January 2013

Variance in Faking in High-Stakes Personality Assessment as an Indication of Job Knowledge

Timothy Ryan Dullaghan
University of South Florida, tdullagh@mail.usf.edu

Variance in Faking in High-Stakes

Personality Assessment as an Indication of Job Knowledge

by

Timothy Ryan Dullaghan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
College of Arts and Sciences
University of South Florida

Major Professor: Walter Borman, Ph.D.
Jennifer Bosson, Ph.D.
Carnot Nelson, Ph.D.
Toru Shimizu, Ph.D.
Stephen Stark, Ph.D.

Date of Approval:
June 18, 2013

Keywords: Response Distortion, Impression Management,
Selection, Testing, Rank-Order

Table of Contents

List of Tables

Abstract

The purpose of this study was to evaluate whether the personality trait elevation between honest and applicant contexts that has been widely seen throughout the personality and selection research is merely universal, blatant trait elevation, or whether something else is underlying this faking behavior. By obtaining both honest and applicant context personality responses in which respondents were provided with focal job knowledge, this study determined that while there is near-universal trait elevation across seven personality traits, there is, in fact, some trait differentiation between jobs. As such, this study provided some evidence of knowledgeable faking, defined as distortion of personality test responses based on knowledge of the job being applied to, within applicant contexts.

Chapter One:

Introduction

Personality tests have been increasingly used in the workplace to make a number of workforce decisions including employee selection, offering of developmental opportunities and coaching programs, and enrollment in high-potential employee career paths, among other uses. Arguably, the most significant advancement in personality theory in recent history has been the development of the Five Factor Model (FFM) of personality (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991), composed of the personality traits emotional stability (or neuroticism), extraversion, openness to experience, agreeableness, and conscientiousness. Since this parsimonious, high-level model of personality was identified, researchers have determined that personality shows incremental validity in predicting job and training performance beyond cognitive ability tests (Schmidt and Hunter, 1998), with little to no adverse impact (Foldes, Duehr, & Ones, 2008), and comparable validity levels to cognitive ability tests (Hogan, 2005; Ones, Hough, & Viswesvaran, 1998; Ones & Viswesvaran 2001a; Ones & Viswesvaran, 2001b).

However, there is continued criticism of the use of personality tests in high-stakes testing environments (e.g., for employee selection) because it appears that personality tests can be readily faked. It is widely believed that faking decreases the criterion-related validity of these tests, yet after two decades of intensive investigation into applicant

faking, it is clear that although some properties of personality tests are affected, the criterion-related validities of personality assessment may be robust to faking (Barrick & Mount, 1996; Hogan, 2005; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Ones, Dilchert, Viswesvaran, & Judge, 2007; Ones, Viswesvaran, & Reiss, 1996; Rosse, Stecher, Miller, & Levin, 1998). As of yet, it is unclear why the predictive validity of personality tests are so robust to faking, despite evidence that the psychometric properties of personality tests under faking conditions can be adversely affected (Douglas, McDaniel, & Snell, 1996; Pauls & Crost, 2004; Schmit & Ryan, 1993; Topping & O'Gorman, 1997), and a larger proportion of those selected under low selection ratios will be fakers than under higher selection ratios (Rosse et al., 1998).

Investigation into job-specific variances in applicant faking and the validity of personality traits for specific jobs is one area that has been lacking in the faking literature, yet this area of research may illuminate why personality assessments are so robust to faking. Based on the accumulated evidence, I propose that faking, most often identified through elevated applicant trait scores (McFarland & Ryan, 2000), should be divided into two categories – blatant faking, in which trait level elevation occurs across all traits, and knowledge-based faking, in which more targeted faking occurs, based on the respondent's understanding of the job's personality-based and behavioral requirements.

I propose that elevated applicant trait scores on traits relevant to the target job, which are presumably faked (McFarland & Ryan, 2000), may indicate the applicant has a high level of job knowledge for the job being applied to, rather than simply indicating a respondent blatantly distorted their responses to a personality assessment. Before

exploring this evidence, however, it is important to concretely establish the utility and validity of personality tests in the workplace.

Utility of Personality Tests

Personality tests have seen increased use in the workplace for several reasons. First, hiring managers care about the personality of the people they hire. Dunn, Mount, Barrick, & Ones (1995) found that hiring managers rated certain personality traits nearly as important as cognitive ability in predicting performance, and more important for predicting counter-productivity across six occupations. Along these lines, for most jobs, it would be hard to imagine a manager wanting to hire someone who is constantly disorganized and not very dependable (low conscientiousness) and has constant emotional fluctuations (low emotional stability), suggesting these two traits may be widely recognized good and bad traits for employees. Conversely, there are some jobs for which a high level of extraversion would be desirable (e.g., an event planner), but others for which high levels could get in the way of completing one's job duties (e.g., a librarian; O*NET, 2013).

Second, personality predicts performance. Several well-conducted meta-analytic and empirical investigations have shown that personality traits are valid predictors of performance on the job (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Dudley, Orvis, Lebiecki, & Cortina, 2006; Hogan & Holland, 2003; Hurtz & Donovan, 2000; Sackett, 2011; Tett et al., 1991). Although cognitive ability measures have shown stronger relationships with performance across jobs, personality traits add incrementally to this predictive validity, better than most other selection methods. In Schmidt and Hunter's (1998) widely discussed meta-analysis investigating the validity of selection

procedures both singularly and in combination in the prediction of job performance across jobs, cognitive ability tests showed a validity of r = .51, integrity tests r = .41, and conscientiousness r = .31. In conjunction with cognitive ability, integrity tests added .14 to the validity of cognitive ability for a Multiple R of .65, and conscientiousness added .09 for a Multiple R of .60. Further, Collins et al. (2003) found in their meta-analysis of personality and overall assessment center ratings (OARs) that although cognitive ability alone predicted much of the variance in OARs (ρ = .65), the addition of personality traits to the model significantly increased the variance accounted for, Multiple R (corrected) = .84. While a Multiple R of .84 is uncommonly high, the investigators showed that in certain contexts, the combination of a set of personality traits and cognitive ability can predict nearly all of the variance in performance ratings.

Personality also predicts other important organizational outcomes, such as organizational citizenship behaviors (correlations between different types of OCBs and the Big Five traits reached up to r = .16; Chiaburu, Oh, Berry, Li, & Gardner, 2011), counter-productive work behaviors (reaching magnitude r = -.36 for agreeableness; Berry, Ones, & Sackett, 2007), and job satisfaction (Multiple R = .41; Judge, Heller, & Mount, 2002). Thus, personality effectively predicts a variety of important organizational outcomes.

Third, personality tests overall show little to no adverse impact based on gender or race, two classes protected by U.S. federal law (Hough, 1998; Hough, Oswald, & Ployhart, 2001; Mount & Barrick, 1995). In general, effect sizes between racial groups on personality traits are small to non-existent. In a recent meta-analysis, Foldes et al. (2008) found that group mean trait score differences were small, and argued that these minor

differences would be unlikely to cause adverse impact in selection. The one exception was the moderately large effect size found between Whites and Asians for Agreeableness, d = .61, which favored the minority group. At the facet level, however, some moderate to large effect sizes were found, as large as d = .50 between Whites and Asians for the order facet of conscientiousness, with the minority group again higher on the trait. However, these facet-level comparisons do not consistently favor one group over another, so with the combination of multiple traits or facets into personality profiles, which is most commonly done in the work place (Hogan, 2005), it is even less likely that these small group differences will accumulate to differential selection rates.

Similarly, researchers have found small to moderate mean group differences for gender, yet again, one group is not consistently favored over the other, and the differences are negligible when using higher-level traits rather than lower-level facets (Hough & Dilchert, 2010; Powell, Goffin, & Gellatly, 2011). For example, Powell et al. (2011) found that women were higher on extraversion's facet of affiliation than men, d = -.26, whereas men were higher on dominance, d = .41. However, for extraversion overall, there was a minimal effect size between the groups, d = .08. The authors found statistically different selection ratios for men and women in only one of the nine hypothetical selection ratios examined for both conscientiousness (when the selection ratio was .30) and extraversion (when the selection ratio was .60).

Despite the clear utility of personality assessment, there are still many detractors of high-stakes personality testing. Critics hold three major arguments against personality testing in the workplace - the comparably low validity of personality traits when predicting performance, the relative ease with which applicants can fake their responses,

and inconsistent relationships found between personality and performance. Each of these criticisms, and recent research on the topics, are addressed below to explore how these criticisms may be unfounded.

Criticism: Low Validity

Although meaningful relationships between individual personality traits and various performance criteria have been identified, these relationships have tended to be small to modest. Table 1 details the predictive validity coefficients found in some foundational meta-analytic investigations (Barrick & Mount, 1991; Hurtz & Donovan, 2000; Tett et al., 1991). The magnitude of relationships found in these early meta-analyses are certainly large enough to improve the quality of selection decisions (ρ = -.03 to .23; Barrick & Mount, 1991), though are small compared to the individual validity of general cognitive ability (r = .51, Schmidt & Hunter, 1998).

Table 1: Correlations between Personality Dimensions and Job Performance

Dimensions	Barrick & Mount (1991)[a]	Tett et al., (1991)[b]	Hurtz & Donovan (2000)[b]	Salgado (2003)[b]
Extraversion	.10	.13	.10	.07
Emotional Stability/ Neuroticism	.07	.19	.14	.16
Agreeableness	.06	.28	.13	.13
Conscientiousness	.23	.16	.22	.28
Openness to experience	-.03	.24	.07	.08

Note: [a] indicates that reported values are ρ for job proficiency, and [b] indicates the reported values are ρ for job performance.

Yet in most circumstances, multiple personality traits, rather than just one, are used to predict performance, with the expectation that each trait will account for some unique variance in the performance criterion. Thus, a combination of personality traits should predict more variance in performance than any single trait alone. The

intercorrelations of the Big Five personality traits are often of a small magnitude. Although varying by study, sample, and the context of testing, intercorrelations among the Big Five traits generally range from r = .07 - .29 for non-applicant, honest response conditions, to r = .13 to .55 (Barrick & Mount, 1996) and r = .06 to .37 (Hogan, Barrett, & Hogan, 2007) for applicant samples. Taking Barrick and Mount's (1991) meta-analytic results as our starting point, we can thus expect the mean relationship between some combination of personality traits and performance to be greater than or equal to ρ = .23, which is the validity of conscientiousness alone, when adding other traits into the predictive model.

Although the research has been limited, as most published research appears to examine the validity of traits individually rather than as a whole, there is evidence that although an individual trait's predictive validity may be low, additional traits add incrementally, resulting in a meaningfully higher validity for personality traits used together (Chiaburu et al., 2011; Hogan & Hogan, 1995; Zimmerman, Carmen Triana, & Barrick, 2010). Chiaburu et al. (2011) examined the incremental validity of the Big Five personality traits in predicting organizational citizenship behaviors (OCB) in their meta-analysis. Although conscientiousness and agreeableness have been widely investigated as predictors of OCBs (Hurtz & Donovan, 2000; Ilies, Fulmer, Spitzmuller, & Johnson, 2009; Organ & Ryan, 1995), Chiaburu et al. (2011) investigated whether emotional stability, extraversion, and openness to experience predicted specific types of OCBs – organization-directed (OCB-O), individual-directed (OCB-I), and change-oriented (OCB-CH). Above and beyond conscientiousness and agreeableness, they found that extraversion and openness predicted OCB-O (ΔR = .08), openness predicted OCB-I (ΔR

= .05), and emotional stability, extraversion, and openness predicted OCB-CH (ΔR = .08). As a group, the Big Five combined to account for much more variance than the strongest single-trait predictor of OCB. The Total R = .28 for OCB-O, compared to mean-r = .13 for conscientiousness alone, Total R = .27 for OCB-I compared to mean-r = .16 for conscientiousness alone, and Total R = .21 for OCB-CH compared to mean-r = .11 for Openness/Intellect alone.

Researchers also have found high validities for so-called compound personality traits, which combine facets of several different personality traits (Ones, Viswesvaran, & Dilchert, 2005). For example, integrity, which is a combination of conscientiousness, agreeableness, and emotional stability (Ones, Viswesvaran, & Schmidt, 1993), shows a moderate correlation with overall performance, ρ = .35. Customer service orientation, violence and aggression, and stress tolerance, all combinations of facets of agreeableness, emotional stability, and conscientiousness, related to performance in the .39-.41 range (Ones & Viswesvaran 2001a; Ones & Viswesvaran, 2001b). Finally, Ones, Hough, & Viswesvaran (1998) found that managerial potential, a combination of emotional stability, extraversion, and conscientiousness, also related to overall performance, ρ = .42. Thus, although individual personality traits account for less variance in performance than general cognitive ability, combinations of multiple personality traits account for incremental variance beyond cognitive ability alone, and a combination of traits can have criterion-related validities nearly as high as cognitive ability.

Criticism: Faking

Another concern with the use of personality tests in high-stakes selection contexts is the issue of faking. Sackett (2011) theorized that for any given score on a personality

test, the test-takers' response is based on at least six factors: the respondent's true trait score and the true trait score within a specific context, along with erroneous self-perception and impression management across contexts, as well as within specific contexts (e.g., the workplace or applicant context). Sackett identified the variance in responding associated with situationally specific impression management as faking. In general, this faking has been defined as intentionally distorting responses on a personality test to appear higher or lower on a trait than the respondent's true trait score (McFarland & Ryan, 2000). Presumably, applicants for a job are motivated to show themselves to be a relatively ideal candidate for the organization. As such, applicants could be tempted to distort their responses toward some conception of the type of person who best fits the target job. Typically, researchers have referred to this distortion as faking, but various researchers have called it impression management, response distortion, intentional distortion, social desirability, and dissimulation (McFarland & Ryan, 2000). Whatever the name used, it is clear that intention to distort responses is core to the construct.

Variance in Faking: A Model of Faking Behavior

Faking is not uniform, as some early researchers assumed (Viswesvaran & Ones, 1999). Rather, there is variance in the extent to which respondents fake their individual trait scores, as well as which traits are faked (McFarland & Ryan, 2000). McFarland and Ryan even found that respondents with certain personality profiles tended to fake to a greater extent than others, though this personality-driven faking is not well understood. The fact that there is variance in faking is important to understand because if applicant faking was uniform, there would be no impact of faking on criterion-related validities.

Several models of applicant faking behavior have been developed (Goffin & Boyd, 2009; McFarland & Ryan, 2000; Snell, Sydell, & Lueke, 1999; Tett & Simonet, 2011). The most comprehensive and widely-cited of these is McFarland & Ryan's (2000) model, which describes the content of faking as well as the way variables interact to drive faking on non-cognitive measures. First in the model are influences on beliefs toward faking, including values, morals, religion, true personality traits, etc. These influences then affect an individual's beliefs toward faking, which in turn determine an individual's intention to fake. The relationship between beliefs toward faking and intention to fake, however, is moderated by situational influences such as the desire for the job and warnings. Intention to fake's relationship with faking behavior is moderated by both the ability to fake (self-monitoring, knowledge of the construct being measured, and item transparency) and the opportunity to fake. Opportunity to fake addresses the limitation for fakers that those already high on the trait may not be able to positively distort their responses. Finally, the model asserts that faking behavior will influence a number of outcomes including validities, test scores, scale reliabilities, and personality's factor structure. Thus, faking is generally accepted as the combined outcomes of the variables in this model, reflecting a person's motives and values, as well as the context of test administration and beliefs about outcomes.

Concern over widespread applicant faking has led to extensive investigations into the effects of faking on the utility of personality tests over the past two decades. Much of the faking research can be organized into three categories – attempts to identify the prevalence of faking, the impact faking has on the predictive validity of personality traits, and the impact of faking on selection decisions.

Prevalence of Faking: Self-Reported Faking

Researchers have often assumed that within a set of applicant responses to a personality test, some applicants faked, but only rarely have investigated to what extent faking actually occurs. Donovan, Dwight, and Hurtz (2003) sought to address this gap in the literature by obtaining a base rate of applicant faking. Donovan et al. created a survey asking recent job applicants whether or not they had engaged in 29 faking activities during a recent job application. Participants reported faking on components relating to personality 27.8% to 53.8% of the time. As a comparison, respondents reported faking on the much more externally verifiable biographical components only 4.7% to 9.5% of the time. The study addressed personality traits such as hardworking, prompt, and thorough (45.2%), dependability and reliability (29%), agreeableness (27.8%), as well as downplaying negative attributes (53.8%; these negative attributes were not specified in the survey). Using another approach, Griffith, Chmielowski, and Yoshita (2007) had those who applied for a job retake a personality test a month after they took the test as applicants. Similar to Donovan et al.'s (2003) results, They found that 30-50% of respondents had elevated their scores as applicants, compared to when they responded more honestly. Thus, two different assessments of the prevalence of faking have both found that a third to half of applicants distorted their responses to some extent on personality assessments, demonstrating that applicant faking is a widespread phenomenon.

Research Methods in Faking

Before the impact of faking can be discussed, it is important to have a clear understanding of the differences in the methods used to investigate faking, as these

methodological differences may explain some of the inconsistent findings in faking research. Researchers have employed two major research methodologies when investigating faking, generally categorized under lab studies and field studies. In lab studies, instructional prompts are used to manipulate an individual's responses to a personality measure, whereas field studies tend to compare naturally occurring groups of respondents (e.g., applicant, non-applicants). The use of instructional prompts (as in lab studies) are employed when the purpose of the investigation is related to examining the maximum limits of faking, and the effect of such blatant distortions, whereas the comparison of different, naturally-occurring, groups (as in field studies) are used to examine the operational level of faking (Smith & Ellingson, 2002; Viswesvaran & Ones, 1999).

In lab studies, researchers typically create faking comparison groups by motivating respondents to elevate (or depress) the expression of the personality traits being measured through instructions and/or incentives. Instructions most often ask participants to blatantly distort their responses, in either a fake-good or fake-bad condition (Topping & O'Gorman, 1997; Hough et al., 1990; Viswesvaran & Ones, 1999; Zickar & Robie, 1999). In fake-good conditions, researchers instruct participants to choose their answers such that they make the most favorable impression they can on those doing the hiring for a generic job (Zickar & Robie, 1999). Similarly, in fake-bad conditions, researchers instruct respondents to try to look their worst through the personality assessment.

Alternatively, researchers have created a faking group by simulating the selection context either through instructions telling respondents to pretend they are applying for a

generic position, and that the assessment is important to the selection decision so it is important for the respondent to do his or her best (Dullaghan & Borman, 2009; Dullaghan & Joseph, 2010; Vasilopoulos, Reilly, & Leaman, 2000; Zickar & Robie, 1999), or through offering monetary or other incentives (e.g., a research assistant position) to the top-scoring respondents to the measure (Cunningham, Wong, & Barbee, 1994; Dwight & Donovan, 2003; Mueller-Hanson, Heggestad, & Thornton, 2003). Typically, in both types of lab studies, the pretend applicant is to respond as if they are applying for a generic job they really want, rather than the researchers specifying a specific job to be applied to. In field studies, the applicant group is classified as the faking group (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006; Rosse et al., 1998; Smith, Hanges, & Dickson, 2001) because applicants are highly motivated by the high-stakes context to make the best impression they can to get the job they are applying for (Rosse et al., 1998).

In both research designs, the faking group's responses are then compared to an honest group's responses. Experimental faking conditions are typically compared to an honest-response condition (or for-research-only condition), where no experimentally induced motivations to distort one's responses are present (Ellingson, Sackett, & Hough, 1999; Pace & Borman, 2006; Vasilopoulos et al., 2000; Zickar, Gibby, & Robie, 2004). Job applicant responses are typically compared to incumbents (Birkeland et al., 2006; Rosse et al., 1998; Smith et al., 2001; Tristan, 2009; Weekly, Ployhart, & Harold, 2003), because as a group, incumbents lack any clear motivation to distort their responses (Rosse et al., 1998). In both research designs, traits in which group means are

significantly elevated or depressed compared to the honest condition are considered to have been faked.

Investigations of faking utilizing the two different approaches have resulted in some consistent as well as discrepant findings. The majority of these findings can be grouped into three outcomes of faking - group mean trait score elevation, the criterion-validity of personality, and changes in the rank-order of respondents.

