

2011

## Sex and Racial Differences in Socially Desirable Responding

Kathryn G. Van Dixhorn  
*Wright State University*

Follow this and additional works at: [https://corescholar.libraries.wright.edu/etd\\_all](https://corescholar.libraries.wright.edu/etd_all)



Part of the [Industrial and Organizational Psychology Commons](#)

---

### Repository Citation

Van Dixhorn, Kathryn G., "Sex and Racial Differences in Socially Desirable Responding" (2011). *Browse all Theses and Dissertations*. 452.

[https://corescholar.libraries.wright.edu/etd\\_all/452](https://corescholar.libraries.wright.edu/etd_all/452)

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

SEX AND RACIAL DIFFERENCES IN SOCIALLY DESIRABLE RESPONDING

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Science

By

KATHRYN GABRIELLE VAN DIXHORN

B.A., University of California, Los Angeles, 2004

2011

Wright State University

WRIGHT STATE UNIVERSITY  
SCHOOL OF GRADUATE STUDIES

OCTOBER 1, 2010

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Kathryn Van Dixhorn ENTITLED Sex and Racial Differences in Socially Desirable Responding BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

---

Corey E. Miller, Ph.D.  
Thesis Director

---

Scott Wantamaniuk, Ph.D.  
Graduate Program Director

---

John Flach, Ph.D.  
Chair, Department of Psychology

Committee on Final Examination

---

Gary Burns, Ph.D.

---

Debra Steele-Johnson, Ph.D.

---

Corey E. Miller, Ph.D.

---

Andrew T. Hsu, Ph.D.  
Dean, Graduate School

## ABSTRACT

Van Dixhorn, Kathryn Gabrielle. M.S., Department of Psychology, Wright State University, 2011. Sex and Racial Differences in Socially Desirable Responding.

The purpose of this study was to examine the magnitude of sex and racial differences in faking behavior, specifically socially desirable responding, in a large ( $N = 295,517$ ), applied sample. Results indicated that females are engaging in more intentional socially desirable responding, whereas males are engaging in more inadvertent socially desirable responding. However, these differences are not likely to influence selection. Caucasians are displaying more intentional socially desirable responding than African Americans ( $d = 0.55$ ), Hispanics ( $d = 0.57$ ), and Asian Americans ( $d = 0.29$ ). Asian Americans - iii - engaged in less inadvertent socially desirable responding than Caucasians ( $d = 0.38$ ), African Americans ( $d = 0.44$ ), and Hispanics ( $d = 0.40$ ). In a simulated hiring situation, employing cut-off scores for both conscientiousness scores and socially desirable responding to eliminate fakers, only Asian Americans were less likely to be selected. This effect was larger as the selection ratio increased.

## TABLE OF CONTENTS

Introduction.....	1
Previous Faking Research.....	2
Previous Research on Sex Differences.....	8
Sex Differences in the Workplace.....	8
Sex Differences in Faking.....	10
Sex Socialization Differences.....	11
Personality and Group Differences.....	14
Previous Research on Racial Differences.....	15
Racial Differences in the Workplace.....	15
Racial Differences in Faking.....	18
Racial Socialization Differences.....	20
Method.....	23
Participants.....	23
Measures.....	23
Demographics.....	23
Personality.....	24
Social Desirability.....	24
Procedure.....	25
Results.....	26
Demographics.....	26
Tests of Hypotheses.....	26
Follow-Up Analyses.....	31

Discussion.....	34
Summary of Results and Implications .....	34
Limitations.....	40
Future Research.....	41
Appendix A.....	44
References.....	49

## LIST OF TABLES

Table	Page
1. Differences between Male and Female Social Desirability Scores.....	63
2. Differences between Caucasian and African American Social Desirability Scores.....	64
3. Differences between Caucasian and Hispanic Social Desirability Scores.....	65
4. Differences between Caucasian and Asian American Social Desirability Scores .....	66
5. Differences between African American and Hispanic Social Desirability Scores.....	67
6. Differences between African American and Asian American Social Desirability Scores .....	68
7. Differences between Hispanic and Asian American Social Desirability Scores.....	69
8. Overall Chi-Square values for Racial Groups with Varying Selection Ratios.....	70
9. Proportion of Observed vs. Expected Selected for Racial Groups with Varying Selection Ratios after a 95 <sup>th</sup> Percentile Cutoff Score Instituted.....	71
10. Proportion of Observed vs. Expected Selected for Sex Groups with Varying Selection Ratios after a 95 <sup>th</sup> Percentile Cutoff Score Instituted.....	72
11. Correlation Matrix of Outcome Variables, Sex, and Age.....	73
12. Adverse Impact Table for Male and Female Applicants .....	74
13. Adverse Impact Table for Caucasian, African American, Hispanic, and Asian	

American Applicants.....75

## INTRODUCTION

Over the past several decades, research regarding faking on non-cognitive measures in employment settings has become increasingly prevalent, underscoring its negative and expensive implications for selection. Faking poses many problems, including squandered resources, wasted time, and decreased predictive and criterion-related validity (e.g., Harold, McFarland, & Weekley, 2003; Rosse, Stecher, Miller, & Levin, 1998). Though precise figures regarding the costs are elusive, Komar, Brown, Komar, and Robie (2008) estimated that faking can cost an organization anywhere from \$520 to \$2,400 per applicant. Clearly, selecting the wrong person for a job due to faking can lead to drastic harm to an organization.

Practitioners commonly use non-cognitive measures in employment settings because they exhibit criterion validity (Weekley, Ployhart, & Harold, 2004), are relatively easy to administer, and are inexpensive. Additionally, there is evidence of less adverse impact than possible alternative measures, such as cognitive ability tests (e.g., Hough, Oswald, & Ployhart, 2001; Sackett & Wilk, 1994). However, previous researchers have established that different subgroups, including various sex, racial, and ethnic groups, significantly differed in their faking behavior in a variety of ways (e.g., Dudley, McFarland, Goodman, Hunt, & Sydell, 2005; Guadagno & Cialdini, 2007; Paulhus, 1991). To determine and investigate these differences, researchers often examine mean differences between various groups. However, practitioners commonly

consider only the most extreme responses to be faking, typically those responses in the ninety-fifth percentile or above. Thus, an investigation of these extreme responses, rather than the mean responses, is warranted. There is currently a lack of research investigating these differences in a large applied sample. The central purpose of this research is to examine faking behavior in a sample of actual applicants, thereby enabling me to not only examine mean differences, but also the more delicate nuances of both individual and group behavior at the most extreme levels.

### Previous Faking Research

Numerous researchers have shown that faking indeed can occur in the workplace (e.g., Hough, 1998; Van Iddekinge, Raymark, Eidson, & Putka, 2003; Weekley et al., 2003), which clearly presents a significant problem in regard to the selection of candidates. Viswesvaran and Ones' (1999) meta-analysis confirmed that respondents could alter their scores on integrity and personality test items by over half a standard deviation when instructed to do so. Furthermore, Viswesvaran and Ones (1999) found that measures of social desirability appeared to be the most sensitive to faking instructions, compared to personality measures. Actual applicants scored significantly higher than incumbents on all of the Big Five personality factors (Ross, Stecher, Miller, & Levin, 1998), suggesting, not surprisingly, that an applicant has a greater tendency to fake than an incumbent. Therefore, this behavior could result in the unfortunate selection of applicants adept at faking, not applicants high on the intended measured construct.

A great debate exists among practitioners on how to best operationalize faking. Inconsistencies proliferate in the research on faking, which obviously impedes, at best, and prevents, at worst, any ability to infer, apply, and generalize findings. Ryan and

Boyce (2006) have called for a clarification on the operationalization of faking, enabling researchers to better base their conclusions on more accurate information and higher quality data by accurately defining and testing for faking. Despite this, socially desirable responding is the most common way to detect faking (e.g., Barrick & Mount, 1996) on employment questionnaire items.

Socially desirable responding occurs when respondents provide answers that make themselves look good (Paulhus, 1984). Paulhus (1984) further divided social desirability into the two subgroups of impression management and self-deception. Impression management occurs when individuals deliberately alter their answers to create a positive impression, which is dependent on the perceived wants of a particular audience. This behavior is purposeful and is more indicative of traditional views of faking, meaning negative, deliberate, and dishonest. Impression management is a response bias that occurs when an individual “purposefully tailors answers to create the most positive social image” (Paulhus, 1991). An example of impression management would be a usually idle employee “looking busy” when his or her boss is present.

Self-deception happens when individuals present themselves in an overly positive light. This behavior is neither blatantly dishonest nor intentional. Individuals who engage in self-deception are not necessarily being dishonest though they might be incorrect. Though somewhat counterintuitive, researchers have found self-deception to be associated with adjustment, optimism, self-esteem, and general capability, making self-deception a desirable attribute for an employee to possess (e.g., McCrae & Costa, 1983; Paulhus, 1984). An example of self-deception would be an employee genuinely believing that he or she is more intelligent than the average person. This person might be

incorrect, but he or she is not being dishonest because he or she genuinely believes that he or she possesses above average intelligence.

High impression managers are motivated by social approval and are more sensitive to the context and the conditions of test administration. High impression managers are also more likely to be female whereas individuals with high self-deception tendencies are more likely to be male and possess inflated perceptions of their self-control, confidence, memory, and religiosity, among other attributes (Paulhus, 1984). Therefore, Paulhus's (1984) previous research informs my first hypothesis. My measures of social desirability include both impression management and self-deception assessments. Thus, I hypothesize that women will score significantly higher than men on my impression management measure.

Hypothesis 1: Women will score significantly higher than men on the impression management dimension of socially desirable responding.

Because of the subtle distinction between these two aspects of social desirability, researchers oftentimes do not differentiate between the two constructs. For example, Ones, Viswesvaran, and Reiss (1996) performed a meta-analysis and found that correcting for social desirability did not lead to an increase in validity. However, they only controlled for social desirability as a whole and did not distinguish between self-deception and impression management. Obviously, the differences between these two constructs are paramount in research on faking, as impression management is deliberate and self-deception is not. This question of whether to assess the sub-facets of social desirability or simply social desirability as a whole has led to further uncertainty on the generalizations of findings.

To control for faking within selection, practitioners typically deem respondents who score high, typically in the ninety-fifth percentile or above, on social desirability measures as fakers, thereby enabling researchers then to statistically control for social desirability and, by extension, eliminate fakers from consideration. Although this seems like a perfectly legitimate notion, Schmitt and Oswald (2006) found that this negatively affected criterion validity, as some social desirability measures are related to personality traits, which in turn are related to performance. Moreover, they found that correcting for faking did not influence performance in a significant way, particularly at the mean levels of performance. They found that the most important determinants of mean performance were test validity, rank ordering, and the selection ratio. Also, Sackett and Elligson (1997) found that adverse impact typically increases and the selection ratio increases.

Another method to operationalize faking is the Minnesota Multi-Phasic Personality Inventory (MMPI) faking subscales. Although the MMPI includes several subscales, three subscales purport to both measure social desirability and to expose fakers, whereas the remainder of the subscales assess substantive content. The L (lie) scale measures self-deception, as discussed previously. The items on this subscale consist of opportunities for a respondent to refute inconsequential flaws that most people do not have a problem admitting. An example of an L scale item might be, "I never lie, not even little white lies" (Miller, 2000). Like common self-deception scales, high scorers on the L scale are not intentionally lying. They simply have more of a tendency to view themselves in a positive light (MMPI manual, 1989). Researchers have criticized this scale both questioning its sophistication and ability to effectively identify fakers (Meehl & Hathaway, 1946), as well as its transparency (Vincent, Linsz, & Greene, 1966).

