A PROCESS MODEL OF APPLICANT FAKING ON OVERT INTEGRITY TESTS

A Dissertation

by

JANIE YU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2008

Major Subject: Psychology

A PROCESS MODEL OF APPLICANT FAKING ON OVERT INTEGRITY TESTS

A Dissertation

by

JANIE YU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee, Daniel A. Newman
Committee Members, Winfred E. Arthur, Jr.
Laura M. Koehly
Michael K. Lindell
Head of Department, Les Morey

December 2008

Major Subject: Psychology

ABSTRACT

A Process Model of Applicant Faking on Overt Integrity Tests.

(December 2008)

Janie Yu, B.A., University of California, Los Angeles;

M.A., California State University, Long Beach

Chair of Advisory Committee: Dr. Daniel A. Newman

To better understand the cognitive processes associated with faking behaviors, Ajzen's Theory of Planned Behavior was adapted to the study of faking on overt integrity tests. This decision-based model is then expanded through the inclusion of a key outcome (counterproductive work behavior) and basic individual differences (conscientious personality and cognitive ability). Results from two student samples (n = 233 and n = 160) demonstrate that conscientiousness negatively predicts attitudes toward faking on employment tests, while cognitive ability predicts the ability to fake. In turn, faking ability moderates the effect of self-reported faking motive on actual test scores, while self-reported faking decreases the validity of integrity tests for predicting counterproductive work behaviors. Implications are discussed.

ACKNOWLEDGMENTS

I would like to thank my committee members Drs. Dan Newman, Laura Koehly, Winfred Arthur and Mike Lindell for their support throughout the dissertation process – I am lucky that they joined me on this scholarly adventure.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

LIST OF TABLES

TABLE Page

INTRODUCTION

Though expanding, our understanding of applicant faking in the context of the selection process remains incomplete. While most researchers agree that job applicants can and do elevate their scores on non-cognitive tests by faking (e.g., Arthur, Woehr, & Graziano, 2000; Berry, Sackett, & Wiemann, 2007; Hough & Oswald, 2007; Ones, Viswesvaran, & Reiss, 1996), the antecedent processes and subsequent consequences of faking are still unclear. Part of our lack of understanding may be attributable to the lack of empirically tested models or appropriate theoretical structures to explain the process (Griffith & McDaniel, 2006; Murphy, 2000). Moreover, there seems to be a limited understanding of possible outcomes associated with applicant faking, such as counterproductive work behaviors (CWB). With mixed evidence on the extent and impact of faking, researchers appear to be divided on how much faking matters in practice, and additional research is needed (e.g., Griffith, Chmielowski, & Yoshita, 2007; Hogan, Barrett, & Hogan, 2007; McFarland & Ryan, 2006; Morgeson et al., 2007).

According to recent studies, approximately 30-50% of job applicants consciously try to elevate their scores (Donovan, Dwight, & Hurtz, 2003; Griffith et al., 2007)[1]. Faking on the part of job applicants represents part of a strategic approach to compete for employment through the use of deliberate attempts to manipulate information and create false beliefs and impressions in others (Masip, Garrido, & Herrero, 2004, as cited

---

This dissertation follows the style and format of the *Journal of Applied Psychology*.

[1] For example, Donovan et al. (2003) surveyed applicants using a randomized response technique (to ensure anonymity), and found that applicants admitted to extensive faking efforts: 56% of respondents admitted to exaggerating positive characteristics, and 17% admitted to outright fabrication.

in Griffith & McDaniel, 2006). Unfortunately, applicant faking has proven difficult to detect and study since, by its very nature, the behavior is hidden. While substantial research has been devoted to detecting and mitigating applicant faking, no techniques appear to solve the problem adequately (Berry et al., 2007; Morgeson et al., 2007). In addition, while several studies have attempted to identify processes and characteristics relevant to faking on measures such as personality tests, much less is known about faking on other kinds of non-cognitive measures such as integrity tests, which differ substantially in both content and application.

Research on faking has focused on two general themes: (1) whether applicants can and do fake (e.g., Cunningham, Wong, & Barbee, 1994; Rosse, Stecher, Miller, & Levin, 1998; Ryan & Sackett, 1987), and (2) the degree to which faking affects the validity of these pre-employment tests (e.g., Barrick & Mount, 1996; Cunningham et al., 1994; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Ones et al., 1996; Ones & Viswesvaran, 1998a; Rosse et al., 1998; Schmitt & Oswald, 2006; Zickar & Drawsgow, 1996; Zickar, 2000). If the extent of faking were relatively constant across individuals, then faking would not generally harm the quality of personnel decisions based on test scores; such decisions are generally based on rank-order (Douglas, McDaniel, & Snell, 1996), not mean-level information. However, research has demonstrated the existence of individual differences in the extent to which applicants fake (Mueller-Hanson, Heggestad, & Thornton, 2006; Morgeson et al., 2007), though our understanding of these influences is limited.

The goal of the current paper is to sketch a theoretical structure for the faking process to explain why and when applicants will intentionally distort their responses (i.e., fake) on integrity tests, a commonly used non-cognitive selection tool. The theory of planned behavior framework (Ajzen, 1985; 1991; and Ajzen & Madden, 1986; Beck & Ajzen, 1991) will be used to build a decision-based model of the faking process. This model will (1) describe the basic decisional processes that relate to faking behavior, (2) differentiate between two primary aspects of faking behavior, (3) assess some of the ways in which individual differences in personality and ability are related to individual differences in faking-related attitudes, behaviors, and outcomes, and (4) examine the effect of faking on individual integrity test scores, and the estimated validity of those scores for predicting workplace deviance, a key organizational outcome.

Integrity in the workplace

Research on integrity subsumes a wide variety of behaviors such as illegal drug use, fraud, petty theft, and unwarranted absenteeism (Berry et al., 2007; Murphy, 2000; Ones, Viswesvaran, & Schmidt, 2003). Tests designed to measure integrity are especially popular in industrial sectors such as retail and finance, where they are used to help select applicants for positions that have unsupervised access to money or merchandise (Berry et al., 2007; Sackett, Burris, & Callahan, 1989), or other positions in which employee theft or dishonesty produce high costs to organizations (Murphy, 1993; Ryan & Sackett, 1987; Sackett & Harris, 1984). These tests are often used as part of a larger selection process to screen out individuals whose dishonesty or CWBs could lead to significant long-term costs if they were to be hired.

Reviews of integrity tests have generally found them to be valid selection tools, capable of predicting a number of important organizational outcomes. Integrity tests demonstrate predictive and concurrent validities for criteria such as job performance, workplace violence, absenteeism, turnover, substance abuse, theft, and other CWBs (Berry et al., 2007; Murphy, 2000; Ones et al.,1993, 2003; Ones & Viswesvaran, 1998a). For example, research has found that integrity test scores show significant predictive validities, such as mean correlations of .21 (.34 correcting for statistical artifacts) with job performance and .33 with measures of counterproductivity (.47 correcting for statistical artifacts; Ones et al., 1993; Schmidt & Hunter, 1998; Wanek, Sackett, & Ones, 2003). The widespread use of integrity tests is likely to continue due to a lack of viable alternative measures (for example, federal legislation drastically limits the use of polygraph testing, one of the more prominent alternatives to paper-and-pencil integrity tests) and a relative lack of adverse impact against protected groups (O'Bannon, Goldinger, & Appleby, 1989; Ones, Viswesvaran, & Schmidt, 1995; Sackett & Wanek, 1996). Specifically, subgroup differences are negligible and there is no correlation of race, sex, or age with integrity scores (Ones & Viswesvaran, 1998b).

Despite broad agreement on the criterion-related validity of integrity tests, there are several potential issues with their use. One major issue is exactly what kind of test is being used; this is a critical issue, as there are multiple types of integrity tests which involve different types of items and ways of measuring integrity. Research suggests these tests might assess slightly different constructs. Tests designed to measure integrity consist of two major types designed to assess individual integrity in different ways,

using different items: (1) *covert* (or "veiled purpose") tests and (2) *overt* (or "clear purpose") tests.

Covert tests are personality-based tests designed to indirectly infer integrity from responses to questions that are not obviously related to integrity. Most covert integrity tests share a number of features, and are composed of items designed to measure one or more of the following areas: (1) general personality traits and beliefs related to integrity, and (2) reactions to hypothetical situations. Some personality–based tests of integrity are related to constructs such as thrill-seeking, socialization, and resistance to authority (Sackett et al, 1989). More recent tests have incorporated broad dimensions of personality that have been empirically linked to theft and CWBs, notably conscientiousness, agreeableness, and emotional stability (Berry et al., 2007).