Impact of Faking on Mean Group Trait Scores

Faking on personality tests has been widely shown to elevate group mean trait levels in comparison to more honest response groups, but the extent of this elevation varies by personality trait (Barrick & Mount, 1996; Dwight & Donovan, 2003; McFarland, 2003; McFarland & Ryan, 2000; Rosse et al., 1998; Viswesvaran & Ones, 1999; Zickar & Robie, 1999). In their meta-analysis of between-subject fake-good studies, Viswesvaran and Ones (1999) found that on average, trait elevation ranged from d = .48 (agreeableness) to .65 (openness to experience). Within-subjects fake-good studies showed somewhat more variability in faking, with effect sizes ranging from .47 (agreeableness) to .93 (emotional stability). Individual studies have shown that there can be much more variance in faking between traits. For example, McFarland and Ryan (2000) found that the group mean trait scores for openness to experience were considerably less elevated (d = .19) under instructions to fake than the other Big Five traits, whereas conscientiousness was faked to the largest extent (d = 1.82).

Field studies likewise have resulted in consistently more positive mean trait scores for applicants than for incumbents, though these differences tend to be of a smaller magnitude than those seen in experimental manipulations. In their meta-analysis of

applicant responses across jobs, Birkeland et al. (2006) found small effect sizes for extraversion, openness to experience, and agreeableness (d = .11, .13, and .16, respectively), and moderate effect sizes for emotional stability and agreeableness (d = .44 and .45, respectively), with substantial variability in faking between studies. Thus, researcher have found that mean personality trait scores are consistently elevated for applicant groups compared to honest groups, but this elevation is not uniform across traits.

However, knowing that applicants can distort their responses on personality tests, and may do so due to motivational pressures to present a positive image of oneself, only matters if this distortion substantially affects the utility of these assessments in practical settings (Donovan et al., 2003; Ones et al., 1996). The two avenues of greatest concern through which faking can impact the utility of personality assessments are the criterion-related validity of distorted responses and the effect on the rank-order of candidates.

Effect of Faking on Criterion-Related Validity

There is a deep-rooted belief that faking reduces the criterion-related validities of personality assessments (Hogan, 2005). Yet much empirical evidence to date has shown that the criterion-related validities of personality assessments are robust to faking (Barrick & Mount, 1996; Ellingson et al., 1999; Hough, 1998; Hough et al., 1990; Ones et al., 1996; Ones et al., 1993; Schmitt & Oswald, 2006). The robustness of personality validity coefficients has been found despite researchers employing several operationalizations and utilizations of faking, including correcting scores for faking, removing fakers, and comparing the validities of personality traits for faking and honest groups.

Corrections for faking. Social desirability measures have been the most widely employed operationalization of applicant faking (Barrick & Mount, 1996; Holden, 2007; Hough et al. 1990; Kurtz, Tarquini, & Iobst, 2008; White, Young, & Rumsey, 2001). Researchers typically have inserted these measures within the personality measure being administered, and elevated scores are taken to mean that the respondent was dishonest on the assessment. Researchers use these measures to partial out the variance in personality responses associated with faking in an attempt to obtain more accurate estimates of the criterion-validity of personality tests (Smith & Ellingson, 2002).

Barrick and Mount (1996) investigated the personality-performance relationship for the Big Five traits utilizing two groups of job applicants in a predictive-validation study. Utilizing a commonly-used and well-validated measure of response distortion, Paulhus' Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984), Barrick and Mount (1996) partialled out the effect of self-deception and impression management from personality-performance validity coefficients. Contrary to the authors' hypotheses that partialling out the variance associated with response distortion would attenuate the correlations between personality and performance criteria (turnover and supervisory performance ratings), the adjustment for response distortion did not result in significantly lower correlations between any of the Big Five traits and job performance. There was no consistent pattern of validity changes, which ranged from Δr = -.08 to .08 across traits and criteria. For all but the agreeableness-turnover relationship, if the relationship was significant before the adjustment, it was still statistically significant, and in the same direction, after the adjustment. O'Connell, Kung, and Tristan (2011) added to this research by examining the impact of different measures of response distortion (i.e., social

desirability, a covariance index, and implausible answers) on the personality-performance relationship. They found that the change in criterion-related validity depended on which response distortion operationalization was used, again demonstrating an inconsistent effect of controlling for faking on the predictive validity of personality traits.

Similarly, Ones et al. (1996) meta-analyzed research on the impact of social desirability on personality-performance relationships. The researchers concluded that across all relevant research, social desirability did not function as a predictor (ρ = .01) nor mediator of job performance, nor did social desirability suppress the validity coefficients across any of the Big Five traits. Thus, the accumulated evidence suggests that faking, as operationalized through social desirability scales, does not significantly affect the criterion-related validity of personality assessments in the prediction of performance.

However, there have been many criticisms of the use of social desirability measures as an operationalization of faking. First, contrary to popular usage, most social desirability measures were not developed and evaluated based on their ability to recover honest trait scores through statistical corrections, but rather on their ability to identify those instructed to fake their responses (Ellingson et al., 1999). Although effective at identifying fakers, these measures may not be an effective measure of the variance in faking, as they were not developed as such.

Second, there is some covariation between personality traits (i.e., conscientiousness, emotional stability, and adjustment) and social desirability scales (Cunningham et al., 1994; Ellingson et al., 1999; Ones et al., 1996). Traditionally, researchers have treated this covariation as indicative of contamination (of faking) in personality scores, and have used social desirability measures to correct scores for faking.

However, more recently, researchers have viewed this covariation as true trait variance between personality and social desirability (Ones et al., 1996), in which case using these corrections would mean mistakenly removing meaningful variance in trait scores.

Investigations into the former view have determined that removing the contamination of socially desirable responding from personality scores by correcting for social desirability adjust the mean trait scores towards honest trait score levels (Ellingson et al., 1999). For some time, researchers took this finding as evidence that social desirability corrections effectively recover honest response scores. However, more recent research has found that the application of a correction, although effective at the group level, does not recover an individual's honest trait scores, nor their applicant rank-order (Ellingson et al., 1999).

Further, the practice of correcting responses for social desirability is based on the assumption that partialling out the variance associated with social desirability removes unwanted trait variance to provide more accurate, and presumably higher, estimates of criterion-related validity (Schmitt & Oswald, 2006). To examine this assumption, Schmitt and Oswald conducted a simulation in which they manipulated five variables - the correlation between the predictor and criterion, the correlation between the predictor and social desirability, the social-desirability-criterion correlation, selection ratio, and proportion of respondents identified as faking. The researchers found that the correlation of a faking measure with the criterion, as well as the proportion of respondents identified as fakers in the dataset each only accounted for about 3.0% of the variance in performance. The faking-predictor correlation accounted for a negligible amount of variance in performance. In contrast, differences in the magnitude of the correlation

between the predictor and criterion accounted for almost 60% of the variance in performance criteria. The minimal impact of the social desirability measure on the variance in criterion performance demonstrates that corrections for social desirability measures will not meaningfully improve the relationship between a predictor and performance. Thus, corrections for social desirability measures are inappropriate when a user's goal is to reduce the contamination of faking in personality scores in order to improve the predictive validity of personality.

In sum, the construct validity of measures of social desirability is in question. Although it is clear that correcting personality test scores for response distortion will bring the applicant group's trait score means down to honest group means, these corrections probably do not result in the correct trait scores for individuals, nor recover honest-response rank-order. As such, researchers have concluded that a social desirability correction will not allow for effectively adjusting faked personality scores (Ellingson et al., 1999), and have argued against using this method to evaluate the outcomes of faking (Burns & Christiansen, 2011). Thus, other approaches to examine the impact of faking on personality's criterion-related validity have been employed.

Removal of fakers. One such alternative approach to examine the effects of faking has been the removal of respondents identified as fakers from the applicant pool. Researchers have varied the methods of identification and removal employed. Hough (1998) administered a social desirability measure to a group of job incumbents (an honest comparison group), and found the score at which 5% of respondents scored at or above. She used this score as the cutoff for identifying which applicants faked their responses. Comparing the criterion-related validity for the full group with the validity when omitting

those scoring above the cutoff score, she found that the removal of fakers changed the concurrent validity of conscientiousness facet scores and the overall composite by only about r = .01 across three independent samples. Similarly, Schmitt & Oswald (2006) simulated a number of faking conditions and determined that the removal of suspected fakers (even up to 30% of the top-scorers) again only minimally influenced criterion-related validity across a large number of selection ratios and validities. In fact, the proportion of those identified as faking across conditions accounted for only about 3% of the variance in average criterion performance, compared to 59% of the variance for the validity of the predictor, and 23% for the selection ratio. Thus, the available evidence suggests that categorizing personality test respondents into fakers and non-fakers, then removing the fakers, has little to no impact on criterion-related validities.

Criterion-validity for applicants vs. non-applicants. A final method utilized to examine the generalizability of criterion-validity coefficients for applicant personality test responses is a direct comparison of the validity of an assessment for applicants to the validity for non-applicants. Although this approach does not allow researchers to directly separate all fakers from non-fakers, the comparison provides a more operational, rather than experimental, comparison of validities.

Hough et al. (1990) compared the criterion-related validity of honest and simulated Army applicant responses for a series of military performance criteria. Despite significant mean differences between honest and simulated applicant conditions comparable to Viswesvaran and Ones' (1999) meta-analytic results for experimental studies (d = .31 - .73), for the non-objective performance criteria, in only four of 22 cases did the criterion-related validities differ significantly between honest and applicant

response groups (Hough et al., 1990). In these four cases, the bivariate correlations differed by no more than Δr = .05. The rest of the honest and applicant validity comparisons differed by an even smaller and nonsignificant amount, Δr = -.01 to .03. Thus, the authors concluded that although response distortion resulted in mean-trait score differences, these differences generally did not affect the criterion-related validities of these scales. In the rare cases where validities were affected, the impact was negligible.

Early meta-analytic research on the impact of faking on the utility of personality in the workplace failed to test applicant/non-applicant status as a moderator of the validity of performance (Barrick & Mount, 1991; Hurtz & Donovan, 2000; Salgado, 1997). Bradley (2003) addressed this gap in the research by recoding all the articles included in these previous meta-analyses, as well as newer publications, including applicant/incumbent status as a moderator in his meta-analysis. Examining the Big Five personality traits, as well as the compound traits of optimism and ambition, he found that although the average validity for all traits but conscientiousness was greater for incumbents than applicants, none of these differences were statistically significant, and there was only a small amount of variance in the magnitude of validity coefficients across studies for both groups. Specifically, only small differences were found for neuroticism ($\Delta\rho$ = -.03), extraversion (($\Delta\rho$ = -.01), openness ($\Delta\rho$ = -.05), agreeableness ($\Delta\rho$ = -.03), conscientiousness ($\Delta\rho$ = .01), optimism ($\Delta\rho$ = -.02), and ambition ($\Delta\rho$ = -.05). Thus, they concluded that applicant/incumbent sample type is not a moderator of the criterion-related validity of personality traits.

Finally, Ones et al. (1993) conducted a targeted meta-analysis on the predictive validity of tests of integrity, which is highly related to conscientiousness, emotional

stability, and agreeableness (Berry, Sackett, & Wiemann, 2007; Murphy & Lee, 1994; Ones et al., 1993), in predicting overall job performance for applicants and incumbents. They found that the mean integrity-performance relationship was r = .24 (ρ = .40) for applicants and r = .17 (ρ = .29) for incumbents, a difference which was statistically significant. For applicant samples, variability in the validity coefficients was entirely explained by statistical artifacts (SD_ρ = .00), but for employees the variability was considerably higher (SD_ρ = .18), with only 42% of variance in the validity explained by the study's statistical artifacts. Although the validity was positive across studies for incumbents, these findings suggest that there may be other statistical artifacts or moderators that were not examined in the study impacting the validity of these assessments for incumbents. The authors did not identify any potential moderators that may be present for the incumbent personality-performance relationship. This trend of a higher validity for applicants than for incumbents was also found when including only supervisory ratings of performance as the criterion (ρ = .42 for applicants, and ρ = .33 for incumbents), as well as when controlling for research strategy (predictive/concurrent), which the authors were concerned could be confounded with the validation sample moderator. The authors viewed this consistent trend as evidence that integrity tests still predict performance despite respondents' motivations to distort their responses. The authors concluded that as integrity tests are more predictive for applicants than for incumbents, faking is not an issue in the selection context for integrity tests, and perhaps the highly related conscientiousness trait.

However, there have been two lab studies which have contradicted the robustness to faking of the personality-performance relationship. Although not their primary aim, in

a simulated applicant scenario where a financial incentive was offered for those who scored highest on the assessment (a variation of the fake-good lab methodology), Schmidt, Ryan, Stierwalt, & Powell (1995) found that the validity of conscientiousness and some of its facets depended on the context of administration. Specifically, there was a significant decrease in the personality-performance relationship in the faking condition compared to the honest condition for overall conscientiousness (r = .25 for honest, r = -.02 for faking), competence (r = .31 vs. r = -.02), achievement striving (r = .25 vs. r = -.10), and deliberation (r = .23 vs. r = .10).

Similarly, Mueller-Hanson et al. (2003) tested for validity differences between faked and non-faked responses using a simulated applicant (lab study) methodology. In their study, the researchers used trait achievement motivation as their predictor, and a simple cognitive performance test as their criterion. Participants were divided into two groups, an honest group and a faking group. Participants in the honest group were told they were completing the assessment for research purposes only. Participants in the faking group were told that high scorers on the assessment would be selected to participate in the second part of the study, and were told exactly which traits were being measured by the assessment (diligence, conscientiousness, and motivation). They were also told only those who qualified for the second part would be eligible for a cash prize, but were warned that dishonest responses would make them ineligible for the prize. Next, for both groups, the cash prize was given to one person in each of the groups of participants. Finally, all participants were told to do the second task, which functioned at the criterion performance measure. They were told that accuracy was most important, they would have as much time as they wanted to complete the task, and that they could

look up their scores on the performance test after they completed the testing session. Mueller-Hanson et al. found that the achievement motivation-performance relationship was stronger for honest respondents (r = .17) than for the fakers (r = .05), though the two correlation coefficients did not differ significantly. They then divided responses in the faking condition into thirds based on the predictor score distribution, and found that predictor-performance correlation was r = .45 for the lower third, but only r = .07 for the upper third. The two correlations differed significantly. Thus, there is limited evidence in lab investigations that blatant faking can attenuate the validity of personality assessments, especially at the high end of the test score distribution.

In summary, much empirical evidence has demonstrated that in general, neither partialling out variance associated with response distortion, removing responses which have been identified as faked, nor separating applicant (faking) from incumbent (honest) groups significantly reduces the criterion-related validity of personality traits. The few exceptions included two lab studies which elicited blatant response distortion, which led to significantly lower assessment validity (Mueller-Hanson et al., 2003; Rosse et al., 1999), as well as a meta-analysis in which integrity tests were found to be better predictors of overall performance for applicants than for incumbents (Ones et al., 1993). In conclusion, most evidence suggests that the criterion-related validity of personality tests are robust to faking, or at least are not negatively impacted by faking, in real world scenarios, as well as the majority of lab studies.

Effect of Faking on the Rank-Order of Respondents

As response distortion is not uniform across respondents (Birkeland et al., 2006; McFarland & Ryan, 2000), faking affects the rank-order of candidates. With the faking

group's means consistently higher than the honest group (Birkeland et al., 2006; Viswesvaran & Ones, 1999), it follows that a larger proportion of candidates from the faking group will be at the top of a rank-ordered list of test-takers than those in the honest group, and thus fakers will be more likely to be selected than non-fakers. Christiansen, Goffin, Johnston, and Rothstein (1994) found with a sample of supervisory participants completing a personality assessment for future selection, developmental, and other purposes, that when statistically correcting personality scores for response distortion, taken as their operationalization of faking, the rank-order changed for over 85% of candidates. This change in rank-order led to what the researchers called discrepancies in hiring decisions. Examining a series of selection ratios, Christiansen et al. determined that up to 16% of those hired would have been rejected if a correction for response distortion had been employed.

Mueller-Hanson et al. (2003) studied this rank-order concern in a lab setting. As discussed previously, for their faking group, the researchers used fake-good instructions and a financial incentive for the top-scorers, and even told respondents which traits were being assessed before test administration. An honest-response condition was created for comparison. Both groups were about equally sized to balance out the proportion of respondents from each group used in their combined analyses. Consistent with previous research, applicants in the faking group scored higher on the assessment than those in the honest condition, $d = .41$. Examining a variety of selection ratios, they found that as the selection ratio decreased, i.e., fewer respondents were selected, the proportion of respondents selected from the faking group increased. Thus, at more restrictive selection

ratios, a much larger proportion of those selected came from the faking group than the honest group (e.g., 64% fakers vs. 36% honest for the .10 selection ratio).

Similarly, using applicant and incumbent responses, Rosse et al. (1998) examined the proportion of applicants with severely elevated response distortion scores, defined as greater than 3SD above the mean for an incumbent sample, who would be selected with a top-down decision rule used on a personality test, under a series of selection scenarios. For their most restrictive selection ratio (.05), 88% of those who would have been hired had elevated response distortion scores. Even if half of all test-takers were selected, nearly a quarter of those hired would have had severely elevated response distortion scores. This increase in fakers selected has been found consistently in other research examining the effect of faking on the rank-order of those selected (Ellingson et al., 1999; Hough, 1998; Komar, Brown, Komar, & Robie, 2008). Thus, in practice, faking consistently affects the rank-order of applicants such that those who faked are selected at a higher rate than those who were honest in their responding.

Effect of Faking on Construct validity

In contrast to the general robustness of the criterion-related validity of personality assessments to faking, there is evidence that the construct validity of personality assessments can be affected by faking. To begin, there have been inconsistencies found in the intercorrelations of personality traits between methods and contexts (Ballenger, Caldwell-Andrews, & Baer, 2002; Douglas et al., 1996; Pauls & Crost, 2004). For example, faked responses tend to relate only loosely to observer ratings of personality (Ballenger et al., 2001; Pauls & Crost, 2005; Topping & O'Gorman, 1997). Ballenger et al. (2001) found that the correlation between self-reported personality under fake-good

conditions and other-rated personality was generally low for all Big Five traits, $r = .13$ for neuroticism, $r = .17$ for extraversion, $r = .23$ for openness, and $r = -.06$ for agreeableness. Of concern, they found a moderate negative relationship between faked and other-rated conscientiousness ($r = -.40$). For comparison, Connelly & Ones (2010) conducted a meta-analysis of the correlations between self and other ratings of personality. They found that the uncorrected mean correlation between sources was $r = .34$ for emotional stability, $r = .41$ for extraversion, $r = .34$ for openness, $r = .29$ for agreeableness, and $r = .37$ for conscientiousness. Thus, faked responses, at least in the lab setting, appear to be much less related to other-ratings of personality compared to more honest responses.

Additionally, some researchers have found that the intercorrelations of traits are consistently higher for fake-good groups compared to honest responses (Douglas et al., 1996; Pauls & Crost, 2005). For example, Douglas et al. (1996) found a statistically significant increase in the correlation between agreeableness and conscientiousness, from $r = .35$ for an honest condition to $r = .63$ for a faking condition. Pauls and Crost (2005) directly tested for increases in trait intercorrelations between honest, fake-good, and applicant responses. Nearly all fake-good traits were significantly more intercorrelated, with the exception of the openness relationship with some of the traits, likely due to openness' unclear relationship to performance for the majority of jobs. In contrast, Bradley (2003) meta-analyzed the intercorrelations among traits treating sample type as a moderator, and found that although in about a third of the comparisons sample type moderated trait intercorrelations, neither applicants nor incumbents showed consistently higher intercorrelations.

Further, some researchers have found differences in the factor-structure of personality assessments for applicant and non-applicant groups (Schmitt & Ryan, 1993; Weekley et al., 2003). There is some evidence that faking affects the psychometric properties of personality tests such that the Five-Factor structure does not always fit faked responses. Schmit and Ryan (1993) hypothesized that personality factor structure would depend on the purpose of the test administration (for-research-only or for selection). They found that the Five Factor structure fit student (for-research-only) samples better than applicant samples. In the applicant sample, an "ideal-employee factor" (Schmit and Ryan, 1993, pp. 971) appeared in the factor analysis, containing elements of all Big Five traits but openness. Similarly, Weekley et al. (2003) compared the factor structure of an assessment for applicant and incumbent groups. Although the same number of factors were extracted for both samples, the factor loadings differed meaningfully across groups. That is, factor loadings for the incumbent sample were nearly all greater than the factor loadings for the applicant sample, and when factor loadings were constrained to be equal across groups, there was a significant decrease in model fit compared to when factor loadings were free to vary ($\Delta\chi^2 = 860.79$, $p < .001$).