Because of these criticisms, the K and F scales were added. The content of the K (correction) scale consists of items that many people, but not all, would deny about themselves. The K scale is more akin to an impression management scale and closely resembles the social desirability scale I employed in my research to measure impression management. The F (infrequency) scale determines random or deviant response biases. Additional methods to conceptualize faking also include the CPI Good Impression Scale and the 16PF Faking Good and Faking Bad Scales, to name a few.

What these measures all have in common is that they all attempt to capture a rather elusive construct that essentially equates faking with social desirability. Because there is a lack of consensus regarding how to best operationalize faking, there is also a lack of consensus on a defined procedure for measuring faking (Lewak, Marks, & Nelson, 1990). There is a growing amount of research that suggests that faking is not really the same construct as social desirability at all. Social desirability scales in general do not correlate well with each other, demonstrating low convergent validity (Paulhus, 1991). Also, as mentioned earlier, social desirability often can be associated with positive personality attributes (McCrae & Costa, 1983; Paulhus, 1984). Therefore, if the top scorer is eliminated because they are faking, there is a good possibility that practitioners are inadvertently eliminating strong candidates for a job.

Recent researchers have put forth several hypotheses regarding faking, social desirability, and methods to prevent and control for its effects. However, many unanswered questions still remain. For example, McFarland and Ryan (2000) examined variance in faking and found that there were individual differences not only in how much people fake but also why they fake. These differences were influenced by the

respondents' personalities. If test-takers are consistently responding differently across the board, then the utility of socially desirable responding measures remains questionable. McFarland and Ryan (2001) further proposed a model of planned behavior that tested an individual's propensity to fake. What they found was that norms, attitudes, and perceived behavioral control all influenced an individual's intention to fake more than situational factors, opportunities to fake, and/or ability to fake. Clearly, these individual differences matter, independent of the situation or method in which faking is measured, perhaps further underscoring the importance of individual differences and subgroup differences that surface in response tendencies.

Thus, in light of the above information, perhaps these individual differences in faking have less to do with both opportunity to fake and situation-specific behavior and have more to do with individual factors. Ryan and Boyce (2006) have called for a moratorium on researching faking and its many measurements and instead called for a focus on further investigating test-takers' actual responses. There is some debate as to whether social desirability is a construct in and of itself or if it is simply an indicator of validity, measuring situation-dependent response styles (see Burns & Christiansen, 2006). However, my focus lies in the naturally occurring differences in response style, rather than in an investigation of the construct of social desirability itself. Social desirability scales do show some evidence of validity as deliberate situational response bias (Burns & Christiansen, 2006), but in the personnel selection arena, the validity is debatable (Schmitt & Chan, 1998). Although investigating the individual differences of ability, locus of control, self-monitoring, integrity, and personality differences and their effects

on faking is reasonably commonplace, inherent differences stemming from a person's sex, race, or ethnicity also warrant an exploration, particularly in an applied setting.

#### *Previous Research on Sex Differences*

*Sex Differences in the Workplace.* There is no shortage of research exemplifying the host of sex differences in the workplace (i.e., Guadagno & Cialdini, 2007; Segal, 1992; Stroh, Brett, & Reilly, 1992). Women comprise 49.8% of the workplace and economists anticipated that women, for the first time in history, would comprise more than half of the labor force by the end of 2009 (Gibbs, 2009). However, women only occupy one-third of management positions (Department of Labor Women's Bureau, 2005). Moreover, women run only 13 of Fortune 500 companies (less than 3%; Jones, 2009). Women are systematically underrepresented in science and math fields despite receiving more Bachelor's degrees than men every year since 1982 (Halpern et al., 2007), more Master's degrees, and an equal number of Doctoral, Law, and Medical degrees (Gibbs, 2009). Clearly, there are prevalent sex disparities in the workplace that do not stem from a lack of education or preparation.

Furthermore, men's salary progression is more accelerated, even when they are matched to female counterparts on both education and experience (Stroh et al, 1992). Women currently earn \$0.77 for every \$1.00 that men earn, and women's earnings fell two percent in 2008, twice as much as men's earnings (Gibbs, 2009). Thacker (1995) found that salary discrepancies are the most pronounced at the highest levels of a given organization. Naturally, individuals who negotiate their salaries at the onset of their employment earn more than those who do not (Gerhart & Rynes, 1991), so perhaps some of these discrepancies stem from sex differences in the salary negotiation process. Some

researchers have found that men have consistently higher pay expectations than women (e.g., Keys, 1985; Latham, Ostrowski, Pavlock, & Scott, 1987). This is true even after women underwent training in negotiation techniques (Stevens, Bavetta, & Gist, 1991). This might result from men expecting more advanced jobs and larger salaries long term than women when men had no information on which to base estimations (Kamen & Hartel, 1994). However, despite this possible explanation, Gerhart and Rynes (1991) found that both male and female MBA graduates used the same negotiation techniques, but that the men still ended up earning more. Moreover, even at the hiring level, Valian (1998) found that small differences in salary can compound to much more substantial differences later on. Thus, if there are small differences within the selection process, those small discrepancies can intensify to larger differences in who gets hired or promoted.

Schmitt, Clause, and Paulakos (1996) investigated adverse impact in ability tests and assessments between males and females. They found that women scored slightly higher in general/cognitive ability tests ( $d = 0.09$ ), personality tests ( $d = 0.07$ ), and accomplishments record ( $d = 0.03$ ) and scored one-quarter standard deviation higher in verbal ability ( $d = 0.24$ ). Men, on the other hand, scored significantly higher in job sample/job knowledge ( $d = 0.38$ ), math ability ( $d = 0.27$ ), and clerical/speed/accuracy ( $d = 0.59$ ). Generally speaking,  $d$  is considered small if it is equal to 0.2 or less, moderate at 0.4, and large if it is equal to 0.8 or larger (Cohen, 1988). These results underscore the notion that, although there are some differences in skills between men and women generally, there is not a significant lack of skills in either group that would justify such exorbitant disparities in position and compensation.

Because of these drastic differences, investigating how these differences play out is a chief concern in selection research. Though Mills (2002) attributed these differences to outright sexism, other contributing factors such as sex-specific behavior might be at work (Guadagno & Cialdini, 2007). Perhaps sex differences in faking behavior might be an additive differential factor not only in selection but also in overall workplace conduct.

*Sex Differences in Faking.* Previous research on sex differences in faking, like much faking research, has been inconsistent. As discussed previously, some of these distinctions might stem from the researchers' definitions of faking, as some include all forms of social desirability (both impression management and self-deception), and some are more refined to include exclusively impression management. Cialdini (1989) and Rosenfeld et al. (1995) both examined specific sex differences in impression management. Although these researchers found either mixed results or no significant sex differences for many behaviors, several behaviors have produced consistent sex dissimilarities.

Researchers using self-reports have found men are more likely to use a wider range of social desirability techniques (e.g., DuBrin, 1991; Karsten, 1994). Males tend to use more self-enhancement (DuBrin, 1994; Lee, Quigley, Nesler, Corbett, & Tedeschi, 1999; Strutton, Pelton, & Lumpkin, 1994; Tannen, 1994), which is the tendency to maintain a positive self-regard under all circumstances (Dunning, 1999). This finding coincides with Paulhus's (1984) finding that men engage in more self-deception than women. Researchers have found also that men engage in more favor-doing or exchange than women (e.g., DuBrin, 1991; Higgins & Snyder, 1989; Strutton et al., 1995), acting with entitlement (Lee et al., 1999), using association with others to gain favor (Lee et al.,

1999), sandbagging or faking inability (Gibson & Sachau, 2000), self-handicapping (Dietrich, 1995; Hirt, McCrea, & Kimble, 2000; Shepperd & Arkin, 1989), blasting or derogating their competitors (DuBrin, 1991), and intimidation (DuBrin, 1991; Lee et al., 1999; Offermann & Schrier, 1985; Smith et al., 1990).

Women, on the other hand, are more likely to engage in the practices of opinion conformity (DuBrin, 1991), modesty (Heatherington, Burns, & Gustafson, 1998; Jones & Wortman, 1973), flattery (Eagly & Carli, 1981; Tannen, 1994), excuses (Konovsky & Jaster, 1989), hedging (Carli, 1990), apologizing (Lee et al., 1999), and supplication (Tannen, 1994). They also are more likely to use acquiescence, particularly in salary negotiation procedures (Mainiero, 1986). Overall, the summation of these findings reveals that men attract attention to their positive actions and are more active in their social desirability behavior, unintentionally or not, whereas women tend to detract from themselves and act more passively.

Why then are these differences present? Why are masculine qualities rewarded with higher salaries and promotions in the workplace, whereas feminine qualities undervalued with slower career progression? Is there a difference in how our society values these qualities and, if so, how do these dissimilarities manifest themselves in selection assessments? This disparity could potentially stem from sex differences in socialization.

*Sex Socialization Differences.* Although biological discussion of sex socialization differences is beyond the scope of this research, there are inherent differences in ways that men and women are taught regarding acceptable behavior for their sex. Social roles theory suggests that men and women possess different positions in society, translating to

behavioral differences in the workplace. Eagly (1987) discovered that men typically possess agenic qualities (i.e., assertive, independent, controlling), whereas women normally behave in a more communal fashion (i.e., emotionally expressive, sensitive, concerned for others). Therefore, if a male behaves in an assertive way, his coworkers view his behavior as congruent with sex roles. Conversely, if a female acts in an assertive manner, it is a violation of typically female roles. That female can suffer also a backlash effect, meaning her peers might view her as competent, yet do not like her personally. This theory is in congruence with the different impression management tactics that men (more agenic behaviors) and women (more communal) use. Similarly, socialization differences might affect what kind of job, workplace, or discipline a person is attracted to.

Schneider's (1987) Attraction, Selection, and Attrition (ASA) framework states that the kinds of individuals that are initially attracted to a given organization are thereby selected. Thus, those individuals begin to define and shape an organization, thereby making the employees more homogenous. For example, many women might have been attracted to elementary education, as it affords some flexibility (shorter school days, summer vacation) that is desirable for many women. This high volume of women are then selected for the jobs and then stay with the job longer than someone who, perhaps, is looking to further their career or interested in a competitive salary. The end result is that by nature of the applicants seeking a flexible environment, elementary education is now known as a flexible and family-friendly field. This is independent of the actual job attributes, but is instead defined by the people in the field. These socialization differences all provide support for the notion that faking is not a situation-specific

activity. Some people, and perhaps some groups, are just more inclined to fake in certain ways than others.

Further socialization differences involve how some intellectual pursuits are either reinforced or discouraged according to sex. A host of researchers (e.g., Halpern & LaMay, 2000; Johnson, Carothers, & Deary, 2009) have investigated sex differences in intellectual ability. Overall, males have a wider range of intelligence, with more males appearing at both the very high and very low ends of the spectrum. Turkheimer and Halpern (2009) found evidence that the X chromosome contributes to some of the disparities at the lower end of the spectrum, potentially resulting in the overrepresentation of males with mental retardation. However, this was not the case for the upper range differences. Benbow and Stanley (1983) found remarkable differences when it came to male and female seventh graders and their mathematical ability. The mean differences were not radically different between males and females. However, at the extremes, there were marked differences in the scores. At the highest levels, boys outnumbered girls 13:1. As to why this difference existed, Benbow and Stanley (1983) suggested that social factors occurring prior to adolescence had an influence. Halpern et al. (2007) also found similar results in that mean differences between sexes were not striking. However, at the extreme levels, or the ninety-fifth percentile, the differences were much more prominent.