Overt integrity tests also share several features. For example, they are generally composed of items that measure the following domains: (1) direct admissions of illegal or questionable activities and (2) attitudes or opinions regarding illegal or questionable behavior. Overt tests usually include direct and indirect measures of perceptions of norms relating to honesty. The rationale for this type of measure is that some individuals believe dishonesty, theft, and similar behaviors are common or widely accepted, and these individuals are more likely to engage in such behaviors themselves. Even questions asking for direct admissions of wrongdoing are probably best understood as measures of perceived norms. For example, individuals who are willing to admit on integrity tests to theft, dishonesty, and rule-breaking might well be individuals who believe that such behaviors are commonplace, or at least likely to be tacitly accepted by others.

Research into the construct validity of integrity tests has produced different views on what these two major types of tests measure. Some researchers have suggested that both types of integrity tests tap into the same underlying construct(s). Other researchers have argued that personality-based tests and overt integrity tests may represent slightly different constructs, and tap into different individual traits and experiences.

Early research by Ones (1994) involved a confirmatory factor analysis (CFA) on three overt tests and four personality-based tests. She found that the relationships between the variables were best described by a hierarchical model in which separate overt and covert factors had strong loadings on a single overall integrity factor, which was labeled "conscientiousness." In a similar study, Hogan and Brinkmeyer (1997) found that scores on two different overt and covert tests loaded on different first-order factors, which loaded on a single higher-order factor again labeled "conscientiousness." Wanek et al. (2003) found that overt and covert tests had similar overall relationships with four separate principal components of integrity (antisocial behavior, socialization, positive outlook, and diligence).

However, other research has found evidence that personality-based tests and overt tests may tap into different constructs. Wooley and Hakstian (1992) found that scores on an overt test correlated very poorly with scores on three different personality-based measures (largest $r = .39$, other rs smaller or not significant). Ones (1994) found that overt tests appeared to be less cohesive in focus (represented by smaller intercorrelations among different measures of the same construct) than personality-based

tests (r = .45, corrected for statistical artifacts, compared to .70). The meta-analysis also found that overt and covert tests tended to have relatively low correlations with one another (r = .39, corrected for statistical artifacts), suggesting that scores on different types of tests may vary considerably (see Ones, 1994). In addition, there is evidence that the two types of tests may differ in their relationships to personality constructs. Ones (1994) also showed that conscientiousness, agreeableness, and emotional stability correlated .45, .44, and .37 with scores on personality-based tests, and .39, .34, and .28 with overt test scores (all correlations corrected for unreliability and range restriction).

While these correlations are of roughly similar magnitude, evidence for consistently different patterns of relationships with one another and other variables led several major reviews of the integrity testing literature to conclude that overt and covert tests tap into broadly similar constructs in dissimilar ways (Cullen & Sackett, 2004) and that research on the two types of tests cannot be treated as interchangeable (Berry et al., 2007). Research has also found that test-takers react differently to the two types of tests – for example, perceiving that overt tests are more procedurally fair (Whitney, Diaz, Mineghino, & Powers, 1999). Finally, the literature suggests that overt tests, given their clear-purpose nature, may be more susceptible to applicant faking than covert tests (e.g., Murphy, 2000; Sacket & Harris, 1984). To date, a significant portion of research into faking and non-cognitive selection tests has focused on personality tests. However, more research on the faking of overt integrity test is needed. Faking on integrity tests is psychologically different from faking on personality tests because of factors such as the strong moral undertones associated with failing or passing integrity tests. In this sense,

the need to manage one's response should be more salient for those taking an integrity tests because of the clear implications of not passing the test (i.e., the possibility of being labeled dishonest).

Faking and selection tests

Past research has identified a number of different types of response distortion that have the potential to negatively affect selection processes. One of the major distinctions, first made by Meehl and Hathaway (1946), was between the concepts of *unconscious self-deception* (social desirability) and *consciously motivated faking* (impression management). Self-deception refers to positively biased responses that respondents believe to be true (Paulhus, 1984). Guion (1965) suggested that self-deceptive responses occur when test-takers lack self-insight or unintentionally use the test to present an idealized self-concept. By contrast, faking (i.e., impression management) involves a deliberate and insincere attempt to create a favorable impression in others (Paulhus, 1984).

Both of these facets of response distortion represent important barriers to accurate measurement in a selection context, leading to bias that affects estimates of applicants' true standing on some construct of interest. Response distortion is seen as a particularly important problem in the domain of non-cognitive selection tests, such as personality tests. Since these tests have no clear right or wrong answer, individuals have much more latitude in the way that they use their responses to present an image (whether consciously or unconsciously). Social desirability represents a dispositional source of honest error (McFarland & Ryan, 2000; Tett et al., 2006); such distortions are not

expected to vary much across different situations (Paulhus, 1984). However, faking represents deliberately dishonest error; this sort of deceptive behavior is especially troubling for employers (Griffith & McDaniel, 2006), who are justifiably worried about the potential for hiring someone who is willing to tell them deliberate falsehoods for personal gain.

To detect and measure applicant faking, researchers have used several different measures. One of the most common approaches is the use of scales to detect faking and other types of response distortion (Hough et al., 1990; Rosse et al., 1998), sometimes referred to as lie scales or social desirability scales (Paulhus, 1991). Lie scales designed to detect faking (e.g., active impression management) represent an attempt to assess faking behavior across different contexts and situations (McFarland & Ryan, 2006). In this sense, it is more like a disposition than a measure of conscious response distortion elicited by the situation (Christiansen, 1998; Smith & Ellingson, 2002). Difference scores based on comparisons between scores from different conditions are another popular way of measuring faking, since they are seen as more objective. However, they are subject to a number of statistical issues (Edwards, 2001). A third way of measuring faking is the use of self-reports (Dwight & Donovan, 2003; Hurtz & Bohon, 1997): simply asking individuals to estimate the extent to which they distorted their responses on a given test. Compared to lie scales, difference scores and self-report measures of faking tend to be very situation-specific and predict who is more likely to fake within a given context (McFarland & Ryan, 2006).

However it is defined (cross-situation or situation-specific), faking has the potential to affect the validity of the selection processes used by organizations, especially integrity tests. Test scores may have long-lasting effects on individual outcomes. False positive decisions could mean the hiring of individuals who are prone to engage in dishonest behaviors in the employment setting; false negatives associated with low integrity test scores not only lead to rejection for the position, but imply the applicant was dishonest, a serious social stigma (Griffith & McDaniel, 2006). For these reasons, deliberate faking is a particularly important issue for applicant selection systems in general, and integrity tests in particular.

Unfortunately, there appears to be little agreement about the impact of faking on the validity of selection tests, perhaps due to inconsistencies in the way faking has been assessed (for example, lie scales, difference scores, and self-reports). Some research has found that response distortion has little to no effect on criterion-related validity (e.g., Ones et al., 1996; Schmitt & Oswald, 2006), and that corrections for it do not change criterion-related validities (Hough, 1998; Ones & Viswesvaran, 1998a; Schmitt & Oswald, 2006). Schmitt and Oswald (2006) concluded that given typical selection scenarios, removing respondents displaying high levels of response distortion has minimal impact on mean performance at the aggregate level. These findings have led some authors to take a dim view of the need for additional faking research (e.g., Morgeson et al., 2007).

However, a growing body of literature suggests that conclusions about the innocuousness of faking are premature, particularly in the case of integrity tests (Griffith

& McDaniel, 2006). In their recent review of the use of personality-based selection tests, Morgeson et al. (2007) cite a number of studies which find that faking can affect hiring decisions, especially in situations with low selection ratios and top-down selection procedures. Other research has found that faking introduces construct-irrelevant variance and can degrade test validity (Griffith, 1998), and that the differences between applicant and non-applicant settings can affect the construct-related validity of selection tests (Stark, Chernyshenko, Chan, Lee, & Drasgow, 2001).

Finally, psychometric studies have shown that tests with similar levels of predictive validity at the test level may have different measurement properties that can affect the accuracy of the selection test and the quality of selection decisions being made. Millsap (2007) has pointed out that regression (predictive) invariance is separate from factorial (measurement) invariance; thus, it is possible to have a construct with the same predictive validity for two groups (for example, faking and non-faking applicants) but for mean differences in the construct between the groups to result in a selection decision which favors one group over another, even when the predictive validity of a test is the same for the two groups. This has the effect of reducing the sensitivity of the selection test and increasing bias in selection rates (Millsap, 2007), and suggests that faking could be a problem even if there was no negative effect on criterion-related validity coefficients.

