). A study conducted demonstrated that organizations that use personality-based human capital resources such as personality measures lead to higher financial performance which enhanced labor productivity (Oh et al., 2015).

However, one of the issues with the use of personality measurements for high-stakes personnel selection testing is “faking”, also known as “faking good”. Faking is when applicants try to present themselves more positively than they truly are to the organization. This can be due to conscious impression management or subconscious socially desirable responding. Even though studies have indicated that personality traits can predict job performance (Barrick & Mount, 1991; Fang et al., 2015; Salgado, 1997) some studies on personality

measurements have also shown that there are validity issues (Murphy & Dzieweczynski, 2005). Therefore, until the last few decades, it was assumed that personality measures should not be used for making employment decisions (Guion & Gottier, 1965). However, more recent studies have determined that personality measurements can have validity for making accurate personnel selection decisions if certain procedures are met (Gatewood et al, 2019; Tett & Jackson, 1991). Specifically, procedures such as using job analysis and using trait-performance linkages to test for these personality traits can improve predictive validity estimates (Tett & Christiansen, 2007).

No matter the stance researchers may have on validity, one of the primary issues of concern is faking, due to applicants' self-reporting. Test takers can distort their answers or respond carelessly threatening the validity of personality measurements (Arthur et al., 2021). Researchers who study faking give several reasons why applicants fake their answers and discuss the consequences and solutions to face this issue. Therefore, the purpose of this study is to explore intentions of faking during personality measurements and determine the best option to address intentions of faking during high stake situations when comparing Likert-Scale items with Multidimensional Forced Choice (MFC) measures. will be to explore intentions to fake during personality measurements used in high stakes situations and determine the best options to address this issue when comparing Likert-Scale items and MFC measures.

History of Personality use in Personnel Selection

There is a rich history of the use of personality measures for personnel selection which goes back to World War 1 in the early 20th century (Gibby & Zickar, 2008). For example, Woodworth's Personal Data Sheet was used during World War 1 to recognize which soldiers would be vulnerable to a nervous breakdown during an enemy attack and to analyze soldiers' emotional ability to determine who would be deployed (Gibby & Zickar, 2008). Since then, many personality measures have been created for use in personnel selection, while existing measures have been updated and improved. One of the most well-known personality measurements is the Myers-Briggs Type Indicator. The measurement is an assessment of personality from Carl Jung's theory (Pittenger, 1993). Jung's theory proposes that an individual uses four psychological functions that are always present during any circumstance. They are sensation, intuition, feeling, and thinking (Jung, 1923). Myers-Briggs personality measurement has been used by large organizations such as Exxon and AT&T to improve decision making and build effective teamwork development programs (Murray, 1990). Myers-Briggs Type indicator has contributed to academia by helping students choose a major based on their personality and their style of learning (Mill, 1984). Uses in the business world include using the MBTI to determine how an applicant handles making decisions due to individuals having different work styles and determining which personality traits lead to faster promotions (Furnham & Cramp, 2015; Moore, 1987). However, the MBTI and

other clinically focused personality measures, such as the Minnesota Multiphasic Personality Inventory (MMPI) and the Rorschach are also famously known to assess the personality of the test-taker (Hathaway & McKinley, 1943), but have been criticized when used for personnel selection purposes (Kottke et al, 2010). Criticism of these measurements have led to the increased use of the Big Five model and the HEXACO model for personnel selections purposes.

In the late 1980s and early 1990s organizations started using personality measurements again due to research demonstrating job performance can be predicted by personality (Gibby & Zickar, 2008). This was due in large part to the Big Five personality traits which include Openness to Experience, Conscientiousness, Extraversion, Neuroticism, and Agreeableness. For example, Barrick and Mount (1991) conducted a study to investigate the relationship between the Big Five personality traits and job performance criteria (e.g., job proficiency, training proficiency, and personnel data) among five occupations. The results from their meta-analytic study indicated that conscientiousness showed a consistent relationship with job performance for all groups. In addition, Extraversion was seen as a good predictor for occupations that have to do with interacting with others and sales. Openness to experience and extraversion were good predictors for training proficiency. The results from their study indicated that personality was a strong predictor of job performance (Barrick & Mount, 1991). Salgado (1997) conducted a similar study using European data which led to results indicating that conscientiousness and emotional stability are strong

predictors of job performance confirming the importance of personality for personnel selection in many jobs (Salgado, 1997).

More recently, a meta-analysis conducted by Ones et al. (2007) demonstrated the relationship of the Big 5 personality traits with performance criteria, leadership criteria, team performance, and work motivation. Findings from Ones et al. (2007) meta-analysis indicated that personality variables demonstrate validity. Therefore, support for the use of personality assessments in organizational settings is encouraged. As a result, the Big Five personality research conducted in the last several decades has contributed to the creation of more work specific personality measurements by having organizations focus on certain personality traits that will lead to success on the job. These traits include conscientiousness and agreeableness (Sackett & Walmsley, 2014). Research has also shifted in looking for personality traits that can lead to deviance toward organizations and personality traits associated with bullying (Neilsen & Knardahl, 2015; O'Neill et al., 2011).

Research on the Big 5 has determined acceptable levels of reliability and validity when using a Big 5 measurement for personal selection. For example, Tett and Jackson (1991) conducted a meta-analysis to assess the validity of personality measures as a predictor of job performance. Based on their meta-analytic findings, there was a .24 relationship between personality and job performance. The researchers determined this to be an acceptable estimate of personality validity by not choosing samples that consist of sampling bias and

included studies that reported all findings. Tett and Jackson determined this to be sufficient evidence to advocate for the use of personality measurements for personnel selection. A notable finding that came out of their meta-analysis was that mean validities from confirmatory studies (.29) were greater than exploratory studies (.12), which has a significant effect on personality measurements' ability to predict job performance.

Using a set of personality traits rather than using many predictors is recommended along with using confirmatory strategies as it leads to higher validity estimates in personality measurements predicting job related outcomes. Therefore, Tett and Jackson (1991) recommend the use of personality measurements during personnel selection disagreeing with Guion and Gottier (1965) having a critical impact on the future of personality research. Since then, there have been many articles published supporting the use of personality measurements in personnel selection (e.g., Rothstein & Goffin, 2006). It has been shown that employee characteristics associated with personality affects job performance (Barrick & Mount, 2005), and that there is a curvilinear relationship between conscientiousness and emotional stability, and job performance (Le et al., 2011). One concern that Tett and Jackson point out with personality measurement is "faking", which researchers determined is a threat to the validity of personality measurements. Morgeson et al. (2007) state that the only thing stopping an applicant from faking is their own honesty. This is an issue that

continues to affect the validity of personality measurements when used in high stakes contexts such as personnel selection, which needs to be addressed.

Why People Fake on Personality Measures

To understand the issue of faking, research has been done on job applicants to better understand why applicants' fake answers during personality measurements used in high stakes personnel selection. Applicants fake their scores during personality measurements due largely to impression management and socially desirable responding. Hogan et al. (2007) defines impression management as the process of controlling one's behavior during any form of social interaction which includes responding to inventory items (Hogan et al., 2007). People typically try to present themselves in a positive way to give a good impression while acting in ways that a person doesn't typically act (Schlenker & Weigold, 1992). Impression management theory states that during social interaction, individuals try to maximize acceptance and status, while minimizing rejection and the loss of status (Hogan, 2006). This is a concern because if an applicant is distorting their answers, then the measurement may be failing to assess future job performance (Ingold et al., 2015).

Socially desirable responding is defined as the likelihood of responding in a manner to make themselves look good (Paulhus, 1991). For example, an applicant who doesn't like to work in a team setting may be pressured to respond that they do like to work in a team setting in order to be hired. This is due to teamwork being seen as a positive trait within the workplace. Paulhus (1986)

distinguished impression management from self-deception in that impression management is referred “to the conscious dissimulation of test responses designed to create a favorable impression in some audience” whereas, self-deception is referred “to any positively biased response that the respondents believe to be true” (Paulhus, 1986 pg. 144).

In some cases, an applicant may be unaware of faking during a personality test due to the action occurring at an unconscious level. For that reason, socially desirable responding is an issue that may affect the validity of personality measurements along with the true score of the personality variable being measured (Backstrom et al., 2009; Salgado, 2016). Concern about self-reporting is that applicants are able to present themselves in a positive way that doesn’t truly represent their underlying personality (Robie et al., 2006). Muller-Hansen et al. (2006) mentions that test-takers typically answer truthfully in a questionnaire in a neutral setting such as a research setting. But when test takers are put in a situation where their test scores have valued consequences, such as applying for a job, then test takers are more likely to provide responses that help them present themselves better than they truly are.

Models on faking have been proposed by researchers to understand the psychological process that leads to faking behavior. For example, Snell et al. (1999) proposed a model that categorized individual differences into two categories on why faking occurred. The first category was the “ability to fake” and the second category was the “willingness to do so”. Characteristics that

consisted of the “ability to fake” included general cognitive ability, emotional intelligence, and the ability to understand what is being measured in the personality measurement. Characteristics of “willingness to do so” consisted of age, gender, integrity, and contextual factors (Snell et al., 1999).

McFarland and Ryan (2000) proposed a second model which explains that faking occurs from an applicant’s intent of faking. Muller-Hanson et al. (2006) proposed a combination of both models along with Ajzen’s “Theory of Planned Behavior”. The model that Muller-Hanson et al. proposed consisted of using antecedents which include perceptions of situations, willingness to fake, ability to fake, conscientiousness, and emotional stability with each individually leading to intentions to fake causing faking behavior. From the study conducted, part of the model was supported. Willingness to fake and ability to fake were not related to intentions to fake. Muller-Hanson et al. determined that there are individual differences on why applicants fake on personality characteristics and in given situations (Muller-Hanson et al., 2006). This can be due to the applicant’s personality and self-interest motivation (McKay et al, 2018).

Ellingson and McFarland (2011) found similar findings on why applicants fake. Results from their study indicated that faking occurs due to different combinations which include valence, expectancy, instrumentality (VIE) and ability to fake (Ellingson & McFarland, 2011). VIE theory has been used to predict why applicants might fake when taking a personality measurement. Vroom (1964) created VIE theory which states that faking tends to be a behavior that occurs

due to an extrinsic reward, which in this case is obtaining the job offer. Dunlop et al. (2022) relate VIE theory on the process that occurs for applicants to fake. First, the applicant must believe that the job offer is worth more than any alternative. Second, the applicant must provide desirable responses to make the situation happen. Lastly, the respondent must believe that they will be successful (Dunlop et al., 2022). Muller-Hanson et al (2006) proposed that due to there being individual differences in whether one will partake in faking, it is recommended that applicants should be informed of the importance of not partaking in faking before starting a personality test that will be used in high stakes situations such as personnel selection..