Perhaps this could also be the case with social desirability. With social desirability measures, fakers are the most extreme respondents. Researchers often use a cut-off at the ninety-fifth percentile (Ellingson, Sackett, & Hough, 1999; Hough et al., 1990). Thus, although there is a small mean difference for men and women in social

desirability, it is conceivable that there is a larger, more marked, difference at the extremes that I will examine.

### *Personality and Group Differences*

Social desirability scales have a positive relationship with several personality factors, such as conscientiousness and emotional stability, as well as with traits such as adjustment and integrity (Ones et al., 1996). This relationship suggests that social desirability, or at least some aspect of it, is related to differences in personality. Because self-deception is largely unconscious and is the product of an excessively positive self-concept, self-deception represents a genuine component of a respondent's personality and is thus an indication of true variance in personality. On the other hand, impression management is deliberate and changes in response to the respondent's audience. As previously discussed, men are more likely to engage in self-deception whereas women are more likely to use impression management techniques (Paulhus, 1991).

Because practitioners often eliminate candidates that score in the top echelon of some personality dimensions, namely Conscientiousness, a brief examination of group differences in personality is warranted. As far as personality differences, women are more likely to be higher in neuroticism, agreeableness, and openness to feelings, whereas men are typically higher in openness to ideas and assertiveness (Costa, Terracciano, & McCrae, 2001). These dissimilarities correspond to the sex variations in faking behavior, as discussed above, and could potentially influence differences in individuals' faking behavior. Additionally, Schmitt, Realo, Voracek, and Allik (2008) found that women typically scored higher on conscientiousness measures across cultures.

There is some debate regarding racial and ethnic subgroup differences in personality. Dudley et al. (2005) did not uncover significant racial and ethnic differences in regard to personality measures, highlighting the notion that correcting personality tests for social desirability only decreases the validity of personality tests for minority individuals (Dudley et al., 2005). However, Hough et al. (2001) found meta-analytic evidence that some small subgroup differences in that minorities scored slightly lower than Caucasians in Conscientiousness measures. Hispanics and Asian Americans scored higher than Caucasians on the Achievement facet of conscientiousness yet scored lower on the Dependability facet of conscientiousness. African Americans scored lower than Caucasians on the Dependability facet of conscientiousness but were similar on the Achievement facet. All differences, however, were minimal and not likely to have practical implications. In addition to social desirability measures, participants also completed a Conscientiousness inventory as part of their assessment. Thus, I was able to further investigate personality variation in conscientiousness within my subgroups.

#### *Previous Research on Racial Differences*

*Racial Differences in the Workplace.* Research on racial and ethnic differences in the workplace and performance is not nearly as prevalent as research on sex differences. However, there were several racial disparities that were consistent within several fields. Some researchers have found that African Americans were more likely than their Caucasian counterparts to be tardy or absent (e.g., Avery, McKay, Wilson, & Tonidandel, 2007; McKay & McDaniel, 2006; Roth, Huffcutt, & Bobko, 2003). However, this difference was more pronounced when African Americans believed that their organizations placed a low value on its commitment to diversity (Avery et al., 2007).

Additionally, objective measures were more likely to produce larger racial differences than subjective measures (Roth et al., 2003) although there is some debate on this issue (see Ford, Kraiger, & Schetchman, 1986).

In a meta-analysis of subgroup differences, significant black-white differences were found for performance utilizing various criteria. Roth et al., (2003) found significant mean differences, with African Americans scoring lower, in regards to quality and quantity ratings, job knowledge, absenteeism, work samples, promotion, and on-the-job training. Hispanic-Caucasian differences were found as well for all criteria although the differences were not as large (Roth et al., 2003). College admissions tests, although not directly relevant to an employment setting, offer insight in the parallels to selection tests. Overall, white men score the highest on these tests in the United States. In tests of mathematics, Asian American men were the top-scoring group, whereas African Americans and Hispanics obtain lower average scores (Halpern, 2000). This can lead to fewer individuals in those groups accepted to top colleges, being less likely to obtain a good job, etc., causing these differences to compound into significant differences over time.

Similarly, McKay and McDaniel (2006) conducted a meta-analysis and found that the overall mean Black-White difference was approximately one-quarter standard deviation for performance though this effect has been less pronounced for research conducted more recently. Dean, Roth, and Bobko (2008) found a similar pattern when examining racial differences in assessment center ratings. However, criterion type and cognitive loading of criteria moderated this relationship in that the difference and cognitive loading had a positive relationship. Schmitt et al. (1996) found the greatest

amount of adverse impact between Caucasians and African Americans occurring in general/cognitive ability tests ( $d = 0.83$ ) and spatial ability ( $d = 0.66$ ), whereas the lowest amount of adverse impact resulted from personality tests ( $d = 0.09$ ), manual dexterity ( $d = 0.14$ ), and clerical/speed/accuracy ( $d = 0.15$ ). In interviews, some evidence of subgroup differences emerges, with African Americans and Hispanics applicants receiving ratings, again, approximately one-quarter standard deviation lower than Caucasian applicants. This difference is influenced and ameliorated by several factors including the use of structured interviews. Additionally, this discrepancy is smaller with more complex work and greater with a larger minority pool of applicants (Huffcutt & Roth, 1998).

Some explanations postulated for these differences include mean racial differences in cognitive ability, rating bias, opportunity bias, and rating purpose (McKay & McDaniel, 2006). For Hispanics, as with Asian Americans, it is oftentimes difficult to generalize findings to this ethnic group, as it is comprised of individuals from many different countries and there are obvious cultural differences within these groups. That being said, in general Hispanic Americans and Caucasian differences typically lie somewhere in between the African American and Caucasian difference. McKay, Avery, and Morris (2008) also found mean differences between Hispanics and Caucasians in sales performance, but the diversity climate moderated the relationship.

Schmitt et al. (1996) found no difference between Hispanics and Caucasians on measures of job sample/job knowledge tests. However, Schmitt et al., (1996) found evidence of moderate mean differences between Hispanics and Caucasians on measures of General/Cognitive ability ( $d = 0.48$ ), math ability ( $d = 0.45$ ), verbal ability ( $d = 0.58$ ),

and interview ratings ( $d = 0.19$ ), with Caucasians scoring higher. On a test battery of verbal, induction, deduction, numbers, and judgment, Asian Americans scored higher on induction, deduction, and numbers than on judgment and verbal (Wing, 1980).

There is a dearth of extensive research investigating Asian American performance or work-related differences. In one of the largest meta-analyses on racial/ethnic differences in performance conducted, Roth et al., (2003) focused on African American and Hispanic differences as there was not sufficient information available assessing Asian Americans. It should also be noted that mean differences are often reported when examining group variation. None of the above previous research reported what group differences appeared to be at the most extreme levels, so the practical implication of these differences remains questionable.

*Racial Differences in Faking.* Dudley et al. (2005) found that there were significant racial differences on social desirability scales. The social desirability scales used by Dudley et al. (2005) were similar to the Balanced Inventory of Desirable Responding, testing for both impression management and self-deception and measuring the extent to which a participant engages in undesirable, yet common behavior. On average, Caucasians scored lower on social desirability scales than African Americans, Asian Americans, and Hispanics, suggesting that Caucasians are faking less than other racial groups. The largest difference in social desirability mean group differences was between Caucasians and Asian Americans, with Asian Americans scoring higher on social desirability. Although there are significant subgroup differences within social desirability measures, there is a possibility that minorities might not want to admit bad behaviors, yet might respond honestly to personality measures. Dudley et al. (2005)

suggested that test administrators should consider the construct validity of social desirability scales across diverse groups before administration. Thus, in light of this information, I hypothesize several subgroup differences in regard to social desirability.

Hypothesis 2: African Americans will score significantly higher than Caucasians on social desirability measures.

Hypothesis 3: Hispanics will score significantly higher than Caucasians on social desirability measures.

Hypothesis 4: Asian Americans will score significantly higher than Caucasians on social desirability measures.

Hypothesis 5: Hispanics will score significantly higher than African Americans on social desirability measures.

Hypothesis 6: Asian Americans will score significantly higher than African Americans on social desirability measures.

Hypothesis 7: Asian Americans will score significantly higher than Hispanics on social desirability measures.

Although Dudley et al. (2005) informs the majority of these hypotheses, there are several important contributions that my research will make that differ from Dudley et al. (2005). Dudley et al. (2005) also used an applied sample, but the sample was significantly smaller than mine. Dudley et al. (2005) performed three studies, with participants of 1,036 sales-job applicants and 534 incumbents (clerical and customer service), but only included 207 African Americans and 148 Hispanics. Asian Americans were not included in that study. Moreover, Dudley et al. (2005) examined mean differences between subgroups on dimensions of personality and social desirability.

Although I investigated mean differences as well, I also examined extreme responses to better ascertain how eliminating high scoring individuals actually manifests in a hiring situation using cut-off scores. If practitioners eliminate individuals who score in the ninety-fifth percentile on social desirability measures from hiring consideration, then this difference translates to practitioners selecting fewer minorities. This could potentially result in adverse impact by increasing subgroup differences, thereby negating one of the chief reasons to use a non-cognitive measure in the first place (i.e., the reduction of subgroup differences). Hypothesis 8 will test if this situation potentially occurs.

Hypothesis 8: Minorities and women would be less likely to be selected based on a Conscientiousness and Social Desirability cut-off score.

*Racial Socialization Differences.* No one really knows why these consistent racial differences persist, although Neisser et al. (1996) posited some hypotheses as to why these differences exist. Socioeconomic factors might have an influence in discrepancies, as African Americans typically have lower incomes than Caucasians. This results, potentially, in poorer nutrition, fewer resources, and lessened prenatal care. However, when African Americans were matched to their Caucasian counterparts on income, this difference still persisted. Ogbu (1978) suggested that individuals belonging to “caste-like” groups did not possess “effort optimism”, which is basically the belief that hard work will be rewarded. In addition, Ogbu (1978) suggested that there are elements of African American culture that might contribute to discrepancies as well as a rather recent legacy of discrimination. Moreover, some researchers have found (e.g., Collins & Gleaves, 1998; Frank, 1992) that African Americans tend to restrict the amount of personal information that they provide to strangers and are more suspicious of testing

situations. Therefore, it is logical to suggest that a restriction of information could be construed as faking on some measures.

Schmitt, Chan, DeShon, Clause, and Delbridge (1997) found that the perceived face validity of a test affected the test-takers' motivations to perform. African Americans found standardized tests less face valid than their Caucasian counterparts, which could explain the differences in performance. Additionally, Chan and Schmitt (1997) found less evidence of adverse impact in a video-based situational judgment test, as opposed to a pencil/paper test, although the differences were small.