Beyond the conflicting evidence in the literature, there are a number of conceptual issues with the current research which suggest that additional research on faking is warranted, especially with regard to overt integrity tests. First, existing faking

research has been based on a wide variety of terms and concepts, including faking, social desirability, response distortion, and impression management. These different terms have sometimes been used interchangeably (e.g., Ones et al., 1995; Ones et al., 1996; Rosse et al., 1998). This confusion makes it difficult to assess the generalizability and value of previous research, given that different researchers have used the same term to refer to fundamentally different response distortion processes. Hence, the lack of consistent terminology and operational definitions of the phenomenon represents a potential source of inconclusive or contradictory results (Griffith & McDaniel, 2006; Rosse et al., 1998). Moreover, most of these results are based on studies of personality tests, not overt integrity tests, which comprise different content and items (Murphy, 2000).

There have also been inconsistencies in the way that faking is measured. Some researchers have used lie scales such as the BIDR (Paulus, 1991) to operationalize faking (Rosse et al., 1998); others have used experimental manipulations of response instructions (Griffith et al., 2006; Ryan & Sackett, 1987) and difference scores (McFarland & Ryan, 2000) or self-reports (Dwight & Donovan, 2003; Hurtz & Bohon, 1997). Given the known limitations and problems with the use of lie scales and related measures (Alliger & Dwight, 2000; Hurtz & Alliger, 2002; McFarland & Ryan, 2000), it is not entirely clear whether the conflicting conclusions of faking research are partly the result of differences in how faking is measured. Research that includes multiple operationalizations could help provide results that generalize across different types of faking, or identify limitations in these measures. Moreover, unlike personality-based

tests, overt integrity tests do not attempt to hide their purpose and do not always contain lie scales (Sackett & Harris, 1984); therefore, the conclusions of research based on covert or personality-based tests may not necessarily apply to overt integrity tests.

A third issue is that deliberate faking, regardless of its effect on test performance and predictive validity, represents a dishonest behavior. An individual willing to engage in one type of dishonest behavior may be more likely to engage in other types of dishonest behaviors – in fact, this is the entire premise behind questions about past behaviors on overt integrity tests. Thus, even if faking has a limited effect on the ability of a test to predict future job performance, it is possible that the increased selection of individuals willing to lie on a test could increase the frequency of behaviors such as theft or other serious CWBs (Patterson, DeBaryshe, & Ramsey, 1989) over an individual's tenure in the organization.

Current findings in faking research

Based on these issues, further research on faking appears to be warranted. In the faking literature, several topics have proven to be of central importance. One such issue is the extent to which applicants are able to fake, and whether or not they actually do so. Another, equally important issue is the extent to which faking on selection tests affects the quality of selection decisions made on the basis of that test.

Overall, research suggests that test takers are both willing and able to fake on integrity measures (Cunningham et al., 1994; Ones et al., 1996; Robie, Brown & Beaty, 2007; Ryan & Sackett, 1987). The dominant method in studies that examine whether people are able to fake are controlled experiments (using both between- and within-

subjects designs) that compare mean integrity test scores obtained under instructions to respond honestly to those obtained under instructions to fake in various ways (e.g., Sackett & Wanek, 1996). Research that compares scores generated from such instructions has been used to demonstrate the extent to which test takers can fake.

Ryan and Sackett (1987) were among the first to show that test takers are able to follow instructions to fake, respond honestly, or to respond as applicants. The results revealed that participants who were instructed to fake good could easily do so, scoring a full standard deviation above the overall mean. However, those instructed to respond as job applicants provided responses more similar to those from the "honest" condition than the "fake good" condition. While a 24% difference was found between the faking good group and the honest group, only an 8% difference was found between the job applicant group and the honest group. A more recent study showed that honest responders take less time to complete and make fewer corrections to their personality inventories than faking responders (Robie et al., 2007), and that respondents may be categorized into one of three different categories: honest responders, slight fakers, and extreme fakers.

According to a current review of integrity testing research (Berry et al., 2007), there appears to be a consensus that respondents are able to fake when instructed to do so. However, there is some concern about the extent to which sample type might affect the degree of faking displayed by test takers (i.e., whether applicants might fake more than a student sample and whether applicants might fake more than incumbents). Research on personality tests suggests that sample type (e.g., applicant versus incumbents) can affect mean differences by as much as .50 *SD* (Hough, 1998; Hough et

al., 1990). In another study on pre-employment personality tests, Rosse et al. (1998) found that job applicants engaged in response distortion on more items than incumbents, and that applicants were able to distinguish between items that did or did not reflect job-relevant information. Among job applicants who were hired, 87% had extremely high response distortion scores on the test, especially on items that were job-related.

In another study of overt integrity test scores of managerial applicants and incumbents, Van Iddekinge, Raymark, Eidson, and Putka (2003) found that applicants only scored .09 *SD* higher than incumbents, suggesting the integrity test was resistant to faking. In a recent meta-analysis, Birkeland, Manson, Kisamore, Brannick, and Smith (2006) found that job applicants scored higher than non-applicants on extraversion, emotional stability, conscientiousness and openness. They interpret these mean differences as measures of faking, and note that the degree of faking on each dimension varied depending on job type.

In a more recent study, Hogan et al. (2007) collected data from job applicants who completed a personality test twice – individuals completed the application process once, were turned down, and reapplied 6 months later. Participants included over 5,200 rejected applicants (from a population of over 260,000 applicants) who applied for a customer service job with a nationwide U.S. employer in the transportation industry. Although the applicants had increased experience with the test and demonstrated motivation to get the job, Hogan et al. (2007) found that about 5% of the rejected applicants improved their scores on the second administration more than would be predicted by chance. The study also found that social skills, social desirability, and a

covert measure of integrity were significant predictors of change in the personality test scores.

However, research by Griffith et al., (2007) found that 30%-50% of applicants elevated their scores when applying for a job, with applicant scores that fell outside the 95% confidence interval established by their honest scores. They also found that faking had substantial effects on hiring decisions: at a .5 selection ratio, 31% of applicants who were hired on the basis of their applicant test scores would not have been hired on the basis of their scores as incumbents. At a .1 selection ratio, 66% of applicants who would have been selected on the basis of their applicant test scores had faked their way to the top. Thus, Griffith et al. (2007) showed that faking can have substantial effects on rank ordering and hiring decisions in applicant settings. Given Griffith et al.'s (2007) finding that a high percentage of applicants are able to alter their test scores as applicants, it is possible that Hogan et al.'s (2007) results might be attributable to the rejected applicants simply being bad at impression management. In fact, some of the rejected applicants received lower scores on the second administration of the same test.

The second major issue in faking research is whether or not faking affects the validity of selection procedures. In this area, there have been a number of studies; several of them have supported the idea that faking harms test validity, but a number of influential studies have found little to no effect of faking.

Some have argued that the effects of faking are minimal, based on studies showing little or no faking-related attenuation in criterion-related validities (e.g., Barrick & Mount, 1996; Hough et al., 1990; Ones & Viswesvaran, 1998a). Cunningham et al.

(1994) found that while people can and do fake, integrity tests possess predictive validity in spite of impression management. In a series of meta-analyses, Ones et al. (1993; 1996) found performance criterion validities of .40 for applicants and .29 for incumbents, and found that response distortion did not attenuate or suppress the predictive validity of integrity tests. This suggests that validities under conditions most associated with the motivation to fake (e.g., applicant settings) were not attenuated, but were actually 38% larger. In a military recruiting sample, Hough et al. (1990) measured faking effects using mean-level personality scores and response distortion scores of applicants and incumbents. Although subsequent analyses by Rynes (1993) found that 29% of participants faked, Hough et al. (1990) found only small differences in criterion-related ability associated with faking on personality tests.

However, a number of researchers have pointed out that faking may have substantive effects on selection system outcomes even in the absence of direct effects on criterion-related validities, and that typical assessments of criterion-related validity are not sufficient to rule out the impact of faking. Instead, the effects of faking should be studied through an examination of the quality of individual hiring decisions (Alliger & Dwight, 2000).

Researchers have noted that the sensitivity of criterion-related validity coefficients to faking is influenced by the size of the applicant pool, whether top down strategies are used, and the selection ratio (e.g., Donovan et al., 2003; Rosse et al., 1998). Under top-down selection strategies, even if only a few individuals fake, those who fake will rise to the top of the distribution and will increase their chances of being

hired, regardless of the overall validity coefficient (Donovan et al., 2003). In this scenario, successful fakers actually possess a lower "true score" on the target construct than their test scores would indicate (Donovan et al., 2003), increasing the probability of a false positive hiring decision. For instance, Rosse et al. (1998) found that the percentage of job applicants with elevated faking scores increased as selection ratios decreased, and that corrections for faking affected the rank order of applicants, and simulation studies by Zickar and Drasgow (1996) and Zickar (2000) found that test takers who faked were disproportionately ranked among the top applicants, despite few changes in the observed validity coefficient.