On the other hand, several researchers have found stability in the factor structure of multiple personality assessments between a number of honest, fake-good, and applicant samples (Ellingson, Smith, & Sackett, 2001; Marshall, de Fruyt, Rolland, & Bagby, 2005; Smith et al., 2001; Vecchione, Alessandri, & Barbarnaelli, 2012). Unexpectedly, in Smith et al.'s (2001) study, the model fit the applicant sample better than the student sample, in direct disagreement with Schmit and Ryan's (1993) findings. Even occasional variance in the factor structure of personality suggests that some other

construct may be operating in personality responses, though what this other construct is composed of has not yet been explicitly identified.

The accumulated criterion and construct validity evidence tells us that although in some contexts personality assessments may be measuring something different between honest and applicant response conditions, as seen through the occasional reduction of the construct-validity of the Big Five model, responses under both conditions are still generally predictive of performance. That is, the robustness of the personality-performance relationship to faking in all but blatant fake-good scenarios suggests that in high-stakes contexts, fakers may perform as well on the job as those with truly high trait levels, yet why fakers seem to perform well is still unclear. The challenge researchers now face is attempting to identify what, exactly, is being measured in applicant conditions. A detailed examination of the research on job-specific personality assessment may shed some light on what this other construct may be.

Job-Specific Personality Research

Although there is abundant evidence in the personality and career counseling literature to the contrary (Costa, McCrae, & Kay, 1995), personality researchers often implicitly assume that there is a specific personality profile that leads to effective performance across all jobs. This assumption is understandable if a researcher makes his or her decisions based solely on the meta-analytic work of Barrick and Mount (1991), Schmidt & Hunter (1998), and other researchers who have only examined faking behavior and the personality-performance relationship across jobs, while failing to examine differences in behavior and validity between jobs. According to these across-jobs meta-analyses, conscientiousness is universally predictive of performance, but the

other Big Five traits are generally only barely or negligibly related to performance. However, focusing on across-job meta-analytic findings masks meaningful differences in which personality traits are relevant, and thus predictive, for specific jobs.

Differences in the Importance of Traits by Job

The Occupational Information Network (O*NET) was developed to systematically examine and report the work and person characteristics of jobs, including personality traits important for effective performance in a given job (referred to as Work Styles in O*NET; Borman, Kubisiak, & Schneider, 1999). Clear differences in which personality traits (or behavioral tendencies) are important for success in different jobs can be readily seen in the database. For example, the most important works styles for police patrol officers are integrity, self-control, stress tolerance, attention to detail, and dependability (which map onto the Big Five traits of emotional stability and conscientiousness; a mapping of O*NET's work styles onto the Big Five traits can be found in Appendix A). In contrast, the most important work styles for preschool teachers are dependability, integrity, concern for others, and cooperation (which map onto conscientiousness and agreeableness). Conscientiousness is important to both jobs, but emotional stability is important for police officers, but not preschool teachers, whereas agreeableness is important for preschool teachers, but not police officers. Thus, personality profiles would clearly differentiate who would be more effective as a police officer from those who would be more effective as a preschool teacher.

Dunlop et al. (2012) obtained ratings from a variety of workers of the desirability of each trait level of a typical personality assessment for a variety of jobs. Although not directly examined in their article, new analyses of the data (P. Dunlop, personal

communication, December 22, 2011) showed that there are clear differences in which are desirable traits and trait levels for different occupations. Further, in many cases, what was desirable for a general worker (without any specified job tasks or job requirements) differed from what was desirable for specific jobs (nurse, firefighter, and car salesman, with a job description provided). For example, although raters thought the general worker should be high on agreeableness, they also thought a used car salesman should be significantly lower on the trait. Similarly, openness to experience was rated significantly less desirable for the general worker than for nurses. Thus, even the average worker recognizes that there are meaningful differences in which traits are important between jobs, and, although not necessarily explicitly aware of it, they also recognize that the general ideal worker profile may not be effective across all jobs.

Evidence of Tactical Faking When Applying for a Specific Job

Although all personality traits can be faked (Birkeland et al., 2006; Viswesvaran & Ones, 1999), individual applicants do not necessarily fake all traits during a given assessment. Although often overlooked in the faking literature, potentially due to the focus on early meta-analytic work, or only addressed as a side note in a study, applicants tend to be more tactical in their faking when given a specific job to apply to, compared to general fake-good investigations (Birkeland et al., 2006; Dullaghan & Joseph, 2010; Pauls & Crost, 2005; Raymark & Tafero, 2009). In directed faking studies, under experimental conditions, respondents are either told to do well, induced to do well, or told exactly what it takes to do well on the assessment. In all three conditions, it is implied that you will do well if you are a top-scorer, and in order to be a top-scorer, you have to score high (i.e., blatantly fake) on all of the positively-oriented traits being

measured, so the participant should select extreme responses to score the highest. On the other hand, in real-world applicant conditions, the goal is to be the ideal candidate, rather than the highest-scorer across dimensions. Rigorous personality-based job analyses have helped employers identify what are the desirable personality profiles for a specific job (Costa et al., 1995; Goffin et al., 2011), which may include high levels on some traits, but low levels on others (Raymark & Tafero, 2009).

Recall that the meta-analytic investigations on faking have demonstrated that respondents elevate all Big Five trait scores for fake-good (Viswesvaran & Ones, 1999) and applicant conditions (Birkeland et al., 2006). However, the few studies available reporting mean-differences between honest and faking groups for specific jobs have shown significant and meaningful mean differences in which traits applicants elevate between jobs. For example, both Dullaghan & Joseph (2010) and Pauls & Crost (2005) found that when students were instructed to complete a personality assessment as part of an application to a nursing position, applicants elevated their trait scores only for the traits rated most important for the nursing profession (as rated in O*NET) - agreeableness, emotional stability, and conscientiousness - with moderate to large effect sizes compared to honest responses. Likewise, Raymark and Tafero (2009) found that applicants for an accountant position depressed their trait scores on extraversion, openness, and agreeableness compared to an honest response condition, whereas general fake-good applicants elevated their scores on these traits. In all three studies, applicants clearly varied their faking in response to the simple presentation of a job title to focus their responding toward a specific job.

Pauls and Crost (2005) also compared the simulated applicant profiles obtained for manager and nurse applicants to a general fake good condition. Like Raymark and Tafero (2009), Pauls and Crost (2005) found clear evidence that job-specific applicant faking differs from general fake-good response distortion. Most often, the job-specific faking was more pronounced than the general, non-job-specific faking. Further, they found evidence that the faked profiles for nursing and managerial positions differed from each other, supporting the idea that respondents can and do differentiate between the personality requirements of different jobs.

Additionally, although Pauls and Crost (2005) found that nearly all traits were significantly more intercorrelated for the fake-good group than the honest group, only the traits rated as most relevant for the two focal positions examined showed higher intercorrelations compared to the honest condition. The greater amount of shared variance for specific traits, rather than all traits as in the fake-good condition, suggests that responses to items related to these traits were driven by some common factor. Thus, the faking seen between the two conditions differs, with one common factor, likely just general response distortion, driving up the intercorrelations between all traits for fake-good condition, but a narrower factor driving up the intercorrelations between only the job-relevant variables.

Using real applicant data, Birkeland et al. (2006) examined the mean differences between applicants to management/non-management and sales/non-sales job groupings. They found that the type of job applied for moderated the mean differences seen in a meta-analysis of applicant/incumbent personality assessment data, and these trait score elevations differed from that seen for the general applicant (that is, applicants to all jobs

for which they obtained effect sizes; Table 2). For example, applicants to sales positions elevated their extraversion scores to a large extent (*d* = .35), whereas applicants to all non-sales positions did not inflate their scores on this trait (*d* = .01). Further, sales applicants significantly depressed their agreeableness scores (*d* = -.20), whereas applicants to non-sales positions elevated their agreeableness scores (*d* = .27). Similarly, for applicants to management as well as non-management positions, emotional stability, extraversion, and openness were faked to the same extent. However, agreeableness scores were elevated for non-management positions (*d* = .22), but not for management positions (*d* = -.07). Finally, applicants to both sales and management positions elevated conscientiousness trait scores to a lesser extent than did the rest of the applicants (*d* = .13 for sales applicants, *d* = .18 for management applicants, and *d* = .45 for the general applicant). Thus, the experimental finding of clear differences in the traits that are faked between jobs can also be seen within real job applicant responses.

Table 2:Birkeland et al.'s (2006) Meta-Analytic Findings in SD Effect Size Units

	ES	E	O	A	C
General Applicant	.44	.11	.13	.16	.45
Management	.38	.16	.17	-.07	.18
Non-Management	.33	.11	.15	.22	.42
Sales	-.01	.35	.16	-.20	.13
Non-Sales	.41	.01	.16	.27	.40

Note: ES = Emotional Stability, E = Extraversion, O = Openness to Experience, A = Agreeableness, and C = Conscientiousness. All values reported are meta-analytically derived mean d effect sizes.

Variance in the Validity of Personality Traits between Jobs

Although there do not often appear to be significant differences between the validities of personality traits across applicant/faking and incumbent/non-faking conditions, there is abundant evidence of variance in the validity of traits between job

groups (Barrick & Mount, 1991; Dudley et al., 2006; Hurtz & Donovan, 2000; see Table 3 for a summary of the meta-analytic validity of the Big Five traits with performance by job type). In their meta-analysis, Barrick and Mount (1991) examined five job groups – professionals, police, managers, sales, and skilled/semi-skilled workers. The only trait which had a relatively stable validity coefficient across these five job groups was conscientiousness (ρ = .20-.23 across jobs). Looking beyond conscientiousness, for police, emotional stability, extraversion, and agreeableness were about equally predictive (ρ = .09-.10). For managers and sales positions, extraversion was the strongest predictor of performance (ρ = .18 and .15, respectively), but for professionals and skilled/semi-skilled workers, emotional stability was the strongest predictor (ρ = -.13 and .12, respectively). Hurtz and Donovan (2000) refined and updated Barrick and Mount's (1991) meta-analysis by only including studies which explicitly measured the Big Five traits, rather than measures of facets or conceptually similar constructs. Using only explicit measures, Hurtz and Donovan (2000) found more variability in the validity of conscientiousness between occupational categories, with higher validity for sales and customer service positions (ρ = .29 and .27, respectively) than for managers and skilled/semi-skilled positions (ρ = .19 and .17, respectively). They also found generally higher validities for all traits for sales positions. For managers, the validity of emotional stability increased (by $\Delta\rho$ = .05 to ρ = .13), but the validity for extraversion, openness, and agreeableness decreased. Results were likewise mixed for skilled/semi-skilled positions, with the validity decreasing slightly for emotional stability (by $\Delta\rho$ = -.03 to ρ = .09), but increasing for agreeableness (by $\Delta\rho$ = .05 to ρ = .11). In sum, clear differences in

the validity of traits to different job types can be seen across several meta-analytic investigations.

Researchers have found similar differences in validity even for the facets of the Big Five traits. For example, in their meta-analysis, Dudley et al. (2006) found that job type (customer service, sales, managerial, and skilled/semi-skilled) moderated the validity of conscientiousness' facets when predicting performance. Specifically, for most groups examined, the order facet was positively correlated with performance (ρ = .12-.21), but for managers, order had a negative relationship with performance (ρ = -.12). They also found that the cautiousness facet was unrelated to performance for sales positions and managers (ρ = -.04 and .01, respectively), but was negatively related to performance for skilled/semi-skilled positions (ρ = -.20). Although not as dramatic, there was also a slight effect of job type on global conscientiousness, such that conscientiousness was more strongly related to performance for sales (ρ = .29) and customer service jobs (ρ = .27), than managerial (ρ = .19) and skilled/semi-skilled positions (ρ = .17).

The variability in the validity of personality traits between job groups is often overlooked by researchers who only want to discuss personality traits which are universally predictive of performance. Such aggregation both understates the predictive validity of personality traits for specific jobs and masks important validity differences. For example, Barrick and Mount's (1991) overall meta-analytic findings tell us that across jobs, conscientiousness (ρ = .22) and extraversion (ρ = .13) are the best predictors of performance. However, selecting a candidate for a professional position based on their high scores on extraversion would be an error, as according to the meta-analytic results

Table 3: Meta-Analytic Findings on the Validity of the Big Five Personality Traits by Job

		r					*ρ*				
Occupational Group	Source	ES	E	O	A	C	ES	E	O	A	C
Across Jobs	Barrick & Mount (1991)	0.05	0.08	0.03	0.04	0.13	0.08	0.13	0.04	0.07	0.22
	Salgado (1997)	0.09	0.05	0.04	0.01	0.10	0.19	0.12	0.09	0.02	0.25
	Hurtz & Donovan (2000)	0.09	0.06	0.04	0.07	0.14	0.14	0.10	0.07	0.13	0.22
Customer Service	Dudley et al. (2006)	--	--	--	--	0.17	--	--	--	--	0.27
	Hurtz & Donovan (2000)	0.08	0.07	0.10	0.11	0.17	0.13	0.11	0.17	0.19	0.27
Managers	Barrick & Mount (1991)	0.05	0.05	0.05	0.05	0.13	0.08	0.18	0.08	0.10	0.22
	Barrick, Mount, & Judge (2001)	0.05	0.10	0.05	0.04	0.12	0.08	0.17	0.07	0.08	0.21
	Dudley et al (2006)	--	--	--	--	0.11	--	--	--	--	0.19
	Hurtz & Donovan (2000)	0.08	0.08	-0.02	-0.03	0.11	0.13	0.13	-0.03	-0.04	0.19
	Salgado (1997)	0.05	0.01	0.01	-0.03	0.06	0.12	0.05	0.03	-0.04	0.16
Police	Barrick & Mount (1991)	0.06	0.05	0.00	0.06	0.13	0.10	0.09	0.00	0.10	0.22
	Barrick, Mount, & Judge (2001)	0.07	0.06	0.02	0.06	0.13	0.11	0.06	0.02	0.10	0.22
	Salgado (1997)	0.10	0.08	0.08	0.06	0.15	0.22	0.20	0.18	0.14	0.39
Professionals	Barrick & Mount (1991)	-0.07	-0.05	-0.05	0.01	0.11	-0.13	-0.09	-0.08	0.02	0.20
	Barrick, Mount, & Judge (2001)	0.04	-0.05	-0.05	0.03	0.11	0.06	-0.05	-0.08	0.05	0.20
	Salgado (1997)	0.19	--	--	0.06	--	0.43	--	--	0.14	--
Sales	Barrick & Mount (1991)	0.04	0.09	-0.01	0.00	0.09	0.07	0.15	-0.02	0.00	0.23
	Barrick, Mount, & Judge (2001)	0.03	0.07	-0.01	0.05	0.11	0.05	0.09	-0.02	0.01	0.21
	Dudley et al (2006)	--	--	--	--	0.18	--	--	--	--	0.29
	Hurtz & Donovan (2000)	0.09	0.10	0.03	0.03	0.18	0.15	0.16	0.04	0.06	0.29
	Salgado (1997)	-0.04	-0.07	--	0.01	0.08	-0.07	-0.11	--	0.02	0.18
Skilled/Semi-skilled	Barrick & Mount (1991)	0.05	0.01	0.01	0.04	0.12	0.12	0.01	0.01	0.06	0.21
	Barrick, Mount, & Judge (2001)	0.06	0.03	0.03	0.05	0.12	0.11	0.05	0.04	0.08	0.19
	Dudley et al (2006)	--	--	--	--	.10	--	--	--	--	.17
	Hurtz & Donovan (2000)	0.06	0.00	-0.01	0.06	0.10	0.09	0.01	-0.02	0.11	0.17
	Salgado (1997)	0.11	0.05	0.08	0.03	0.09	0.25	0.08	0.17	0.05	0.23

for professional positions (Barrick & Mount, 1991; Barrick et al., 1995), professionals higher on extraversion would be expected to perform worse than those who were more introverted (ρ = -.09). Thus, failing to recognize and understand between-job differences in validity can lead to errors in selection decisions.

The one factor common to all of these job-specific studies was that a specific job context was provided before test administration to focus the participants' responses. Several researchers stated that providing this job title or description gave participants a context for responses, but did not attempt to delve deeper into how providing this context affected the respondent. Carefully reviewing the evidence across studies, the following conclusions can be made. First, in non-job-specific fake-good studies, all personality traits are faked. Second, in contrast, job-specific faking studies led to only certain traits being faked. Third, the specific traits faked varied by job, such that only job-relevant traits, identified through expert ratings, were elevated when a specific job context was provided to respondents. Thus, it appears that for all jobs examined, applicants used some sort of knowledge or stereotype (which may be considered a limited level of job knowledge; Mahar et al., 2006) of the jobs to generate ideal-worker profiles.

Different Types of Faking

Pauls and Crost (2005) attempted to set the foundation for future investigations into faking. In honest contexts, Pauls and Crost argued that a person's true personality is being measured, in which case assessment scores should effectively predict performance. In the applicant setting, however, they argued that part of what is being measured is true personality, but a substantial portion of the variance in personality scores is due to some personality-irrelevant variable. They propose that in contrast to the general fake-good

condition, which calls for the applicant to attempt to look good on all traits, job-specific applicant responses would be more differentiated, and in line with the specific job's requirements. This proposition has been supported by a growing body of evidence in the job-specific faking literature (Birkeland et al., 2006; Dullaghan & Joseph, 2010; Pauls & Crost, 2005; Raymark & Tafero, 2009).

Building upon Pauls and Crost's (2005) proposition, I propose that evidence has accumulated suggesting that there are two types of response distortion seen in the faking literature. First, there is blatant faking, in which testing conditions motivate respondents to look as generally socially desirable as they can across all traits being assessed. This type of faking is elicited by non-job-specific fake-good studies, as well as studies in which incentives are offered for the highest scorers on the assessments. Second, there is knowledgeable faking, in which faking is driven by some level of knowledge of the requirements of the specific job for which the assessment is being taken. This type of faking is elicited in job-specific applicant simulations as well as real-world applicant settings.

Job knowledge has been defined as the technical information, facts, principles, and procedures required to do a job (Palumbo, Miller, Shalin, & Steele-Johnson, 2005; Schmidt, Hunter, & Outerbridge, 1986). Understanding of a specific job's requirements may be acquired through personal experience, education, or training. Such knowledge should tell a person what sorts of behaviors need to be exhibited on the job for the effective execution of job duties. As most personality assessments are behavior-based, applicants knowledgeable about specific positions should be better able than others to recognize which behaviors, and which levels of these behaviors, will lead to the effective

completion of the target role's job duties (Goffin & Boyd, 2009; Levashina & Campion, 2007; Snell et al., 1999).

Applicants may use their knowledge of the requirements of a specific job to guide their responses to a personality assessment to present themselves as the ideal candidate for the position. In this sense, personality assessments are a type of performance test, where the criterion for performance is how well the respondent's profile matches that of the ideal candidate for a specific job (Johnson & Hogan, 2006; Schmit & Ryan, 1993; Tett & Simonet, 2011). What is required to perform well on the test is the knowledge of how one needs to behave on the job to be successful, and then the test taker answers the assessment accordingly.

Across different research contexts, respondents appear to utilize different response processes to answer the personality test. Some are immune to or ignore the context of administration, and respond based on their personality across contexts, just as in an honest response context (honest respondents). Others respond to experimental manipulations to blatantly distort their responses across all traits as they were instructed to do, as in fake-good and incentive studies (blatant fakers). Still others appear to provide more targeted responses that reflect their knowledge (either through stereotypes, training, or experience) of effective personalities on the job (knowledgeable fakers).

Different research contexts are likely composed of different combinations of honest respondents, blatant fakers, and knowledgeable fakers. In honest conditions, in which there is no apparent motivation for respondents to fake, it is likely that all respondents are honest. In non-job-specific fake-good contexts, it is likely that a small proportion of the respondents ignore instructions and are honest responders, but the

majority of respondents will be affected by the experimental manipulation and fake across all traits assessed, as evidenced by the large mean trait elevations seen in fake-good studies (Viswesvaran & Ones, 1999). In job-specific response contexts, responses are likely composed of a few honest responders, who again are unaffected by the context instructions, a few blatant fakers who try to score highest rather than appear to be the ideal candidate for the specific job, and a large number of knowledgeable fakers, who use their stereotypes or real knowledge of effective behavior for the job to guide their responses. Finally, in real-world applicant contexts, there will be some honest responders who are still unmotivated to fake their responses (for any number of the reasons provided in McFarland & Ryan's (2000) model of faking). There will also be some blatant fakers, who try to look as generally desirable as they can, as well as many knowledgeable fakers who provide more targeted faking, based on their knowledge of effective personalities in their roles. With sophisticated and valid blatant faking detection methods incorporated within, or added to, the majority of the personality assessments currently in use, it is probable that many of the blatant fakers will be identified as faking, and will have had their assessment scores invalidated, thus removing them from consideration. The remaining group of respondents should be composed of honest respondents and knowledgeable fakers, both of whom should be effective on the job as they either naturally exhibit job-effective behaviors, or know how they should act on the job.