Children in China and Japan are more familiar with math and typically perform better on tests of cognitive ability, yet their IQs are not dissimilar to Caucasians (Neisser et al, 1996). Herrnstein and Murray (1994) found that the average IQ differential between Asian Americans and Caucasians was only zero to ten points. Neisser et al. (1996) suggested that the difference in academic ability between Asian Americans and Caucasians might be due to cultural emphases. Asian-American culture places a higher importance on spatial ability, which might contribute to Asian Americans' more accomplished math performance. Parental expectations also might have more of an influence with Asian Americans than they do with other racial groups (Hieshima & Schnieder, 1994). Most Asian countries are collectivist societies, placing more importance on groups rather than the individual. Thus, Asian Americans might have a greater propensity to conform to group norms. However, because all of my participants live within the same country, culture might have less of an impact.

Neisser et al. (1996) suggested that cultural differences between Hispanics and Caucasians might be due to a language barrier issue. Ramos (1981) administered a

mostly non-verbal employment test battery to bilingual Hispanic participants and found that those individuals who received instructions in Spanish performed better than those who received their instructions in English. Ross and Mirowsky (1984) found that some Hispanic individuals responded to tests with a higher level of acquiescence. Furthermore, they also have a tendency to distrust the fairness of both career advancement opportunities (US Merit Systems Protection Board, 1997) and of tests in general (Fouad, 1994). Again, this propensity to be suspicious of tests could easily be misconstrued as skewed social desirability. Regardless, any group differences are a function of not only shared group characteristics, but criterion characteristics as well.

In light of the above information, an examination of an applied sample is paramount to discern furthered knowledge regarding subgroup differences. The current study fills a void in the current literature, with its major strength lying in the applied setting and, because of the large number of participants, the ability to investigate the behavior of many individuals at the extreme levels of socially desirable responding.

## Method

### *Participants*

Participants included 295,517 job applicant and incumbent employees who completed a web-based assessment. Data was compiled over the past 15 years via a large private management consultant organization located in the Midwestern United States. Participants completed a variety of measures in a test battery as part of either the application process or their ongoing professional development. To be included in this study, participants had to be fluent English speakers, had to indicate the United States as their country of origin, and had to solely identify themselves as one of the indicated races (Caucasian, African American, Hispanic, or Asian American). Although it is a worthwhile notion to investigate outcomes on an international and multi-cultural level, it is beyond the scope of this research. Furthermore, Dean et al. (2009) found that cross-cultural variance in response distortion did not account for incremental variance. Therefore, non-Americans, American expatriates, or multi-racial individuals were not included as participants.

### *Measures*

*Demographics.* A measure was administered to all participants that requested them to voluntarily provide their age, sex, race, and years of professional experience in a job comparable to the position they were either currently working in or applying for.

*Personality.* Conscientiousness was measured via a 19-item measure that has an alpha of .694. This measure correlates with the IPIP conscientiousness measure ( $r = .57$  (129),  $\alpha = .78$ ; Culbertson, 2007). This approach was developed on the basis of prior research as a proxy for conscientiousness using applicants who completed the IPIP measure of conscientiousness as well as a proprietary pool of items geared toward a sales population (Culbertson, 2007). Example items include “In a test, I am generally one of the first to finish.” and “Many people see me as a perfectionist.”

*Social desirability.* To measure social desirability, I used two of the consulting firm’s social desirability scales that correspond to the MMPI’s K scale and the MMPI’s F scale. To measure intentional social desirability, or impression management, a nine-item validity scale was used ( $\alpha = .696$ ). Again, this scale measures the impression management aspect of social desirability, thus responding in a socially desirable manner was intentional. Given that prior research has found significant sex and racial differences in impression management behavior, this informs my hypotheses. Example items of this scale include “Criticism never bothers me” and “Sometimes I wish I could take back what I say”.

I also used the consulting firm’s social desirability scale that corresponds to the MMPI’s F scale, or the unguardedness scale. This scale measures unintentional response distortion, also known as self-deception (Dean et al., 2009). This measure consists of 13 items ( $\alpha = .699$ ). This scale measures self-deception by gauging an individual’s propensity to generate deviant responses. An example item is “I can express my feelings easily.”

All scales were scored dichotomously, with the participants indicating if the given statement applied to them (“true”) or did not (“false”). Providing a “true” response (or “false” in the case of reverse-scored items) indicated that the individual was responding in a socially desirable manner. Thus, a higher score indicated more faking was occurring. Taken together, the K and F scale and their respective consulting firm derivations tap both conscious and unconscious motivated response distortion, as suggested by Paulhus (1991).

### *Procedure*

Participants completed all of the measures included in this study via a web-based questionnaire, in fulfillment of the application of a client company. All participants were assigned a login identification and were able to complete the assessment from any computer at their leisure. After logging on, participants were given one of the consulting firm’s standard set of instructions, as shown in Appendix A (Tristan, 2008). Once the participants completed the evaluation, the consulting firm appraised them and made a recommendation to their client company. Individual firms’ instructions were dispensed before the completion of my measures.

Because this is an applied setting, the very nature of this research’s strength lies in the participants believing, rightly so, that the assessment they completed was for an actual selection situation. Thus, informed consent was not collected. However, given that all participants completed these measures anyway, regardless of this research, the participants were not under any additional stress or harm. Thus, a lack of informed consent was not incongruent with ethical standards.

## Results

### *Demographics*

Because not every participant indicated both race and gender, some of the individual breakdowns of participants do not necessarily add up to the total participants. The sample consisted of 90,962 females (30.8%) and 160,489 males (54.3%). The remaining participants either did not indicate their sex by leaving the question blank or did not wish to respond. As far as the participants' races, there were 200,222 Caucasians (67.8%), 20,377 African Americans (6.9%), 12,253 Hispanics (4.1%), and 6,898 Asian Americans (2.3%). The remaining participants either did not indicate their race, left the question blank, or belonged to a different racial category (i.e., American Indian, Indian, Middle Eastern, Native Hawaiian or Pacific Islander, or were two or more racial categories). Applicants comprised 50.09%, whereas incumbents comprised 49.91% of participants. The average age for 31.9 years old and the average tenure of all participants was 1.83 years. The sample had a majority of participants in sales positions (59.4%) and management positions (7.5%). "Unknown" or "None" comprised 6.3% of the sample.

### *Tests of Hypotheses*

Hypothesis 1 predicted that women would score significantly higher than men on the impression management measure of social desirability. In order to test this hypothesis, an independent-measures t-test was used to compare men versus women mean scores on the social desirability measures. As shown in Table 1, females did

indeed score significantly higher than men on the social desirability measure focusing on impression management,  $t(247392) = -4.59, p < .001$ . Conversely, men scored significantly higher than women on the self-deception measure of social desirability,  $t(247392) = 33.90, p < .001$ , confirming Paulhus's (1991) findings.

Additionally, because my sample was so large, Cohen's  $d$  was calculated to determine effect size. For impression management,  $d = -0.02$ , indicating women scored higher and therefore engaging in more impression management. For self-deception,  $d = 0.14$ , indicated men engaged in more self-deception. A one-way ANOVA confirmed significant group differences in impression management,  $F(1, 247392) = 21.63, p < .001$  and self-deception behavior,  $F(1, 247392) = 1179.07, p < .001$ . Thus, hypothesis 1 was supported.

Hypothesis 2 predicted that African Americans would score significantly higher than Caucasians on measures of social desirability. In order to test this hypothesis, an independent-measures  $t$ -test was used to compare African American versus Caucasian mean scores on the social desirability measures. As shown in Table 2, African Americans did score significantly higher than Caucasians on the self-deception measure,  $t(216864) = -6.45, p < .001$ . However, Caucasians actually scored significantly higher than African Americans on the impression management facet of the measures,  $t(216864) = 74.28, p < .001$ , which is contradictory to Dudley et al.'s (2005) findings. For self-deception, Cohen's  $d = -0.05$ , indicating that African Americans were engaging in more self-deception. For the impression management scale,  $d = 0.55$ , indicating that Caucasians were engaging in considerably more impression management than African Americans, by more than half a standard deviation. A one-way ANOVA confirmed these

significant group differences in both impression management,  $F(1, 216864) = 5546.56$ ,  $p < .001$  and self-deception,  $F(1, 216864) = 35.73$ ,  $p < .001$ ). Thus, hypothesis 2 was partially supported.

Hypothesis 3 predicted that Hispanics would score significantly higher than Caucasians on measures of social desirability. In order to test this hypothesis, an independent-measures t-test was used to compare Hispanic versus Caucasian mean scores on the social desirability measures. As shown in Table 3, Hispanics did score significantly higher than Caucasians on the self-deception measure,  $t(208887) = -2.59$ ,  $p < .01$ . However, Caucasians actually scored significantly higher than Hispanics on the impression management facet of the measures,  $t(208887) = 60.21$ ,  $p < .001$ , which is again contradictory to Dudley et al.'s (2005) findings. For self-deception,  $d = -0.02$ , indicating that Hispanics were engaging in more self-deception. However, for the impression management scale,  $d = 0.57$ , indicating that Caucasians, once again, were engaging in significantly more impression management than Hispanics by more than half a standard deviation. A one-way ANOVA also indicated significant group differences for both impression management ( $F(1, 208887) = 3729.49$ ,  $p < .001$ ) and self-deception ( $F(1, 208887) = 7.07$ ,  $p < .01$ ). Thus, hypothesis 3 was partially supported.

Hypothesis 4 predicted that Asian Americans would score significantly higher than Caucasians on measures of social desirability. In order to test this hypothesis, an independent-measures t-test was used to compare Asian American versus Caucasian mean scores on the social desirability measures. As shown in Table 4, Caucasians scored significantly higher than Asian Americans on both the self-deception measure,  $t(203557) = 28.85$ ,  $p < .01$ , and the impression management measure,  $t(203557) = 23.18$ ,  $p < .001$ .

Thus, in both facets of social desirability, Caucasians were faking more than Asian Americans. These results are also contrary to previous research. For self-deception,  $d = 0.38$ , indicating that Caucasians were engaging in more self-deception. For the impression management scale,  $d = 0.29$ , indicating that Caucasians were engaging in more impression management than Asian Americans. A one-way ANOVA indicated significant group differences in both impression management,  $F(1, 203557) = 605.47$ ,  $p < .001$  and self-deception,  $F(1, 203557) = 1172.48$ ,  $p < .001$ . Thus, hypothesis 4 was not supported.

Hypothesis 5 predicted that Hispanics would score significantly higher than African Americans on measures of social desirability. In order to test this hypothesis, an independent-measures t-test was used to compare Hispanic versus African American mean scores on the social desirability measures. As shown in Table 5, there were not significant differences between Hispanics and African Americans on measures of social desirability. For self-deception,  $d = 0.02$ , indicating that African Americans engaged in more self-deception. For impression management,  $d = 0.02$ , again indicating that African Americans engaged in more impression management. A one-way ANOVA did not yield significant group differences. Thus, hypothesis 5 was not supported.

Hypothesis 6 predicted that Asian Americans would score significantly higher than African Americans on measures of social desirability. In order to test this hypothesis, an independent-measures t-test was used to compare Asian American versus African American mean scores on the social desirability measures. As shown in Table 6, African Americans scored significantly higher than Asian Americans on the self-deception measure,  $t(27079) = 29.47$ ,  $p < .01$ . However, Asian Americans scored

significantly higher than African Americans on the impression management facet of the measures,  $t(27079) = -16.93, p < .001$ . For self-deception,  $d = 0.44$ , indicating that African Americans were engaging in much more self-deception. However, for the impression management scale,  $d = -0.24$ , indicating that Asian Americans were engaging in significantly more impression management than African Americans. A one-way ANOVA confirmed significant group differences in both impression management,  $F(1, 27079) = 304.19, p < .001$  and self-deception,  $F(1, 27079) = 1135.55, p < .001$ . Thus, hypothesis 6 was partially supported.