Mueller-Hanson, Heggestad, and Thornton (2003) used an experiment in an applied setting to examine the effect of personality test faking on criterion-related validity and the quality of selection. They found that the incentive condition (in which test takers were given a $20 incentive) had more prediction errors among those with scores at the high end of the selection test score distribution. The mean difference between the control and incentive groups on the personality test was .41 *SD*. This same study also indicated that, as the selection ratio decreased, examinees from the "honest" condition were selected proportionally less often than examinees from the "incentive" group (36% versus 64% at a selection ratio of .10). Finally, their results showed that when the selection ratio was small, individuals in the "incentive" condition were also likely to have lower mean levels of performance than those in the "honest" condition.

Overall, these studies suggest that faking will continue to play an important role in selection test research and applications – including integrity testing – contrary to

recent statements about the potential futility of faking research (see Griffith & McDaniel, 2006; Morgeson et al., 2007). Specifically, research on faking has shown that individuals can and do fake, and that faking can affect selection decisions. However, less is known about the specific processes involved in individual decisions to engage in impression management behaviors on selection tests, or how that process is affected by individual differences like personality or cognitive ability. Thus, an important next step in faking research is the development of models of faking behavior for various non-cognitive selection measures, such as integrity tests. Some initial attempts to develop such models have been proposed in the case of personality tests; these will be reviewed in the next section, and will serve as the basis for a model of some of the major causes and consequences of faking on overt integrity tests.

A MODEL OF FAKING DECISIONS

Recently, a number of authors have suggested that attitudes towards faking might be used as a proxy for the intent to fake and the likelihood of engaging in faking (Rees & Metcalfe, 2003) and play an important role in the decision to fake (McFarland, 2000). According to this perspective, faking is the result of individual decisions or intentions based on individual attitudes and cognitions towards faking (Mueller-Hanson et al., 2006; Snell, Sydell, & Lueke, 1999). Empirical support for this perspective has been found in at least two prior studies. Mueller-Hanson et al. (2006) found that attitudes and cognitions towards faking predicted willingness to fake, and McFarland (2000) found that faking on a Big Five measure of personality (NEO-FFI; Costa & McCrae, 1989, cf. McFarland, 2000) could be explained by individual intentions to fake, which were related to individual cognitions and attitudes towards faking.

However, individual cognitions towards faking, faking behaviors, and even the end result of such faking may be influenced by a number of individual differences; and faking itself is not necessarily a unitary construct. Thus, the proposed model integrates a cognitive/decision-theoretic model of faking on overt integrity tests with research on key individual differences, taking into account critical distinctions among faking criteria. At the heart of the proposed model is Ajzen's (1985; 1991) theory of planned behavior (TPB), which has been used for decades to study the link between cognitions and behavior.

Theory of planned behavior

According to the TPB, behaviors are a function of three behavior-specific attitudes and cognitions, mediated by intentions. Although widely used in other behavioral contexts, the TPB has rarely been used to study faking in the context of selection tests, with two relatively recent exceptions. McFarland (2000) found that faking behaviors on four of the five NEO-FFI scales (neuroticism, extraversion, agreeableness, and conscientiousness) were predicted by intention to fake, and that intention to fake was significantly (albeit weakly) related to individual attitudes towards faking and perceived behavioral control, but not social norms. More recently, Mueller-Hanson et al. (2006) found that a common factor represented by the three TPB-related cognitions was a strong predictor of intentions to fake ($\beta = .99$, $p < .01$), which was itself a strong predictor of faking behavior on a five-factor model test of personality ($\beta = .70$, $p < .01$).

These results confirm that faking cognitions can be a significant factor in the decision to fake. However, the results of the larger study by McFarland (2000) are slightly mixed. While intentions to fake were related to applicant-honest difference scores on some personality measures (especially conscientiousness and agreeableness), relationships were generally small, with large variation in the correlation between intentions and faking behavior across the different measures of personality (r from .00 to .20).

These results also suggest that intentions to fake are not equally good predictors of applicant faking behavior across different types of tests, and that the appropriateness

of a TPB-based model of faking may be moderated by the content of the test. One major difference in test content is the nature of the items used on the test: meanings of items on overt (i.e., clear-purpose) tests are more obvious than those on covert (i.e., personality-based) tests, by definition. This implies that, compared to personality-based tests, applicant decisions to fake could be even more important for overt integrity tests.

A second difference between overt integrity tests and other non-cognitive measures is the nature of the construct: integrity has clear overlap with conscientiousness and neuroticism (Ones, 1994), but has particularly powerful meanings attached to high and low test performance. A person who receives low scores on an overt integrity tests receives strongly negative value-based labels like "thief," "liar," or "unethical." By contrast, low scores on personality-based tests (even on related constructs like conscientiousness) do not always carry such clearly negative connotations; responses to certain items could be interpreted as either "disorganized" or "big picture person."

Together, these two factors suggest the possibility that intentions to fake will be more easily translated into increased scores for overt integrity tests. If this is true, then the TPB would be even more appropriate for such tests than for other non-cognitive measures. This would explain the mixed findings of McFarland (2000) and expand on the results of Mueller-Hanson et al. (2006).

In this proposed study, the TPB (Figure 1) was used to predict intentions to distort responses (i.e., fake) and subsequent applicant faking behavior on an overt integrity test. The framework identifies three cognitive components (attitudes, social

norms and perceived behavior control) that are believed to jointly predict the intention to perform a given behavior, with intention mediating the effects of the three cognitive components on the actual behavior (Ajzen, 1985; 1991; and Ajzen & Madden, 1986). Beck and Ajzen (1991) found that attitudes regarding several types of dishonesty are related to intentions to behave in a dishonest manner. Positive attitudes toward a behavior are linked to stronger intentions to perform that behavior. This study will assess whether individual attitudes towards faking, subjective faking norms, and perceived behavioral control over faking explain variance in faking on pre-employment integrity tests, and whether these effects are mediated by individual intentions to fake. A description of these individual components is presented below.

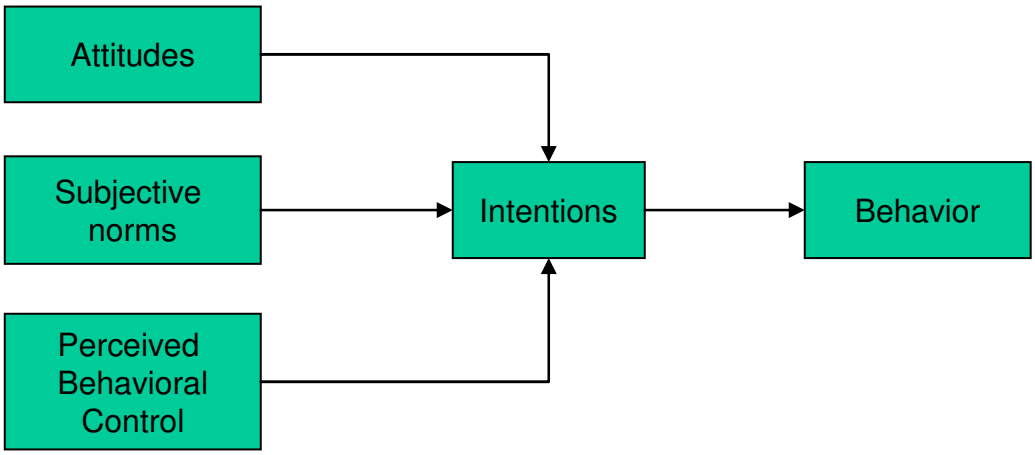

Figure 1. Theory of Planned Behavior

Attitudes refer to an individual's evaluative appraisal of the behavior in question. It has been found that positive attitudes toward a behavior lead to greater intentions to perform that behavior. For example, Beck and Ajzen (1991) found that attitudes toward

dishonest actions predicted intentions to cheat on a test, shoplift, and lie. It is expected that these results will generalize to faking on an integrity test.