Although job knowledge has been previously been identified as one of many factors influencing applicant faking in two models of applicant faking behavior put forth (Goffin & Boyd, 2009; Snell et al., 1999), job knowledge has only indirectly, or

superficially, been tested as an antecedent of faking. This study will seek to further explore and test this antecedent of faking.

"Faking" Personality on the Job

The primary driver of the use of personality tests is their utility in predicting later performance. If an applicant can behave consistent with a successful worker's profile, the validity evidence suggests that that applicant should perform and behave just as effectively on the job as someone who honestly had that personality profile, with the exception of experimentally manipulated extreme cases of faking. Replication of the personality profile of a successful worker (often called faking) for any specific position would require solid knowledge of effective on-the job behaviors, regardless of the applicant's trait scores in non-work contexts. Therefore, in contrast to some researcher's concerns (Christiansen et al., 1994; Mueller-Hanson et al., 2003; Rosse et al., 1998) selecting a knowledgeable faker (though perhaps not a blatant faker) based on rank-order may not be an error in terms of expected performance.

Understandably, many readers will be concerned with the proposition that any type of faker could perform as well on the job as a non-faker whose true trait scores match the desirable profile for the job. However, knowledgeable fakers would not necessarily have much difficulty expressing the desirable traits for a specific role on the job for two main reasons. First, effective performance on the job may not require constant expression of desirable levels of desirable traits. Desirable trait expression may only be required in situations which demand the expression of these traits in order to effectively accomplish the task. Second, as all job-specific personality and faking research has shown, fakers would not need to have a completely different personality on the job.

Instead, the faker would have to express higher (or lower) levels of the few most important traits for performance on the job. For example, for manager roles, research suggests that only higher levels of conscientiousness and extraversion have an impact on performance (Barrick & Mount, 1991; Hurtz & Donovan, 2000). Thus, the manager could still express his or her true trait levels on openness, agreeableness, and neuroticism, without having an impact on performance.

The Current Study

This study answers Birkeland et al., (2006), Christiansen et al. (2010), and Raymark and Tafero's (2009) call for additional research to be conducted examining the variance in faking between jobs, as well as for a deeper investigation into what is driving job-specific faking. Much of the job-specific research to date has been exploratory in nature with job types as moderators of the personality-job performance relationship (e.g., Birkeland et al., 2006), with only a few researchers hypothesizing specific trait elevation for specified jobs (Raymark & Tafero, 2009; Vasilopoulos et al., 2000). I examined job-specific applicant faking in a more direct and empirical manner by examining whether the level of job knowledge an applicant has directly impacts the traits faked to further the investigation of the substance of applicant faking.

Applicant Faking between Jobs

The current body of literature on job-specific faking is limited by a lack of research investigating a variety of jobs. Using several applicant scenarios, I compared personality trait elevation during the selection process for a variety of positions. To increase the likelihood that participants would be naïve of the personality requirements of the focal jobs, and thus only be able to draw upon the job knowledge manipulation for

information about the jobs, I selected three jobs with which participants were unlikely to be familiar.

I used the widely used FFM-based Hogan model of personality (Hogan Assessments, 2002) in this study. In the Hogan model, extraversion is represented by sociability and ambition, agreeableness by interpersonal sensitivity, conscientiousness by prudence, neuroticism by adjustment, and openness by inquisitive and learning approach. Borman et al. (1999) discussed how the O*NET's Work Styles map onto the Hogan traits. For this investigation, I selected jobs listed in the O*NET database based on how different the personality requirements of the jobs were, as well as how likely it would be that a respondent would already have significant knowledge of the personality requirements of the jobs. I then used these selected jobs in the first part of this investigation, Study 1, with the purpose of identifying a set of three jobs that respondents would be unfamiliar with for inclusion in Study 2, which simulated a portion of the applicant scenario for these jobs.

Study 1

Study 1 sought to identify obscure jobs for inclusion in this study by having participants rate their familiarity with ten different jobs – accountant, air traffic controller, compliance manager, computer systems analyst, industrial/organizational psychologist, intelligence analyst, legal secretary, marketing manager, police patrol officer, and property claims insurance examiner. These jobs were selected because, although participants may be somewhat familiar with the jobs, they are unlikely to have substantial job knowledge about the jobs, nor be readily able to identify the personalities needed to be successful on the job. Further, the selected jobs varied on which traits were

identified as most important in O*NET. As expected, based on the extant personality research, prudence was important for all jobs. Interestingly, sociability was not identified as important for any of the jobs. Participant familiarity ratings were used to identify the three jobs with which participants were least familiar. Participants also rated the importance of each trait for Study 1's jobs to provide an empirical check against O*NET's aggregate importance ratings (reported in Tables 4-5).

*Table 4: O*NET's Work Style Ratings for Study 1's Jobs*

	Accountant	Air Traffic Controller	Compliance Manager	Computer Systems Analyst	Industrial/ Organizational Psychologist
Achievement/Effort	81	83	76	72	90
Adaptability/Flexibility	70	84	85	88	88
Analytical Thinking	84	77	85	93	93
Attention to Detail	95	94	96	95	80
Concern for Others	67	53	55	70	73
Cooperation	81	78	80	83	89
Dependability	85	90	95	90	92
Independence	78	65	72	75	86
Initiative	80	79	82	78	90
Innovation	67	66	63	84	76
Integrity	94	74	99	90	91
Leadership	68	67	70	60	82
Persistence	78	81	82	76	73
Self Control	74	82	72	73	82
Social Orientation	54	62	47	66	78
Stress Tolerance	78	96	84	77	80

*Table 4 (continued): O*NET's Work Style Ratings for Study 1's Jobs*

	Intelligence Analyst	Legal Secretaries	Marketing Manager	Police Patrol Officer	Property Claims Insurance Examiner
Achievement/Effort	81	76	81	71	73
Adaptability/Flexibility	81	80	81	81	80
Analytical Thinking	98	67	71	77	80
Attention to Detail	97	95	87	92	87
Concern for Others	51	74	68	82	73
Cooperation	77	88	86	86	83
Dependability	85	92	89	89	88
Independence	75	81	80	79	81
Initiative	86	83	83	83	79
Innovation	77	58	77	65	63
Integrity	95	93	85	95	96
Leadership	60	67	84	80	71
Persistence	78	76	82	78	81
Self Control	66	84	76	93	85
Social Orientation	50	68	72	74	58
Stress Tolerance	72	84	80	92	87

*Note: The table above reports all Work Style scores for each focal job from the O*NET.*

Table 5: Most Important HPI Traits for Each Focal Position

HPI Trait	Accountant	Air Traffic Controller	Compliance Manager	Computer Systems Analyst	Industrial/ Organizational Psychologist
Adjustment	74.00	87.30*	80.30*	79.30	83.30*
Ambition	74.00	73.00	76.00	69.00	86.00*
Inquisitive	75.50	71.50	74.00	88.50*	84.50*
Interpersonal Sensitivity	74.00	65.50	67.50	76.50	81.00*
Learning Approach	81.00*	83.00*	76.00	72.00	90.00*
Prudence	86.00*	80.80*	88.80*	85.20*	84.40*
Sociability	54.00	62.00	47.00	66.00	78.00

Table 5 (continued): Most Important HPI Traits for Each Focal Position

HPI Trait	Intelligence Analyst	Legal Secretaries	Marketing Manager	Police Patrol Officer	Property Claims Insurance Examiner
Adjustment	73.00	82.70*	79.00	88.70*	84.00*
Ambition	73.00	75.00	83.50*	81.50*	75.00
Inquisitive	87.50*	62.50	74.00	71.00	71.50
Interpersonal Sensitivity	64.00	81.00*	77.00	84.00*	78.00
Learning Approach	81.00*	76.00	81.00*	71.00	73.00
Prudence	86.00*	87.40*	84.60*	86.60*	86.60*
Sociability	50.00	68.00	72.00	74.00	58.00

*Note: O*NET importance ratings were averaged after matching O*NET Work Styles to the higher-level HPI traits to determine the rank-order of importance for each of the traits for each job. A '*' indicates a given trait was identified as most important, using a cutoff score of 80 on a 100-point rating scale of trait importance.*

Study 2

Similar to Vasilopoulos et al. (2000), I conducted an experiment in which I manipulated the amount of job information provided to respondents with little to no experience in the focal jobs. This manipulation involved providing participants with a job description for the unfamiliar jobs, then asking them to complete a personality assessment as if they were applying for that job. Providing respondents with a job description increased the amount of job-relevant information respondents had, making them more expert on the position, and thus more similar to an applicant who has had experience or training in the position. I expected that those who read a detailed job description will elevate the trait scores most relevant to the position when compared to an honest response condition.

For Study 2, O*NET's Work Style ratings (as detailed in Tables 4 and 5, above) guided hypothesized mean trait elevation for the three jobs with which participants in Study 1 were least familiar.

Hypothesis 1a: For applicants to compliance manager positions, adjustment and prudence will be elevated for applicants to the focal position compared to honest responses.

Hypothesis 1b: For applicants to computer systems analysts positions, inquisitive and prudence will be elevated for applicants to the focal position compared to honest responses.

Hypothesis 1c: For intelligence analyst positions, inquisitive, learning approach, and prudence will be elevated for applicants to the focal position compared to honest responses.

In order to ensure that participants used the information provided in the job descriptions to guide their responding to the personality assessment, and thus engaged in knowledgeable faking, I inserted some counter-intuitive job information into one of the focal job descriptions. Table 6 details which trait O*NET reported as least important for each of Study 1's ten jobs. For most of the jobs (eight of the ten), sociability was to be the least important trait, followed by inquisitive (two of the ten), and learning approach (tied with inquisitive for police patrol officers). For all three jobs chosen for Study 2, sociability was the least important trait. Among the jobs, sociability was rated lowest for compliance managers (47), so task information related to sociability was inserted into the compliance manager job description.

Table 6: Identification of the Least Important Trait for Each of Study 1's Jobs

Job	Lowest Rated Trait	Trait Rating Value
Accountant	Sociability	54.00
Air Traffic Controller	Sociability	62.00
Compliance Manager	Sociability	47.00
Computer Systems Analyst	Sociability	66.00
Industrial/ Organizational Psychologist	Sociability	78.00
Intelligence Analyst	Sociability	50.00
Legal Secretaries	Inquisitive	62.50
Marketing Manager	Sociability	72.00
Police Patrol Officer	Inquisitive/Learning Approach	71.00
Property Claims Insurance Examiner	Sociability	58.00

Note: The importance rating for all traits was on a 100 point importance rating scale.

Hypothesis 2: The counter-intuitive trait of sociability will be elevated for respondents given the compliance manager job description, in addition to the other job-relevant traits.

Relationship between Familiarity with Job and Trait Importance

Additionally, I used the data collected in Study 1 and Study 2 to examine the relationship between familiarity with a job and the reported importance (from Study 1) or trait elevation (from Study 2) of personality traits for that job. In Study 1, I expected that those familiar with each job would rate the most important traits (based on O*NET ratings) as more important for the job than those unfamiliar with the job. However, those unfamiliar with the job should not be sure how necessary a trait is for the job. Thus, I expected some amount of heteroscedasticity in the relationship between job familiarity and the importance of personality traits for the job such that the lower end of the familiarity rating scale would have more variance in importance ratings, whereas the higher end of the familiarity scale would have less variance in importance ratings. If

heteroscedasticity was found for trait ratings between those familiar and unfamiliar with the job, I would utilize a transformation for the personality trait importance before analyzing relationships. Hypothesized relationships were based on the O*NETs trait importance ratings for each of the jobs in the Study 1 (detailed above in Tables 4 and 5) with the expectation that there would be a strong relationship between familiarity and the importance rating of important traits for each job. Expected relationships are hypothesized below, and summarized in Table 7. As prudence was considered important across nine of the ten jobs examined, it can be considered generally job-relevant. As such, no relationship between familiarity and prudence was expected nor hypothesized.

Hypothesis 3a: For accountants, there should be a positive relationship between job familiarity and learning approach importance.

Hypothesis 3b: For air traffic controllers, there should be a positive relationship between job familiarity and adjustment and learning approach.

Hypothesis 3c: For compliance manager, there should be a positive relationship between job familiarity and adjustment.

Hypothesis 3d: For computer systems analysts, there should be a positive relationship between job familiarity and inquisitive.

Hypothesis 3e: For industrial/organizational psychologists, there should be a positive relationship between job familiarity and learning approach, ambition, inquisitive, adjustment, and interpersonal sensitivity.

Hypothesis 3f: For intelligence analysts, there should be a positive relationship between job familiarity and inquisitive and learning approach.

Hypothesis 3g: For legal secretaries, there should be a positive relationship between job familiarity and adjustment and interpersonal sensitivity.

Hypothesis 3h: For marketing managers, there should be a positive relationship between job familiarity and ambition and learning approach.

Hypothesis 3i: For police patrol officers, there should be a positive relationship between job familiarity and adjustment, ambition, and interpersonal sensitivity.

Hypothesis 3j: For property claims insurance examiner, there should be a positive relationship between job familiarity and adjustment.

Table 7: Summary of Hypotheses 3a-j

H:	Adjustment	Ambition	Inquisitive	Interpersonal Sensitivity	Learning Approach
3a: Accountants					+
3b: Air Traffic Controllers	+				+
3c: Compliance Manager	+				
3d: Computer Systems Analyst			+		
3e: Industrial/ Organizational Psychologist	+	+	+	+	+
3f: Intelligence Analyst			+		+
3g: Legal Secretary	+			+	
3h: Marketing Manager		+			+
3i: Police Patrol Officer	+	+		+	
3j: Property Claims Insurance Examiner	+				

Note: Hypotheses tested using a Bonferroni correction to control for Type 1 error.

For Study 2, I expected that there would be a relationship between familiarity with the focal job and trait elevation for each job's important traits. Hypothesized relationships between job familiarity and traits would be identical to Hypothesis 2, above, with the expectation that job familiarity for the three focal jobs would be related to actual trait elevation, rather than importance ratings. That is:

Hypothesis 4a: Familiarity with compliance managers will result in higher trait elevation for adjustment.

Hypothesis 4b: Familiarity with computer systems analysts will result in higher trait elevation for inquisitive.

Hypothesis 4c: Familiarity with intelligence analysts will result in higher trait elevation for inquisitive and learning approach.

Examination of Rank-Order

In line with recent research on faking (Christiansen et al., 1994; Ellingson et al., 1998; Komar et al., 2008; Mueller-Hanson et al., 2003; Rosse et al., 1998), I also examined the effect of faking on the rank-order of candidates to assess the fairness of hiring decisions based on applicant personality data. I expected that the rank order of respondents completing a personality assessment in a simulated applicant scenario and provided with job knowledge (through either a job title or a job description) would differ meaningfully from their rank-order when given honest responses. This rank-order investigation allowed for comparison of the present study's findings with previous research examining the effect of faking on the rank-order of applicants.

Exploratory Question: What effect does the direct manipulation of job knowledge have on the rank-order of applicants to a specific position?

Chapter Two:

Method

Study 1

Participants. Participants (N = 185; 80% female) in Study 1 were recruited through the SONA online survey system at a southeastern US university, of which five were omitted from analyses due to uniform responding. All students registered for the SONA system were eligible to participate. The only restriction within SONA is participants have to be at least 18 years old.

Measures. *Familiarity Scale:* Participants rated their familiarity with each of the ten jobs in Study 1 by answering the question "How familiar are you with the job of *[focal job title]*?" on a 5-point Likert-type scale with anchors 1 – Not at all Familiar to 5 – Extremely familiar. All participants rated their familiarity with all ten jobs.

Personality ratings: Participants rated the importance of each of the HPI's six personality traits. For a random subset of six of Study 1's jobs, participants were asked "How important are the following personality traits for the job of *[focal job specified here]?*" for each personality trait, and responded on a five-point Likert-type scale with response options 1 – Not at all important to 5 – Extremely important. Participants were presented with only six of the ten jobs to be rated in order to avoid survey fatigue. In order to inform their ratings, participants were provided with definitions of each trait (detailed in Table 8) for each job they provide ratings for. To further reduce the effect of

survey fatigue, the order in which jobs were presented for personality ratings was randomized for all participants.

Table 8: Definitions for Each HPI Trait

Traits and Definitions
Adjustment: confidence, self-esteem, and composure under pressure
Ambition: initiative, competitiveness, and desire for leadership roles
Sociability: extraversion, gregarious, and need for social interaction
Interpersonal Sensitivity: tact, perceptiveness, and ability to maintain relationships
Prudence: self-discipline, responsibility and conscientiousness
Inquisitive: imagination, curiosity, and creative potential
Learning Approach: achievement-oriented, stays up-to-date on business and technical matters

Procedure. Participants for Study 1 signed up for the study through the SONA online data-collection system. The system required participants to sign into the website, and then select the study. Before beginning the study, participants were presented with an electronic informed consent. Participants were informed that they were free to end participation in the study at any time. Consent was indicated if participants continued to the study. Upon agreeing to participate, participants were rewarded points in the SONA system, which the Psychology department's instructors use to assign course credit to students.

Participants were then presented with the study's ten jobs and asked to rate their familiarity with each job. Next, they were presented with a random selection of six of the ten jobs and asked to rate how important they thought each of the seven HPI personality traits were for each job.

Study 2

Participants. I recruited participants for Study 2 from the MTURK system hosted by Amazon. This system is used by a growing number of researchers and companies to reach a wide group of participants, and has been found to be more demographically diverse than typical college samples (Buhrmester, Kwang, & Gosling, 2011). 399 participants completed the study, of which 49 were screened out due to random or uniform responding, resulting in 350 being included for analysis. Participants were compensated $.50 per response as an incentive for participation. As of Jan 12, 2013, only 483 of the total 1819 studies (25.55%) on MTURK compensated participants more than $.50 per response, suggesting that the compensation given in this study was substantially larger than the compensation offered for the majority of studies in the MTURK system.

Measures. *Personality Test IPIP-HPI:* I used the International Personality Item Pool (IPIP) (Goldberg et al., 2006) as my personality assessment tool for Study 2. The IPIP has a number of scales that Goldberg et al. (2006) developed which correspond to a number of widely-used personality assessments. The Hogan Personality Inventory (HPI; Hogan Assessments, 2002) is a widely-used measure of personality in the workplace. The IPIP-HPI has seven high-level constructs that the researchers developed to correspond to the HPI scales (Table 9). The correlations between IPIP-HPI and HPI scales are quite high (ranging from r = .66 - .77), especially after correcting for unreliability (ρ = .83 - .99). The developers selected the IPIP items included in the IPIP-HPI assessment based on rank-ordered correlations with scores on the HPI traits. Items showing the highest correlations with each trait were rank-ordered for inclusion in the final measure. The developers performed a visual content analysis to identify items that addressed the same

construct. If they found two items that were too similar in content, they removed the item with the lower correlation, and the next item in the rank-order was included. The top ten items which remained after this two-step process were included in the final measure for the trait, resulting in ten items for each of the IPIP-HPI traits. Participants responded to the IPIP on a five-point Likert-type scale, with anchors 1 = Very Inaccurate, 3 = Neither Inaccurate nor Accurate, and 5 = Very Accurate.

Table 9: Comparison of IPIP and HPI Scales

IPIP-HPI Scale	IPIP Alpha[a]	HPI Scale	HPI Alpha[a]	Correlation	Study 2 IPIP-HPI Alpha[b]
Stability	.86	Adjustment	.87	.74 [.86]	.79
Leadership	.82	Ambition	.87	.77 [.91]	.81
Sociability	.75	Sociability	.83	.73 [.93]	.77
Friendliness	.86	Interpersonal Sensitivity	.74	.67 [.84]	.84
Dutifulness	.78	Prudence	.69	.67 [.91]	.73
Creativity	.83	Inquisitive	.78	.67 [.83]	.79
Quickness	.82	Learning Approach	.76	.64 [.81]	.80

Note: [a] indicates calculations in the table come from the Eugene-Springfield Community Sample (Goldberg et al., 2006). Numbers in brackets are correlations when corrected for unreliability. [b] indicates reliability calculations based on the honest response condition in this study.