Hypothesis 7 predicted that Asian Americans would score significantly higher than Hispanics on measures of social desirability. In order to test this hypothesis, an independent-measures t-test was used to compare Asian American versus Hispanic mean scores on the social desirability measures. As shown in Table 7, Hispanics scored significantly higher than Asian Americans on the self-deception measure,  $t(19102) = 25.99, p < .01$ . However, Asian Americans did score significantly higher than Hispanics on the impression management facet of the measures,  $t(19102) = -16.94, p < .001$ . For self-deception,  $d = 0.40$ , indicating that Hispanics were engaging in much more self-deception. However, for the impression management scale,  $d = -0.26$ , indicating that Asian Americans were engaging in more impression management than Hispanics. A one-way ANOVA indicated significant group differences in both impression management,  $F(1, 19102) = 294.68, p < .001$  and self-deception,  $F(1, 19102) = 736.00, p < .001$ . Thus, hypothesis 7 was partially supported.

Hypothesis 8 predicted that minorities and women would be less likely to be selected on the basis of a conscientiousness and social desirability score. To decipher

whether these differences manifest themselves in an actual hiring setting, I conducted a Chi-Square analysis by comparing the number of individuals expected to be selected and the number of individuals still in consideration based on their scores on social desirability and conscientiousness measures and the use of a ninety-fifth percentile cutoff score for both conscientiousness and social desirability. From this simulation and the resulting statistics, I determined if there was a significant difference between the expected number of minorities and women hired compared to the actual number of minorities and women hired. For this analysis, I included only applicant participants.

Additionally, I employed various selection ratios to examine if these differences were dependent on how selective a hiring situation was. This enabled me to discern whether minorities and women are less likely to be hired if a cut-off score is used for social desirability and/or conscientiousness. When the top five percent of both Conscientiousness scorers and Social Desirability scorers were eliminated, there were not significant results between racial groups overall. However, when I instituted several different selection ratios, there were significant differences. The Chi-Square values for the various racial groups are reported in Table 8 within the different selection ratios. A Chi-Square analysis of male versus female applicants yielded no significant results. The proportions of observed versus expected individuals in various racial groups are shown in Table 9, whereas the proportions of observed versus expected individuals in different sex groups are shown in Table 10. The results indicate that Asian Americans were less likely to be selected, but that African Americans, Hispanics, and females were no less likely to be selected. Thus, Hypothesis 8 was partially supported.

#### *Follow-Up Analyses*

Because of the large effect on potential hiring for Asian Americans, I performed follow-up analyses to further investigate if the effect was due to race or other factors, such as age or sex. Pearson's correlation coefficients were calculated to assess whether age or sex were significantly correlated with measures of conscientiousness, self-deception, and impression management. Asian American participants were younger than the other participants, which might account for some of the differences in responses on the various measures. The Asian American average age was 30.15 ( $SD = 8.29$ ), whereas the average age was 36.00 ( $SD = 10.11$ ) for Caucasians, 33.18 ( $SD = 8.75$ ) for African Americans, and 33.03 ( $SD = 8.79$ ) for Hispanics. As Table 11 indicates, age was significantly and negatively correlated ( $r = -.09, p < .001$ ) with conscientiousness. Age was significantly and positively correlated with both impression management ( $r = .14, p < .001$ ) and self-deception ( $r = .03, p < .001$ ). Using hierarchical regression, I found that race had an incremental effect above and beyond sex and age on both impression management ( $\Delta R^2 = .002, p < .001$ ) and conscientiousness ( $\Delta R^2 = .008, p < .001$ ), but did not offer any additional variance for self-deception ( $\Delta R^2 = .000, p < .001$ ).

Sex was found to be negatively and significantly correlated with both conscientiousness ( $r = -.01, p < .001$ ) and self-deception ( $r = -.07, p < .001$ ) in that men scored higher on both measures. For impression management, sex was significantly and positively correlated ( $r = .01, p < .001$ ), in that women scored higher on this measure. Sex did not offer additional variance for conscientiousness ( $\Delta R^2 = .000, p < .001$ ) or impression management ( $\Delta R^2 = .000, p < .001$ ), but did for self-deception ( $\Delta R^2 = .006, p < .001$ ). After controlling for sex and age, effects for race were much less pronounced and likely no longer practically significant for conscientiousness ( $\Delta R^2 = .001, p < .001$ ),

self-deception ( $\Delta R^2 = .002$ ,  $p < .001$ ), or impression management ( $\Delta R^2 = .014$ ,  $p < .001$ ).

From a practitioner's viewpoint, social desirability measures appeared to have adverse impact in regards to the Asian subgroup. However, subsequent analyses revealed that the effect was largely due to subgroup differences in age, rather than racial or cultural differences between racial groups. Perhaps the conclusion that I can most safely draw is that researchers or practitioners looking at subgroup differences need to control for demographic variables such as age and sex as differences that appeared to be due to racial or cultural differences were in fact due to demographic differences.

## Discussion

The purpose of this research was to examine subgroup differences in faking behavior in a large and applied setting and to further inspect how these differences could potentially have a bearing on an actual hiring situation. Although it corroborated some previous findings in the faking literature, it also made some unique contributions as well.

### *Summary of Results and Implications*

The findings from Hypothesis 1 confirm Paulhus's (1991) assertion that females do indeed engage in more intentional faking behavior than men do, whereas men engage in more inadvertent faking behavior than women do. Moreover, because the sample size was so large, Cohen's  $d$  determined a more meaningful comparison between the groups. Although the  $d$  for intentional faking was relatively small, the effect size for unintentional faking was more pronounced, with men scoring .14 of a standard deviation more than women.

This finding suggests that women are not intentionally faking as much as men inadvertently do. Some researchers have found that, perhaps, women are better chameleons than men, better able to adjust their responses to a given situation. Jean and Reynolds (1984) found that, when asked to appear either liberal or conservative in regard to sex roles, women were better than men at altering their responses accordingly. Moreover, Joseph and Newman (2010) produced meta-analytic findings that women were higher in measures of emotional intelligence, perhaps suggesting that women are better

able to tailor their answers to a given audience. Furthermore, because of the lagging promotion and equality of women in the workplace, perhaps females are more inclined to employ intentional faking techniques because they feel like if they do not, they will not be as competitive and their success might continue to be hampered.

This difference might also be attributable to differences in confidence between men and women. Even amongst gifted students, Ziegler and Heller (2000) found that females were less likely to possess strong self-confidence in their abilities, as compared to their male counterparts. Perhaps men are more confident in their abilities, thus faking is not something they have to do intentionally. Furthermore, individuals with high self-confidence were more likely to attribute praise to their own successes, whereas those with low self-esteem attributed the praise to more extraneous or coincidental factors (Ziegler & Heller, 2000). This potentially makes them more susceptible to methods to catch fakers in pre-employment screening. Men, on the other hand, are consistent in their greater propensity to engage in self-deception enhancement, on average. This response style could potentially stem from the previously mentioned socialization differences and employment expectations. However, the hiring simulation executed in the current study did not predict adverse impact excluding either. Adverse impact occurs when standards that apply to all applicants inadvertently exclude certain groups. For example, a height requirement might unintentionally exclude women. Thus, when instituting a standard for all applicants, practitioners must justify that the standard is job-related. The *Uniform Guidelines* (1978) contend that to decipher whether adverse impact is occurring, practitioners can conduct the 80-percent rule. This rule states that the proportion of minorities selected should be at least 80% of the proportion of the majority group

selected. Table 12 indicates the adverse impact ratios for males versus females, which do not indicate that adverse impact is taking place.

Hypotheses 2 through 7 led to some interesting conclusions, some that contradicted prior research. Although African Americans did score slightly higher in some facets of social desirability, namely the more inadvertent facet, Caucasians scored over half a standard deviation higher than African Americans on intentionally faking. Hypothesis 3 compared Hispanics and Caucasians and resulted in similar results as a comparison of Caucasians to African Americans. Hispanics were engaging in very slightly more unintentional social desirability, but Caucasians were engaging in a great deal more intentional faking behavior by more than half a standard deviation.

Caucasians were also faking more than Asian Americans on both dimensions of social desirability. Again, this is contrary to prior research findings. The difference in intentional social desirability behavior was not quite as pronounced as the differences between African Americans and Hispanics, but was still over one-quarter of a standard deviation for both impression management and for social desirability. Hypothesis 5 compared Hispanics and African Americans and predicted that Hispanics would be engaging in more social desirability behavior. On the basis of Hypothesis 2 and 3, both Hispanics and African Americans were similar when both groups were compared to Caucasians and Asian Americans. There was no meaningful difference between Hispanics and African Americans on either intentional or unintentional social desirability.

Hypothesis 6 examined differences between Asian Americans and African Americans. African Americans were engaging in significantly more unintentional social

desirability behavior (almost one-half a standard deviation), whereas Asian Americans were engaging in more intentional social desirability behavior by about one-quarter standard deviation. When Asian Americans were compared to Hispanics, similar results emerged with Hispanics engaging in more unintentional social desirability behavior and Asian Americans engaging in more intentional social desirability behavior to similar degrees.

In comparing the various sex and racial groups, several notable questions emerge. Although sex comparison yielded both established and expected results, the racial group comparisons were more unexpected. This is particularly unsettling as workplaces become more diverse and global. Frei, Yoshita, and Isaacson (2006) remarked a void in the literature examining cross-cultural differences in response distortion. Dean et al. (2010) noted that many of the current response distortion practices stem from Western paradigms and participants. Thus, these practices are not necessarily appropriate for research applied across cultures. Li and Reb (2009) found that the BIDR, a common social desirability measure that assesses both impression management and self-deception, was not necessarily equivalent across various countries or motivational conditions.

Overall, Caucasians were by far engaging in the most intentional social desirability behavior, followed by Asian Americans, and then African Americans and Hispanics. The reverse order was loosely true for unintentional social desirability behavior. African Americans and Hispanics employed the most unintentional social desirability behavior, followed by Caucasians and then Asian Americans.

The most interesting overall results emerge from the findings on Asian American individuals. As Table 13 indicates, Asian Americans were the only group for which the

80-percent rule was violated at the majority of selection ratios. Some researchers have found that centrality is more indicative of the response styles of Asian individuals, suggesting, perhaps, that Asians are less likely to have extreme responses (i.e., Chun, Campbell, & Yoo, 1974; Stening & Everett, 1984; Zax & Takahashi, 1967), as compared to their Caucasian, African American, and Hispanic counterparts. Jang and Kim (2009) found that Asian students had lower levels of explicit life satisfaction (or at least responded indicative of that result). On the other hand, Caucasians more often have a positive view of themselves, as evidenced by several previous research findings. The College Board (1976-1977) found that 100% of high school Caucasians rated themselves as above average on their ability to get along with others and Heine, Lehman, Markus, and Kitayama (1999) found that a whopping 93% of European-Canadians rated themselves as better than average when it came to self-esteem.