Subjective norms refer to perceived social pressure to engage in a given behavior. Studies on the TPB have consistently shown that when subjective norms toward a behavior are positive, the intention to perform that behavior will be greater (Ajzen, 1991, Schifter & Ajzen, 1985). For example, when individuals perceive that important others (e.g., parents or friends) approve of the behavior under consideration, they should be more likely to intend to engage in that behavior (Ajzen, 1991). It is expected that these results will generalize to faking on an integrity test, such that those who have more positive subjective norms toward faking will be more likely to have the intention to fake on the test.

The third variable in the TPB framework is perceived behavioral control. According to Ajzen (1991), this variable is similar to Bandura's perceived self-efficacy (Bandura, 1986; 1991). Perceived behavioral control refers to an individual's perceptions of their ability to perform a given behavior. If the person believes they can perform the behavior with ease, he or she would be more likely to intend to behave in that manner. In this study, it is expected that those who perceive themselves to have greater control over faking behavior will have a greater intention to fake.

Finally, greater intentions to perform a behavior increase the likelihood that the behavior is performed (Ajzen, 1991; Ajzen & Fishbein, 1980; Beck & Ajzen, 1991; Fishbein & Ajzen, 1975). Intentions predict a wide variety of behaviors, from losing weight (Shifter & Ajzen, 1985) to stealing (Beck & Ajzen, 1991). Intentions to fake will

be positively related to attempts to engage in faking behavior, so people with a greater intention to fake will fake more extensively on the test.

It is expected that the more favorable the attitude and the subjective norm, and the greater the perceived control, the more a person should be likely to intend to engage in a behavior. Taken together, the current research suggests that:

H1: Intentions to fake will mediate the relationship of applicant faking behaviors with attitudes, social norms, perceived behavioral control.

Specifically, attitudes towards faking, perceived social norms towards faking, and perceived behavioral control over faking will predict intentions to fake on an overt integrity test. Intentions to fake will predict actual faking behavior on overt integrity tests. In addition, this study will rely on multiple operationalizations of faking, which will provide information on the extent to which the model may be generalized across different measures of faking, including difference scores, self-reports, and lie scale measures.

Faking behaviors and overt integrity test performance

With the TPB, the model provides a potential explanation for the way in which cognitions about faking on selection tests translate into actual faking behaviors. However, faking behaviors are primarily of interest to employers and researchers because of the potential effect they have on observed applicant test performance. In order to understand how individual faking behaviors on integrity tests translate into applicant test performance, research suggests that two separate aspects of faking play an

important role: *applicant faking behavior* (which might also be referred to as *motivated faking* or *typical faking*) and *faking ability*.

Applicant faking behavior reflects "typical" or "will do" faking, of the sort expected to arise from the level of motivation induced by a selection context. This is the aspect of faking most studies of applicant faking measure. More recently, studies of faking have begun to pay attention to a second aspect of faking – ability to fake. Faking ability is the extent to which an individual is able to successfully fake on the test, regardless of their actual inclination to do so in a job application context. Thus, it represents "can do" or "maximal" faking on a selection test. Faking ability has been operationalized two primary ways: first, through measures of test knowledge (e.g., McFarland, 2000; Mueller-Hanson et al., 2006); second, through the use of instructions to applicants to "beat" the test or maximize their score.

The distinction between will-do and can-do faking has previously been acknowledged as conceptually important. It has appeared in recent research, although studies have taken different views about their relationship and measured them in different ways. In a conceptual article, Snell et al. (1999) argued that successful selection test faking required both the motivation and ability to fake, and that different characteristics of the test, test-taker, or situation would ultimately influence the amount of successful faking observed on a given test. In their model, faking motivation and faking ability are expected to interact, since ability without motivation or motivation without ability are not expected to lead to successful faking. In other words, Snell et al. (1999) argue that faking on selection tests should display a similar relationship to the

classic "performance = motivation x ability" formula: increased ability should maximize the relationship between the applicant's faking behaviors and their resulting test performance.

The relationship between "can-do" faking (faking ability) and "will-do" faking (applicant faking) has been discussed in a small number of studies (for example, Mueller-Hanson et al., 2006), but the specific relationship between the two has rarely been tested. To the extent that empirical tests of their relationship have appeared in the literature, the results have been mixed. For example, McFarland (2000) found that increased knowledge about a selection test (by providing a description of the test constructs) actually *weakened* the relationship between faking intentions and faking behavior. Mueller-Hanson et al. (2006) found no evidence for any main effects of faking ability on faking intentions or faking behavior, or any relationship with faking motivation, but failed to test for any interactions of the sort predicted by earlier conceptual work on non-cognitive faking (McFarland, 2000; Snell et al., 1999). Despite the lack of empirical evidence, the basic premise articulated by Snell et al. (1999) remains: both ability and motivation are conceptually necessary in order to successfully fake on a selection test. Thus, existing research leads to the following two hypotheses:

H2: Increased applicant faking will result in elevated integrity test scores.

H3: Faking ability will moderate the effects of applicant faking on test performance, such that individuals who engage in faking and have a *high* ability to do so will outperform those who have a *low* ability to do so.

Individual differences and faking

Another area the faking literature suggests requires additional attention is the role played by individual differences in explaining observed individual-level variation in faking. The effect of individual differences on faking may occur at a number of different stages in the faking process. For example, individuals who are more conscientious or display greater levels of trait honesty may have different perceptions of the frequency or acceptability of faking behavior among their peers, due to either self-selection or social influence effects. Individuals who are less honest or conscientious may differ in the extent to which they believe such behaviors have negative consequences. There are a number of individual differences which theoretical and empirical research suggest might lead to individual variance in faking and, ultimately, integrity test scores.

Conscientiousness. One of the most potentially relevant predictors of integrity test faking behavior is conscientiousness, as research has reported links to both faking and integrity. For example, McFarland and Ryan (2000) found that individuals with higher conscientiousness scores had lower rates of faking (r = - .57) in an applicant instruction condition. Conscientiousness is also one of the core constructs measured by personality-based integrity tests (Hogan & Brinkmeyer, 1997; Sackett & Wanek, 1996; Wanek et al., 2003). However, research also shows that conscientiousness predicts attitudinal-based (i.e., "overt") integrity tests (Berry et al., 2007), not just trait-based integrity. In other words, individuals higher in conscientiousness are less likely to report attitudes consistent with low levels of integrity. In the context of the current model, this implies that:

H4: Higher levels of conscientiousness will predict more negative attitudes towards integrity test faking behaviors.

Cognitive ability. Successful faking on self-report measures such as integrity tests imply several conditions (Morgeson et al., 2007): that test-takers have sufficient self-insight to accurately respond to an item, that they understand the item as intended, and that they consciously deviate from the truth so they can create a specific impression. That is, successful faking is presumed to involve some level of ability (1) to understand written items on the test and identify the key concepts they contain, and (2) to reason about those concepts and their relationship to employer desires. In terms of individual differences which might be expected to predict these factors, the most obvious is general mental ability.

Cognitive ability has significant links to a number of lower-order factors that may affect the likelihood of successful faking attempts, such as language comprehension and lexical ability, memory and learning, inductive reasoning, among other factors (Carroll, 1993; Deary, 2000). There is also empirical evidence to support this relationship. For example, Brown and Cothern (2002) found that cognitive ability was a significant predictor of faking success, and Alliger, Lilenfeld, and Mitchell (1996) found that intelligence and the ability to fake an overt integrity test were significantly correlated (r = .36). More recently, the review by Berry et al. (2007) described recent research that has found a relationship between faking and cognitive ability.

Beyond its direct effect on the ability to fake, cognitive ability also has the potential to affect individual differences in perceived behavioral control. For instance,

Bell and Kozlowski (2002) found that cognitive ability is related to individual's perceptions of their ability and self-efficacy. That is, people with greater levels of ability tend to have greater beliefs in their ability to perform a task. Based on this research, it is therefore expected that:

H5: Higher levels of cognitive ability will be associated with a higher ability to fake, and

H6: Higher levels of cognitive ability will be associated with increased perceptions of behavioral control over faking.

Faking behaviors and overt integrity test validity

Ultimately, any process model of integrity test faking needs to take into account the effect that faking has on the validity of inferences drawn from the observed test scores. In the most recent review of the literature on integrity testing, Berry et al. (2007) review research showing that integrity tests are valid predictors for a number of criteria, such as CWBs – including rule breaking, theft, and cheating in a laboratory setting. In addition, meta-analytic approaches (Ones et al., 1993) have shown that integrity test scores are substantial predictors of supervisor ratings of performance (ρ = .35), self-admissions of narrow and broad criteria of counterproductive work behavior (ρ = .58), and various narrow criteria of theft (ρ = .52). However, research on the effect of faking on integrity test validity has been less clear.