Social Desirability Scale: In order to be able to compare the effects of my manipulation of job knowledge with previous research operationalizing faking through the use of social desirability measures, I included the IPIP's-Unlikely Virtues (IPIP-UV) scale in my experimental investigation. The scale includes a number of positive and negative attributes, to which extreme positive responding is suspect. The measure is 17 items (α = .83), with the same response format as the IPIP-HPI, enabling the measure to be embedded in the personality assessment. Sample items include "Never give up hope" and "Will do anything for others."

Job Descriptions: Job descriptions for the job knowledge manipulation were developed from the O*NET's job descriptions of the positions selected from the applicant dataset. For each, a brief job description was developed, followed by a listing of the key tasks performed in the positions (Appendix B). The three developed for Study 2 were based on the jobs selected from Study 1.

Rather than providing detailed personality information in a brief job description as some researchers have done when simulating an applicant context (Krahe, Becker, & Zollter, 2008; Mueller-Hanson et al., 2003), my study sought to better replicate a real-world applicant condition by providing a more standard job description, providing a brief description of the job, and detailing the job's core tasks, rather than explicitly stating the personality needed for the job. By providing only job task and competency information, applicants would have had to make the same job-task to personality inferences as actual applicants do, thus increasing the external validity of the investigation.

Familiarity Scale: After completing the personality assessment, participants rated their familiarity with the focal jobs on a two item assessment of familiarity. Participants were first asked "How familiar were you with the job of *[focal job title]* before reading the job description?", then "How familiar were you with the job of *[focal job title]* after reading the job description?" Participants responded to both items on a 5-point Likert-type scale, with anchors 1 – "Not at all Familiar" and 5 "Extremely Familiar".

Counter-Intuitive Trait Information: Counter-intuitive trait information was also placed into one of the job descriptions in Study 2. The condition counter-intuitive information was added to was determined by examining the O*NET ratings for each of Study 2's selected jobs. As all three jobs selected for Study 2, compliance manager,

computer systems analyst, and intelligence analyst, had sociability as the lowest-rated traits in O*NET, information related to sociability was added to a job description. The compliance manager job description was chosen because of the three jobs selected for Study 2 based on Study 1's familiarity ratings, sociability was rated lowest for compliance managers, and thus was viewed as the least relevant for the job. The text added was: "The compliance manager works with peers and clients on a daily basis in a team-based environment requiring ongoing social interaction."

Procedure. To assess whether job knowledge influenced which traits are faked, job information was manipulated through an experiment. Participants signed up for the study through an online data-collection system, MTURK. The system required participants to sign into the website, and then select the study in which to participate. After being presented with the informed consent form, all participants completed the personality assessment under honest conditions first, and then were randomly assigned to one of the three job description conditions. All participants who consented to participate were granted credit, and compensation, for participating.

For the Honest condition, participants read the following prompt:

Honest condition:

> *You are about to take a personality test. As you answer the following questions, please be as honest as you can. Your responses will be used for research purposes only. There will be no identifying information kept with your responses, and all responses will be kept strictly confidential. Honest answers will help us to get an idea of the typical person's true personality.*

For the job applicant conditions, participants read the following:

Job Description condition:

> *Pretend you are a job applicant trying to get your ideal job as a [insert focal job title here]. The personality test you are about to take is a very important part of the job selection process, so it is important that you do well. Please respond to the test as you would if you were applying for the [focal job] position. This test's results will be used in the decision to hire all job candidates.*

Participants were then provided with the focal job's job description and instructed to read it, then to complete the personality test. The job description was present on every page of the assessment for easy reference. The job descriptions that were provided can be found in Appendix B.

Chapter Three:

Results

Study 1

I examined the mean familiarity ratings for each of the 10 jobs rated. A one-way ANOVA with job as the independent variable and familiarity rating as the dependent variable was significant, $F(9, 1780) = 23.81$, $p < .001$. Tukey's post hoc tests (Appendix C), using a conservative Bonferroni correction to control for Type 1 error, showed that participants were least familiar with intelligence analysts ($M = 1.84$, $SD = 1.13$), compliance managers ($M = 1.96$, $SD = 1.09$), and computer systems analysts ($M = 2.18$, $SD = 1.17$; Table 10 provides a focused summary of the post hoc findings for the focal jobs; a more detailed table summarizing how all means compared can be found in Appendix B). Across jobs, participants were less familiar with intelligence analysts than seven other jobs, less familiar with compliance manager than five other jobs, and less familiar with computer systems analysts than four other jobs. The seven remaining jobs were rated lower than from only zero to two other jobs. Thus, based on the mean familiarity ratings, participants were clearly the least familiar with three jobs, and thus these were chosen to be the focal jobs in Study 2.

Table 10: Familiarity Ratings of Jobs

				Tukey's Post Hoc Significance (*p*)		
Job Title	N	*M*	*SD*	Compliance Manager	Computer Systems Analyst	Intelligence Analyst
Accountant	184	3.17	1.24	< .001	< .001	< .001
Air Traffic Controller	184	2.35	1.23			< .001
Compliance Manager	183	1.96	1.09			
Computer Systems Analyst	184	2.18	1.17			
Industrial/ Organizational Psychologist	184	2.77	1.39	< .001	< .001	< .001
Intelligence Analyst	137	1.84	1.13			
Legal Secretary	184	2.54	1.24	< .001		< .001
Marketing Manager	183	2.74	1.29	< .001	< .001	< .001
Police Patrol Officer	183	3.17	1.34	< .001	< .001	< .001
Property Claims Insurance Examiner	184	2.28	1.29			< .001

Note: Tukey's post hoc significance is reported in this table only for the three lowest-rated jobs for clarity. A full reporting of post hoc significance can be found in Appendix C.

Study 2

Equivalence of samples. Before testing Study 2's hypotheses, I first examined honest responses under each of the study's three job conditions to assess whether participants were initially equivalent on personality traits. Participant honest personality scores did not differ on any of the traits between any of the job conditions (Table 11), Wilks' $\Lambda = .97$, $F(12, 684) = .892$, $p = .56$, partial $\eta^2 = .02$.

Table 11: Descriptive Statistics for Honest Responses for Each Job Condition

		N	*M*	*SD*
Adjustment	Compliance Manager	117	3.39	.71
	Computer Systems Analyst	117	3.26	.79
	Intelligence Analyst	116	3.29	.79
	Total	350	3.31	.76
Ambition	Compliance Manager	117	3.50	.69
	Computer Systems Analyst	117	3.46	.73
	Intelligence Analyst	116	3.53	.74
	Total	350	3.50	.72
Inquisitive	Compliance Manager	117	3.74	.62
	Computer Systems Analyst	117	3.55	.65
	Intelligence Analyst	116	3.70	.65
	Total	350	3.66	.64
Interpersonal Sensitivity	Compliance Manager	117	3.58	.75
	Computer Systems Analyst	117	3.55	.73
	Intelligence Analyst	116	3.62	.75
	Total	350	3.58	.74
Learning Approach	Compliance Manager	117	3.96	.59
	Computer Systems Analyst	117	3.79	.58
	Intelligence Analyst	116	3.79	.69
	Total	350	3.85	.63
Prudence	Compliance Manager	117	3.56	.65
	Computer Systems Analyst	117	3.47	.57
	Intelligence Analyst	116	3.44	.62
	Total	350	3.49	.61
Sociability	Compliance Manager	117	3.24	.73
	Computer Systems Analyst	117	3.20	.73
	Intelligence Analyst	116	3.26	.71
	Total	350	3.23	.72

Manipulation check. As an initial manipulation check to ensure the response conditions impacted responses in a manner consistent with previous research, I examined elevation of social desirability scores, as measured by the unlikely virtues scale, between the honest and each job description. The unlikely virtues score was elevated significantly for all three job conditions compared to honest response conditions (Table 12), in line with previous research on faking conditions (Hough, 1998; Rosse et al., 1998).

Table 12: Manipulation Check: Within-Subjects t-Test on the Unlikely Virtues Scale

		Honest		Job Condition						
Job Condition	N	*M*	*SD*	*M*	*SD*	*r*	*t*	*df*	*p*	Cohen's *d*
Compliance Manager	117	3.49	0.59	3.98	0.54	0.35	8.35	116	< .001	0.76
Computer Systems Analyst	117	3.45	0.50	3.80	0.56	0.42	6.65	116	< .001	0.61
Intelligence Analyst	116	3.45	0.60	3.79	0.60	0.33	5.43	115	< .001	0.49

Note: Cohen's d is reported for within-subjects designs using Morris & DeShon's (2002) correction for dependence between means.

A second manipulation check involved an examination of the effect of the counter-intuitive information being added into the job description for compliance manager. As sociability was rated as the least important personality trait for compliance managers in O*NET, it was expected that sociability would not be elevated compared to the honest condition based on the content of the job itself. However, sociability was elevated for compliance managers under the job condition ($M = 3.31$, $SD = .70$) compared to the honest condition ($M = 3.18$, $SD = .70$), $t(116) = -2.21$, $p = .029$, suggesting the job knowledge manipulation had the expected effect on participants.

As a final manipulation check, participants reported their familiarity with the job they were presented with before and after reading the job description. Participant familiarity ratings increased meaningfully for all three job conditions (Table 13).

Table 13: Manipulation Check: Change in Self-Reported Job Familiarity After Job Description

		Before		After						
Job Condition	N	*M*	*SD*	*M*	*SD*	*r*	*t*	*df*	*p*	Cohen's *d*
Compliance Manager	117	2.32	1.08	4.05	.76	.47	18.90	116	<.001	1.83
Computer Systems Analyst	115	2.85	1.16	3.97	.82	.48	11.29	114	<.001	1.11
Intelligence Analyst	116	2.25	1.14	3.89	.79	.37	15.72	115	<.001	1.51

Note: Cohen's d is reported for within-subjects designs using Morris & DeShon's (2002) correction for dependence between means.

The three manipulation checks provided evidence that the study's manipulations were effective at eliciting responses comparable to other studies with similar research conditions (Hough, 1998; Rosse et al., 1998), that participants used information from the job descriptions to inform their responding to the personality test under the job condition manipulations, and that the job descriptions successfully increased participants' knowledge about the focal jobs.

Trait elevation by job: Mixed-model factorial MANOVA. To test for hypothesized trait elevation, I conducted a mixed-model factorial MANOVA with one within-subjects variable, context (honest vs. applicant), one between-subjects variable, job condition (compliance manager, computer systems analyst, and intelligence analyst), and the seven IPIP-HPI personality traits as the dependent variables. A MANOVA is preferable in this scenario over a series of ANOVAs because personality variables tend to show small to moderate intercorrelations across research contexts (Barrick & Mount,

1996; Hogan et al., 2007). The MANOVA accounts for the correlations between dependent variables, and is thus a more powerful omnibus statistics than a series of ANOVAs.

Examining the assumptions underlying the MANOVA, Box's test of the equality of covariance matrices was significant, Box's $M = 263.40$, F (210, 317783) = 1.18, p = .04, suggesting that the covariance matrices may differ between conditions in this study, though the MANOVA is robust to violations of this assumption. Examining the data for equality of variances across job conditions, Levene's test of the equality of variances showed differences in variances only for learning approach under the honest response context, $F(2,347) = 3.51$, $p = .03$, adjustment under the applicant response context, $F(2,347) = 5.26$, $p = .01$, and prudence under the applicant context, $F(2,347) = 3.17$, p = .04. However, the F statistic, which underlies the MANOVA, is generally robust to small amounts of heterogeneity of variance (Lindman, 1974).

The full mixed-model repeated measures MANOVA was significant, Wilks' Λ = .93, $F(12, 684) = 2.16$, $p = .01$, partial $\eta^2 = .04$, demonstrating that trait elevation varied by job condition within the study (see Table 14 for Study 2 descriptives and Table 15 for the MANOVA results for all possible models).

Trait elevation by job: Mixed-model factorial MANCOVA, controlling for job familiarity. In order to isolate the effects of this study's experimental job condition manipulations on participant responding, I ran a mixed-model factorial MANCOVA identical to the above MANOVA, but controlled for familiarity with the focal job before the study. Adding in this control assessed whether initial familiarity impacted trait elevation between contexts, and detailed the impact of the experimental manipulations

Table 14: Summary of Study 2 Descriptives

Trait	Job	N	Honest Context		Applicant Context	
			M	*SD*	*M*	*SD*
Adjustment	Compliance Manager	117	3.39	.71	3.81	.72
	Computer Systems Analyst	117	3.26	.79	3.67	.73
	Intelligence Analyst	116	3.29	.79	3.63	.86
	Total	350	3.31	.76	3.70	.77
Ambition	Compliance Manager	117	3.50	.69	4.06	.72
	Computer Systems Analyst	117	3.46	.73	3.91	.65
	Intelligence Analyst	116	3.53	.74	3.89	.77
	Total	350	3.50	.72	3.95	.72
Inquisitive	Compliance Manager	117	3.74	.62	4.03	.70
	Computer Systems Analyst	117	3.55	.65	3.96	.69
	Intelligence Analyst	116	3.70	.65	4.08	.74
	Total	350	3.66	.64	4.02	.71
Interpersonal Sensitivity	Compliance Manager	117	3.58	.75	4.08	.72
	Computer Systems Analyst	117	3.55	.73	3.85	.66
	Intelligence Analyst	116	3.62	.75	3.83	.69
	Total	350	3.58	.74	3.92	.70
Learning Approach	Compliance Manager	117	3.96	.59	4.23	.65
	Computer Systems Analyst	117	3.79	.58	4.10	.75
	Intelligence Analyst	116	3.79	.69	4.15	.79
	Total	350	3.85	.63	4.16	.73
Prudence	Compliance Manager	117	3.56	.65	3.96	.72
	Computer Systems Analyst	117	3.47	.57	3.64	.70
	Intelligence Analyst	116	3.44	.62	3.68	.73
	Total	350	3.49	.61	3.76	.73
Sociability	Compliance Manager	117	3.24	.73	3.34	.77
	Computer Systems Analyst	117	3.20	.73	3.32	.70
	Intelligence Analyst	116	3.26	.71	3.43	.70
	Total	350	3.23	.72	3.36	.72

Table 15: Summary of Repeated-Measures Mixed-Model MANOVA

	Wilk's Λ	F	H df	Error df	p	Partial η^2
Response Context	.70	145.96	1	347	< .001	.30
Job Condition	.92	2.05	14	682	.013	.04
Trait	.44	71.88	6	342	< .001	.56
Response Context x Job Condition	1.00	.60	2	347	.552	.00
Trait x Job Condition	.97	.94	12	684	.511	.02
Response Context x Trait	.86	8.91	6	342	< .001	.14
Response Context x Job Condition x Trait	.93	2.16	12	684	.012	.04

(the context of responses and the job knowledge provided through the job descriptions) independent of initial familiarity with the job. Controlling for familiarity did not change the results of the MANOVA (Table 16). The context by trait by job condition interaction term was still significant, Wilks' $\Lambda = .92$, $F(12, 678) = 2.27$, $p = .008$, partial $\eta^2 = .04$.

Table 16: Summary of Repeated-Measures Mixed-Model MANCOVA

	Wilks' Λ	F	H df	Error df	p	Partial η^2
Response Context	.85	60.51	1	344	.000	.15
Response Context x Familiarity	.97	9.24	1	344	.003	.03
Trait	.75	18.69	6	339	.000	.25
Trait x Familiarity	.95	3.18	6	339	.005	.05
Response Context x Job Condition	1.00	.71	2	344	.495	.00
Trait x Job Condition	.96	1.02	12	678	.428	.02
Response Context x Trait	.93	4.29	6	339	.000	.07
Response Context x Trait x Familiarity	.98	1.41	6	339	.212	.02
Response Context x Job Condition x Trait	.92	2.27	12	678	.008	.04

Trait elevation by job: Post hoc repeated-measures tests. I approached my post-hoc analyses to the mixed-model repeated measures MANOVA in three ways. First, I ran a series of within-subjects t-tests by job for the traits hypothesized to be elevated for each job (for comparison's sake, trait elevation was calculated for all seven IPIP-HPI traits in Tables 17-19), using a Bonferroni correction to control for Type 1 error. As hypothesized (H1a), for the compliance manager job condition, trait score elevation was seen for adjustment, $t(116) = 5.83$, $p < .001$, and prudence, $t(116) = 7.30$, $p < .001$. As hypothesized for computer systems analysts (H1b), trait score elevation was seen for inquisitive, $t(116) = 7.11$, $p < .001$, and prudence, $t(116) = 4.23$, $p < .001$. Also as hypothesized, for intelligence analysts, traits scores were elevated for inquisitive, $t(115) = 5.36$, $p < .001$, learning approach, $t(115) = 5.36$, $p < .001$, and prudence, $t(115) = 5.25$, $p < .001$. All hypothesized mean differences showed medium effect sizes using Cohen's (1988) guidelines for interpretation. However, although all hypothesized traits were elevated, supporting H1a-c, all other traits except sociability were also elevated between honest and applicant contexts. Thus, in line with previous research, there is evidence of near-universal trait elevation between honest and applicant contexts.

However, universal trait elevation between honest and applicant conditions does not preclude the possibility that there was differential trait elevation between job conditions, which is the focus of the study. As the full model MANOVA was significant, further analyses were required to detail the trait elevation differences between conditions. Differential trait elevation was examined using three complementary methods.

Table 17: Post Hoc t-tests for the Compliance Manager Job Condition

		Honest		Job Condition						
Trait	N	*M*	*SD*	*M*	*SD*	*r*	*t*	*df*	*p*	*d*
Adjustment	117	3.39	.64	3.78	.68	.41	-5.83	116	< .001	.54
Ambition	117	3.53	.63	4.07	.64	.33	-7.97	116	< .001	.74
Inquisitive	117	3.72	.60	4.01	.61	.62	-6.00	116	< .001	.55
Interpersonal Sensitivity	117	3.61	.66	4.02	.65	.38	-6.01	116	< .001	.56
Learning Approach	117	3.93	.55	4.22	.61	.67	-6.62	116	< .001	.62
Prudence	117	3.49	.60	3.95	.67	.42	-7.30	116	< .001	.67
Sociability	117	3.18	.70	3.31	.70	.59	-2.21	116	.029	.21

*Note: Significance was determined using a Bonferroni correction for 7(trait) * 3 (job condition) post hoc analyses to control for Type 1 error; p < .002 was used to determine significance.*

Table 18: Post Hoc t-tests for the Computer Systems Analyst Job Condition

		Honest		Job Condition						
Trait	N	*M*	*SD*	*M*	*SD*	*r*	*t*	*df*	*p*	*d*
Adjustment	117	3.25	.72	3.64	.67	.51	-6.21	116	< .001	.57
Ambition	117	3.43	.67	3.85	.61	.45	-6.82	116	< .001	.63
Inquisitive	117	3.54	.61	3.94	.65	.53	-7.11	116	< .001	.66
Interpersonal Sensitivity	117	3.54	.69	3.78	.63	.48	-3.84	116	< .001	.36
Learning Approach	117	3.78	.56	4.10	.67	.66	-6.62	116	< .001	.63
Prudence	117	3.45	.57	3.66	.64	.58	-4.23	116	< .001	.38
Sociability	117	3.15	.65	3.24	.65	.61	-1.72	116	.088	.16

*Note: Significance was determined using a Bonferroni correction for 7(trait) * 3 (job condition) post hoc analyses to control for Type 1 error; p < .002 was used to determine significance.*

Trait elevation by job: Post hoc repeated-measures tests, controlling for familiarity. Although controlling for job familiarity before the study in the mixed-model factorial MANCOVA still resulted in a significant full model interaction, it did not

Table 19: Post Hoc t-tests for the Intelligence Analyst Job Condition

		Honest		Job Condition						
Trait	N	*M*	*SD*	*M*	*SD*	*r*	*t*	*df*	*p*	*d*
Adjustment	116	3.24	0.73	3.65	0.75	0.49	-5.91	115	< .001	.55
Ambition	116	3.46	0.71	3.88	0.69	0.36	-5.75	115	< .001	.53
Inquisitive	116	3.69	0.63	4.02	0.72	0.53	-5.36	115	< .001	.50
Interpersonal Sensitivity	116	3.56	0.73	3.85	0.64	0.38	-4.06	115	< .001	.38
Learning Approach	116	3.81	0.64	4.11	0.73	0.6	-5.36	115	< .001	.49
Prudence	116	3.41	0.59	3.7	0.69	0.58	-5.25	115	< .001	.49
Sociability	116	3.24	0.68	3.42	0.6	0.54	-3.25	115	0.001	.29

*Note: Significance was determined using a Bonferroni correction for 7(trait) * 3 (job condition) post hoc analyses to control for Type 1 error; p < .002 was used to determine significance.*

necessarily follow that post hoc results would be the same after introducing the control. As such, post hoc tests controlling for initial job familiarity were conducted to isolate the effect of this study's experimental manipulations. Adding in the control changed the conclusions substantially. Instead of near-universal trait elevation across jobs, controlling for initial job knowledge demonstrated differential trait elevation due to the study's manipulation (Table 20). Specifically, for the compliance manager job condition, adjustment and inquisitive were no longer significantly elevated between honest and applicant contexts. For the computer system analyst condition, learning approach and prudence were no longer significantly elevated. For the intelligence analyst condition, adjustment, interpersonal sensitivity, prudence, and sociability were no longer significantly elevated. Thus, the experimental manipulation within this study appears to have led to at least some differential trait elevation across jobs independent of initial job familiarity, although the elevated traits did not always match those hypothesized in H1a-c based on O*NET's work style ratings.