Studies that Demonstrate Applicants' Fake

Studies have been conducted to determine if applicants actually partake in faking while taking personality measurements. For example, a study was conducted by Mersman and Shultz (1997) determined if faking is caused by individual differences. It was demonstrated that participants in their study significantly increased their scores during the second administration of the measure which consisted of faking compared to the first time which was them answering honestly. The results of their study determined that participants still managed to fake and not fake their scores regardless of their social desirability, self-monitoring, and impression management level. Therefore, determining that faking occurs due to individual differences.

Applicants are more willing to fake in a personality measure if the stakes are higher. Blatant Extreme Responding (BER) is known to be when applicants respond to measures by choosing extreme responses such as 1s and 5s. In a study conducted to investigate Blatant Extreme Responding, it was determined that applicants who were applying for managerial positions were more likely to score higher on BER than applicants who were not applying for these managerial positions (Levashina et al., 2014). More recently, a meta-analysis conducted by Hu and Connely (2021) compared applicants' responses on personality measures between high stakes contexts and low stakes context using a within-subjects design. The results from their study indicated that applicants had higher means, reduced variability and stronger rank order in high stakes settings compared to low stakes settings. The findings were consistent with studies that determined that individuals are more willing to exaggerate their answers in high stake situations (Hu & Connely, 2021). Theories of faking predict that in order to fake, one must believe that faking is necessary to obtain a new job (Keifer & Benit, 2016). In this case, high stakes settings usually lead to higher salary and benefits which can encourage applicants to fake in order to obtain the job.

Other prominent studies investigating faking have instructed participants to fake their responses in order to investigate differences between applicants who fake and applicants who answer truthfully. In studies conducted within a laboratory setting, participants are instructed to “fake good” their responses in order to appear more socially desirable compared to applicants who are

responding honestly (Viswesvaran & Ones, 1999). When looking at the means between these two groups, applicants who were instructed to fake good answers had higher means than those who were instructed to answer honestly. Results from other studies have determined that participants tailor their answers to what they believe an organization is looking for in an applicant. But even so, applicants understand that responses should appear credible and not too good to be true (Goffin & Boyd, 2009).

In a study conducted by Furnham (1990) participants were asked to determine what type of personality an ideal candidate (faking) would have for three different positions (advertising executive, banker, and librarian). The participants were then asked again to choose honestly. The results from their study indicated that personality measurements are susceptible to faking (Furnham, 1990). Studies that instruct applicants to fake their scores typically have more social desirability mean scores (Viswesvaran & Ones, 1999) and a reduction in variability (Hooper, 2007). This causes applicants who fake to stand out more and get hired when personality measurements are used as the deciding factor to hire someone. That's why it's important to implement different selection methods to detect the best candidates.

MacCann (2013) conducted a study to see if faking also occurred when using the HEXACO as a personality measurement. The argument made was that most of the research conducted on whether participants fake used the Big Five model. Results from the study indicated that participants who take the HEXACO

measurements can fake their answers similar to taking a Big Five measurements (MacCann, 2013). MacCann suggested that faking isn't necessarily a problem of the Big Five nor the HEXACO model, it's simply an issue with self report personality measurements more broadly. As Grieve et al. (2011) study indicated, there is no difference in faking when a personality measurement is conducted online versus paper and pen.

Consequences of Faking

One of the main issues that faking portends in personality measurements is its effects on the validity of personality measures. The concern is that it affects the validity of the measure to the point that the measure won't be able to predict future job performance, as well as affecting hiring decisions. Validity is defined as the degree to which a test measures what it is supposed to measure. In a study conducted by Converse et al. (2009) it was found that faking in a single predictor selection setting has a negative effect on validity big enough to cause concern. Tett and Simonet (2021) wrote an article arguing that faking threatens "trait based-inferences" drawn from test scores and argues against the side that believes faking is good because it contributes to predictor-criterion correlations. Research on faking has shown that faking causes damage to the construct validity by causing an unwanted source of test score variance (Tett & Simonet, 2021).

Construct validity is important in personality measurements because it provides valid interpretations of personality test scores meaning that the

instrument is measuring what it's supposed to be measuring (Ellingson et al., 2001). Researchers have mentioned faking does occur among job applicants and that it affects the construct and criterion-related validity (Barrick & Mount, 1996; Ellingson, et al., 2001) For example, Lee et al. (2019) conducted a study to see the effects of faking on construct validity using the Big 5 personality measure through a Monte Carlo simulation. The goal was to investigate the prevalence of fakers, percentage of faked items, and the magnitude of faking effect. Results from the study indicated that construct validity was negatively affected by large amounts of faking. However, when small amounts of faking occurred there wasn't a substantial impact (Lee et al., 2019). Tonkovic et al. (2012) were able to find similar findings in their study by indicating that faking reduces construct validity when respondents fake in greater numbers (Tonkovic et al., 2012).

Some studies have indicated that faking helps criterion validity and others have indicated that it damages it. Studies from scores in lab settings that simulated job applications showed lower criterion validity compared to those in low-stakes assessments (Bing et al., 2011). Donovan et al. (2014) found similar results when a group of participants scored relatively high during the selection process but scored lower during the training assessment indicating faking had occurred and that the criterion-validity was lowered.

Alternatively, studies that demonstrate that faking helps criterion-validity were from Buehl et al. (2019) where the results indicated that participants had improved their interview responses from when they took the measures in a low

stake setting. The interview ratings were a stronger predictor for academic performance as a result of the faking that occurred by the test takers. Huber et al. (2021) found similar findings where participants were instructed to fake during personality measurements which led to stronger correlations. Hogan et al. (2007) conducted a study where applicants completed a 5-factor model personality measure when they first applied to a job. After being rejected they reapplied and completed the 5-factor model personality measure. Hogan et al. wanted to see if the applicants were motivated to score better in their personality scores during the second administration. The results from the studies demonstrated that 5.2% or fewer, improved their scores on the second try. However, construct validity remained the same, as did the acceptable fit to the scale score.

Other consequences associated with faking include a rank order change among applicants (Hartman & Grubb, 2011). Hartman and Grubb conducted a study to investigate if faking in personality measurements has an effect in the rank orders. Results from the study suggested that faking did change the rank order among those who fake. Personality traits that were faked the most were conscientiousness and emotional stability (Hartman & Grubb, 2011). A change in rank order was also shown in a study conducted by Griffith and Yoshita (2007) which can have an impact in hiring decisions. This is a concern because applicants who partake in faking may increase their chances of being hired by having their application among the top, thus defeating the purpose of what

personality measurements are supposed to do, which is to accurately measure the personality of test takers (Fluckinger et al., 2008).

Proposed Solutions to Reduce Faking

Methods have been designed to help prevent and correct faking among test takers. One of the more well-known preventions is Multidimensional Forced Choice Measures (MFC). MFC is a forced-choice format, where two or more items are shown to the test-taker on the computer screen. Similar measurements were first used by Edwards (1954) and Kuder (1960). An example of an MFC measure is provided in Figure 1.

	Most like me	Least like me
I am relaxed most of the time	X	
I start conversations		
I catch on to things quickly		X

Figure 1 – An Example of a MFC Personality Item

One of the benefits of using MFC is that it eliminates extreme answer responding and response biases (Wetzel et al., 2020). However, research has been conducted to analyze if MFC is able to prevent faking. Comparisons between Likert-scale items and MFC measures have been the main focus of research. This is due to Likert-scale items increasing the chance of halo, central

tendency, and faking (Wetzel & Greiff, 2018). One of the benefits in the use of MFC in measurements is that it organizes different dimensions statements into blocks forcing the test-taker to have to choose a statement (Lee et al., 2019). This makes it impossible for respondents to choose a statement that represents all possible answers which Brown and Maydeau-Olivares (2011) suggest, theoretically reduces faking. Most research on MFC measures has shown that it helps reduce faking, score inflation, and maintains validity (Heggestad et al., 2006; O'Neill et al., 2017).

Early research on MFC indicated that MFC was able to help combat faking more than Likert-scale items due to mean differences between honest and faking scores (Fisher et al., 2019). Lee and Joo (2021) wanted to investigate MFC measures and how they help combat faking through a new methodology. The researchers hoped to accomplish this by using differential item functioning (DIF) and differential testing functioning (DTF). The results from the study indicated that MFC measures were better at preventing faking than Likert-scale items. A key finding from the study was that MFC measures can cause faking if there are many blocks that contain positively and negatively worded statements. For that reason, it is important to limit the number of blocks used (Lee & Joo, 2019). A study by Lee et al. (2019) was conducted to investigate whether the type of personality measure used (Likert type vs MFC) had an impact on the stability of personality trait solutions. The results from the study indicated that MFC

measures produced more stable personality profiles and determined that Likert-scales were more fakable than MFC measures (Lee et al., 2019).

Another method uses to detect faking among test takers is the use of a Social Desirability (SD) scale. The objective of a SD scale is to detect, minimize and correct for social desirability responding (Van de Mortel, 2008). One of the most used social desirability scales to detect faking is the Marlowe-Crown Social Desirability scale. The scale is used in disguise during the test for test-takers to answer the questions which can detect if applicants have a high need for social approval which may lead them to answer more positively (Van de Mortel, 2008). High scores in this scale are associated with wanting to be seen in an unrealistically favorable manner which may be associated with faking.

Another scale used to detect faking is The Balanced Inventory of Desirable Responding (BIDR). The scale measures an individual's tendency to provide socially desirable responses on self-report inventories. The BIDR contains two scales which are the Self-Deceptive Enhancement (SDE) subscale, which measures the tendency to subconsciously give unrealistic favorable descriptions, and Impression Management (IM), which measures when applicants give conscious unrealistic positive self-descriptions (Paulhus, 1998). A study conducted by Lambert et al. (2016) wanted to compare Marlowe-Crowne Social Desirability Scale (MCSDS) and BIDR to determine which of these measures is more effective in detecting faking. Results from the study indicated

that MCSDS was more effective at identifying faking than BIDR (Lambert et al., 2016).