Ones and colleagues (1996; 1998a) provided evidence suggesting that scores on social desirability scales did not moderate the ability of integrity tests to predict supervisory ratings of job performance. They also found that including social desirability

scores did not appear to reduce the ability of integrity test scores to successfully predict training performance (ρ = .36) or CWBs (ρ = .32). However, these results were based on 20 studies, used only social desirability scale scores (i.e., the meta-analysis included no measures of actual faking behavior), and did not differentiate between the effects of social desirability scores on overt integrity tests or personality-based tests. In a subsequent study, Jackson, Wroblewski, and Ashton (2000) found that "typical" (i.e., will-do) faking reduced the criterion-related validity coefficients of an integrity test with self-report CWBs from r = .48 to r = .18. Thus, the overall picture with regard to faking behaviors and the validity of integrity tests is unclear.

Although the existing evidence is unclear, from a psychometric perspective, factors that affect the quality of measurement are expected to have an effect on a measure's criterion-related validity, and faking behaviors represent a significant source of response distortion and bias. Even in the case where predictive validity is not substantially harmed, Millsap (1997; 2007, see also Borsboom, 2006) has noted that measurement issues may still affect selection criteria, by reducing the sensitivity of the selection test – and, as a result – the accuracy of the decisions made using that instrument. In the area of personality testing, these conclusions have been supported by others (e.g., Morgeson et al., 2007). Thus, to the extent that individuals are able to successfully fake on an overt integrity test, support for the following hypothesis is expected:

H7: Increased levels of applicant faking are expected to moderate the relationship between observed integrity test scores and CWBs, with

increased levels of faking reducing the strength of the relationship.

Additional potential effects of faking

Another way of looking at the validity of integrity tests is through the quality of hiring decisions being made (Christiansen, Goffin, Johnston, & Rothstein, 1994; Rosse et al., 1998). Morgeson et al. (2007) note that faking on personality tests may do more than simply raise scores – to the extent that individuals may fake differently on the basis of their individual differences and their situations, and that these individual differences in faking may result in changes to the rank order of applicants. In the case of overt integrity tests, even if measures of predictive validity are robust to applicant faking, the quality of employment decisions being made on the basis of the test will be a function of many factors such as selection ratios, the type of decision being made (e.g., top down selection versus other systems), and the extent to which individuals differentially increase their scores on the test.

To add to the complexity of this area of study, research on faking suggests that not only do individuals fake, but that that some individuals fake more than others. For example, Griffith et al. (2007) found substantial individual variance in the magnitude of changes in test scores between applicants versus incumbents who were instructed to respond honestly. Looking at just the top ten applicants out of a total of 60, effect sizes associated with applicant responding (i.e., motivated faking) ranged from .11 to 1.54. This broad range of scores suggests that there is substantial variance in individual faking, even for a narrow band of scores on a measure of conscientiousness.

Other studies have found that – regardless of mixed findings on the effects of faking with respect to criterion-related validity – faking can still have substantial effects on the measurement properties of the test. Stark et al. (2001) found that personality and impression management scales both exhibited substantial differential item functioning (DIF), and some degree of differential test functioning (DTF), leading them to conclude that each test contained items that favored one group (applicants or non-applicants) over another, which could lead to observable differences at the test level. In addition to mean level differences, they found that the situation (applicant versus non-applicant) affected the construct-related validity of the personality and impression management scales, and that applicant samples showed lower levels of reliability, consistent with previous research.

Additional research by Zickar (2000) and Zickar and Drasgow (1996) came to similar conclusions; in a series of simulations, they found that even though faking may not necessarily affect the overall validity coefficients, individuals who faked tended to be overrepresented in the top scores of the test. They also found that even low levels of faking could potentially have large effects on rank orders – when 5% of a hypothetical sample consisted of "slight fakers", such fakers made up 15% of the top scorers on the test. Fakers also appeared to have lower "true levels" of the target trait than honest responders who were also top scorers.

According to Zickar (2000), individuals who endorsed less positive response options when responding honestly were likely to endorse the same option when responding in a situation in which they were asked to fake. In other words, honest

individuals may fake less than others, even when asked to do so. This result is aligned with previous results showing that higher levels of conscientiousness are associated with a lower willingness to fake (Mueller-Hanson et al., 2006) and that impression management scores across different instructional conditions (often used as an indicator of faking; see McFarland & Ryan, 2006) reflect more of an individual trait than a situational variable. Thus, there appears to be evidence for significant individual differences in faking, even in the face of similar situational pressures introduced by different instructional conditions. To the extent that there are individual differences in responses to situational pressures to fake, research suggests the following hypothesis:

H8: Increased levels of applicant faking should be associated with differences in the rank order of individual integrity test scores, such that instructions to fake will alter the rank order of individual integrity test scores.

Figure 2 provides an overview of the model based on the different hypotheses, describing how the various cognitive, behavioral, individual differences, and outcome variables are hypothesized to relate to one another.

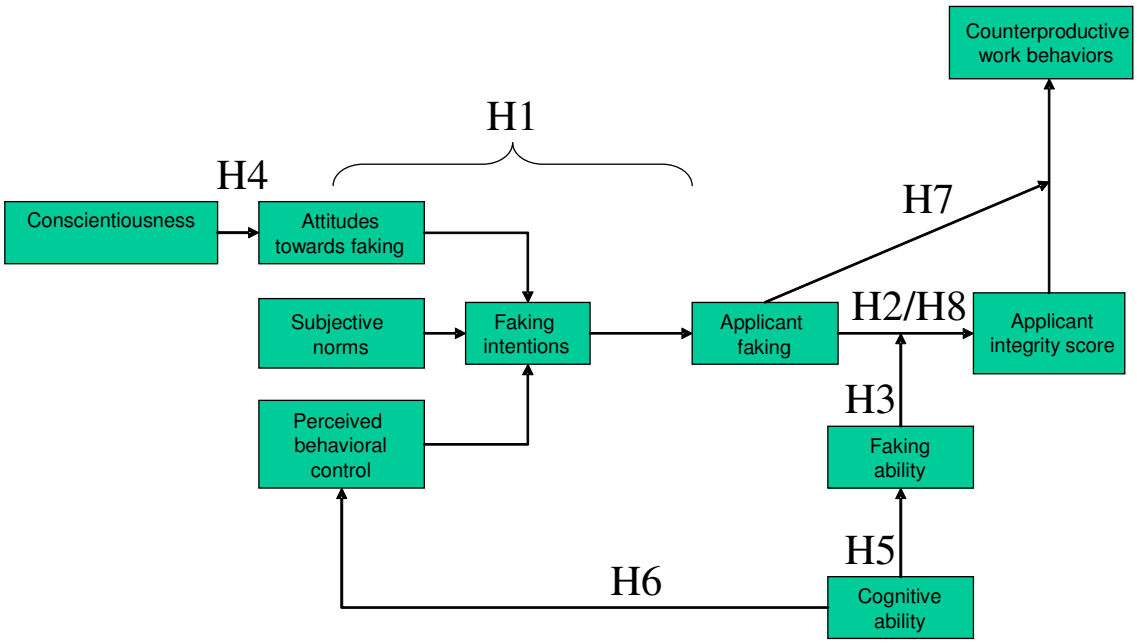

Figure 2. Overall model of proposed relationships and associated hypotheses

METHOD

To test the various hypotheses, two separate studies were conducted, using slightly different sets of measures and administrations (see below for details). The inclusion of a second study allowed for the use of an additional measure of cognitive ability that was not available in Study 1. The variations in administrations also provided some information on potential issues related to priming and temporal stability.

Participants

Study 1. A total of 289 undergraduates from a large southwestern university were recruited for Study 1. Of these individuals, 233 returned approximately 1 week later to complete the study at Time 2, so the attrition rate from Time 1 to Time 2 was 19%. The majority of participants were Caucasian (75%) and female (67%), with some minority representation (Latino, 12%; Asian, 6%; African American, 5%). The majority of the sample contained first or second year students (29% freshmen; 35% sophomore), but substantial numbers of third and fourth years students also participated (23% junior; 14% senior).

Many (67%) participants reported working at least 10 hours per week. This indicates that most had some recent experience in the workplace. Most (62%) of participants completed the applicant measures with a specific job in mind, of which roughly 40% involved retail or food service positions. Almost half of participants reported having taken similar employment tests in the past, but the use of such tests was not generally endorsed - only 38% agreed with the statement that “These sorts of tests should be used to select applicants.” The vast majority (92%) of respondents said they

responded honestly to items when instructed to do so, and a majority of students reported being motivated to take the test seriously by the rewards offered to participants [i.e., being the recipient of a $20 monetary award (56%) and earning extra credit toward their class (51%)].