Trait elevation by job: Post hoc tests, between-subjects MANOVA.

Differential trait elevation was seen in a within-subjects design which accounted for participants' initial trait levels. However, much faking research has used between-subjects designs to examine trait elevation. In order to more directly compare this study's manipulations to this between-subjects research, as well as reduce the impact of taking the same personality test twice, as all participants did in this study, I conducted a between-subjects MANOVA with personality test response condition (Honest, Job 1, Job 2, and Job 3) as the independent variable, and personality trait scores (the seven HPI traits) as the dependent variables. To create an honest condition independent of the job description conditions, I randomly selected a fourth of responses (N = 29) from each of the three job description conditions to comprise the honest comparison. This method left each of the four response conditions with balanced sample sizes (N = 87-88). The honest trait scores for those sampled to compose the honest condition did not differ from the honest responses that were not sampled for any of the personality traits measured, nor the unlikely virtues scale, Wilks' $\Lambda = .98$, $F(8,341) = .82$, $p = .589$, partial $\eta^2 = .02$.

As expected based on the within-subjects results, the between-subjects MANOVA showed there were significant mean differences between conditions, Wilks' $\Lambda = .84$, $F(21,977) = 2.95$, $p < .001$, partial $\eta^2 = .06$. Post hoc ANOVAs (Table 21) demonstrated that response conditions differed for all traits. Tukey's post hoc tests were then conducted to test for H1a-1c (Table 21) using a conservative Bonferroni correction to control for Type 1 error. For the compliance manager condition, adjustment and prudence were elevated compared to the honest condition, supporting H1a. However, ambition, inquisitive, interpersonal sensitivity, and learning approach, which were not rated above

Table 20: Post Hoc Repeated-Measures ANCOVA, Controlling for Familiarity

	Compliance Manager			
	Wilks' Λ	F (1,115)	p	Partial η^2
Adjustment	.97	3.69	.057	.03
Ambition	.84	22.42	<.001	.16
Inquisitive	.94	7.38	.008	.06
Interpersonal Sensitivity	.95	5.90	.017	.05
Learning Approach	.84	22.41	<.001	.16
Prudence	.85	20.35	<.001	.15
Sociability	1.00	.01	.925	<.001

	Computer Systems Analyst			
	Wilks' Λ	$F(1,113)$	p	Partial η^2
Adjustment	.90	12.35	.001	.10
Ambition	.89	14.63	<.001	.12
Inquisitive	.85	19.59	<.001	.15
Interpersonal Sensitivity	.90	13.03	<.001	.10
Learning Approach	.98	2.40	.124	.02
Prudence	.98	2.66	.105	.02
Sociability	.99	1.45	.232	.01

	Intelligence Analyst			
	Wilks' Λ	$F(1,114)$	p	Partial η^2
Adjustment	.97	4.02	.047	.03
Ambition	.88	15.82	<.001	.12
Inquisitive	.86	18.96	<.001	.14
Interpersonal Sensitivity	.97	3.70	.057	.03
Learning Approach	.84	22.15	<.001	.16
Prudence	.93	9.17	.003	.07
Sociability	.94	7.45	.007	.06

*Note: Significance was determined using a Bonferroni correction for 7 (trait) * 3 (job condition) post hoc analyses to control for Type 1 error; $p < .002$ was used to determine significance.*

80 in O*NET were also elevated compared to the honest condition. For computer systems analysts, inquisitive was elevated, but prudence was not, only partially supporting H1b. Adjustment and ambition were also elevated. For intelligence analysts, inquisitive and learning approach were elevated, but prudence was not, providing only partial support for H1c. Adjustment, ambition, and sociability were also elevated for intelligence analysts. Although not hypothesized, job conditions did not differ from each other for any personality traits, using a Bonferroni correction.

Trait elevation by job: ANCOVAs controlling for honest trait scores. A final manner in which differential trait elevation can be evaluated is to isolate the effect of the job condition manipulation by comparing trait scores under the applicant scenarios, controlling for participants' honest scores. To do this, I ran a series of one-way ANCOVAs, controlling for honest trait scores, for each IPIP-HPI trait with job condition as the independent variable (Table 22). This method isolated trait elevation between jobs by accounting for the near-universal trait elevation found between contexts through the control of baseline honest scores. Controlling for the honest trait score, means differed for interpersonal sensitivity, $F(2,346) = 5.43$, $p = .005$, such that the trait score for the compliance manager condition was greater than that of the computer systems analyst condition, $p = .031$, and intelligence analyst condition, $p = .007$. Job conditions also differed on prudence, $F(2,346) = 5.86$, $p = .003$, such that the score for the compliance manager condition was again greater than the computer systems analyst condition, $p = .004$, and the intelligence analyst condition, $p = .034$. Using a Bonferroni correction to control for Type 1 error, the significance level for interpersonal sensitivity for computer

systems analysts, as well as prudence for intelligence analysts, did not make the cutoff for significance.

Table 21: Post Hoc Tests for Differences in Trait Elevation by Response Condition

		N	*M*	*SD*	*Cohen's d*	*F*(3,346)	*p*
Adjustment	Honest	87	3.30[a]	.84		5.80	< .001
	Compliance Manager	88	3.76[b]	.74	.58		
	Computer Systems Analyst	88	3.69[b]	.75	.49		
	Intelligence Analyst	87	3.64[b]	.86	.40		
Ambition	Honest	87	3.40[a]	.75		12.19	< .001
	Compliance Manager	88	4.01[b]	.75	.81		
	Computer Systems Analyst	88	3.92[b]	.65	.74		
	Intelligence Analyst	87	3.87[b]	.76	.62		
Inquisitive	Honest	87	3.61[a]	.65		8.43	< .001
	Compliance Manager	88	4.00[b]	.71	.57		
	Computer Systems Analyst	88	3.97[b]	.72	.53		
	Intelligence Analyst	87	4.11[b]	.72	.73		
Interpersonal Sensitivity	Honest	87	3.57[a]	.77		5.47	0.001
	Compliance Manager	88	4.01[b]	.77	.57		
	Computer Systems Analyst	88	3.88[ab]	.67	.43		
	Intelligence Analyst	87	3.85[ab]	.71	.38		

Table 21: Post Hoc Tests for Differences in Trait Elevation by Response Condition (Continued)

		N	*M*	*SD*	*Cohen's d*	*F*(3,346)	*p*
Learning Approach	Honest	87	3.85[a]	.60		4.60	.004
	Compliance Manager	88	4.20[b]	.68	.55		
	Computer Systems Analyst	88	4.03[ab]	.75	.27		
	Intelligence Analyst	87	4.17[b]	.77	.47		
Prudence	Honest	87	3.53[a]	.57		5.04	.002
	Compliance Manager	88	3.92[b]	.75	.59		
	Computer Systems Analyst	88	3.63[ab]	.73	.15		
	Intelligence Analyst	87	3.64[ab]	.75	.17		
Sociability	Honest	87	3.13[a]	.74		2.87	.036
	Compliance Manager	88	3.32[a]	.78	.25		
	Computer Systems Analyst	88	3.33[a]	.66	.29		
	Intelligence Analyst	87	3.45[a]	.74	.43		

Note: Different letters indicate a significant difference between means using a Bonferroni correction with significance level = .05/4 jobs for each ANOVA post hoc = .0125 to control for Type 1 error. Cohen's d reported compares each job condition to the honest condition.

Table 22: ANCOVA Results for Differential Trait Elevation Controlling for Honest Scores

		N	Estimated Marginal Mean	S.E.	Score Level of Honest Response	*F*(2, 346)	*p*
Adjustment	Compliance Manager	117	3.78[a]	.07	3.31	1.19	.310
	Computer Systems Analyst	117	3.69[a]	.07			
	Intelligence Analyst	116	3.64[a]	.07			
Ambition	Compliance Manager	117	4.06[a]	.06	3.50	2.22	.111
	Computer Systems Analyst	117	3.93[a]	.06			
	Intelligence Analyst	116	3.88[a]	.06			
Inquisitive	Compliance Manager	117	3.98[a]	.06	3.66	.47	.624
	Computer Systems Analyst	117	4.02[a]	.06			
	Intelligence Analyst	116	4.06[a]	.06			
Interpersonal Sensitivity	Compliance Manager	117	4.08[a]	.06	3.58	5.43	.005
	Computer Systems Analyst	117	3.86[b]	.06			
	Intelligence Analyst	116	3.82[b]	.06			
Learning Approach	Compliance Manager	117	4.16[a]	.06	3.85	.16	.854
	Computer Systems Analyst	117	4.14[a]	.06			
	Intelligence Analyst	116	4.18[a]	.06			
Prudence	Compliance Manager	117	3.92[a]	.06	3.49	5.86	.003
	Computer Systems Analyst	117	3.65[b]	.06			
	Intelligence Analyst	116	3.71[b]	.06			
Sociability	Compliance Manager	117	3.34[a]	.06	3.23	.66	.520
	Computer Systems Analyst	117	3.33[a]	.06			
	Intelligence Analyst	116	3.42[a]	.06			

Note: Significant ANCOVAs were followed by Bonferroni post hoc tests to determine which job conditions differed on each trait. Different letters on marginal means indicate that groups differed significantly on trait, $p < .017$.

Alternative Approaches to Identifying Importance Ratings

As the traits that were elevated in Study 2 did not unequivocally match O*NET ratings, two alternative, empirically-based, methods of identifying important traits for each job were considered. As was discussed in the initial discussion of O*NET, Work Style ratings were provided for each job in the system in the aggregate. As such, significance testing could not be used to identify which Work Styles were rated significantly more important than others, so a cutoff of 80 on the Work Styles' 100 point scale was used to identify important traits for each job. However, Study 1 was designed so that an empirical comparison of importance ratings could be made. An examination of trait ratings from Study 1 through a series of repeated measures ANOVAs demonstrated that respondents viewed certain traits as more important that other traits for each of the three focal jobs - compliance managers, Wilks' $\Lambda = .53$, $F(6,84) = 12.11$, $p < .001$, partial $\eta^2 = .46$; computer systems analyses, Wilks' $\Lambda = .27$, $F(6,92) = 41.43$, $p < .001$, partial $\eta^2 = .73$; and intelligence analysts, Wilks' $\Lambda = .55$, $F(6, 83) = 11.55$, $p < .001$, partial $\eta^2 = .46$ (*M* and *SD* in Tables 23-25) - showing conscious trait importance differentiation. To identify the set of traits rated most important for each job, trait scores were rank-ordered in a top-down manner based on the ratings of all participants. Traits that did not differ significantly from the highest-rated (most important) trait (using Tukey's post hoc tests with a Bonferroni correction) were identified as the most important traits for each job (Table 26). For compliance managers, prudence was rated as the most important ($M = 4.03$, $SD = 1.00$). Post hoc tests showed that learning approach ($M = 3.93$, $SD = 1.07$) and adjustment ($M = 3.86$, $SD = 1.08$) did not differ significantly from prudence. For computer systems analysts, learning approach was rated the most important trait ($M =$

4.63, *SD* = .74). All other traits were rated as less important. For intelligence analysts, learning approach was also rated the most important trait (*M* = 4.27, *SD* = 1.02), followed by prudence (*M* = 4.18, *SD* = .91).

Table 23: Comparison of Trait Importance Ratings in Study 1, Compliance Manager

	Compliance Manager (N = 90)	
	M	*SD*
Prudence	4.03[a]	1.00
Learning Approach	3.93[a]	1.07
Adjustment	3.86[a]	1.08
Interpersonal Sensitivity	3.66[b]	1.03
Ambition	3.53[b]	1.07
Sociability	3.38[b]	1.09
Inquisitive	2.84[b]	1.15

Note: Letters indicate differences compared to the highest-rated trait, so 'a' indicates trait ratings did not differ from the most important, and 'b' indicates trait ratings differed from most important.

Table 24: Comparison of Trait Importance Ratings in Study 1, Computer Systems Analyst

	Computer Systems Analyst (N = 98)	
	M	*SD*
Learning Approach	4.63[a]	.74
Prudence	4.15[b]	.88
Inquisitive	3.82[b]	1.14
Adjustment	3.55[b]	1.05
Ambition	3.33[b]	1.15
Interpersonal Sensitivity	2.86[b]	1.17
Sociability	2.36[b]	1.17

Note: Letters indicate differences compared to the highest-rated trait, so 'a' indicates trait ratings did not differ from the most important, and 'b' indicates trait ratings differed from most important.

Table 25: Comparison of Trait Importance Ratings in Study 1, Intelligence Analyst

	Intelligence Analyst (N = 89)	
	M	*SD*
Learning Approach	4.27[a]	1.02
Prudence	4.18[a]	.91
Adjustment	3.91[b]	1.16
Ambition	3.66[b]	1.15
Interpersonal Sensitivity	3.61[b]	1.19
Inquisitive	3.47[b]	1.29
Sociability	3.22[b]	1.28

Note: Letters indicate differences compared to the highest-rated trait, so 'a' indicates trait ratings did not differ from the most important, and 'b' indicates trait ratings differed from most important.

Comparison of Alternate Trait Importance Ratings

For compliance managers and intelligence analysts, O*NET and Study 1 trait ratings were generally aligned. For compliance managers, learning approach was identified as important, in addition to O*NET's adjustment and prudence. For intelligence agents, learning approach and prudence were rated important in both O*NET and Study 1, but inquisitive was only identified as important in O*NET. However, the Study 1 and O*NET ratings did not identify any of the same traits as important for computer systems analysts. Thus, although there is a fair amount of alignment across the two sets of ratings, there are still some discrepancies in trait importance.

Unfortunately, as Study 2's jobs were selected because they were the ones participants were least familiar with, a similar set of significance tests could not be conducted including only expert ratings (defined as a familiarity rating of four or five) because the sample size for comparison was too small, ranging from only seven experts for compliance managers, to 17 experts for computer systems analysts. However, a

Table 26: Traits Elevated by Job Using the Different Post Hoc Methods

	Compliance Manager					
	Hypothesized Trait Elevation	Study 1 Most Important Traits	Post Hoc t-tests	Control for Familiarity Before the Study	Between Analyses	Hypothesis Supported?
Adjustment	H1a	X	Elevated	ns	Elevated	Partial
Ambition			Elevated	Elevated	Elevated	NA
Inquisitive			Elevated	ns	Elevated	NA
Interpersonal Sensitivity			Elevated	Elevated	Elevated	NA
Learning Approach		X	Elevated	Elevated	Elevated	NA
Prudence	H1a	X	Elevated	Elevated	Elevated	Supported
Sociability			ns	ns	ns	NA

	Computer Systems Analyst					
	Hypothesized Trait Elevation	Study 1 Most Important Traits	Post Hoc t-tests	Control for Familiarity Before the Study	Between Analyses	Hypothesis Supported?
Adjustment			Elevated	Elevated	Elevated	NA
Ambition			Elevated	Elevated	Elevated	NA
Inquisitive	H1b		Elevated	Elevated	Elevated	Supported
Interpersonal Sensitivity			Elevated	Elevated	ns	NA
Learning Approach		X	Elevated	ns	ns	NA
Prudence	H1b		Elevated	ns	ns	Partial
Sociability			ns	ns	ns	NA

Table 26: Traits Elevated by Job Using the Different Post Hoc Methods (Continued)

	Intelligence Analyst					
	Hypothesized Trait Elevation	Study 1 Most Important Traits	Post Hoc t-tests	Control for Familiarity Before the Study	Between Analyses	Hypothesis Supported?
Adjustment			Elevated	ns	Elevated	NA
Ambition			Elevated	Elevated	Elevated	NA
Inquisitive	H1c		Elevated	Elevated	Elevated	Supported
Interpersonal Sensitivity			Elevated	ns	ns	NA
Learning Approach	H1c	X	Elevated	Elevated	Elevated	Supported
Prudence	H1c	X	Elevated	ns	ns	Partial
Sociability			Elevated	ns	ns	NA

comparison of the rank-order of O*NET and expert ratings from Study 1 does show some differences in which traits were identified as the most important traits (Table 27). For example, for compliance managers and computer systems analysts, Study 1 experts rated learning approach most important. In O*NET, however, learning approach was rated and ranked much lower. Thus, preliminary examination of expert ratings still suggests some disagreement between sources as to which traits are truly important.

Summary of Post Hoc Analyses, Compared to Alternate Trait Importance Ratings

The results of the post hoc analyses are also summarized in Table 26, above. This table shows that there was an effect for response context, such that nearly all traits were elevated between honest and applicant conditions using a within-subjects t-test. However, controlling for initial familiarity with the job, or using a between subjects design to compare honest to applicant responses, resulted in a more differentiated picture of faking.

*Table 27: Comparison of Importance Ratings between O*NET and Study 1*

Compliance Manager			
O*NET Ratings		Study 1 - Expert N = 7	
	M		*M*
Prudence	88.80	Learning Approach	4.86
Adjustment	80.30	Prudence	4.57
Ambition	76.00	Adjustment	4.00
Learning Approach	76.00	Ambition	3.86
Inquisitive	74.00	Interpersonal Sensitivity	3.14
Interpersonal Sensitivity	67.50	Sociability	3.00
Sociability	47.00	Inquisitive	2.29

Computer Systems Analyst			
O*NET Ratings		Study 1 - Expert N = 17	
	M		*M*
Inquisitive	88.50	Learning Approach	4.82
Prudence	85.20	Prudence	4.35
Adjustment	79.30	Inquisitive	4.12
Interpersonal Sensitivity	76.50	Adjustment	3.76
Learning Approach	72.00	Ambition	3.71
Ambition	69.00	Interpersonal Sensitivity	3.12
Sociability	66.00	Sociability	2.71

Intelligence Analyst			
O*NET Ratings		Study 1 - Expert N = 12	
	M		*M*
Inquisitive	87.50	Adjustment	4.00
Prudence	86.00	Ambition	3.75
Learning Approach	81.00	Inquisitive	3.75
Adjustment	73.00	Interpersonal Sensitivity	4.33
Ambition	73.00	Learning Approach	4.17
Interpersonal Sensitivity	64.00	Prudence	4.33
Sociability	50.00	Sociability	3.75

For the compliance manager condition, hypothesized trait elevation was supported for prudence across all three post hoc methods, but was only supported for two post hoc methods for adjustment, resulting in only partial support of H1a. Learning approach, identified as important in Study 1's ratings, was also elevated across all analyses. For computer systems analysts, hypothesized trait elevation occurred for inquisitive across all methods, but elevation for prudence was only seen in the post hoc t-tests. Similarly, learning approach, which was identified as important in Study 1, was only elevated in the post hoc t-tests. Finally, although hypothesized trait elevation from O*NET was found for both inquisitive and learning approach for intelligence analysts, prudence was only elevated in the post hoc t-tests. In sum, Study 1's trait importance ratings did not do a better job predicting elevated traits than O*NETs ratings, suggesting neither method may be completely accurate for identifying the important traits for different jobs.

Relationship between Familiarity with Job and Trait Importance Ratings

To test the hypothesis that job-relevant traits would be rated highly by those more familiar with the focal jobs than by those less familiar, I examined the correlations between job familiarity and trait importance ratings between all ten jobs in Study 1 (Table 28). Results did not support any of H3a-3j. Although before using the Bonferroni correction, for accountants, there was a positive relationship between learning approach and job familiarity, and for air traffic controllers, there was as positive relationship between familiarity and adjustment, after using the correction, only learning approach and prudence were significant for compliance managers (which were not hypothesized). No other familiarity/trait relationships were significant.