Other techniques such as mouse tracking have been used in research studies to detect any patterns that test applicants may show when taking a test. Mouse tracking has been used to follow the movement of the cursor from the beginning to the end (Freeman & Ambady, 2010). Findings associated with tracking an applicant's response time have been focused on the response time, trajectory, and the velocity of the mouse movement as results from a study showed that there were differences between truthful responses and lies that were instructed (Duran et al., 2010). A study by Monaro (2020) investigated if mouse tracking could improve the detection of fake-good responses. Results from the study indicated that participants who answered fake-good were slower in responding compared to participants who answered honestly (Monaro, 2020). Using warnings to caution applicants against faking is another technique used to reduce faking. Dwight and Donovan (2003) suggest letting the participant know that there are consequences with faking answers. When warnings are used studies have indicated that scores are different between those who are given a warning and those who are not (Dwight & Donovan, 2003). Robson et al. (2008) study had similar results in that their findings indicated that warning statements resulted in lower mean scores. However, it negatively impacted convergent validity (Robson et al., 2008).

Current Study

The goal of this study is to further explore why applicants' intent to fake during personality measurements and investigate the best solutions to prevent faking. As mentioned, applicants are willing to influence the perception of others during a social interaction to create a positive impression from others. This also applies when taking inventory items (Hogan, 2007). Since individuals want to create a positive perception about themselves to others, it is hypothesized that individuals partaking in impression management will have more intentions to fake when taking personality measurements to create a positive perception to the organization. The same goes for self-deceptive enhancement, in that respondents will tend to exaggerate on certain positive items in personality measurements due to respondents tending to think high of themselves during personality measurements. For that reason, individuals partaking in self-deceptive enhancement will have more intentions to fake in personality measurements. The hypotheses are listed below:

H1a: Impression management will be positively correlated with intentions to fake.

H1b: Self-deceptive enhancement will be positively correlated with intentions to fake.

Research on personality has indicated that personality is a consistent, albeit modest, predictor for job performance. The personality traits that organizations typically look for within applicants is the willingness to do one's

work well and emotional stability. As a result, the personality measurements used for personnel selection tend to measure conscientiousness, agreeableness, and neuroticism. Since applicants have an understanding of what organizations typically look for and applicants want to create a positive impression due to impression management, it is hypothesized that applicants will have more intentions to fake when completing personality measurements with regard to the conscientiousness, agreeableness and neuroticism scale. The hypotheses are listed below.

H2a: Conscientiousness will be negatively correlated with intentions of faking.

H2b: Neuroticism will be positively correlated with intentions of faking.

H2c: Agreeableness will be negatively correlated with intentions of faking.

Based on the research done comparing MFC and Likert scale, findings have indicated that the use of MFC measures reduces faking, score inflation, and maintains validity compared to when Likert-items scales are used. Another benefit of a MFC measure is it forces a respondent to typically choose between three items on which describes them the most. This makes it difficult for a respondent to determine which personality trait an organization is looking for. For those reasons this study will be focusing on comparing MFC measures with a Big 5 Likert-scale to see which causes less faking when focusing on certain personality traits. The hypothesis is listed below.

H3: There will be less intention to fake on a MFC measure compared to a Big 5 Likert-Scale measure when testing for Conscientiousness, Neuroticism, Open to Experience and Agreeableness.

H3a: The difference between the “Dream Job” group mean and the “Answer Honestly” group mean will be larger for Likert response scale items compared to MFC response items on conscientiousness.

H3b: The difference between the “Dream Job” group mean and the “Answer Honestly” group mean will be larger for Likert response scale items compared to MFC response items on neuroticism.

H3c: The difference between the “Dream Job” group mean and the “Answer Honestly” group mean will be larger for Likert response scale items compared to MFC response items on agreeableness.

CHAPTER TWO

METHOD

Participants

Individuals participating in the study had to be at least 18 years old to be eligible for the study. The sample size consisted of 618 participants who completed the survey. The average age for the sample was 31.23 years old. When looking at the demographic characteristics of the sample, 47.5% identified as men and 51.2% identified as women. For ethnicity, the sample reported as 62.3% identifying as White, 23.5% identifying as Hispanic/LatinX, 4.9% identifying as Asian and 4.4% identifying as African American/Black. 47.2% identified as having a Bachelor's degree, 17.7% identified as having a Master's degree and 16.9% identified as having an Associate degree. When reporting employment status. 67.3% identified as working Full-Time, 16.6% as Part-Time and 8.9% as just being a student. Other demographics the survey asked was how many jobs participants had applied within the past year and if they knew what the survey was about.

Procedures

The study was administered online via Qualtrics. Participants were recruited primarily online via Mturk and SONA where the Qualtrics link was distributed to participants to encourage them to participate in the study. 31.3% of the participants were from SONA and 68.6% were from Mturk. Participants who

were recruited via Mturk got a compensation of $1.50 for taking the survey. Participants who were recruited by SONA were compensated with extra credit. The Qualtrics link was distributed to CSUSB students via SONA where they were given the option to participate in the study. The study was completely voluntary, and participants were given the option to withdraw at any given moment.

Participants were given the questionnaire online and had to complete the “Intentions to Fake” measurement as well as the BIDR Version 6 scale. Once that was completed, participants were given written instructions to determine if participants were randomly placed in the “Answer honestly” group or "The answer as if you are applying for your dream job” group. A randomized function was placed in Qualtrics to determine what group participants were placed in. The function was set so that it equally placed participants into one of these groups. Once the instructions were given, participants had to self-report their answers when filling out the Big-Five MFC measures and the Big-Five Likert-Scale Items.

When looking over the data, there were a few participants who did not complete the attention check questions or answered them incorrectly. The data from these individuals were removed from the sample due to this being an indication of careless responding. Ultimately, six of these careless responders were removed. Other data was removed from the sample due to respondents inputting their age which consisted of being under 18 or due to inserting the year they were born which was an indication of not following directions. Six of these

types of situations were removed from the sample size. In total 12 surveys were removed from the sample.

Measures

Specific demographic items were included in the measurement which included race/ethnicity, gender, education, and employment status. This was necessary in order to have an idea of the nature of the sample. The questions being asked were “I identify my ethnicity as:”, “The gender I identify with is:”, “What is the highest degree or level of school you have completed?”, and “What is your current employment status?”. Participants were given the option of selecting their answers from the menu provided. (Please see Appendix A for the complete scale). Attention check items were also used throughout the measurement to detect careless respondents (Please see Appendix B for the complete scale).

The BIDR version 6 scale was used to test Hypothesis 1 (Paulhus, 1984). Paulhus’ BIDR scale is a self-report measure that measures social desirability responding in two components: impression management and self-deceptive enhancement (Chung, 2012). The scale contained 40-items and uses a 7-point Likert response scale. Example questions that the measures consist of include, “I have not always been honest with myself” (SDE) and “I never cover up my mistakes” (IM). Typical alpha reliability for this measurement is between .67-.77 when testing for self-deceptive enhancement and between .77-.85 when testing for impression management (see Appendix C for the complete scale).

Grieve's (2012) intention to fake scale was used to measure intention to fake. The scale is a 5-point Likert response scale that contains 4 items. The alpha reliability of the measurement is .84 and consists of questions such as "I intend to fake on future psychological tests" (see Appendix D for the complete scale).

For Hypothesis 2 and 3 the MFC measure that was designed to measure the Big Five factors of personality was used in this study (Lee et al., 2018). The MFC measure contains items for each dimension of the Big 5. That same MFC measure was then reconstructed to a Likert-Scale item response format to compare the responses of the MFC measure to the Likert-Scale responses to determine which measure resulted in less intentions to fake. (See Appendix E and F for the complete scale.) The MFC measures were scored by giving participants two points for the item they ranked that "best describes them", one point for the item they ranked as their second option, and zero points for the item they ranked as "least describe them". There were some negatively keyed items such as "I would describe my experience as dull" that were testing for extraversion. In this case, if participants described this item as "least like me", the participant was given 1 point for extraversion. Participants who described this item as "most like me" or ranked this item as their second option were given 0 points since it did not represent extraversion. Once the points were determined for each item, all of them were added and a score was given for each of the Big 5 Personality traits.

Analyses

For the first hypothesis, two correlations were conducted independently to measure the relationship between impression management and intentions to fake, along with the relationship between self-deceptive enhancement and intentions to fake. For Hypothesis 1a to be supported, there must be a positive correlation between impression management and the intention to fake. Hypothesis 1b will be supported if there is a positive correlation between self-deceptive enhancement and intentions to fake.

For the second hypothesis, three different bivariate correlations were conducted. Hypothesis 2a tested the relationship between conscientiousness with intentions of faking. Hypothesis 2b tested the relationship between neuroticism and intentions of faking. Hypothesis 2c tested the relationship between agreeableness with intentions of faking. Hypotheses 2a will be supported if there is a negative correlation between conscientiousness and intentions of faking. Hypothesis 2b will be supported if there is a positive correlation between neuroticism and intentions of faking. Hypothesis 2c will be supported if there is a negative correlation between agreeableness and intentions of faking.

For the third hypothesis, we compared the means of the “Answer honestly” group versus "The answer as if you are applying for your dream job” group when looking at conscientiousness, agreeableness, and neuroticism. A t-test was then used to compare the means to test if there is a significant mean

difference between the groups when looking at these personality traits. A follow-up Cohen’s d statistic was computed to determine the effect size. Hypothesis 3a will be supported if the means of conscientiousness are higher for the "The answer as if you are applying for your dream job group” versus the “Answer honestly group”. Hypothesis 3b will be supported if the means of neuroticism are lower for the "The answer as if you are applying for your dream job group” versus the “Answer honestly group”. Hypothesis 3a will be supported if the means of agreeableness are higher for the "The answer as if you are applying for your dream job group” versus the “Answer honestly group”.

CHAPTER THREE

RESULTS

Assumptions of homoscedasticity were met when looking at the Regression Standardized Residual graphs, in that they have equal variance among the predicted value. Assumptions of Normality of Residual were also met, due to the residuals being small (falling between 3.3 and -3.3), symmetrical, and centered around zero with the exception of two plots not falling between 3.3 and -3.3.

When testing the substantive hypotheses, the results indicated that impression management did not have a significant correlation with intentions to fake ($r=.013$, $p=.756$). Therefore, this fails to support Hypothesis 1A which states that impression management will be positively correlated with intentions to fake.

When conducting a second bivariate correlation between self-deceptive enhancement and intentions to fake, the results indicated that there was a significant, yet small, negative correlation between self-deceptive enhancement and intentions to fake ($r=-.112$, $p=<.05$). This result indicates that as self-deceptive enhancement increases, intentions to fake decreases. Therefore, Hypothesis 1B which states that self-deceptive enhancement will be positively correlated with intentions to fake was not supported.