Study 2. For Study 2, a total of 187 participants from a large southwestern university were initially recruited. There was a 14% attrition rate, with 160 participants completing the full study at time 2. This sample was slightly more diverse in terms of ethnicity (70% Caucasian; 19% Latino; 5% Asian; 5% African American) and gender (54% female). The sample was more stratified in terms of student classification, containing nearly 60% freshmen, and only 17% sophomores.

The sample generally had less current work experience (45% reported working at least 10 hours per week). Fewer participants also reported that their responses to items were given with a specific job in mind (34%), but those who did tended to use similar jobs as a reference point for responses (44% thought of a retail/food job when answering items). As with the Study 1 sample, a large proportion of participants reported having some previous experience with these sorts of tests (48%), and less than half agreed that such tests should be used (42%). Most (98%) of participants reported responding to items truthfully when instructed to do so, and most were motivated by the opportunity to obtain a $20 monetary award (55%) and earn extra credit toward their class (97%).

Measures

Conscientiousness. Conscientiousness was assessed using a 10-item scale from Goldberg (1999). This short public domain scale has been shown to correlate .79 with

the NEO conscientiousness scale (Goldberg, 1999). Coefficient alpha for Study 1 was .77; alpha for Study 2 was .80. Higher scores represent greater levels of conscientiousness.

Cognitive ability. Cognitive ability (or general mental ability, GMA) was measured in two different ways. First, it was measured using archival Scholastic Aptitude Test (SAT) scores (combined verbal and math) as a proxy for cognitive ability. Although imperfect, several authors have found substantial relationships between SAT scores and cognitive ability. Frey and Detterman (2004) found that composite SAT scores had a correlation of .72 with the Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998a, cf. Frey & Detterman, 2004) test, corrected for range attenuation. They also found a correlation of .82 with the Armed Services Vocational Aptitude Battery, another widely used proxy for cognitive ability. Rohde and Thompson (2007) found that composite SAT scores correlated .39 (uncorrected) with the Raven's Advanced Progressive Matrices and .69 with the Mill Hill Vocabulary Scales (Raven, Raven, & Court, 1998b, cf. Frey & Detterman, 2004), a verbal-based test of general cognitive ability. SAT scores are used in both Study 1 and Study 2.

In Study 2, cognitive ability was also measured using the Wonderlic Personnel Test, a short test designed to offer measures of *g* for use in a variety of applicant settings. The internal consistency of the Wonderlic reported by the manual varies from .82 to .94 (Wonderlic, 1992). Reliability in the current study was estimated at $\alpha = .80$.

Integrity. Integrity was measured using the *Employee Reliability Index* (ERI), a multiple-choice employment-based overt integrity measure developed by Ryan and

Sackett (1987). The ERI is a research tool designed to closely mirror existing overt integrity measures. It consists of 74 items divided into three subscales: 52 items measuring integrity-related attitudes, 11 items measuring admission to theft, and 11 items measuring social desirability. The social desirability scale is used to detect response distortion, and is not considered a part of the actual construct. The remaining 63 items were used to measure overall integrity. In the honest condition, coefficient alpha reliabilities were .93 for Studies 1 and 2. For applicant conditions, alpha was between .92 and .94. For fake-good conditions, alpha was .96.

Typical faking behavior. Applicant faking behavior is measured in a variety of different ways, including difference scores (applicant condition minus honest condition), self-reports, and lie scale scores.

The most direct method of measuring faking is to use difference scores between scores produced in an applicant-response condition from scores produced in an honest-response condition (see Ellingson, Sackett, & Hough, 1999; McDaniel, Douglas, & Snell, 1997; McFarland & Ryan, 2000). Difference scores are known to have a number of potential problems (Edwards, 2001), but have been used in the context of faking research.

Edwards (2001) cited two of the main issues of difference scores as possibly low reliability and unclear validity. Reliability of difference scores is a major issue since the reliability of the difference score is lower than either of the component tests, difference scores have the potential for low levels of internal consistency (for example, see

McFarland & Ryan, 2000). Reliabilities of the difference score were computed using the method described by McFarland and Ryan (2000):

$$r_{dd} = (\sigma_d{}^2 - \sigma_{ed}{}^2) / \sigma_d{}^2$$
$$\sigma_{ed}{}^2 = \sigma_{honest}{}^2 (1 - r_{honest}) + \sigma_{fake}{}^2 (1 - r_{fake})$$

where r_{dd} is the reliability of the difference score and σ^2_d is the variance of the difference score. An analysis of the reliabilities showed that each of the difference score measures of typical faking behavior had acceptable reliabilities in both Study 1 (alpha = .78 for Time 1 and .89 for Time 2) and Study 2 (alpha = .87).

Edwards (2001) also acknowledged that the validity of difference scores could be checked empirically. This was done by conducting an exploratory factor analysis (EFA) which showed that difference scores and self-report faking behaviors had similarly high loadings on an overall faking factor (see below for details). This suggests that the difference scores demonstrated convergent validity. When combined with the acceptable levels of reliability, and the fact that difference scores may be conceptually appropriate when expecting a participant x treatment interaction (McFarland & Ryan, 2000), this suggests that difference scores may be an acceptable measure of faking.

Another way of measuring faking is the use of self-report questionnaires (Dwight & Donovan, 2003). Items for the self-report measure were obtained from McFarland (2000), who used the items as a manipulation check for an applicant test-taking condition. The items match closely to those used on similar self-report questionnaires designed to assess faking: for example, an item from McFarland (2000) reads “I intended to make myself look as good as possible on the employment selection test,” and

an item from Hurtz and Bohon (1997) reads "I purposefully tried to portray myself as an ideal applicant." The measure consisted of seven items (alpha = .87) on a 5-point scale (strongly agree to strongly disagree). Higher scores represented greater levels of faking.

A third way of measuring typical faking behavior is the use of measures to detect or control faking, sometimes referred to as lie scales or social desirability scales. Scores were provided by the Impression Management subscale of the BIDR (Paulhus, 1991), which consists of 12 items on a 5-point scale (not true to very true); alpha reliability was .85. Lie scale scores are coded such that *low* scores represented dishonest behavior, and *high* scores represented honest behavior.

As a construct validity check, the three separate measures of faking behavior were subjected to an exploratory factor analysis (1 factor extracted, using maximum likelihood estimation), to identify the extent to which each of the measures was associated with a single overall factor of typical faking behavior. Results of Study 1 find that difference scores, self-reports, and lie scale scores have factor loadings of .79, .56, and -.29 on a single factor that explains over 52% of the variance in the three scores. In Study 2, these measures had factor loadings of .88, .58, and -.15, and explained 52% of the variance in the observed measures. These results indicate that the generalized faking behaviors measured by lie scale scores, although widely used to assess individual levels of typical or "will do" faking, do not appear to be capturing the same aspects of faking measured by the two situation-specific measures of faking, although the differences appear more substantial in Study 2. These patterns are consistent with the view that the

two types of measures capture different aspects of typical faking behavior than the lie scale.

Faking ability. Individual levels of faking ability were indexed by calculating integrity scores under maximal faking conditions minus integrity scores under honest conditions. Reliabilities of the difference scores were calculated as above, with adequate reliabilities for both Study 1 and Study 2 (alpha = .93).

Attitudes towards faking. Participants were asked to assess their attitudes toward faking by responding to a 5-item measure based on five different semantic differential-type response scales: good-bad, pleasant-unpleasant, foolish-wise, useful-useless, and unattractive-attractive. An example item is "I think lying on an employment selection tests is: [good or bad]". Reliability was .81 in Study 1 and .84 in Study 2. High scores represent favorable attitudes toward faking.

Social norms for faking. Social norms for faking were assessed using a 4-item measure on a 5-point scale (strongly agree to strongly disagree). An example item is "Most people who are important to me would look down on me if I lied on a selection test." Reliability was .78 for Study 1 and .73 for Study 2. High scores represent social norms in favor of faking.

Perceived behavioral control over faking. Perceptions of behavioral control over faking were measured using three items on a 5-point response anchor ranging from strongly agree to strongly disagree. An example item is "It would be easy for me to lie on a selection test." Reliability was .82 for Study 1 and .85 for Study 2. High scores represent high perceptions of behavioral control.

Intention to fake. Intentions to fake on future employment selection tests were measured using ten items on a 5-point scale, ranging from strongly agree to strongly disagree. An example of this item is: "I would not lie on a pre-employment selection test." Reliability was .85 for Study 1 and Study 2. High score represents greater intention to fake.