Table 28: Summary of Correlations between Job Familiarity and Trait Importance Ratings

Hypothesis:	Adjustment	Ambition	Inquisitive	Interpersonal Sensitivity	Learning Approach	Sociability	Prudence
3a: Accountants	.15	.06	.10	.01	.22*	-.01	.16
3b: Air Traffic Controllers	.21*	.09	.001	.02	.09	.07	.10
3c: Compliance Manager	.18	.18	-.03	.05	.35**	-.03	.26**
3d: Computer Systems Analyst	.10	.21*	.10	.21*	.14	.16	.18
3e: Industrial/ Organizational Psychologist	.04	.14	.24*	-.07	.13	.05	.15
3f: Intelligence Analyst	.04	.07	.10	.19	.005	.19	.09
3g: Legal Secretary	.07	.20	.16	-.02	-.04	.24*	-.02
3h: Marketing Manager	.23*	.17	.16	.24*	.10	.22*	.07
3i: Police Patrol Officer	.10	-.05	-.06	.09	.08	-.11	.12
3j: Property Claims Insurance Examiner	.06	-.08	.12	.02	-.01	-.05	.18

*Note: * indicates $p < .05$, ** indicates significance using a Bonferroni correction, at the $p < .007$ level.*

I assessed whether job familiarity led to greater trait elevation in Study 2 by examining the correlations between job familiarity and applicant trait scores (Table 29). For compliance managers, there was a negative relationship between prudence and familiarity before the study, $r = -.28$, $p = .002$, but no relationship for adjustment. Thus, H4a was not supported. For computer systems analysts, there was no relationship between familiarity and inquisitive nor prudence, so H4b was not supported. For intelligence analysts, all traits but sociability were significantly related to engagement. However, the relationships were all negative, such that familiarity before the study was

negatively related to respondent trait scores. As H4c hypothesized positive relationships between familiarity and trait ratings, H4c was not supported.

Table 29: Summary of Correlations between Job Familiarity Rating and Trait Scores

	CM		CSA		IA	
	Familiarity Before	Familiarity After	Familiarity Before	Familiarity After	Familiarity Before	Familiarity After
	N = 117	N = 117	N =115	N = 117	N = 116	N = 116
Adjustment	-.10	.00	-.03	-.19*	-.22*	.09
Ambition	-.15	.03	-.09	-.12	-.20*	.12
Inquisitive	-.05	.16	-.09	-.06	-.20*	.15
Interpersonal Sensitivity	-.05	.20*	-.19*	-.06	-.24*	.06
Learning Approach	-.11	-.01	-.02	-.04	-.21*	.12
Prudence	-.28**	-.11	-.12	-.17	-.23*	.00
Sociability	.21*	.16	-.04	.18*	.05	.34*

*Note: * indicates significance at the p < .05 level, and ** indicates significance at the p < .017 level.*

It is important to note that, as intended, most participants were unfamiliar with the jobs in Study 1 (see Table 10 for average familiarity; Table 30-31 for frequency of ratings), so for compliance managers, computer systems analysts, intelligence analysts, and property claims insurance examiners, the correlations provided likely do not reflect a balanced spectrum of familiarity ratings, and thus may not accurately reflect the true relationship between familiarity and trait ratings. As the jobs in Study 2 were selected because participants in Study 1 were the least familiar with them, few of Study 2's participants reported being familiar with any of the jobs prior to the study. Further, few participants reported being unfamiliar with the jobs after being presented with the focal job's job description. As such, the relationships between familiarity and trait ratings may not be accurate for neither the before nor after perspective.

Table 30: Frequencies of Familiarity Ratings for Study 1's Jobs

	Not at All Familiar				Extremely Familiar
	1	2	3	4	5
Accountant	19	31	52	48	29
Air Traffic Controller	55	52	38	23	12
Compliance Manager	79	50	30	14	5
Computer Systems Analyst	65	50	39	18	8
Industrial/ Organizational Psychologist	41	44	35	32	27
Intelligence Analyst	73	27	18	11	4
Legal Secretary	44	47	48	27	14
Marketing Manager	35	48	43	31	22
Police Patrol Officer	25	31	41	45	36
Property Claims Insurance Examiner	61	55	32	14	18

Table 31: Frequencies of Familiarity Ratings for Study 2's Jobs

		Not at All Familiar				Extremely Familiar	
	Familiarity	1	2	3	4	5	Total
Compliance Manager	Before	31	39	27	18	2	117
	After	2	1	16	68	30	117
Computer Systems Analyst	Before	20	21	36	32	6	115
	After	1	4	23	59	30	117
Intelligence Analyst	Before	38	33	27	14	4	116
	After	1	5	22	66	22	116

Rank-Order Effects

In line with previous research and Rosse et al.'s (1998) call for personality research to examine the impact of faking on decision-making, not just validity, I examined the effect of faking under the different job conditions on the rank-order of

respondents. I examined the proportion of total respondents hired from the job knowledge conditions in a series of hypothetical selection ratios, ranging from 100% to 5% of respondents being selected based on the O*NET's most important traits for each of the three jobs in Study 2. To obtain a composite personality score for each respondent, the scores of the important traits identified from the O*NET ratings were summed. This method of forming a composite assumes that the job-relevant traits are compensatory in nature, which may not necessarily be appropriate - a minimum level of each trait may be needed to be successful on the job. However, the specified traits in Table 5, above, were all rated as relevant to the job, and thus should be examined when evaluating an applicant for a position. As such, a composite personality score was preferable to examining each trait individually for this study's purposes. To examine the fairness of the selection decision, I examined the proportion of respondents under the applicant context who also would have been selected in the honest context using a top-down selection method, based on their composite scores (Table 32).

In general, as the selection ratio decreased, the proportion of those selected using their applicant context rank-order who would also have been selected using their honest rank-order also decreased. In all scenarios, however, except for the 5% and 20% selection ratios for intelligence analysts, over a third of respondents ranked highest on their applicant scores would also have been selected based on their honest scores. For computer systems analysts, there was no selection ratio in which fewer than half of those selected would also have been selected based on their honest scores. Furthermore, if personality tests are used to screen out the bottom portion of applicants, faking has only a minimal impact on who would be selected.

Table 32: Rank-Order Effects

		Compliance Manager		Computer Systems Analyst			Intelligence Analyst		
Selection Ratio	N Selected	N Selected Under Both Contexts	%	N Selected	N Selected Under Both Contexts	%	N Selected	N Selected Under Both Contexts	%
5%	6	2	34.19%	6	4	68.38%	6	0	0.00%
10%	12	4	34.19%	12	6	51.28%	12	4	34.48%
20%	23	11	47.01%	23	12	51.28%	23	6	25.86%
30%	35	25	71.23%	35	21	59.83%	35	19	54.60%
40%	47	31	66.24%	47	30	64.10%	46	34	73.28%
50%	59	40	68.38%	59	42	71.79%	58	48	82.76%
60%	70	51	72.65%	70	59	84.05%	70	58	83.33%
70%	82	62	75.70%	82	70	85.47%	81	73	89.90%
80%	94	81	86.54%	94	87	92.95%	93	87	93.75%
90%	105	100	94.97%	105	102	96.87%	104	98	93.87%
100%	117	117	100.00%	117	117	100.00%	116	116	100.00%

Chapter Four:

Discussion

This study answered Birkeland et al. (2006), Christiansen et al. (2010) and Raymark & Tafero's (2009) call to examine differences in faking between jobs. The focus of this study was to evaluate whether the personality trait elevation between honest and applicant contexts that has been widely seen throughout the personality and selection research (Barrick & Mount, 1996; Dwight & Donovan, 2003; McFarland, 2003; McFarland & Ryan, 2000; Rosse et al., 1998; Viswesvaran & Ones, 1999; Zickar & Robie, 1999) is simply universal, blatant trait elevation, or whether some other factor influences applicant faking behavior. By obtaining personality responses under both honest and applicant contexts, this study determined that although there is wide-spread trait elevation across the HPI's seven personality traits, between honest and applicant contexts, there is also some trait differentiation between jobs. As such, this study provided evidence that a piece of faking behavior within applicant contexts can be considered knowledgeable faking, defined as distortion of personality test responses based on knowledge of the job being applied to.

Differential Trait Importance between Jobs

In Study 1, participants rated the importance of each of the IPIP-HPI's personality traits. Trait ratings indicated that participants viewed some traits as more important for success for a specified job than other traits, in line with Dunlop et al.'s (2012) findings.

Contrary to the robust validation research demonstrating that conscientiousness has the highest validity with job performance of any of the personality traits across jobs (Schmidt & Hunter, 1998), participants in Study 1 (as well as informed ratings from O*NET) did not consistently rate conscientiousness as the most important trait for every job examined. As such, although important across jobs, conscientiousness may not always be the most important trait for a given job. On the other hand, even if conscientiousness is important across all jobs, there may be some jobs in which conscientiousness is significantly more important, above and beyond for other jobs, as seen in this study's compliance manager condition.

Differential Trait Elevation between Jobs

Using a number of between-subjects and within-subjects, as well as univariate and multivariate methods, this study identified meaningful trait differentiation between jobs. This differentiation was found beyond the universal trait elevation typically seen between honest and applicant response conditions. Rather than examining a general, non-job-specific applicant context, this study detailed the tasks and duties for three specific jobs that respondents applied to by providing participants with a job description prior to completing the applicant personality assessment. This job knowledge manipulation impacted trait elevation between jobs, such that only one trait, ambition, was elevated for all three jobs, across methods of examination. Elevation for all other traits depended on the job examined.

However, in no case were all the hypothesized traits, based on O*NET's work style ratings, all elevated nor were these the only traits elevated. Instead, although trait elevation between honest and applicant conditions depended on the job condition

provided, only some of the elevated traits were hypothesized, failing to provide wide support for H1a-c. However, there are several potential issues with the way in which important traits were identified in this study. In an effort to hypothesize elevation only for those traits which were most important for jobs, an artificial cutoff of 80 on a 100-point rating scale was chosen to identify which traits would be elevated. Such an artificial cutoff ignored that traits rated in the upper 70s may also have been viewed as essential for the job, and thus elevated in applicant responses.

A more desirable method for determining trait importance to drive hypothesized trait elevation would have been to use statistical testing of the differences in trait ratings. Unfortunately, as personality-related information provided in O*NET is only in aggregate, significance testing was not possible on the O*NET ratings. However, the data collected in this study offered some limited opportunities to conduct such significance testing. Analyzing the trait ratings from Study 1 for the three focal jobs used for Study 2 demonstrated that certain traits were rated as significantly more important than other traits.

However, as with the O*NET ratings, the traits which were rated as significantly more important for the focal jobs did not entirely identify the traits elevated in the job knowledge condition. One possible explanation for this discrepancy may be that all participants in Study 1, including those familiar with the focal jobs as well as those unfamiliar with the jobs, rated each trait for each job. This likely resulted in quite a bit of noise in the ratings, as it is improbable that those unfamiliar with a job could determine the important traits for the job from the job title alone. It would have been better to look at the trait ratings only for those who reported being familiar with the jobs. However, as

these focal jobs were chosen because participants, on average, reported being unfamiliar with the jobs, there were too few ratings from those familiar with the jobs to conduct significance testing within this study. Yet even a comparison of the average trait ratings from only these experts showed differences in the rank-order of traits compared to the O*NET ratings, again making it unclear which source of ratings are most appropriate for basing hypothesized trait elevation.

Another explanation for the discrepancy between hypothesized trait elevation and actual trait elevation draws on how experts are defined. In O*NET, Work Styles ratings were obtained from the ratings of job analysts (Borman et al., 1999). In Study 1, familiarity ratings were used as a proxy for expertise on the job. In Study 2, expertise on each job was manipulated by providing respondents with a job description detailing the tasks in each job. Experts from each of these three sources have declarative knowledge about the jobs they rated. However, within all three sources, it is unknown, and quite doubtful considering the lack of familiarity with the jobs, whether these "experts" had any real experience in the focal jobs. Therefore, O*NET's and this study's experts may have based their ratings and responses entirely on this declarative knowledge of the jobs. It is possible that the cognitive demand of successfully making the inferential leap between job tasks and personality-driven behaviors is dependent on experience performing those tasks, and therefore a deeper understanding of what it takes for successful execution of those tasks. As such, there is the potential that none of these sources of "expert" ratings reflect accurate knowledge of the personality requirements of the jobs. Likewise, as Study 2 was a simulated applicant scenario, and respondents, on average, reported being unfamiliar with the jobs, it is possible that the traits that are

generally elevated due to job knowledge acquired from experience differ from those elevated with task information.

A final potential explanation for the discrepancy between hypothesized trait elevation and the actual trait elevation seen is that there may be an ideal level of each relevant personality trait on the job, rather than a linear relationship in which higher trait levels are invariably better. If experts thought there was an ideal trait level for a given personality trait, recognizing, for example, that being too high on conscientiousness might be counterproductive, but novices thought that a higher score was always better, the truly important traits may not have been the most highly elevated ones, and thus mean comparisons may not be the most appropriate method for analyzing expert trait ratings nor applicant data.

Relationships between Familiarity and Trait Importance Ratings

Hypothesized relationships between job familiarity and trait importance ratings were not supported for any job examined in Study 1 or Study 2. Although for one of the jobs participants reported being most familiar with in Study 1, accountants, hypothesized trait elevation was supported prior to using a Bonferroni correction, for the other most familiar job, police patrol officer, not a single correlation between familiarity and the traits was significant. Thus, familiarity may not relate to trait elevation.

Several explanations for findings discussed in the differential trait elevation section, above, can also apply to the familiarity-trait relationships as well. Specifically, familiarity with each job does not necessarily mean that the participants knew how the job was performed, and thus they may not have been able to make accurate ratings of the importance of traits for each job. Perhaps an assessment of actual knowledge about each

job would have been a more effective method of determining familiarity or expertise with each job.

Further, even if those familiar with the jobs did, in fact, accurately identify the personality requirements of each job, then again, trait ratings may have operated in an ideal point manner. There is some potential evidence of an ideal point mentality in Study 2's ratings for intelligence analysts. Unexpectedly, there was a negative correlation between familiarity and all personality traits except sociability. This may suggest that those more familiar with the job thought that there was an ideal level of each trait for intelligence analysts. On the other hand, those unfamiliar with the job may have overcompensated their trait elevation and thus over-elevated the traits, thereby explaining this negative correlation. However, as only a few participants reported being familiar with intelligence analysts in the study, the negative correlations found may not be robust. Further research on intelligence analysts and like jobs, seeking a larger sample of those familiar with the jobs, is warranted to identify the stability of these unexpected relationships.

The lack of relationships between familiarity and trait importance or trait elevation could also have been due to several characteristics of the studies. First, as already discussed, few participants reported being familiar with the majority of the jobs examined in either study, constricting the range of ratings that were examined. Second, there may have been widespread heteroscedasticity in ratings. Specifically, as those familiar with the jobs were expected to be best able to identify the most important traits for each job, their trait ratings would be expected to have little variance, as all expert ratings should be aligned. On the other hand, those unfamiliar with each job would be

expected to be less sure which traits were more relevant for the job than other traits. As such, their trait ratings should be less aligned, resulting in greater variance in ratings. Unfortunately, the sample sizes of those reporting being familiar with each job were too small for a robust analysis of heteroscedasticity.

Additionally, as the experimental job condition manipulation in Study 2 gave each participant knowledge about each job before they completed the personality assessment, unknowledgeable personality responses were not obtained. Thus, the relationships reported between familiarity and trait scores in Study 2 do not reflect a lack of job knowledge, but rather the effect of the job knowledge manipulation, and thus may not effectively test the familiarity-trait hypotheses.

Rank-Order Effects

In line with previous research on the effects of response distortion on hiring decisions (Mueller-Hanson et al., 2003; Rosse et al., 1998), faking between honest and applicant conditions had a meaningful impact on which participants would have been selected under a series of top-down selection ratios. Rosse et al. speculated that top-scoring applicants may not be respondents who, in fact, have truly high scores, due to their response distortion. However, as Rosse et al. simply reported the proportion of respondents selected who had elevated response distortion scores, they did not directly test this proposition. This study, however, examined Rosse et al.'s proposition by examining the proportion of those selected under the applicant context who also would have been selected in the honest context.

The selection ratios uncovered using this within-subjects methodology suggested that, to an extent, the proportion of faking that is due to trait elevation between honest

and applicant contexts makes selection decisions considerably more fair than previous researchers have suggested. That is, rather than assuming that nearly all who elevated trait scores between honest and applicant contexts would not have been selected otherwise, this study showed that at least a third of respondents would, in fact, have been selected using their honest scores under even the most restrictive selection ratios. Further, this impact varied dramatically by job. Thus, independent of whether knowledgeable faking occurs, a substantial portion of the larger pool of applicants would be selected based on both their honest and (distorted) applicant scores, thereby supporting at least some fairness in personality-based selection decisions.

Implications

The foremost implication of this research is that applicant faking is not composed of just blatant distortion. Faking is more nuanced, in that while there is widespread elevation between honest and applicant contexts, there is also some job-specific trait elevation beyond such near-universal elevation. As this elevation occurred due to job knowledge being conveyed to applicants through the use of a job description detailing the tasks required in the focal job, there is evidence that this job-specific trait elevation was driven by job knowledge. If such differential trait elevation was driven by job knowledge, then hiring decisions based on applicant scores may not be as problematic as some practitioners believe. Instead, such distortion may merely reflect the respondent's knowledge of the personality-based requirements of the job.

As the traits elevated for each job differed from those hypothesized to be elevated, it is possible that the differential trait elevation found in this study was due to some uncontrolled-for variable. Perhaps the job descriptions created for this study over-

emphasized certain aspects of these jobs, thereby inaccurately over-weighting certain personality traits. Alternatively, especially for computer systems analysts and intelligence analysts, television shows and movies could have influenced participants' perceptions of the type of personality needed for the jobs. Such preconceptions could have influenced trait elevation instead of the job descriptions, though if participants had had these preconceptions, they likely would have reported at least some familiarity with the jobs, which was only infrequently the case.

Regardless of the accuracy of trait importance perceived by applicants, Study 1 provided evidence that participants perceive meaningful differences in the importance of traits for each of several different jobs, as well as differences in trait importance across jobs. Thus, there is evidence that participants consciously recognize that in order to be successful in a given job, certain personality traits are needed. These perceptions of the desirable personality for a job may then come through in their response patterns for a personality assessment, as they attempt to demonstrate to a hiring manager that they are the ideal candidate for the position.

Future Research

This study opened up several opportunities to reconsider and reevaluate the nuances of applicant faking behavior. First and foremost, it is important to look into differential faking for jobs other than the three included in this study. Each of these jobs was conceptually similar in that all were rather specialized, technical jobs. Evidence of the similarity of these jobs was found when comparing the average trait levels for each job condition controlling for honest responses, in which the compliance manager was the only condition in which trait scores were elevated compared to the other jobs (as opposed

to comparing to honest responses). Thus, the personality traits needed for each of these jobs may be more similar than, say, a job which relies heavily on interpersonal sensitivity. Similarly, to better compare job-specific faking to the wide body of research on general, non-job-specific applicant faking, another response condition could be examined which does not specify a job being applied to. Such a manipulation would provide a baseline applicant faking condition, which could be used as a comparison for job-specific applicant responses to further isolate the effects of general, blatant faking from job-specific, knowledgeable faking.

To examine the accuracy of trait elevation, future research could have only those with a significant amount of experience in one of the focal jobs complete the assessment under honest and applicant conditions. An experienced response group would help assess the accuracy of faking due to a manipulation of job knowledge, as was done in this study. The ideal level of personality on the job could be assessed by this experienced group as well. Using a methodology similar to Dunlop et al. (2012), experienced participants could rate the desirability of different trait levels for a focal job. These ratings may show, first, which traits are more important for each job, and second, if a linear model is an inaccurate representation of the importance of traits for the job.

Conclusion

In conclusion, in this study, participants demonstrated that they did not view all personality traits as relevant for a given job. Instead, a select few traits are viewed as most important for a specific job. Similarly, accounting for near-universal trait elevation between honest and applicant contexts, differences in which traits were faked were seen under an applicant condition in which detailed job task knowledge was conveyed to

participants. Furthermore, although response distortion may be high under the applicant conditions, selection decisions based on applicant personality scores may not differ as substantially from selection decisions based on honest responses as previously thought. If some of the variance in faking is informed by real job knowledge, then it would seem that faking is not necessarily a bad thing, or at least not as detrimental as previously thought.

List of References

Ballenger, J. F., Caldwell-Andrews, A., & Baer, R. A. (2002). Effects of positive impression management on the NEO Personality Inventory-Revised in a clinical population. *Psychological Assessment, 13,* 254-260.

Barrick, M. R., & Mount, M. K. (1991). The big-five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1–26.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81,* 261–272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *Personality and Performance, 9*(1/2), 9 – 30.