For the second hypothesis, we computed another bivariate correlation to test if conscientiousness was negatively correlated with intentions to fake. The

results indicated that there was a small positive correlation between conscientiousness and intentions to fake (r=.087, p=.03). Therefore, Hypothesis 2a was not supported in that it was hypothesized that conscientiousness will be negatively correlated with intentions to fake.

When examining a bivariate correlation between neuroticism and intentions to fake the results indicated that there was a moderate significant positive correlation between neuroticism and intentions to fake (r=.234, p<.001). Therefore, Hypothesis 2b was supported in that neuroticism is positively correlated with intentions of faking.

When examining the bivariate correlation between agreeableness and intentions to fake, the results indicated that there was a moderate significant negative correlation between agreeableness and intentions to fake (r=-.480, p<.001). Therefore, Hypothesis 2c is supported in that agreeableness is negatively correlated with intentions of faking. Meaning those who scored high in agreeableness will have less intention to fake. Please refer to Table 1 for the correlations of hypothesis one and hypothesis two.

When testing the third hypothesis we compared the means of the “answer honestly” group versus "the answer as if you are applying for your dream job” group when looking at conscientiousness, neuroticism, and agreeableness between both Likert-scale items and MFC measures. When computing an independent sample t-test to see if there is a difference between the answer honestly group and faking group within conscientiousness Likert-scale items, the

results indicated that there was a statistically significant difference t (613) =-2.258, p=.024; d=.996. between the answer honestly group (M=3.44, SD=.48) and the applying for your dream job group (M=3.53, SD=.50).

When computing an independent sample t-test to see if there is a difference between the answer honestly group and faking group within conscientiousness MFC measures, the results indicated that there was a statistically significant difference t (613) =4.008, p<.001; d=.987. between the answer honestly group (M=11.82, SD=3.18) and the applying for your dream job group (M=10.84, SD=2.82).

Cohen's d was used to determine the difference between both groups to determine the differences between MFC measures and Likert-scales to determine the effect size. The scores were converted to z-scores to put them within the same scale since the scoring between the Likert-scale measures and MFC measures was different. When looking at the effect sizes of conscientiousness between MFC measures and Likert-scale items, the results indicate that there was not a substantial difference between the conscientiousness MFC measures (d=.987) compared to the conscientiousness Likert scale (d=.996).

When computing an independent sample t-test to see if there was a difference between the answer honestly group and the faking group within neuroticism Likert-scale items, the results indicated that there was not a significant difference t (613) =1.694, p=.091; d=.998. between the answer

honestly group (M=3.22, SD=.62) and the applying for your dream job group (M=3.14, SD=.60).

When computing an independent sample t-test to see if there was a difference between the answer honestly group and the faking group within neuroticism MFC measures, the results indicated that there was not a significant difference t (613) =1.227, p=. 220; d=.999 between answer honestly group (M=9.79, SD=3.44) and the applying for your dream job group (M=9.46, SD=3.14).

Cohen's d was used to estimate the differences between MFC measures and Likert-scales to determine the effect size. The scores were converted to z-scores to put them within the same scale since the scoring between the Likert-scale measures and MFC measures were different. When looking at the effect sizes of neuroticism between MFC measures and Likert-scale items, the results indicated that there was not a substantial difference between neuroticism Likert measures (d=.998) compared to "neuroticism" MFC measures (d=.999).

When computing an independent sample t-test to see if there was a difference between the answer honestly group and the applying for your dream job group within agreeableness Likert-scale items. Results indicated that there was not a statistically significant t (613) =-.812, p=.417; d=1.00 difference between the answer honestly group (M=3.62, SD=.48) and the applying for your dream job group (M=3.65, SD=.48).

When looking at the differences between the answer honestly group and the answer as if you are applying to your dream job group within MFC measures, an independent sample t-test was conducted to see if there was a difference between the answer honestly group and the applying for a dream job group within agreeableness MFC measures. Results indicated that there was a statistically significant difference t (613) =-2.431, p=.015; d=.996. between the answer honestly group (M=10.19, SD=2.70) and the applying for your dream job group (M=10.73, SD=2.78).

Cohen's d was used to compute the difference between MFC measures and Likert-scales to determine the effect size. The scores were converted to z-scores to put them within the same scale since the scoring between the Likert-scale measures and MFC measures was different. When looking at the effect sizes of agreeableness between MFC measures and Likert-scale items, the results indicate that there was not a substantial difference between agreeableness MFC measures (d=.996) compared to the agreeableness Likert scale (d=1.000). Please refer to Table 2 for the results of hypothesis three.

Table 1. Bivariate Correlation of Intentions to Fake

Variable	n	M	SD	1	2	3	4	5
1. Intention to fake	618	2.74	1.04	-				
2. Agreeableness	618	3.56	.496	-.480**				
3. Conscientiousness	618	3.62	.444	.087**	.287**			
4. Neuroticism	618	3.17	.611	.234**	-.250**	-.034		
5. SDE	617	4.04	.471	-.122**	.227**	.224**	-.564**	
6. IM	617	3.94	.555	.013	.125**	.190**	-.266**	.363**

** p< .05

Note: SDE is Self-Deceptive Enhancement. IM is Impression Management.

Table 2. MFC Measures and Likert Scale Table

Variable	n	M	SD	t-test	Cohen's d
Agreeableness (H)	307	3.61	.447	-.812	1.00
Agreeableness (F)	311	3.64	.483	-.812	1.00
Agreeableness (MFC-H)	307	10.194	.444	-2.431**	.996
Agreeableness (MFC-F)	311	10.733	2.78	-2.431**	.996
Conscientiousness (H)	307	3.43	.478	-2.258**	-.996
Conscientiousness (F)	311	3.52	.496	-2.258**	.996
Conscientiousness (MFC-H)	307	11.815	3.18	4.008**	.987
Conscientiousness (MFC-F)	311	10.842	2.82	4.008**	.987
Neuroticism (H)	307	3.22	.621	1.694	.998
Neuroticism (F)	311	3.13	.602	1.694	.998
Neuroticism (MFC-H)	307	9.78	3.44	1.227	.999
Neuroticism (MFC-F)	311	9.46	3.14	1.227	.999

** $p < .05$

Note: H is Honestly for Likert-Scale Items. F is Faking for Likert-Scale Items. MFC-F is Faking for Multidimensional Force Choice items. MFC-H is Honestly for Multidimensional Force Choice Items.

CHAPTER FOUR

DISCUSSION

General Discussion

The primary purpose of this study was to explore intentions of faking to determine if Impression Management (IM) and Self-Deceptive Enhancement (SDE) influenced intentions of faking during personality measurements. Secondly, we wanted to test which of the Big 5 personality traits between agreeableness, conscientiousness, and neuroticism would be correlated with intentions of faking. Lastly, we also examined which of the Big 5 personality traits are exaggerated when there are high stakes situations and the best option to address this issue when comparing Likert-scale items with Multidimensional Forced Choice (MFC) measures.

The first hypothesis explored intentions of faking when completing a personality measurement having to do with impression management and self-deceptive enhancement. Past studies have indicated that applicants fake their scores during personality measurements due to impression management and socially desirable responding to create a positive perception of themselves to others (Schlenker & Weigold, 1992). Therefore, hypothesis 1a predicted that "impression management will be positively correlated with intentions to fake" (h1a). The results from computing a correlation indicated that impression management did not have a significant correlation with intentions to fake. The

result from this correlation failed to support hypothesis 1a which predicted that "impression management will be positively correlated with intentions to fake".

It was hypothesized that impression management would lead to intentions to fake due to past research coming up with models to better understand the psychological process that leads to faking. The Muller-Hanson et al. (2006) model explains that faking occurs from an applicant's intent of faking which consists of perceptions of situations, willingness to fake, ability to fake, conscientiousness, and emotional stability with each individually leading to intentions to fake causing faking behavior. For this reason, it was predicted that impression management would be positively correlated with faking. However, the results do not indicate this happened. A possible explanation for why this hypothesis was not supported is due to individual differences in whether one will partake in faking which Muller-Hanson et al. (2006) mentions being one of the reasons one may or may not partake in faking behavior.

When testing Hypothesis 1B, it was predicted that "Self-deceptive enhancement will be positively correlated with intentions to fake". This was due to believing that individuals would exaggerate certain positive items in personality measurements due to respondents tending to think highly of themselves. When computing the correlation, the results indicated that there was a significant small negative correlation between self-deceptive enhancement and intentions to fake. Indicating that as self-deceptive enhancement increases, intentions to fake

decrease. This goes against what was originally hypothesized for Hypothesis 1B that states self-deceptive enhancement will be positively correlated with intentions.

A possible explanation for why both hypotheses were not supported might be due to part of the sample consisting primarily of college students. When looking at the “intentions to fake” scale and running descriptive statistics, the mean for the scale was 2.74. If the four items of each scale are broken down to see the mean of each item. The mean for “faking in future psychological measurements” is 2.79, “benefiting from faking in psychological measurements” is 3.06, “I would never fake in psychological measurements” is 3.87, and “expect to fake in psychological measurements” is 2.99. The reason for reporting this information is due to part of the sample size consisting of college students (N=188) who received an incentive of extra credit for participating in the survey. College students are more accustomed to filling out personality measurements and there might have been a case that students scored low when filling out the intentions to fake scale due to fear of disqualification from the survey for admitting to faking believing it may cause them not to receive extra credit.

In conclusion, both Hypothesis 1A and 1B were rejected in that impression management is not positively correlated with intentions to fake, and self-deceptive enhancement is negatively correlated with intentions to fake.

For the second hypothesis, we tested which of the personality traits within the Big Five would be correlated with intentions of faking. The reason for testing this hypothesis is because we wanted to determine what makes an individual partake in faking when taking a personality measurement during personnel selection.

For Hypothesis 2a the results indicated that there was a small positive correlation between conscientiousness and intentions to fake causing Hypothesis 2a to not be supported. The reason it was expected for conscientiousness to be negatively correlated with intentions to fake was that conscientiousness is commonly described as a trait that is associated with self-control and being caring workers. A meta-analysis by Barrick and Mount (1991) determined that conscientiousness showed a consistent relationship with job performance for all groups, which is one of the reasons why conscientiousness is a desirable personality trait for organizations when it comes to personnel selection (Salgado, 1997). Therefore, it was expected for conscientiousness to be negatively correlated with intentions to fake based on the positive characteristics of this personality trait. However, this hypothesis was not supported.