In this study, Caucasians interestingly did not score higher on measures of conscientiousness, contrary to previous findings. Perhaps Caucasians are just more likely to fake on non-conscientiousness items, particularly for sales positions. On the other hand, Asian cultures are less likely to make favorable self-evaluations. This is known as the modesty norm, which is when an individual downplays his or her accomplishments, as regulated by cultural norms (Kim et al., 2010). Bond et al. (1982) found that Chinese students who were more self-effacing were perceived more favorably than those who behaved in a more self-enhancing manner. Furthermore, Kim et al. (2010) found that Asian individuals were not only more comfortable making favorable self-evaluations in lower-pressure situations (i.e., with friends) than a higher-pressure situation (i.e., applying for a job), but that Asian respondents are more likely to deny negative traits than

to endorse positive ones. Triandis and Suh (2002) conceded that lying is more acceptable in collectivist cultures in order to avoid confrontation for the good of the group. However, individuals in a collectivist culture typically view self-enhancement as detrimental to the group. My results provide support for this contention, as Asian Americans scored higher than African Americans and Hispanics on the intentional aspect of socially desirable responding, but score lower than African Americans and Hispanics on the unintentional socially desirable responding.

Although a discussion of the adaptability of an overly positive image of self is beyond the scope of this research, perhaps these findings indicate that overall Asian individuals are culturally less motivated to endorse a given item. Furthermore, Chun et al. (1974) even conjecture that when Asian respondents do endorse extreme responses, they are more likely to be reliable. Perhaps Asian Americans in general are simply too hard on themselves. This could potentially explain why out of all the racial groups, Asian Americans scored the lowest on unintentional social desirability. When age and sex were controlled for, race had less of an influence on socially desirable responding. In fact, age actually had the largest impact on both conscientiousness and self-deception over race and sex, whereas sex had the largest impact on impression management over race and age. However, Hypothesis 8 offers unique insight to the magnitude of practical significance from this investigation. What appeared to be a racial subgroup difference issue, was in fact largely due to demographic differences among groups (age), not cultural or racial issues.

Hypothesis 8 involved a simulation of an actual hiring situation by examining applicants only and the observed amount of minorities versus the expected amount of

minorities based on a cutoff score for both conscientiousness and intentional social desirability at the ninety-fifth percentile. In a situation where everyone would be hired after implementing the cutoff score, there were not any differences between any groups. However, after implementing a 75%, 50%, 25%, and 10% selection ratios for the racial groups, there were significant differences between groups. What was most interesting about this pattern was that the most significant differences between observed versus expected within groups lied, again, with the Asian American groups. Furthermore, they were in the reverse direction that was unexpected, given that Asian Americans typically exhibit higher performance on many criterion measures. In this case, however, there were actually fewer Asian Americans selected than would be expected.

Hough et al. (2001) did find significant differences between Asian Americans and Caucasian counterparts on the Achievement aspect of Conscientiousness and Folds, Duehr, and Ones (2008) further found that Asian Americans scored significantly higher than Caucasians in the Conscientiousness facet of Order. As my results indicated, Asian Americans were intentionally faking more so than African Americans and Hispanics, but not more than Caucasians. Therefore, if practitioners are eliminating candidates due to suspected faking on Conscientiousness measures, they seem to be inadvertently eliminating a significant portion of Asian American applicants that are not faking. Practitioners commonly use personality measures to combat adverse impact. However, this exclusion of many Asian applicants offers support that the use of personality in selection does not uniformly eliminate adverse impact concerns.

### *Limitations*

Like any research, this study has some limitations. One of the most salient limitations concerns the high number of salespeople. Sales positions might vary from other positions, as sales positions tend to rely heavily on impression management. Oftentimes an individual's success in a sales position depends heavily on their ability to make him or herself look good. Thus, salespeople might have a slightly skewed view as to what constitutes faking versus taking advantage of opportunities to look good (Tristan, 2008). Moreover, some researchers (e.g., Velicer & Weiner, 1975; Wesman, 1952) found that a person will portray him or herself differently when applying for a sales job versus a non-sales job. Lastly, Tristan (2008) found very little difference between sales applicants and sales incumbents, suggesting that social desirability measures are not as effective with salespersons.

Another potential limitation concerns the scales employed. Their reliabilities, although possessing acceptable psychometric properties for research purposes, were not inordinately high. Thus, improved measures could potentially yield more generalizable results. Finally, a limitation also stems from this research's status as a field study. Although the use of an applied sample provides unique and important insight into an actual hiring situation, there is some loss of experimental control that would be present in a laboratory study. This led to decreased control in some control variables and random assignment was not plausible.

#### *Future Research*

Future research could potentially integrate more understudied racial groups. This obviously includes Asian Americans, but also American Indian, Middle Eastern, and biracial individuals. Additionally, more research on socially desirable responding and

racial and ethnic differences is warranted. There is relatively limited research investigating these differences, especially with understudied ethnic groups. Moreover, because my results were not entirely consistent with prior research, more study on this subject is warranted to discern a more established interpretation of the results. Finally, future research could integrate the use of item response theory to determine if various subgroups are interpreting the questionnaire items differently.

The strength of this study resided in its very large sample size, ability to investigate some understudied racial groups, and its applied nature, advancing the research on personnel selection. Thus, several noteworthy implications stem from this research. First, I found that there were indeed group differences and that different forms of faking yielded various group differences. This is significant because practitioners commonly use personality measures, as previously mentioned, to combat adverse impact. However, if adverse impact is still potentially occurring with some groups, practitioners should use personality measures with caution. Based on these findings, Asian Americans are less likely to be hired in an actual hiring situation. Asian Americans scored lower on both Conscientiousness and Self-deception measures, suggesting that Asian Americans might be more critical about their strengths. As a group, Asian Americans are not self-enhancing as much as other groups. Thus, the use of personality measures might result in adverse impact for Asian Americans, as the results of Hypothesis 8 suggest. This could be a meaningful problem for practitioners.

On the basis of this research, I suggest that practitioners proceed cautiously when interpreting the results of pre-employment screenings and carefully consider contextual differences when making hiring decisions. Further examinations of the reasons behind

socially desirable responding will undoubtedly result in an improved knowledge base as to what exactly the social desirability scales are measuring for a target population.

## Appendix A

### Assessment Instructions

#### *Instructions for Primary Consulting Firm*

Note: Since most business functions include some customer contact, the questions in the Sales Orientation Assessment are required for all positions. If you have never had direct sales experience, answer the questions according to how you believe they would apply if you were in sales. **Note:** a "No Opinion" response is acceptable and will not substantially affect your results.

**Instructions: Be sure to answer every required question.** If you skip any required questions, you will be prompted to complete them before you can go on to the next section.

**If you change your mind about an answer,** scroll back up on the page and change your answer. Do this **before** you click on the "Click Here to Continue" button at the bottom of each page, because you won't be able to come back to that page of questions.

**If you have to stop** before you have completed the assessment, don't worry: when you log back in to resume it will return you to where you need to continue. **However, if you have finished a page of questions, those answers cannot be saved until you click on the button at the bottom of the page, and the following page of questions appears.** If your browser times out before the next page appears (it will say "The page cannot be

found", or something similar), refresh the screen using the **Refresh and Retry** commands if you are using Microsoft Internet Explorer as your browser. If you are using Netscape, use the **Reload and OK** commands.

Give yourself every advantage: **Your best bet is to be as honest as possible, and give your first-impulse answer.** This survey is not timed, but please answer the statements as quickly as possible for the most accurate results. There are no "right" or "wrong" answers.

**Caution:** Don't make the mistake of trying to out-think the questions and make yourself appear different from what is really true. There are validity scales built into this assessment.

**[Consulting firm's name] do not report individual answers.** Only a summarization and profile of your scores will be reported.

Assessment Instructions

*Instructions for Secondary Consulting Firm*

Dear [FirstName] [LastName],

Congratulations, you have been chosen to complete the SalesPro assessment. This assessment asks you to answer questions about yourself how you approach your sales job. The information collected will be used by your organization to better understand their salesforce and to help make better hiring decisions.

Please complete this assessment at your earliest convenience. We appreciate you taking the time participate in this assessment process and help your organization to improve and succeed.

The assessment will take approximately 60 to 90 minutes to complete. We suggest that you take the assessment in a quiet place where you can focus and concentrate. Please be honest in your responses. The more honest and accurate participants are when taking the assessment, the more accurate the results will be.

Before completing the assessment, please make sure that you have a high speed internet connection (e.g., Cable, DSL, T1). The assessment may not be able to run on a dial up connection.

To complete the assessment, click on the following link:

<https://selectintl.selectrakonline.com/recruitment/portal/myjobs.aspx>

If there is no response when clicking the link above, please copy and paste the link into your browser and try again. Then, follow these steps:

Log in by entering your email address (work email) and clicking “Continue”

Click on the link with your position title (HM Insurance: Consultant or Director)

Click on the “Launch Assessment” button and the assessment will soon begin

NOTE: if nothing happens after clicking “Launch Assessment”, you likely have a pop-up blocker still activated. A quick shortcut to bypass pop up blockers is to hold down the CONTROL key while you click on the “Launch Assessment” button, and continue holding the Control key for 2-3 seconds. This usually bypasses any blockers you may have turned on (e.g., Yahoo or Google task bars have their own pop up blockers).

To successfully view and complete the assessment, certain technical parameters must be met. If the assessment does not load after clicking the Launch Assessment button, please close all windows and re-enter the assessment. Please be sure to check your system for the requirements below.

## **TECHNICAL SPECIFICATIONS**

Windows 98/NT/2000/XP

Mouse

Monitor screen resolution set to 800x600 with color quality set to high color (thousands of colors) or better 4. High-speed Internet Access (DSL connection acceptable; T1 or better is ideal)

Internet browser pop-up blockers need to be disabled (you can also hold Control key while launching assessment)

Internet Explorer 6 or higher

Flash Player Plug-in 7.0.19.0 or higher

If you continue to experience technical difficulties, please contact your organization's HR representative.

Thank you.

## REFERENCES

- Avery, D.R., McKay P.F., Wilson, D.C., & Tonidandel, S. (2007). Unequal attendance: The relationships between race, organizational diversity cues, and absenteeism. *Personnel Psychology, 60*, 875-902.
- Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261-272.
- Benbow, C.P. & Stanley, J.C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, New Series, 22*, 1029-1031.
- Bond, M.H., Leung, K., & Wan, K.C. (1982). How does cultural collectivism operate? The impact of task and maintenance contributions on reward distribution. *Journal of Cross-Cultural Psychology, 13*, 186-200.
- Burns, G.N. & Christiansen, N.D. (2006). Sensitive or senseless: On the use of social Desirability measures in selection and assessment. In R.L. Griffith & M.H. Peterson (Eds.), *A Closer Examination of Applicant Faking Behavior* (pp. 115-150). *Greenwich, CT: Information Age Publishing.*
- Carli, L.L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology, 59*, 941-951.
- Chan, D. & Schmitt, N. (1997). Video-based versus paper-and-pencil method of

- assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82, 143-159.
- Chan, D., Schmitt, N., DeShon, R.P., Clause, C.S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology*, 82, 300-310.
- Chun, K.T., Campbell, J.B., & Yoo, J.H. (1974). Extreme response style in cross-cultural Research: A reminder. *Journal of Cross-Cultural Psychology*, 5, 465-480.
- Cialdini, R.B. (1989). Indirect tactics of image management: Beyond basking. In R.A. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 45-56). Hillsdale, NJ: Erlbaum.
- Cohen, J.(1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Earlbaum Associates. College Board. (1976-1977) *Student descriptive questionnaire*. Princeton, NJ: Educational Testing Services.
- Collins, J.M. & Gleaves, D.H. (1998). Race, job applicants, and the five-factor model of personality: Implications for black psychology, industrial/organizational psychology, and the five-factor theory. *Journal of Applied Psychology*, 83, 531-544.
- Costa, P., Terracciano, A, & McCrae, R.R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322-331.
- Culbertson, J.D. (2007). Predicting sales performance: Considering nonlinear

relationships between GMA, performance, and behavior. Unpublished Master's thesis proposal, Wright State University, OH.