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology, 92*(2), 410 – 424.

Berry, C. M., Sackett, P. R., & Wiemann, S. (2007). A review of recent developments in integrity test research. *Personnel Psychology, 60,* 271-301.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14*(4), 317-335.

Borman, W. C., Kubisiak, U. C., & Schneider, R. J. (1999). Work styles. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret, & E. A Fleishman (Eds.), *An occupational information system for the 21st century: The development of O*NET* (p.213-226). Washington, D.C.: American Psychological Association.

Bradley, K. M. (2003). Personality test validation research: Present-employee and job applicant samples. *Dissertation Abstracts International, 68*(2-B)*,* 1346. Retrieved January 5, 2012, from PsychInfo.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1)*,* 3-5.

Burns, G. N., & Christiansen, N. D. (2011). Methods of measuring faking behavior. *Human Performance, 24*(4)*,* 358-372.

Chiaburu, D. S., Oh, I.-S., Berry, C. M., Li, N., Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis.

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

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

Collins, J. M., Schmidt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., & Le, H. (2003). Can basic individual differences shed light on the construct meanings of assessment center evaluations? *International Journal of Selection and Assessment, 11*(1)*,* 17 – 29.

Costa, P. T., McCrae, R. R. & Kay, G. G. (1995). Persons, places, and personality: Career assessment using the Revised NEO Personality Inventory. *Journal of Career Assessment, 3*(2)*,*123-139.

Cunningham, M. R., Wong, D. T., & Barbee, A. P., (1994). Self-presentation dynamics on overt integrity tests: Experimental studies of the Reid Report. *Journal of Applied Psychology, 79*(5)*,* 643-658.

Donovan, J. J., Dwight, S. A., & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance, 16*(1)*,* 81-106.

Douglas, E. F., McDaniel, M. A., & Snell, A. F. (1996). The validity of non-cognitive measures decays when applicants fake. *Academy of Management Proceedings,* 127–131.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of Conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91,* 40–57.

Dullaghan, T. R., & Borman, W. C. (2009). Effects of warnings against faking on personality tests for selection. Poster presented at the 24th annual conference of the Society for Industrial and Organizational Psychology in New Orleans, LA.

Dullaghan, T. R., & Joseph, N. (2010). Do warnings on personality tests result in honest responses? Poster presented at the 25th annual conference of the Society for Industrial and Organizational Psychology in Atlanta, GA.

Dunlop, P. D., Telford, A. D., & Morrison, D. L. (2012). Not too little, but not too much: The perceived desirability of responses to personality items. *Journal of Research in Personality, 46*(1)*,* 8-18.

Dunn, W. S., Mount, M. K., Barrick, M. R., & Ones, D. S. (1995). Relative importance of personality and general mental ability in managers' judgments of applicant qualification. *Journal of Applied Psychology, 80*(4)*,* 500-509.

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

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*(2)*,* 155-166.

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

Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group differences in personality: Meta-analyses comparing five U.S. racial groups. *Personnel Psychology, 61,* 579 – 616.

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

Goffin, R. D., Rothstein, M. G., Rieder, M. J., Poole, A., Krajewski, H. T., Powell, D. M., & Mestdagh, T. (2011). Choosing job-related personality traits: Developing valid personality-oriented job analysis. Personality and Individual Differences, 51(5), 646-651.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

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

Hogan Assessments (2002). *The Hogan Personality Inventory Interpretive Report.* Hogan Assessment Systems, Inc. Tulsa, OK.

Hogan, R. (2005). In defense of personality measurement: New wine for old whiners. *Human Performance, 18*(4), 331-341.

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92*(5)*,* 1270-1285.

Hogan, R., & Hogan, J. (1995). *Hogan personality inventory manual.* Tulsa, OK: Hogan Assessment Systems.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88,* 100–112.

Holden, R. R. (2007). Socially desirable responding does moderate personality scale validity both in experimental and in non-experimental contexts. *Canadian Journal of Behavioural Science, 39*(3)*,* 184-201.

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

Hough, L. M. (1998). Personality at work: Issues and evidence. In M. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131–166). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hough, L., & Dilchert, S. (2010). Personality: Its measurement and validity for employee selection. In: Handbook of employee selection. Farr, James L. (Ed.); Tippins, Nancy T. (Ed.); New York, NY, US: Routledge/Taylor & Francis Group, 2010. pp. 299-319.

Hough, L. M., Eaton, N. L., 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.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85,* 869–879.

Ilies, R., Fulmer, I. S., Spitzmuller, M., & Johnson, M. D. (2009). Personality and citizenship behavior: The mediating role of job satisfaction. *Journal of Applied Psychology, 95,* 945-959.

Johnson, J. A., & Hogan, R. (2006). A socioanalytic view of faking. In R. Griffith (Ed.), *A closer examination of applicant faking behavior* (pp. 209-231). Greenwich, CT: Information Age Publishing.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*(3)*,* 530 – 541.

Judge, T. A., Bono, J. E., Ilies, R., & Gerdardt, M. W. (2002). Personality and leadership: A qualitiative and quantitative review. *Journal of Applied Psychology, 84*(4)*,* 765 – 780.

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*(1)*,* 140-154.

Krahe, B., Becker, J., & Zollter, J. (2008). Contextual cues as a source of response bias in personality questionnaires: The case of the NEO-FFI. *European Journal of Personality, 22,* 655-673.

Kurtz, J. E., Tarquini, S. J., Iobst, E. A. (2008). Socially desirable responding in personality assessment: Still more substance than style. *Personality and Individual Differences, 45,* 22 – 27.

Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology, 92*(6), 1638-1656.

Lindman, H. R. (1974). Analysis of Variance in Complex Experimental Designs. W. H. Freeman & Co Ltd.

Mahar, D., Coburn, B., Griffin, N., Hemeter, F., Potappel, C., Turton, M., & Mulgrew, K. (2006). Stereotyping as a response strategy when faking personality questionnaires. Personality And Individual Differences, 40(7), 1375-1386.

Marshall, M. B., De Fruyt, F., Rolland, J., & Bagby, R. (2005). Socially desirable responding and the factorial stability of the NEO PI-R. *Psychological Assessment, 17*(3), 379-384.

McFarland, L. A. (2003). Warning against faking on a personality test: Effects on applicant reactions and personality test scores. *International Journal of Selection and Assessment, 11*(4)*,* 265-276.

McFarland, L. A, & Ryan, A. M. (2000). Variance in faking across non-cognitive measures. *Journal of Applied Psychology, 85*(5)*,* 812-821.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105-125

Mount, M. K., & Barrick, M. R. (1995). The big five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resources Management, 13,* 153–200.

Mueller-Hanson, R. A., Heggestad, E. D., & Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology, 88*(2)*,* 248-355.

Murphy, K. R., & Lee, S. L. (1994). Does conscientiousness explain the relationship between integrity and job performance? *International Journal of Selection and Assessment, 2*(4)*,* 226-233.

O'Connell, M., Kung, M.-C., & Tristan, E. (2011). Beyond impression management: Evaluating three measures of response distortion and their relationship to job performance. *International Journal of Selection and Assessment, 19*(4)*,* 340-351.

O*NET OnLine (2013). Retrieved January 15, 2013, from http://www.onetonline.org/

Ones, D., S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational setting. *Personnel Psychology, 60,* 995 – 1027.

Ones, D. S., Hough, L. M., & Viswesvaran, C. (1998, April). Validity and adverse impact of personality-based managerial potential scales. In R. Page (Chair), *Personality determinants of managerial potential, performance, progression and ascendancy.* Symposium conducted at the annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

Ones, D. S., & Viswesvaran, C. (2001a). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. *International Journal of Selection and Assessment, 9,* 31–39.

Ones, D. S., &Viswesvaran, C. (2001b). Personality at work: Criterion-focused occupational personality scales used in personnel selection. In B. W. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 63–92). Washington, DC: American Psychological Association.

Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Personality at work: Raising awareness and correcting misconceptions. *Human Performance, 18*(4)*,* 389 – 404.

Ones, D. S., Viswesvaran, C., & Reiss, A. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81,* 660–679.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology Monograph, 78,* 679–703.

Pace, V. L., & Borman, W. C. (2006). The use of warning to discourage faking on non-cognitive inventories. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 283-304). Greenwich, Connecticut: Information Age Publishing.

Palumbo, M. V., Miller, C. E., Shalin, V. L., & Steele-Johnson, D. (2005). The impact of job knowledge in the cognitive ability-performance relationship. *Applied Human Resource Management Research, 10*(1)*,* 13-20.

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

Pauls, C. A., & Crost, N. W. (2004). Effects of faking on self-deception and impression management scales. *Personality and Individual Differences, 37,* 1137 – 1151.

Pauls, C. A., & Crost, N. W. (2005). Effects of differential instructional sets on the construct validity of the NEO-PI-R. *Personality and Individual Differences, 39,* 297 – 308.

Powell, D. M., Goffin, R. D., & Gellatly, I. R. (2011). Gender differences in personality scores: Implications for differential hiring rates. *Personality and Individual Differences, 50,* 106 – 110.

Raymark, P. H., & Tafero, T. L. (2009). Individual differences in the ability to fake on personality measures. Human Performance, 22(1), 86-103.

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

Sackett, P. R. (2011). Integrating and prioritizing theoretical perspectives on applicant faking of personality measures. *Human Performance, 24,* 379-385.

Salgado, J. F. (1997). The five factor mode of personality and job performance in the European Community. *Journal of Applied Psychology, 82*(1)*,* 30 – 43.

Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology, 76,* 323-346.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2)*,* 262-274.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71*(3), 432-439.

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology, 78,* 966-974.

Schmit, M. J., Ryan, A., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology, 80*(5), 607-620.

Schmitt, N., & Oswald, F. L. (2006). The impact of corrections for faking on the validity of noncognitive measures in selection settings. *Journal of Applied Psychology, 91*(3)*,* 613 – 621.

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

Smith, D. B., Hanges, P. J., & Dickson, M. W. (2001). Personnel selection and the Five-Factor Model: Reexamining the effects of applicant's frame of reference. *Journal of Applied Psychology, 86*(2)*,* 304-315.

Snell, A. F., Sydell, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review, 9*(2)*,* 219 – 242.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44,* 703-742.

Tett, R. P., & Simonet, D. V. (2011). Faking in personality assessment: A "multisaturation" perspective on faking as performance. *Human Performance, 24,* 302 – 321.

Topping, G. D., & O'Gorman, J. G. (1997). Effects of faking set on validity of the NEO-FFI. *Personality and Individual Differences, 23*(1)*,* 117 – 124.

Tristan, E. (2009). Measuring applicant faking with job desirability: Prevalence, selection, and measurement issues in an applied sample. *Dissertation Abstract International, 70*(*4*-B), 2605. Retrieved January 7, 2011, from PsychInfo.

Vasilopoulos, N. L., Reilly, R. R., & Leaman, J. A. (2000). The influence of job familiarity and impression management on self-report measure scale scores and response latencies. *Journal of Applied Psychology, 85*(1)*,* 50 – 64.

Vecchione, M., Alessandri, G., & Barbaranelli, C. (2012). The five factor model in personnel selection: Measurement equivalence between applicant and non-applicant groups. *Personality and Individual Differences, 52*(4)*,* 503-508.

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*(4), 433 – 461.

White, L. A., Young, M. C., & Rumsey, M. G. (2001). ABLE implementation issues and related research. In J. P. Campbell, D. J. Knapp, J. P. Campbell, D. J. Knapp (Eds.) , *Exploring the limits in personnel selection and classification* (pp. 525-558). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

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

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

Zimmerman, R. D., del Carmen Triana, M., & Barrick (2010). Predictive criterion-related validity of observer ratings of personality and job-related competencies using multiple raters and multiple performance criteria. *Human Performance, 23,* 361 – 378.

Appendices

Appendix A: Mapping O*NET Work Styles onto HPI Traits

*Table A1: Mapping of the O*NET Work Styles onto the Big Five Personality Traits and the HPI Traits*

Big Five Trait	ONET First-level Work Style	ONET Second-Level Work Style	HPI Trait
Conscientiousness	Achievement Orientation	Achievement/Effort	Learning Approach
Emotional Stability	Adjustment	Adaptability/Flexibility	Adjustment
Openness	Practical Intelligence	Analytical Thinking	Inquisitive
Conscientiousness	Conscientiousness	Attention to Detail	Prudence
Agreeableness; Extraversion	Interpersonal Orientation	Concern for others	Interpersonal Sensitivity
Sociability; Agreeableness	Interpersonal Orientation	Cooperation	Interpersonal Sensitivity
Conscientiousness	Conscientiousness	Dependability	Prudence
Extraversion	Social Influence	Energy	Sociability
Neuroticism; Agreeableness	Independence	Independence	Prudence
Conscientiousness	Achievement Orientation	Initiative	Ambition
Openness	Practical Intelligence	Innovation	Inquisitive
Conscientiousness	Conscientiousness	Integrity	Prudence
Extraversion	Social Influence	Leadership Orientation	Ambition
Conscientiousness	Achievement Orientation	Persistence	Prudence
Neuroticism	Adjustment	Self-Control	Adjustment
Extraversion	Interpersonal Orientation	Social Orientation	Sociability
Neuroticism	Adjustment	Stress Tolerance	Adjustment

Appendix B: Job Descriptions

Compliance Manager

Compliance managers plan, direct, and coordinate the activities of an organization to ensure compliance with ethical or regulatory standards. Responsibilities include ensuring that the company complies with its outside regulatory requirements and internal policies by leading internal audits of procedures. A compliance manager may review and set standards for outside communications by requiring disclaimers in emails, or may examine facilities to ensure that they are accessible and safe. Compliance managers may also design or update internal policies to mitigate the risk of the company breaking laws and regulations. Compliance managers are expected to provide an objective view of company policies. Influence by other employees, including management and executives, to overlook infractions may result in significant fines or even business closure.

Tasks

- Maintain documentation of compliance activities, such as complaints received or investigation outcomes.
- File appropriate compliance reports with regulatory agencies.
- Conduct or direct the internal investigation of compliance issues.
- Identify compliance issues that require follow-up or investigation.
- Report violations of compliance or regulatory standards to duly authorized enforcement agencies as appropriate or required.
- Disseminate written policies and procedures related to compliance activities.
- Conduct periodic internal reviews or audits to ensure that compliance procedures are followed.
- Serve as a confidential point of contact for employees to communicate with management, seek clarification on issues or dilemmas, or report irregularities.
- Provide employee training on compliance related topics, policies, or procedures.
- Verify that all firm and regulatory policies and procedures have been documented, implemented, and communicated.
- Updates job knowledge on an ongoing basis by participating in educational opportunities, reading professional publications, maintaining personal/professional networks, and participating in professional organizations.

*Note: Developed from O*NET Online (2013)*

Appendix B (Continued)

Computer Systems Analyst

Computer systems analysts study an organization's current computer systems and procedures and make recommendations to management to help the organization operate more efficiently and effectively. They bring business and information technology (IT) together by understanding the needs and limitations of both. They analyze science, engineering, business, and all other data processing problems for application to electronic data processing systems and analyze user requirements, procedures, and problems to automate or improve existing systems and review computer system capabilities, workflow, and scheduling limitations. They may also be asked to analyze or recommend commercially available software to provide client solutions.

Tasks

- Expand or modify system to serve new purposes or improve work flow.
- Test, maintain, and monitor computer programs and systems, including coordinating the installation of computer programs and systems.
- Develop, document and revise system design procedures, test procedures, and quality standards.
- Provide staff and users with assistance solving computer related problems, such as malfunctions and program problems.
- Review and analyze computer printouts and performance indicators to locate code problems, and correct errors by correcting codes.
- Consult with management to ensure agreement on system principles.
- Confer with clients regarding the nature of the information processing or computation needs a computer program is to address.
- Read manuals, periodicals, and technical reports to learn how to develop programs that meet staff and user requirements.
- Coordinate and link the computer systems within an organization to increase compatibility and so information can be shared.
- Determine computer software or hardware needed to set up or alter system.

*Note: Developed from O*NET Online (2013)*

Appendix B (Continued)

Intelligence Analyst

Intelligence Analysts use information from a variety sources to determine changes in enemy capabilities, vulnerabilities, and probable courses of action. The Intelligence Analyst is primarily responsible for supervising, organizing, and participating in the analysis, processing and distribution of intelligence. Some intelligence analysts design their own databases and compile criminal profiles in order that they may link individuals with criminal organizations. Analysts will at times operate in the field, using surveillance equipment to record criminal actions and intercept criminal communiques.

Tasks

- Validate known intelligence with data from other sources.
- Gather, analyze, correlate, or evaluate information from a variety of resources, such as law enforcement databases.
- Prepare comprehensive written reports, presentations, maps, or charts based on research, collection, and analysis of intelligence data.
- Study activities relating to narcotics, money laundering, gangs, auto theft rings, terrorism, or other national security threats.
- Collaborate with representatives from other government and intelligence organizations to share information or coordinate intelligence activities.
- Evaluate records of communications, such as telephone calls, to plot activity and determine the size and location of criminal groups and members.
- Gather intelligence information by field observation, confidential information sources, or public records.
- Link or chart suspects to criminal organizations or events to determine activities and interrelationships.
- Study the assets of criminal suspects to determine the flow of money from or to targeted groups.
- Design, use, or maintain databases and software applications, such as geographic information systems (GIS) mapping and artificial intelligence tools.

*Note: Developed from O*NET Online (2013)*

Appendix C: Post Hoc Comparison of Study 1 Familiarity Ratings

Table A2: Post Hoc Comparison of Study 1 Familiarity Ratings

Job Title	N	*M*	*SD*	Accountant	Air Traffic Controller	Compliance Manager	Computer Systems Analyst	Industrial/ Organizational Psychologist	Intelligence Analyst	Legal Secretary	Marketing Manager	Police Patrol Officer	Property Claims Insurance Examiner
Accountant	184	3.17	1.24	--	.000	.000	.000	.070	.000	.000	.037	1.000	.000
Air Traffic Controller	184	2.35	1.23	.000	--	.089	.955	.037	.011	.906	.073	.000	1.000
Compliance Manager	183	1.96	1.09	.000	.089	--	.811	.000	.997	.000	.000	.000	.288
Computer Systems Analyst	184	2.18	1.17	.000	.955	.811	--	.000	.316	.151	.001	.000	.999
Industrial/ Organizational Psychologist	184	2.77	1.39	.070	.037	.000	.000	--	.000	.737	1.000	.069	.007
Intelligence Analyst	137	1.84	1.13	.000	.011	.997	.316	.000	--	.000	.000	.000	.053
Legal Secretary	184	2.54	1.24	.000	.906	.000	.151	.737	.000	--	.860	.000	.624
Marketing Manager	183	2.74	1.29	.037	.073	.000	.001	1.000	.000	.860	--	.036	.015
Police Patrol Officer	183	3.17	1.34	1.000	.000	.000	.000	.069	.000	.000	.036	--	.000
Property Claims Insurance Examiner	184	2.28	1.29	.000	1.000	.288	.999	.007	.053	.624	.015	.000	--

Note: Values reported are p values from Tukey's post hoc significance testing.

Appendix D: USF IRB Approval Letter

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-7091

April 16, 2013

Timothy Dullaghan, M.A.
Psychology
Tampa, FL 33612

RE: **Exempt Certification**
IRB#: Pro00011480
Title: Variance in Faking in High-Stakes Personality Assessment as an Indication of Job Knowledge

Study Approval Period: 4/15/2013 to 4/15/2018

Dear Mr. Dullaghan:

On 4/15/2013, the Institutional Review Board (IRB) determined that your research meets USF requirements and Federal Exemption criteria as outlined in the federal regulations at 45CFR46.101(b):

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Approved Documents:
TRDullaghan - Dissertation Proposal Version #1 4.3.13

Internet IC - MTURK
Internet IC - SONA

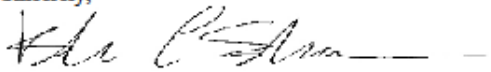
As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF IRB policies and procedures. Please note that changes to this protocol may disqualify it from exempt status. Please note that you are responsible for notifying

the IRB prior to implementing any changes to the currently approved protocol.

The Institutional Review Board will maintain your exemption application for a period of five years from the date of this letter or for three years after a Final Progress Report is received, whichever is longer. If you wish to continue this protocol beyond five years, you will need to submit a new application at least 60 days prior to the end of your exemption approval period. Should you complete this study prior to the end of the five-year period, you must submit a request to close the study.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,



Kristen Salomon, Ph.D., Vice Chairperson
USF Institutional Review Board