Another reason why this hypothesis was not supported may have been because as mentioned, conscientiousness is a desirable personality trait among organizations. Based on the extensive research on the Big 5 personality traits, applicants tend to understand what an organization typically looks for which

might have led to participants exaggerating their score on the conscientiousness items. This relates back to the topic of faking among personality traits in that participants may not necessarily be faking but they may be participating in self-deceptive enhancement in that respondents may think highly of themselves within this particular trait. When looking at the mean of conscientiousness and comparing it with the mean of agreeableness, the mean of conscientiousness tends to be slightly greater.

Yet another explanation for this result is based on Snell et al.'s (1999) model of faking which categorized individual differences on why individuals fake. The first one is the "ability to fake" and the second one is the "willingness to fake". For someone to acknowledge they have the "ability to fake" one must need to be conscientious to acknowledge that. Muller-Hanson et al. (2006) model of faking also listed conscientiousness as one of the antecedents for faking. Therefore, these models support the findings on why conscientiousness was positively correlated with intentions of faking.

For Hypothesis 2b the result indicated that there was a moderate significant positive correlation between neuroticism and intentions to fake. This supported Hypothesis 2b which hypothesized that neuroticism is positively correlated with intentions of faking. Neuroticism is seen as an unfavorable trait within the Big 5 due to it being associated with anxiety, hostility, and impulsiveness. Barrick and Mount's (1991) study on personality measures and

job performance indicated that conscientiousness, extraversion, and openness to experience were strong predictors of job performance, whereas neuroticism typically leads to high turnover rates. Since neuroticism is associated with bad work, it was hypothesized that it would also be associated with intentions of faking. This was supported by the Snell et al. (1999) model on faking which categorized individual differences on why individuals fake. The first one is the "ability to fake" and the second one is the "willingness to fake". It seems like an individual must first identify that they have the ability to fake which may be why the results from Hypothesis 2a indicate that there was a small positive correlation between conscientiousness and intentions of faking. Once an individual can identify that they have the ability to fake, they must be willing to partake in faking which is where the personality trait of neuroticism comes into play. This demonstrates a possible explanation of why the result from this hypothesis indicates that there is a moderate significant positive correlation between neuroticism and intentions of faking.

For Hypothesis 2c the results indicated that there was a moderate significant negative correlation between agreeableness and intentions to fake. This supported Hypothesis 2c which hypothesized that agreeableness is negatively correlated with intentions of faking. Agreeableness is seen as a personality trait associated with kindness, cooperation, and putting others' needs before one's own needs. A possible explanation for these results is based on the Snell et al. (1999) model of faking. As mentioned, Snell's model categorizes

individual differences on why individuals fake. The first one is the “ability to fake” and the second one is the “willingness to fake” which is an explanation of why conscientiousness and neuroticism are positively correlated with intentions of faking. The personality trait of agreeableness does not reflect anything from Snell’s et al. model that indicates why individuals may have intentions of faking. Muller-Hanson et al. (2006) had an initial model on faking that was similar to Snell et al. model in that willingness to fake and ability to fake were in the model along with the perception of situations, conscientiousness, and emotional stability as being the cause of intentions of faking which eventually leads to causing faking behavior. For this reason, agreeableness was negatively correlated with intentions of faking.

For the third hypothesis, we tested to see if there were any significant differences when comparing the means of the “Answer honestly” group versus "The answer as if you are applying for your dream job” group when looking at conscientiousness, neuroticism and agreeableness within both Likert-scale items and MFC measures. Along with comparing Cohen’s d to determine the differences between Likert-scale items and MFC measures.

Hypothesis 3a examined the difference between both groups when looking at conscientiousness. The results indicated that there was a statistically significant difference between the “answer honestly group” and “applying for your dream job group” between Likert-scale items in that the “applying for your dream

job” group had a higher mean compared to the “answer honestly group”. When looking at both groups between MFC measures the results indicated that there was a statistically significant difference between both groups in that the “answer honestly group” had a higher mean compared to the “answer as if you are applying for your dream job”. There were smaller effect sizes for “conscientiousness” MFC measures compared to “conscientiousness” Likert scale items.

This hypothesis was constructed based on the VIE Theory in that participants not placed in the “answer honestly group” would score higher in conscientiousness. This was supported in the Likert scale items in that those participants placed in the “answer as if you are applying to your dream job group” had a higher mean that participants placed in the “answer honestly group” which supports the VIE theory. However, the findings from the MFC measures for conscientiousness did not support the VIE theory due to candidates placed in the “answer honestly group” having a higher mean compared to the “answer as if you are applying for your dream job”. An explanation for this finding is that conscientiousness is the most desired personality trait that organizations look for among candidates due to studies demonstrating that conscientiousness has shown a consistent relationship with job performance (Barrick & Mount, 1991). Therefore, it is not a surprise that the answer honestly group had a higher mean than the experimental group. This can be supported by studies on impression management and socially desirable responding in that participants may try to

present themselves positively by responding in a manner that makes them look good or exaggerating (Paulhus, 1991;Hogan, 2007) As Mueller-Hanson et al. (2006) mentioned there tend to be individual differences in whether one will partake in faking. When comparing the Cohen's d between conscientiousness Likert-scale items and conscientiousness MFC measures the findings suggested that there wasn't a substantial difference between the Cohen's d estimate of Likert-scale items and Cohen's d of MFC measures . But since the control group had a higher mean than the experimental group within MFC measures, the theory that MFC measures would be more effective in combating Likert-scale items was not supported.

Hypothesis 3b examined the differences between both groups when looking at neuroticism. The results indicated that there was not a statistically significant difference between the answer honestly group and the answer as if you are applying for your dream job group when looking at Likert-scale items. When looking at both groups within MFC measures the results also indicated that there was not a significant difference between both groups as well. When looking at both Cohen's d to determine the effect sizes between Likert-scale items and MFC measures for neuroticism, the results indicated that there wasn't a substantial difference between the effect sizes of "neuroticism" Likert-scale items compared to "neuroticism" MFC measures.

The findings from this hypothesis are supported by studies done on the Big 5 personality traits in that neuroticism is not the ideal trait that organizations typically look for among candidates due to the trait being associated with negative actions (Barrick & Mount, 1991). Therefore, applicants tend to understand that when filling out personality measures during a selection process it is not ideal to score high in items that measure neuroticism. Therefore, impression management or socially desirable responding could have been active among participants when answering items that are measuring neuroticism. This can explain why there was not a difference between both groups within both Likert-scale items and MFC measures. This can be supported by studies done within high stakes situations and low stakes situations in that applicants in high stakes situations usually answer more favorably which is why it would be ideal to not score high in neuroticism (Hu & Connelly, 2021).

Hypothesis 3c examined the difference between both groups when looking at agreeableness. The results indicated that there was not a statistically significant difference between “the answer honestly group” and the “applying for your dream job” group between Likert-scale items. When looking at the difference between the MFC measures between the “answer honestly” group versus “answer as if you are applying for your dream job” within agreeableness. Results indicated there was a statistically significant difference in that there was a higher mean for the “answer as if you are applying for your dream job” group compared to the group that was going to “answer honestly”. When looking at

both Cohen's d to determine the effect sizes between Likert-scale items and MFC measures for agreeableness after being converted to z-scores. There was not a substantial difference in the effect size between "agreeableness" MFC measures and "agreeableness" Likert-scale items.

This hypothesis was constructed on the basis that participants who were placed in the "answer as if you are applying to your dream job" group would have higher scores of agreeableness compared to participants who were asked to "answer honestly" due to the VIE Theory stating that faking tends to happen when there is an extrinsic reward. In this case, the opportunity for participants to have the opportunity to land their dream job would be the extrinsic reward (Vroom, 1964). As the results suggested, there was no difference between the answer honestly group and the answer as if you are applying to your dream job group when looking at agreeableness within Likert-scale items. This supports Fischer's et al. (2019) research on the drawbacks of Likert-scale items in that may be difficult to detect faking when Likert-scale items are used. The results from the MFC measures suggest that MFC measures can detect differences between both the "applying for your dream job group" and "answer honestly group" when looking at agreeableness in that it was able to detect differences between the "applying for your dream job group" and the "answer honestly group" in that there was a greater mean for the "applying for your dream job group". This supported Vroom's (1964) VIE Theory that faking tends to happen when there is

an extrinsic reward in place which in this case would be the opportunity for participants to land their dream job.

The result is also supported by Hu and Connelly's (2021) study which compared personality measurements between high stake situations and low stakes situations in that there were higher means for those placed in a high-stake setting. When comparing the Cohen's d between agreeableness Likert-scale items and agreeableness MFC measures, the findings suggested that the Cohen's d estimate for Likert-scale items was larger than the Cohen's d for MFC measures. This demonstrates that MFC measures are slightly effective at reducing faking than Likert-scale items. An explanation for this is due to the differences between MFC measures groups and Likert-scale items. Studies on MFC measures have suggested that MFC measures are more helpful in combating faking than Likert-scale items (Fisher et al., 2019). The reason being that MFC measures can eliminate extreme answer responding and response biases based on their structure of organizing different dimensions statements into blocks and forcing candidates to choose one that best describes them and one that least describes them (Lee et al., 2019; Wetzel et al., 2020). Whereas Likert-scale items typically increase the chances of halo, central tendency, and faking (Wetzel & Grief, 2018).

Participants were asked towards the end of what they thought the measurement was testing. Majority of the participants answered the question

whereas others provided minimal responses. There were few participants who mentioned faking, however, that could have been spoiled due to the title of the survey mentioning “faking” when clicking on the link to take the survey. Other participants mentioned topics such as personality and ethics whereas, others provided short statements. Those who answered with short statements could have been due to survey fatigue.

Theoretical Implications

Theoretical implications of the results of this study indicate that the personality trait of conscientiousness tends to be scored higher in general for both the control group and experimental group. Studies on why applicants fake have indicated that applicants fake their scores during personality measurements largely via impression management and socially desirable responding since they want to present a positive image of themselves to the organization (Paulhus, 1991; Hogan, 2006). This is more common when the stakes are higher such as applying for managerial positions (Levashina et al., 2014). Vroom’s (1964) VIE Theory states that faking tends to be a behavior that occurs due to an extrinsic reward, which in this case is participants potentially obtaining their dream job. Since past studies have indicated conscientiousness as being a key predictor of job performance (Barrick & Mount, 1991; Salgado, 1997). It only makes sense that the personality trait of conscientiousness is highly scored among the

“applying for your dream job” and the “answer honestly group” for both Likert-scale and MFC measures.