Dean, M.A., Roth, P.L., & Bobko, P. (2008). Ethnic and subgroup differences in assessment center ratings: A meta-analysis. *Journal of Applied Psychology, 93*, 685-691.

Dean, S.L. & Leasher, M.K. (2009). Assessments in a global workforce: Cross-cultural variation in response distortion. Symposium submitted to the 25<sup>th</sup> annual conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Dean, S.L., Leasher, M.K., Miller, C.E., Filipkowski, J.N., & Culbertson, J.D. (April, 2010). Assessments in a global workforce Cross-cultural variation in response distortion. Symposium presented at the 25<sup>th</sup> meeting of the Society for Industrial/Organizational Psychologists, Atlanta, GA.

Department of Labor Women's Bureau. (2005). Women in the labor force in 2004. Retrieved 2009 for the U.S. Department of Labor via <http://www.dol.gov/wb/factsheets/Qf-laborforce-04.htm>.

Dietrich, D. (1995). Gender differences in self-handicapping: Regardless of academic or social competence implications. *Social Behavior and Personality, 23*, 403-410.

DuBrin, A.J. (1991). Sex and gender differences in tactics of influence. *Psychological Reports, 68*, 635-646.

DuBrin, A.J. (1994). Sex differences in the use and effectiveness of impression management. *Psychological Reports, 74*(2), 531-544.

Dudley, N.M., McFarland, L.A., Goodman, S.A., Hunt, S.T., & Sydell, E.J. (2005).

- Racial differences in socially desirable responding in selection contexts: Magnitude and consequences. *Journal of Personality Assessment*, 85, 50-64.
- Dunning, D. (1999). A newer look: Motivated social cognition and the schematic representation of social concepts. *Psychological-Inquiry*, 10, 1-11.
- Eagly, A.H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.
- Eagly, A.H. & Carli, L.L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis comparing women and men. *Psychological Bulletin*, 129, 569-591.
- Ellingson, J.E., Sackett, P.R., & Hough, L.M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology*, 84, 155-166.
- Foldes, H.J., Dueher, E. E., & Ones, D.S. (2008). Group differences in personality: Meta-analyses comparing five U.S. racial groups. *Personnel Psychology*, 61, 579-616.
- Ford, J.K., Kraiger, K. & Schechtman, S.L. (1986). Study of race effects in objective indices and subjective evaluations of performance: A meta-analysis of performance criteria. *Psychological Bulletin*, 99, 330-337.
- Fouad, N.A. (1994). Career assessment with Latinos/Hispanics. *Journal of Career Assessment*, 2, 226-239.
- Frank, G. (1992). The response of African-Americans to the Rorschach: A review of the research. *Journal of Personality Assessment*, 59, 317-325.
- Gerhart, B. & Rynes, S. (1991). Determinants and consequences of salary negotiations by male and female MBA graduates. *Journal of Applied Psychology*, 76, 256-262.

- Gibbs, N. (2009, 26 October). What Women Want Now. *Time*, 174, 24-33.
- Gibson, B., & Sachau, D. (2000). Sandbagging as a self-presentation strategy: Claiming to be less than you are. *Personality and Social Psychology Bulletin*, 26, 56-70.
- Guadagno, R.E. & Cialdini, R.B. (2007). Gender differences in impression management in organizations: A qualitative review. *Sex Roles*, 56, 483-494.
- Halpern, D.F. (2000). Validity, fairness, and group differences: Tough questions for selection testing. *Psychology, Public Policy, and Law*, 6, 56-62.
- Halpern, D.F., Benbow, C.P., Geary, D.C., Gur, R.C., Hyde, J.S., & Gernsbacher, M.A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, 8, 1-51.
- Halpern, D.F. & LaMay, M.L (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, 12, 229-246.
- Harold, C. M., McFarland, L. A., & Weekley, J. A. (2003, April). *The validity of verifiable and nonverifiable biodata items: An examination across applicants and incumbents*. Paper presented at the 18<sup>th</sup> Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Heatherington, L., Burns, A.B., & Gustafson, T.B. (1998). When another stumbles: Gender and self-presentation to vulnerable others. *Sex Roles*, 38, 889-913.
- Heine, S.J., Lehman, D.R., Markus, H.R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106, 766-794.
- Herrnstein, R.J. & Murray, C.A. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Hieshima, J.A. & Schnieder, B. (1994). Intergenerational effects on the cultural and

- cognitive socialization of third- and fourth-generation Japanese Americans. *Journal of Applied Developmental Psychology, 15*, 319-327.
- Higgins, R.L. & Snyder, C.R. (1989). The business of excuses. In R.A. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 73-84). Hillsdale, NJ: Erlbaum.
- Hirt, E.R., McCrea, S.M., & Kimble, C.E. (2000). Public self-focus and sex differences in behavioral self-handicapping: Does increasing self-threat still make it “just a man’s game?” *Personality and Social Psychology Bulletin, 26*, 1131-1141.
- Hough, L. M. (1998). Effects of Intentional Distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11*, 209-244.
- Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D., & McCloy, R.A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581-595.
- Hough, L.M., Oswald, F.L., & Ployhart, R.E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9*, 152-194.
- Huffcutt, A.I. & Roth, P.L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology, 83*, 179-189.
- Jang, D., & Kim, D.Y. (2009). Two faces of human happiness: Explicit and implicit life satisfaction. *Asian Journal of Social Psychology, 12*, 185-198.
- Jean, P.J., & Reynolds, C.R. (1984). Sex and attitude distortion: Ability of females and

- males to fake liberal and conservative positions regarding changing sex roles. *Sex Roles, 10*, 805-815.
- Johnson, W., Carothers, A., & Deary, I.J. (2009). A role for the X chromosome in sex differences in variability in general intelligence? *Perspectives on Psychological Science, 4*, 598-611.
- Jones, D. (2009, January 1). Women CEOs slowly gain on corporate America. *USA Today*, Retrieved from [http://www.usatoday.com/money/companies/management/2009-01-01-women-ceos-increase\\_N.htm](http://www.usatoday.com/money/companies/management/2009-01-01-women-ceos-increase_N.htm).
- Jones, E.E. & Wortman, C. (1973). *Ingratiation: An attributional approach*. New York: General Learning.
- Joseph, D.L., & Newman, D.A. (2010). Emotional intelligence: An integrative meta-analysis cascading model. *Journal of Applied Psychology, 95*, 54-78.
- Kamen, V.S. & Hartel, C.E.J. (1994). Gender differences in anticipated pay negotiation strategies and outcomes. *Journal of Business and Psychology, 9*, 183-197.
- Karsten, M.F. (1994). *Management and gender: Issues and attitudes*. Westport, CT: Quorum Books.
- Keys, D.E. (1985). Gender, sex role, and career decision making of certified management accountants. *Sex Roles, 13*, 33-46.
- Kim, Y.H., Chiu, C.Y., Peng, S., Cai, H., & Tov, W. (2010). Explaining east-west differences in the likelihood of making favorable self-evaluations: The role of evaluation apprehension and directness of expression. *Journal of Cross-Cultural Psychology, 41*, 62-75.

- King, E.B., Madera, J.M., Hebl, M.R., Knight, J.L. & Mendoza, S.A. (2006). What's in a name? A multiracial investigation of the role of occupational stereotypes in selection decisions. *Journal of Applied Social Psychology, 36*, 1145-1159.
- Komar, S., Brown, D.J., Komar, J.A., & Robie, C. (2008). Faking and the validity of conscientiousness: A Monte Carlo investigation. *Journal of Applied Psychology, 93*, 140-154.
- Konovsky, M.A., & Jaster, F. (1989). "Blaming the victim" and other ways business men and women account for questionable behavior. *Journal of Business Ethics, 8*, 391-398.
- Latham, M.H., Ostrowski, B.A., Pavlock, E.J., & Scott, R.A. (1987, January). Recruiting entry level staff: Gender differences. *The CPA Journal, January*, pp. 30-42.
- Lee, S.J., Quigley, B.M., Nesler, M.S., Corbett, A.B., & Tedeschi, J.T. (1999). Development of a self-presentation tactics scale. *Personality and Individual Differences, 26*, 701-722.
- Lewak, R.W., Marks, P.A., & Nelson, G.E. (1990). Therapist guide to the MMPI & MMPI-2: Providing feedback and treatment. Muncie, IN: Accelerated Development.
- Li, A. & Reb, J. (2009). Across-nations, cross-cultures, and cross-conditions analysis on the equivalence of the Balanced Inventory of Desirable Responding. *Journal of Cross-Cultural Psychology, 40*, 214-233.
- Mainiero, L.A. (1986). Coping with powerlessness: The relationship of gender and job dependency to empowerment strategy usage. *Administrative Science Quarterly, 31*, 633-653.

- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51*, 882-888.
- McFarland, L.A. & Ryan, A.M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85*, 812-821.
- McFarland, L.A. & Ryan, A.M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*, 979-1016.
- McKay, P.F. & McDaniel, M.A. (2006). A reexamination of black-white mean differences in work performance: More data, more moderators. *Journal of Applied Psychology, 91*, 538-554.
- McKay, P.F., Avery, D.R., & Morris, M.A. (2008). Mean racial-ethnic differences in employee sales performance: The moderating role of diversity climate. *Personnel Psychology, 61*, 349-374.
- Meehl, P.E. & Hathaway, S.R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30*, 525-564.
- Miller, C.E. (2000). The susceptibility of personality selection tests to coaching and faking. Unpublished doctoral dissertation, University of Akron, OH.
- Mills, A.J. (2002). Studying the gendering of organizational culture over time: Concerns, issues and strategies. *Gender, Work, and Organization, 9*, 286-307.
- Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.
- Offermann, L.R., & Schrier, P.E. (1985). Social influence strategies: The impact of sex,

- role, and attitudes toward power. *Personality and Social Psychology Bulletin*, 11, 286-300.
- Ogbu, J.U. (1978). *Minority education and caste: The American system in cross-cultural perspective*. New York: Academic Press.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for Personnel Selection: The red herring. *Journal of Applied Psychology*, 81, 660-679.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp.17-59). San Diego, CA: Academic Press.
- Ramos, R.A. (1981). Employment battery performance of Hispanic applicants as a function of English or Spanish test instructions. *Journal of Applied Psychology*, 66, 291-295.
- Rosenfeld, P., Giacalone, R.A., & Riordan, C.A. (1995). *Impression management in organizations*. London: Routledge.
- Ross, C.E. & Mirowsky, J. (1984). Socially-desirable response and acquiescence in a cross-cultural survey of mental health. *Journal of Health and Social Behavior*, 25, 189-197.
- Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on pre-employment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634-644.