Another theoretical implication from this study is that it supports Snell et al. (1999) model of individual differences of why faking occurs and part of Muller-Hanson et al. (2006) model in that faking occurs by an individual first identifying “the ability to fake” and the “willingness to do so”. Muller-Hanson et al. model goes more into detail by including the perception of situations, conscientiousness, and emotionally stability as being the cause of faking. The result from this study supports their models in that it was demonstrated that conscientiousness and neuroticism are positively correlated with intentions of faking.

Practical Implications

We compared the Likert-Scale response format and MFC response format measures to determine which one is associated with participants having less intentions of faking based on the estimated effect sizes when looking at Cohen’s d. The results of the present study have practical implications in that it was determined that MFC measures for agreeableness and conscientiousness had smaller effect sizes than Likert scale items for agreeableness and conscientiousness. This demonstrated that MFC measures help reduce faking compared to Likert scale items. Organizations will have the option to select which type of measurement they would want to use in the future based on past research and the results from this study. Some of the potential drawbacks of

Likert-scale items are that it increases the chance of halo, central tendency, and faking (Wetzel & Greiff, 2018). Organizations will have to determine those are issues that they should be concerned about. Selection analysts would need to convince the organization on implementing that change if an organization is concerned about faking. Studies comparing MFC and Likert-Scale items have determined that personality profiles within Likert-scales were more fakable than MFC measures (Lee et al., 2019). If organizations determine that MFC measures should be used, then selection analysts would need to be trained on how to score MFCs since they vary in how they are scored. Overall, the findings from the study bring new light on if MFC measures can prevent faking compared to Likert-Scale items which will contribute to already existing research.

Limitations and Direction of Future Research

There were some limitations in this study. The first one being that participants had to self-report their answers when answering the survey. This is a limitation due to the possibility of participants not answering truthfully to some of the items. One of the instructions from the survey was for applicants to answer as if “they are applying to their dream job”. Even though the instructions mentioned for participants to place themselves in this situation, it is not the same thing as taking a personality exam for a job position one may be interested in. Therefore, the replication of using high stakes situations could not be done appropriately. Another limitation was that a previous MFC measure was used from a previous study.

There are different directions future research can go based on the findings from this study. One of the findings from the study conducted was that conscientiousness is positively correlated with intentions of faking. A possible future direction future studies can explore is to interview individuals who scored high in conscientiousness in testing and ask them questions about potentially faking in a personality measurement. It is important we get an understanding of the potential thought process of individuals high in conscientiousness regarding potentially having intentions of faking.

Regarding MFC measures, future directions that studies on MFC measures can go based on our findings is to continue to conduct studies on the comparisons to MFC measures and Likert-scale items. Our findings indicated that there were smaller effect sizes for MFC measures than Likert-scale item, for two out of three of the personality traits. However, the difference was closer than expected. This speaks of the continuous studies that should be conducted to replicate findings on this subject.

Conclusion

The focus of this study had various points. First, we wanted to explore intentions of faking to determine if Impression Management (IM) and Self-Deceptive Enhancement (SDE) had an influence on intentions of faking during personality measurements. Secondly, we wanted to test which of the Big 5 personality traits between agreeableness, conscientiousness, and neuroticism would be correlated with intentions of faking. Lastly, which of the Big 5

personality traits are exaggerated when there are high stakes situations and the best option to address this issue when comparing Likert-scale items with Multidimensional Forced Choice (MFC) measures. Findings from this study determined that Impression Management and Self-Deceptive Enhancement were not positively correlated with intentions of faking rejecting the hypothesis. Another finding was it was determined that agreeableness was negatively correlated with intentions of faking and both conscientiousness and neuroticism were positively correlated with intentions of faking. Lastly, when looking at which of the personality traits were exaggerated when placed in high stakes situations. It was determined that the personality trait of agreeableness was scored higher under the “applying for your dream job” group than the “answer honestly group” under MFC measures. Whereas conscientiousness was scored higher under the “applying for your dream job” group than the “answer honestly group” under Likert-scale items but vice versa under MFC measures. When comparing the differences between the Likert-scale items and MFC measures through Cohen’s d estimate there was not much of a substantial difference between the measures. The current study demonstrates that intentions to fake can be caused by conscientiousness and neuroticism and that personality traits such as conscientiousness and agreeableness can be exaggerated when placed in high stakes situations.

APPENDIX A

DEMOGRAPHIC ITEMS

Demographic items

I identify my ethnicity as:
- o Asian
- o Black or African
- o Hispanic/Latinx
- o Middle Eastern
- o Native American
- o Pacific Islander
- o White
- o Prefer not to answer
- o Not listed: _____

The gender I identify with is:
- o Female
- o Male
- o Nonbinary
- o Prefer not to answer
- o Other_____

What is the highest degree or level of school you have completed?
- o 12th grade or less
- o Graduated high school or GED
- o Some college, no degree
- o Associate degree
- o Trade School
- o Bachelor's degree
- o Master's degree
- o Doctorate

What is your current employment status?
- o Employed full time (40+ hours a week)
- o Employed part-time (less than 40 hours a week)
- o Unemployed
- o Student
- o Retired
- o Self-employed
- o Intern

APPENDIX B

ATTENTION CHECK QUESTIONS

Attention Check Questions

The following question is an attention check item to make sure you are still participating attentively in the study. When answering this question please select blue as your answer.
Based on the information provided in the instructions, what color have you been asked to enter?

- o Green
- o Red
- o Blue
- o Brown
- o Black
- o Yellow
- o Orange

The following question is an attention check item to make sure you are still participating attentively in the study. When answering this question please select pizza as your answer.

- o Hamburger
- o Cheeseburger
- o Pizza
- o Hot Dog
- o Tacos
- o Sandwich

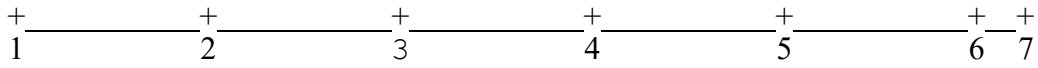
APPENDIX C

BIDR VERSION- FORM 40A

BIDR Version 6 - Form 40A

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

Using the scale below as a guide, write a number beside each statement to indicate how true it is.



____ 1. My first impressions of people usually turn out to be right.

____ 2. It would be hard for me to break any of my bad habits.

____ 3. I don't care to know what other people really think of me.

____ 4. I have not always been honest with myself.

____ 5. I always know why I like things.

____ 6. When my emotions are aroused, it biases my thinking.

____ 7. Once I've made up my mind, other people can seldom change my opinion.

____ 8. I am not a safe driver when I exceed the speed limit.

____ 9. I am fully in control of my own fate.

____ 10. It's hard for me to shut off a disturbing thought.

____ 11. I never regret my decisions.

____ 12. I sometimes lose out on things because I can't make up my mind soon enough.

____ 13. The reason I vote is because my vote can make a difference.

____ 14. My parents were not always fair when they punished me.

____ 15. I am a completely rational person.

____ 16. I rarely appreciate criticism.

____ 17. I am very confident of my judgments

____ 18. I have sometimes doubted my ability as a lover.

____ 19. It's all right with me if some people happen to dislike me.

____ 20. I don't always know the reasons why I do the things I do.

____ 21. I sometimes tell lies if I have to.

____ 22. I never cover up my mistakes.

____ 23. There have been occasions when I have taken advantage of someone.

____ 24. I never swear.

____ 25. I sometimes try to get even rather than forgive and forget.

____ 26. I always obey laws, even if I'm unlikely to get caught.

____ 27. I have said something bad about a friend behind his/her back.

____ 28. When I hear people talking privately, I avoid listening.

____ 29. I have received too much change from a salesperson without telling him or her.

____ 30. I always declare everything at customs.

____ 31. When I was young I sometimes stole things.

____ 32. I have never dropped litter on the street.

____ 33. I sometimes drive faster than the speed limit.

____ 34. I never read sexy books or magazines.

____ 35. I have done things that I don't tell other people about.

____ 36. I never take things that don't belong to me.

____ 37. I have taken sick-leave from work or school even though I wasn't really sick.

____ 38. I have never damaged a library book or store merchandise without reporting it.

____ 39. I have some pretty awful habits.

____ 40. I don't gossip about other people's business.

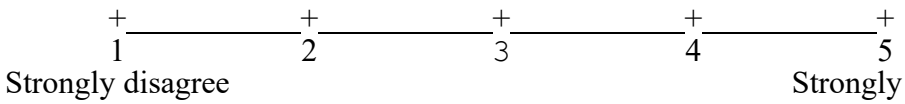
APPENDIX D

INTENTION TO FAKE SCALE

Intention to fake scale

Grieve, R. (2012). The role of personality, psychopathy, and previous experience with assessment in intentions to fake in psychological testing. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 31*(4), 414-422. https://doi.org/10.1007/s12144-012-9158-x

Using the scale below as a guide, write a number beside each statement to indicate how true it is.



____ 1. I intend to fake on future psychological test.

____ 2. If I could see a benefit, I would fake responses on a psychological test.

____ 3. I would never fake on a psychological test (reverse coded).

____ 4. I expect to fake on future psychological tests.

APPENDIX E

20-TRIPLET MFC MEASURE OF BIG-FIVE PERSONALITY

Block	Items	Ranking
1	Respect others. Have a rich vocabulary. Follow through with my palns.	
2	Get excited by new ideas. Warm up quickly to others. Don't put my mind on the task at hand.	
3	Feel threatened easily. Do not enjoy going to art museums. Know how to captivate people.	
4	Am always prepared. Accept people as they are. Do not worry about things.	
5	Am the life of the party. Cut others to pieces. Am filled with doubts about things.	
6	Rarely look for a deeper meaning in things. Cheer people up. Carry out my plans.	
7	Fear for the worst. Keep in the background. Am not interested in abstract ideas.	
8	Have a good word for everyone. Am exacting in my work. Have a vivid imagination.	
9	Do things according to a plan. Get back at others. Feel comfortable around people.	
10	Find it difficult to get down to work. Am often down in the dumps Enjoy thinking about things.	

Block	Items	Ranking
11	Treat all people equally. Would describe my experiences as somewhat dull. Am not easily frustrated.	
12	Waste my time. Find it difficult ot approach others. Trust what people say.	
13	Am comcerned about others. Believe in the importance for art. Feel comfortable with myself.	
14	Complete tasks successfully. Don't like to draw attention to myself. Am relaxed most of the time.	
15	Have a sharp tongue. Get stressed out easily. Carry the conversation to a higher level.	
16	Am interested in theoretical discussions. Talk to a lot of different people at parties. Have frequent mood swings.	
17	Finish what l start. Insult people. Panic easily.	
18	Enjoy wild flights of fantasy. Get chores done right away. Sympathize with others' feelings.	
19	Believe that others have good intentions. Often feel blue. Make friends easily.	
20	Do not like art. Start conversations. Need a push to get started.	

APPENDIX F

LIKERT SCALE MEASURE OF BIG 5 PERSONALITY

Likert-Scale Measure of Big-Five personality

Lee, P., Lee, S., & Stark, S. (2018). Examining validity evidence for multidimensional forced choice measures with different scoring approaches. *Personality and Individual Differences, 123,* 229–235. https://doi.org/10.1016/j.paid.2017.11.031

Using the scale below as a guide, write a number beside each statement to indicate how true it is.

1	2	3	4	5
Least like me				Most Like me

____ 1. Respect others.

____ 2. Have a rich vocabulary.

____ 3. Follow through with my plans.

____ 4. Get excited by new ideas.

____ 5. Warm up quickly to others.

____ 6. Don't put my mind on the task at hand.

____ 7. Feel threatened easily.

____ 8. Do not enjoy going to art museums.

____ 9. Know how to captivate people.

____ 10. Am always prepared.

____ 11. Accept people as they are.

____ 12. Do not worry about things.

____ 13. Am the life of the party.

____ 14. Cut others to pieces.

____ 15. Am filled with doubts about things.

____ 16. Rarely look for a deeper meaning in things.

____ 17. Cheer people up.

____ 18. Carry out my plans.

____ 19. Fear for the worst.

____ 20. Keep in the background.

____ 21. Am not interested in abstract ideas.

____ 22. Have a good word for everyone.

____ 23. Am exacting in my work.

____ 24. Have a vivid imagination.

____ 25. Do things according to a plan.

____ 26. Get back at others.

____ 27. Feel comfortable around people.

____ 28. Find it difficult to get down to work

____ 29. Am often down in the dumps.

____ 30. Enjoy thinking about things.

____ 31. Treat all people equally.

____ 32. Would describe my experiences as somewhat dull.

____ 33. Am not easily frustrated.

____ 34. Waste my time.

____ 35. Find it difficult to approach others.

____ 36. Trust what people say.

____ 37. Am concerned about others.

____ 38. Believe in the importance of art.

____ 39. Feel comfortable with myself.

____ 40. Complete tasks successfully.

____ 41. Don't like to draw attention to myself.

____ 42. Am relaxed most of the time.

____ 43. Have a sharp tongue.

____ 44. Get stressed out easily.

____ 45. Carry the conversation to a higher level.

____ 46. Am interested in theoretical discussions.

____ 47. Talk to a lot of different people at parties.

____ 48. Have frequent mood swings.

____ 49. Finish what I start.

____ 50. Insult people.

____ 51. Panic easily.

____ 52. Enjoy wild flights of fantasy.

____ 53. Get chores done right away.

____ 54. Sympathize with others' feelings.

____ 55. Believe that others have good intentions.

____ 56. Often feel blue.

____ 57. Make friends easily.

____ 58. Do not like art.

____ 59. Start conversations.

____ 60. Need a push to get started.

APPENDIX G

IRB APPROVAL

November 15, 2022

CSUSB INSTITUTIONAL REVIEW BOARD
Administrative/Exempt Review Determination
Status: Exempt
IRB-FY2023-115

Kenneth Shultz
CSBS - Psychology
California State University, San Bernardino
5500 University Parkway
San Bernardino, California 92407

Dear Kenneth Shultz :

Your application to use human subjects, titled "The Effect of Response Format on Faking in Personality Measurement used for Personnel Selection" has been reviewed and determined exempt by the Institutional Review Board (IRB) of California State University, San Bernardino under the federal regulations at 45 CFR 46. As the researcher under the exempt category, you do not have to follow the requirements under 45 CFR 46 which requires annual renewal and documentation of written informed consent which are not required for the exempt category. However, exempt status still requires you to attain consent from participants before conducting your research as needed.

Your IRB proposal is approved. This approval is valid from November 15, 2022.

This approval notice does not replace any departmental or additional campus approvals which may be required including access to CSUSB campus facilities and affiliate campuses. Investigators should consider the changing COVID-19 circumstances based on current CDC, California Department of Public Health, and campus guidance and submit appropriate protocol modifications to the IRB as needed. CSUSB campus and affiliate health screenings should be completed for all campus human research related activities. Human research activities conducted at off-campus sites should follow CDC, California Department of Public Health, and campus guidance. See CSUSB's COVID-19 Prevention

Plan for more information regarding campus requirements.

Your responsibilities as the investigator include reporting to the IRB Committee the following three requirements highlighted below. Please note, failure of the investigator to notify the IRB of the below requirements may result in disciplinary action.

- **Submit a protocol modification (change) form if any changes (no matter how minor) are proposed in your study for review and approval by the IRB before being implemented in your study to ensure the risk level to participants has not increased,**
- **Submit an unanticipated/adverse events form if harm is experienced by subjects during your research, and**
- **Submit a study closure through the Cayuse IRB submission system when your study has ended.**
- **Ensure your CITI human subjects training is kept up-to-date and current throughout the study for all investigators.**

The protocol modification, adverse/unanticipated event, and closure forms are located in the Cayuse Human Ethics (IRB) System. If you have any questions regarding the IRB decision, please contact Michael Gillespie, the Research Compliance Officer. Mr. Michael Gillespie can be reached by phone at (909) 537-7588, by fax at (909) 537-7028, or by email at mgillesp@csusb.edu. Please include your application approval identification number (listed at the top) in all correspondence.

If you have any questions regarding the IRB decision, please contact Dr. Jacob Jones, Assistant Professor of Psychology. Dr. Jones can be reached by email at Jacob.Jones@csusb.edu. Please include your application approval identification number (listed at the top) in all correspondence.

Best of luck with your research.

Sincerely,

King-To Yeung

King-To Yeung, Ph.D., IRB Chair
CSUSB Institutional Review Board

KY/MG

REFERENCES

Arthur, W., Hagen, E., & George, F. (2021). The lazy or dishonest respondent: Detection and prevention. *Annual Review of Psychology and Organizational Behavior, 8(1),* 105–137.

Bäckström, M., Björklund, F., and Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, 43 (3), 335-344.

Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., & Mount, M. K. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance, 18*(4), 359–372. https://doi.org/10.1207/s15327043hup1804_3

Bing, M. N., Kluemper, D. H., Davison, H. K., Taylor, S., & Novicevic, M. (2011). Overclaiming as a measure of faking. *Organizational Behavior and Human Decision Processes*, 116(1), 148–162. https://doi.org/10.1016/j.obhdp.2011.05.006

Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, 71(3), 460–502. https://doi.org/10.1177/0013164410375112

Buehl, A.-K., Melchers, K. G., Macan, T., & Kühnel, J. (2019). Tell me sweet little lies: How does faking in interviews affect interview scores and interview validity? *Journal of Business and Psychology*, 34(1), 107–124. https://doi.org/10.1007/s10869-018-9531-3

Chung J. M. (2012). The contribution of self-deceptive enhancement to display rules in the United States and Japan. *Asian journal of social psychology*, *15*(1), 69–75. https://doi.org/10.1111/j.1467-839X.2011.01358.x

Converse, P. D., Peterson, M. H., & Griffith, R. L. (2009). Faking on personality measures: Implications for selection involving multiple predictors. *International Journal of Selection and Assessment, 17*(1), 47–60. https://doi.org/10.1111/j.1468-2389.2009.00450.x

Donovan, J. J. (2014). The impact of applicant faking on selection measures, hiring decisions, and employee performance. *Journal of Business and Psychology, 29*(3), 479–493. https://doi.org/10.1007/s10869-013-9318-5.

Dunlop, P.D., Holtrop, D., Ashby, L.M., Bharadwaj, A., & Donovan J.J. (2022) Valence, instrumentality, expectancy, and ability as determinants of faking, and the effects of faking on criterion-related validity. *Journal of Business and Psycholo*gy. https://doi.org/10.1007/s10869-022-09797-0

Duran, N. D., Dale, R., & McNamara, D. S. (2010). The action dynamics of overcoming the truth. *Psychonomic Bulletin & Review*, 17(4), 486–491.

Dwight, S. A., & Donovan, J. J. (2003). Do warnings not to fake reduce faking? *Human Performance, 16*(1), 1–23.

Edwards, A. L. (1957). Manual for the Edwards personal preference schedule. New York, NY: *Psychological Corporation.*

Ellingson, J. E., & McFarland, L. A. (2011). Understanding faking behavior through the lens of motivation: An application of VIE theory. *Human Performance, 24*(4), 322–337. https://doi.org/10.1080/08959285.2011.597477

Ellingson, J. E., Smith, D. B., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology, 86,* 122-133.

Fang, R., Landis, B., Zhang, Z., Anderson, M. H., Shaw, J. D., & Kilduff, M. (2015). Integrating personality and social networks: A meta-analysis of personality, network position, and work outcomes in organizations. *Organization Science, 26*(4), 1243–1260.

Fisher, P. A., Robie, C., Christiansen, N. D., Speer, A. B., & Schneider, L. (2019). Criterion-related validity of forced-choice personality measures: A cautionary note regarding Thurstonian IRT versus classical test theory scoring. *Personnel Assessment and Decisions, 5(1),* 49-61.

Fluckinger, C.D., McDaniel, M., & Whetzel, D. (2008) *Review of faking in personnel selection*, 90-105. The University of Akron.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mouse-tracking method. *Behavior Research Methods, 42(1),* 226–241.

Furnham, A. (1990). Faking personality questionnaires: Fabricating different profiles for different purposes. *Current Psychology, 9(1)*, 46-55.

Furnham, A. and Crump, J. (2015) The myers-briggs type indicator (MBTI) and promotion at work. *Psychology*, *6(12)*, 1510-1515. doi: 10.4236/psych.2015.612147.

Gatewood, R. D., Feild, H. S., & Barrick, M. (2019). *Human Resource Selection* (9th ed). Wessex Publishing Company.

Gibby, R. & Zickar, M. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. *History of Psychology, 11(3),* 164-184. 10.1037/a0013041.

Goffin, R. D., & Boyd, A. C. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology/Psychologie canadienne, 50 (3),* 151–160.