- Roth, P.L., Huffcutt, A.I., & Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology, 88*, 694-706.
- Ryan, A.M. & Boyce, A.S. (2006). What do we know and where do we go? Practical directions for faking research. In R.L Griffith & M.H. Peterson (Eds.), *A Closer Examination of Applicant Faking Behavior* (pp. 357-371). Greenwich, CT: Information Age Publishing.
- Sackett, P.R., & Ellingson, J.E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology, 50*, 707-721,
- Sackett, P.R. & Wilk, S.L. (1994). Within group norming and other forms of score adjustment in preemployment testing. *American Psychology, 49*, 929-954.
- Schmitt, D.P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*, 168-182.
- Schmitt, N. & Chan, D. (1998). *Personnel Selection: A theoretical approach*. Thousand Oaks, CA: Sage Publications.
- Schmitt, N. & Oswald, F. (2006). The impact of corrections for faking on the validity of noncognitive measures in selection settings. *Journal of Applied Psychology, 91*, 613-621.
- Schmitt, N., Clause, C.S., & Pulakos, E.D. (1996). Subgroup differences associated with different measures of some common job-relevant constructs. *International Review of Industrial and Organizational Psychology, 11*, 115-139.

- Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40, 437-453.
- Segal, A.T. (1992, June 8). Corporate Women. *Business Week*, pp. 74-83.
- Shepperd, J.A. & Arkin, R.M. (1989). Self-handicapping: The moderating roles of public self-consciousness and task importance. *Personality and Social Psychology Bulletin*, 15, 252-265.
- Smith, S.W., Cody, M.J., Lovette, S., & Candry, D.J. (1990). Self-monitoring, gender, and compliance-gaining goals. In M.J. Cody & M.L. McLaughlin (Eds.), *The psychology of tactical communication* (pp. 91-135). Clevedon, UK: Multilingual Matters.
- Stening, B.W., & Everett, J.E. (1984). Response styles in a cross-cultural managerial study. *The Journal of Social Psychology*, 122, 151-156.
- Stevens, C.K., Bavetta, A.G., & Gist, M.E. (1991). Paper presented at the annual meeting of the Academy of Management, Miami, FL.
- Stroh, L.K., Brett, J.M., & Reilly, A.H. (1992). All the right stuff: A comparison of female and male managers' career progression. *Journal of Applied Psychology*, 77, 251-260.
- Strutton, D., Pelton, L.E., & Lumpkin, J.R. (1995). Sex differences in ingratiation behavior: An investigation of influence tactics in the salesperson-customer dyad. *Journal of Business Research*, 34, 35-45.
- Tannen, D. (1994). *Talking from 9 to 5: How women's and men's conversational styles affect who gets heard, who gets credit, and what gets done at work*. New York: Morrow.
- Thacker, R.A. (1995). Gender, influence tactics, and job characteristics preferences:

- New insights into salary determination. *Sex Roles*, 32, 617-638.
- Tristan, E. (2008). Measuring applicant faking with job desirability: Prevalence, selection, and measurement issues in an applied sample. Unpublished doctoral dissertation, Wright State University, OH.
- Turkheimer, E. & Halpern, D.F. (2009). Sex differences in variability for cognitive measures: Do the ends justify the genes? (Commentary of Johnson et al., 2009). *Perspectives on Psychological Science*, 4, 612-614.
- Uniform guidelines on employee selection procedures. (1978). 43 Fed. Reg. 38, 290-38, 315.
- University of Minnesota Press. (1989). Minnesota Multi-Phasic Personality Inventory-2 Manual for Administration and Scoring. Minneapolis, MN: Hathaway, S.R. & McKinley, J.C.
- U.S. Merit Systems Protection Board. (1997). *Fair and equitable treatment: A progress report on minority employment in the federal government*. Washington, DC: Office of Policy and Evaluation.
- Valian, V. (1998). *Why so slow? The advancement of women*. Cambridge, MA: The MIT Press.
- Van Iddekinge, C. H., Raymark, P. H., Eidson Jr., C. E., Putka, D. J. (2003, April). *Applicant-incumbent differences on personality, integrity, and customer service measures*. Paper presented at the 18<sup>th</sup> Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Velicer, W. F., & Weiner, B. J. (1975). Effects of sophistication and faking sets on the Eysenck Personality Inventory. *Psychological Reports*, 37, 71-73.

- Vincent, N.M, Linsz, N.L., & Greene, M.I. (1966). The L scale of the MMPI as an index of falsification. *Journal of Clinical Psychology, 22*, 214-215.
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*, 197-210.
- Weekley, J.A., Ployhart, R.E., & Harold, C.M. (2004). Personality and situational judgment tests across applicant and incumbent settings: An examination of validity, measurement, and subgroup differences. *Human Performance, 17*, 433-461.
- Wesman, A. G. (1952). Faking personality test scores in a simulated employment situation. *Journal of Applied Psychology, 36*, 112-113.
- Wing, H. (1980). Profiles of cognitive ability of different racial/ethnic and sex groups on a multiple abilities test battery. *Journal of Applied Psychology, 65*, 289-298.
- Zax, M., & Takahashi, S. (1967). Cultural influences on response style: Comparisons of Japanese and American college students. *The Journal of Social Psychology, 71*, 3-10.
- Ziegler, A. & Heller, K.A. (2000). Conditions for self-confidence among boys and girls achieving highly in chemistry. *Journal of Secondary Gifted Education, 11*, 144-151.

Table 1

*Differences Between Male and Female Social Desirability Scores*

Measure	Male		Female		<i>t</i> (247392)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.70	3.28	13.62	3.37	5.77***	0.02
Self-Deception	10.74	2.39	10.39	2.50	33.90***	0.14
Impression Management	4.13	2.19	4.18	2.31	-4.59***	-0.02

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A negative *d* indicates that females score higher on that particular measure.

Table 2

*Differences Between Caucasian and African American Social Desirability Scores*

Measure	Caucasian		African American		<i>t</i> (216864)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.69	3.25	13.96	3.22	-11.05***	-0.08
Self-Deception	10.65	2.41	10.76	2.19	-6.45***	-0.05
Impression Management	4.36	2.18	3.16	2.19	74.28***	0.55

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A negative *d* indicates that African Americans score higher on that particular measure.

Table 3

*Differences Between Caucasian and Hispanic Social Desirability Scores*

Measure	Caucasian		Hispanic		<i>t</i> (208887)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.39	3.25	13.85	3.47	-4.91***	-0.05
Self-Deception	10.65	2.41	10.71	2.48	-2.59**	-0.02
Impression Management	4.36	2.18	3.11	2.22	60.21***	0.57

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A negative *d* indicates that Hispanics score higher on that particular measure.

Table 4

*Differences Between Caucasian and Asian American Social Desirability Scores*

Measure	Caucasian		Asian American		<i>t</i> (203557)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.69	3.25	12.61	3.83	23.17***	0.31
Self-Deception	10.65	2.41	9.63	2.90	28.85***	0.38
Impression Management	4.36	2.18	3.70	2.32	23.18***	0.29

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A positive *d* indicates that Caucasians score higher on all measures.

Table 5

*Differences Between African American and Hispanic Social Desirability Scores*

Measure	African American		Hispanic		<i>t</i> (32409)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.96	3.22	13.85	3.47	2.72**	0.03
Self-Deception	10.76	2.19	10.71	2.48	1.68	0.02
Impression Management	3.16	2.19	3.11	2.22	1.68	0.02

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A positive *d* indicates that African Americans score higher on all measures.

Table 6

*Differences Between African American and Asian American Social Desirability Scores*

Measure	African American		Asian American		<i>t</i> (27079)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.96	3.22	12.61	3.83	26.17***	0.38
Self-Deception	10.76	2.19	9.63	2.90	29.47***	0.44
Impression Management	3.16	2.19	3.70	2.32	-16.93***	-0.24

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A negative *d* indicates that Asian Americans score higher on that particular measure.

Table 7

*Differences Between Hispanic and Asian American Social Desirability Scores*

Measure	Hispanic		Asian American		<i>t</i> (19102)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Conscientiousness	13.85	3.47	12.61	3.83	22.24***	0.34
Self-Deception	10.71	2.48	9.63	2.90	26.00***	0.40
Impression Management	3.11	2.22	3.70	2.32	-16.94***	-0.26

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

Note: A negative *d* indicates that Asian Americans score higher on that particular measure.

Table 8

*Overall Chi-Square values for Racial Groups with Varying Selection Ratios*

	100%	75%	50%	25%	10%
$X^2$	1.695	29.412***	30.624***	24.246***	9.122*
	(3, N = 30049)	(3, N = 24222)	(3, N = 15542)	(3, N = 6781)	(3, N = 2751)

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

Table 9

*Proportion of Observed vs. Expected Selected for Racial Groups with Varying Selection Ratios after a 95<sup>th</sup> Percentile Cutoff Score Instituted*

	100%	75%	50%	25%	10%
Caucasians	.997	1.004	1.003	.999	1.001
African Americans	1.017	1.043	1.068	1.106	1.007
Hispanics	1.015	.954	.946	.978	1.095
Asian Americans	1.021	.786***	.736***	.642***	.626***

\*\*\*  $p < .001$

Note: A value greater than one indicates more from the given group were selected than expected, whereas a value less than one indicates fewer from the given group were selected than expected.

Table 10

*Proportions of Observed vs. Expected Selected for Sex Groups with Varying Selection Ratios after a 95<sup>th</sup> Percentile Cut-off Score Instituted*

	100%	75%	50%	25%	10%
Males	1.003	1.005	1.001	.990	.990
Female	.995	.992	.998	1.006	1.016

Note: A value greater than one indicates more from the given group were selected than expected, whereas a value less than one indicates fewer from the given group were selected than expected.

Table 11

*Correlation Matrix of Outcome Variables, Sex, and Age*

	Sex	Conscientiousness	Self-Deception	Impression Management
Conscientiousness	-.014**			
Self-Deception	-.072**	.795**		
Impression Management	.010**	-.555**	-.465**	
Age	-.114**	-.086**	.025**	.138**

\*\*p < .001

Note: A negative value on the sex variable indicates that the value is more associated with males than with females.

Table 12

Adverse impact table for male and female applicants

Selection Ratio	100%	75%	50%	25%	10%
Percent Selected	92.97%	74.52%	47.61%	20.70%	8.36%
Males					
Percent Selected	92.28%	73.53%	47.47%	20.88%	8.57%
Females					
Adverse Impact	.993	.987	.997	1.01	1.03
Ratio					

Table 13

Adverse impact table for Caucasian, African American, Hispanic, and Asian American Applicants

<b>Selection Ratio</b>	<b>100%</b>	<b>75%</b>	<b>50%</b>	<b>25%</b>	<b>10%</b>
<b>Percent Selected Caucasian</b>	92.30%	74.94%	48.04%	20.87%	8.49%
<b>Percent Selected African American</b>	94.16%	77.84%	51.04%	23.11%	8.54%
<b>African American Adverse Impact Ratio</b>	1.02	1.04	1.06	1.11	1.01
<b>Percent Selected Hispanic</b>	93.98%	71.19%	45.33%	20.45%	9.28%
<b>Hispanic Adverse Impact Ratio</b>	1.02	.950	.889	.980	1.093
<b>Percent Selected Asian American</b>	94.54%	58.66%	35.26%	13.42%	5.30%
<b>Asian American Adverse Impact Ratio</b>	1.02	.783	.734	.643	.624