Grieve, R. (2012). The role of personality, psychopathy, and previous experience with assessment in intentions to fake in psychological testing. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 31*(4), 414–422. https://doi.org/10.1007/s12144-012-9158-x

Grieve, R., & de Groot, H. T. (2011). Does online psychological test administration facilitate faking? *Computers in Human Behavior, 27*(6), 2386–2391. https://doi.org/10.1016/j.chb.2011.08.001

Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review, 36*(3), 341–355. https://doi.org/10.1108/00483480710731310

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*(2), 135–164. https://doi.org/10.1111/j.1744-6570.1965.tb00273.x

Hartman, N. S., & Grubb, W. L. III. (2011). Deliberate faking on personality and emotional intelligence measures. *Psychological Reports, 108*(1), 120–138. https://doi.org/10.2466/03.09.28.PR0.108.1.120-138

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

Heggestad, E. D., Morrison, M., Reeve, C. L., & McCloy, R. A. (2006). Forced-choice assessments of personality for selection: Evaluating issues of normative assessment and faking resistance. *Journal of Applied Psychology, 91*(1),9-24.

Henle, C., Giacalone, R., & Jurkiewicz, C. (2005). The role of ethical ideology in workplace deviance. *Journal of Business Ethics, 56(3),* 219-230. 10.1007/s10551-004-2779-8.

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92,* 1270-85. 10.1037/0021-9010.92.5.1270.

Hogan, R. (2006). *Personality and the fate of organizations.* Psychology Press.

Hooper, A. C. (2007). *Self-presentation on personality measures in lab and field settings: A meta-analysis* (dissertation).

Hu, J., & Connelly, B. S. (2021). Faking by actual applicants on personality tests: A meta-analysis of within-subjects studies. *International Journal of Selection and Assessment*, *29(3-4),* 412– 426. https://doi.org/10.1111/ijsa.12338

Huber, C.R., Kuncel, N.R., Huber, K.B., & Boyce, A.S. (2021). Faking and the validity of personality tests: An experimental investigation using modern forced choice measures. *Public Administration and Development, 7(1)*, 3.

Ingold, P. V., Kleinmann, M., König, C. J., & Melchers, K. G. (2015). Shall we continue or stop disapproving of self-presentation? Evidence on impression management and faking in a selection context and their relation to job performance. *European Journal of Work and Organizational Psychology, 24*(3), 420-432. https://doi.org/10.1080/1359432X.2014.915215

Jung, C.G. (1923). *Psychological types.* Harcourt, Brace.

Kottke, J.L., Olson, D.A., & Shultz, K.S. (2010). The devil is in the details: A comment on what might this be? Rediscovering the Rorschach as a tool

for Personnel Selection in Organizations (Del Guidice, 2010). *Journal of Personality Assessment*, *92,*610-612. doi: 10.1080/00223891.2010.513313

Kuder, G. F. (1960). *Administrator's manual, Kuder Preference Record, Vocational, Form C*. Chicago, IL: Science Research Associates.

Lambert, C. E., Arbuckle, S. A., & Holden, R. R. (2016). The Marlowe–crowne social desirability scale outperforms the BIDR impression management scale for identifying fakers. *Journal of Research in Personality, 61,* 80–86. https://doi.org/10.1016/j.jrp.2016.02.004

Le, H., Oh, I.-S., Robbins, S. B., Ilies, R., Holland, E., & Westrick, P. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance. *Journal of Applied Psychology, 96*(1), 113–133. https://doi.org/10.1037/a0021016

Lee, P., & Joo, S-H (2021). A new investigation of fake resistance of a multidimensional forced-choice measure: An application of differential item/test functioning. *Personnel Assessment and Decisions*: 7(1). DOI: https://doi.org/10.25035/pad.2021.01.004

Lee, P., Joo, S.-H., & Lee, S. (2019). Examining stability of personality profile solutions between likert-type and multidimensional forced choice measure. *Personality and Individual Differences, 142,* 13–20. https://doi.org/10.1016/j.paid.2019.01.022

Lee, P., Lee, S., & Stark, S. (2018). Examining validity evidence for multidimensional forced choice measures with different scoring approaches. *Personality and Individual Differences, 123,* 229–235. https://doi.org/10.1016/j.paid.2017.11.031

Levashina, J., Weekley, J. A., Roulin, N., & Hauck, E. (2014). Using blatant extreme responding for detecting faking in high-stakes selection: Construct validity, relationship with general mental ability, and subgroup differences. *International Journal of Selection and Assessment, 22*(4), 371–383. https://doi.org/10.1111/ijsa.12084

MacCann, C. (2013). Instructed faking of the HEXACO reduces facet reliability and involves more Gc than Gf. *Personality and Individual Differences, 55(7),* 828-833. 10.1016/j.paid.2013.07.007.

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

McKay, A. S., Garcia, D. M., Clapper, J. P., & Shultz, K. S. (2018). The attentive and the careless: Examining the relationship between benevolent and malevolent personality traits with careless responding in online surveys. *Computers in Human Behavior, 84,* 295–303. https://doi.org/10.1016/j.chb.2018.03.007

Mersman, J. L., & Shultz, K. S. (1998). Individual differences in the ability to fake on personality measures. *Personality and Individual Differences, 24*(2), 217–227. https://doi.org/10.1016/S0191-8869(97)00160-8

Mill, J. (1984) High and low self-monitoring individuals: Their decoding skills and empathic expression. *Journal of Personality, 52,* 372-388

Monaro, M., Mazza, C., Colasanti, M., Ferracuti, S., Orrù, G., di Domenico, A., Sartori, G., & Roma, P. (2021). Detecting faking-good response style in personality questionnaires with four choice alternatives. *Psychological Research*, *85*(8), 3094–3107. https://doi.org/10.1007/s00426-020-01473-3

Moore, T. (1987) Personality tests are back. *Fortune*, 115, 76-81

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*(3), 683–729. https://doi.org/10.1111/j.1744-6570.2007.00089.x

Mueller-Hanson, R. A., Heggestad, E. D., & Thornton, G. C. III. (2006). Individual differences in impression management: An exploration of the psychological processes underlying faking. *Psychology Science, 48*(3), 288–312.

Murphy, K. R., & Dzieweczynski, J. L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18,* 343–358.

Murray, J. B. (1990). Review of research on the myers-briggs type indicator. *Perceptual and Motor Skills*, *70*(3_suppl), 1187–1202. https://doi.org/10.2466/pms.1990.70.3c.1187.

Oh, I-S., Kim, S & Van Iddekinge, C.H. (2015). Taking it to another level: Do personality-based human capital resources matter to firm performance?. *Journal of Applied Psychology, 100.* 10.1037/a0039052.

O'Neill, T. A., & Hastings, S. E. (2011). Explaining workplace deviance behavior with more than just the “Big Five”. *Personality and Individual Differences, 50*(2), 268–273. https://doi.org/10.1016/j.paid.2010.10.001

O'Neill, T. A., Lewis, R. J., Law, S. J., Larson, N., Hancock, S., Radan, J., Lee, N., & Carswell, J. J. (2017). Forced-choice pre-employment personality assessment: Construct validity and resistance to faking. *Personality and Individual Differences, 115,* 120–127. https://doi.org/10.1016/j.paid.2016.03.075

Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology, 60*(4), 995–1027. https://doi.org/10.1111/j.1744-6570.2007.00099.x

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

Paulhus, D. L. (1986). Self-Deception and impression management in test Responses. *Personality Assessment via Questionnaires* (pp. 143-165). http://dx.doi.org/10.1007/978-3-642-70751-3_8

Paulhus, D. L. (1991). Measurement and control of response bias. *Measures of Personality and Social Psychological Attitudes*,12-59.

Paulhus, D. L. (1998) Paulhus Deception Scales (PDS) user's manual. North Tonawanda, NY: *Multi-Health Systems.*

Pittenger, D. J. (1993). The utility of the myers-briggs type indicator. *Review of Educational Research*, 63(4), 467–488. https://doi.org/10.3102/00346543063004467

Robie, C., Tuzinski, K. A., & Bly, P. R. (2006). A survey of assessor beliefs and practices related to faking. *Journal of Managerial Psychology*, 21, 669–681. doi:10.1108/02683940610690204

Robson, S. M., Jones, A., & Abraham, J. (2008). Personality, faking, and convergent validity: A warning concerning warning statements. *Human Performance, 21*(1), 89–106. https://doi.org/10.1080/08959280701522155

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review, 16*(2), 155–180. https://doi.org/10.1016/j.hrmr.2006.03.004

Sackett, P. R., & Walmsley, P. T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science, 9*(5), 538–551. https://doi.org/10.1177/1745691614543972

Salgado, J.F. (1997). The five factor model of personality an job performance in the European community. *The Journal of applied psychology*. 82. 30-43. 10.1037/0021-9010.82.1.30.

Salgado, J. F. (2016). A theoretical model of psychometric effects of faking on assessment procedures: Empirical findings and implications for personality at work. *International Journal of Selection and Assessment*, 24(3), 209-228.

Schlenker, B. R., & Weigold, M. F. (1992). Interpersonal processes involving impression regulation and management. *Annual Review of Psychology, 43,* 133–168. https://doi.org/10.1146/annurev.ps.43.020192.001025

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.

Stabile, S.J. (2002). The use of personality tests as a hiring tool: Is the benefit worth the cost? *University of Pennsylvania Journal of Business Law, 4*, 279-313.

Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology, 60*(4), 967–993. https://doi.org/10.1111/j.1744-6570.2007.00098.x

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*(4), 703–742. https://doi.org/10.1111/j.1744-6570.1991.tb00696.x

Tett, R.P., & Simonet, D. (2021). Applicant faking on personality tests: Good or bad and why should we care? *Personnel Assessment and Decisions, 7,* 6-19. 10.25035/pad.2021.01.002.

Tonković G., Maša., Jerneić, Ž., & Galic, Z. (2012). The effects of faking on the construct validity of personality questionnaires: A direct faking measure approach. *Psihologijske Teme*. 21. 1-18.

Van de Mortel, Thea. (2008). Faking it: Social desirability response bias in self-report research. Thea van de Mortel. 25.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*(2), 197–210. https://doi.org/10.1177/00131649921969802

Vroom, V. H. (1964). *Work and motivation*. Wiley

Wetzel, E., & Greiff, S. (2018). The world beyond rating scales: Why we should think more carefully about the response format in questionnaires. *European Journal of Psychological Assessment, 34,* 1-5. 10.1027/1015-5759/a000469.

Wetzel, E., Frick, S., & Greiff, S. (2020). The multidimensional forced-choice format as an alternative for rating scales: Current state of the research. *European Journal of Psychological Assessment, 36,* 511-515. 10.1027/1015-5759/a000609.