Self-report faking measure. Items for the self-report faking measure were obtained from McFarland (2000), who used the items as a manipulation check for an applicant test-taking condition. The items match closely to those used on similar self-report questionnaires designed to assess faking (e.g., Dwight & Donovan, 2003; Hurtz & Bohon, 1997). An example is "I attempted to fake my responses on the employment selection test." The measure consisted of seven items with alpha of .87, with high scores represented greater levels of faking attempted.

The distinction between intention to fake and self-report faking measure is that the intention items asked respondents about their intention to fake on a selection test in the future, while the self-report faking measure asked participants about their faking effort in reference to a test that they just took. Thus, the self-report faking items can be seen as an index of faking behavior on a specific test, while intention items refer to general faking on an abstract future test.

Counterproductive work behaviors. Counterproductive work behaviors (CWBs) were measured using a measure of workplace deviance from Bennett and Robinson (2000). The scale captures two different facets of deviations from organizational norms that harm the organization or other individuals (Robinson & Bennett, 1995):

organizational deviance and interpersonal deviance. The measure consists of 19 items on a 7-point frequency scale (never, once a year, twice a year, several times a year, monthly, weekly, or daily). Reliability for Study 1 was .90, and .91 for Study 2. Higher scores represent more frequent CWBs.

Procedure

Both studies (Study 1 and Study 2) required applicants to participate in two separate sessions; these two sessions were spaced roughly 1 week apart. During Time 1, participants were asked to sign a consent form and a waiver of release form that allowed the researcher to access their Scholastic Aptitude Test (SAT) score from the university, complete demographic items, and items measuring TPB variables (attitudes, social norms, perceived behavioral control, and intentions). In Study 1, participants were also asked to complete an overt integrity test under three different conditions: (1) as an applicant, (2) honestly-as-possible, and (3) fake-good[2]. They were also asked to respond to a conscientiousness measure and the impression management aspect of the BIDR (the lie scale, though it was not labeled as such). Participants in Study 2 were asked to complete the Wonderlic Personnel Test, but were not asked to take the integrity test at Time 1. To improve participant motivation in the applicant and fake-good conditions, participants in each condition were told they would receive $20 if they were in the top 15% of scorers on that test. (See Appendix A for instruction sets, Appendix B for details on specific measures, and Appendix C for a table describing the order of administration.)

---

[2] McFarland (2000) reported that order effects in a TPB model of faking were not a concern, but meta-analytic results (Edens & Arthur, 2000) showed that having participants respond to honest conditions first, *then* fake conditions makes them more effective distorters than using the reverse condition, which suggests that honest conditions ought not be given first.

Participants were invited back for Time 2, about a week later. In Time 2, they were asked to complete the overt integrity test under three different conditions: (1) as an applicant, (2) honestly-as-possible, and (3) fake-good. They were also asked to complete another copy of the BIDR (lie scale), a questionnaire (self-report) measure of their faking behavior immediately after the integrity test in each condition. At the end of the study, participants were debriefed on the purpose and implication of the study, and told that they would be entered in a drawing for $20 prizes, instead of rewarded based on test scores. This was to avoid labeling winners as high on integrity and others as lower on the construct.

According to the TPB, the sooner behaviors are measured after attitudes and intentions, the stronger the relationship between them (Ajzen & Madden, 1986). Asking about faking immediately before the test might prompt some respondents to think about faking and thus increase the likelihood of faking; therefore a time delay such as one week is generally used as a means to prevent potential effects associated with priming (see McFarland & Ryan, 2006). However, there were additional integrity test administrations used in the Time 1 session of Study 1, which were administered immediately following the completion of the TPB and intention measures. While care was taken in the wording of instructions to avoid demand characteristics, the "respond as an applicant" condition was placed before the "respond honestly" condition, because the model under investigation is focused on applicant faking behavior; placing the applicant condition before the honest condition avoids diagnosticity effects or demand characteristics that could result in inflated applicant faking (Feldman & Lynch, 1988).

Integrity scores were measured twice during Study 1. Separate analyses were conducted based on integrity scores at Time 1 and Time 2, and the overall results were very similar for the different hypotheses, suggesting that transient demand characteristics and priming were not likely to be a major factor in the results of the study. In addition, there were very similar levels of self-report faking across conditions for both Time 1 and 2, suggesting that prior exposure to the test at Time 1 in Study 1 did not influence the extent to which respondents would fake at Time 2. Based on these results, and to ease interpretation, only the results of Time 2 are presented in the Results section[3].

[3] Results for Time 1 are available on request from the author.

STUDY 1 RESULTS

Descriptive statistics

Means, standard deviations, and correlations for variables used in Study 1 are presented in Table 1.

Manipulation checks

Self-report measures of faking were administered in each condition as a way to differentiate the degree of faking attempts across test-taking conditions. To the extent that individuals followed instructions for each condition, self-reported faking scores should vary significantly across conditions. A repeated measures analysis of variance (RM-ANOVA) and follow-up comparisons found significant differences in the mean level of self-reported faking behaviors across different conditions, $F(2, 474) = 126.1$, $p < .01$; $\eta^2 = .35$, with significant differences between each condition ($p < .01$).

Table 2 shows there is at least half of a *SD* difference in the mean of each pair of self-reported faking behaviors. These differences suggest that participants understood the instruction sets, and engaged in different levels of faking across the three conditions. In addition, self-reported levels of faking for the applicant condition were not related to self-reported faking in the honest condition ($r = -.01$, $p > .05$), but were moderately related to self-reports of faking behavior in the fake-good condition ($r = .28$, $p < .01$) . This further supports that instructional sets elicited different faking efforts across conditions.

Table 1
Study 1 correlation matrix

Variable	N	Mean (SD)	1	2	3	4	5	6	7	8	9	10	11	12	13
(1) SAT	286	1028 (367)													
(2) CON	286	3.6 (.55)	-.11	.77											
(3) ATT	286	1.8 (.69)	-.01	-.29**	.81										
(4) SN	286	2.1 (.80)	-.02	-.31**	.58**	.78									
(5) PBC	286	2.6 (.99)	.02	-.28**	.57**	.49**	.82								
(6) INTENT	286	2.8 (.64)	.04	-.28**	.70**	.55**	.60**	.85							
(7) HON (T1)	286	3.5 (.45)	-.08	.32**	-.40**	-.41**	-.47**	-.44**	.92						
(8) APP (T1)	286	3.8 (.40)	-.02	.25**	-.36**	-.35**	-.39**	-.35**	.64**	.92					
(9) FAKE (T1)	286	4.3 (.60)	.14*	.04	-.29**	-.20**	-.24**	-.22**	.25**	.42**	.97				
(10) HON (T2)	231	3.6 (.45)	.01	.29**	-.38**	-.34**	-.43**	-.39**	.82**	.57**	.26**	.93			
(11) APP (T2)	231	4.0 (.42)	.04	.20**	-.23**	-.20**	-.25**	-.15*	.33**	.63**	.47**	.38**	.94		
(12) FAKE (T2)	231	4.3 (.59)	.12	.14*	-.19**	-.17*	-.15**	-.12	.21**	.29**	.62**	.31**	.46**	.96	
(13) MF (T1)	286	.27 (.36)	.07	-.12*	.10	.12*	.14**	.16**	-.54**	.31**	.15**	-.37**	.30**	.06	.78
(14) MF (T2)	231	.42 (.48)	.02	-.10	.16*	.14*	.19**	.24**	-.48**	.02	.17**	-.59**	.52**	.12	.61**
(15) SR-H	231	2.1 (.75)	-.06	-.08	.22**	.16*	.17**	.22**	-.18**	-.25**	-.31**	-.21**	-.38**	-.29**	-.04
(16) SR-A	230	2.7 (.86)	.07	-.18**	.21**	.16*	.25**	.31**	-.29**	-.06	-.01	-.32**	.16*	-.06	.29**
(17) SR-F	231	3.5 (1.3)	.06	.00	-.03	-.01	.08	.05	-.17*	-.02	.11	-.16*	.13*	.14*	.18**
(18) IM	286	3.2 (.79)	.03	.28**	-.25**	-.19**	-.26**	-.31**	.43**	.31**	.19**	.39**	.16*	.18**	-.19**
(19) FA (T1)	286	.77 (.65)	.18*	-.19**	.01	.09	.10	.10	-.46**	-.05	.74**	-.32**	.21**	.42**	.51**
(20) FA (T2)	231	.72 (.62)	.11	-.08	.10	.09	.17**	.17**	-.39**	-.14*	.40**	-.43**	.16*	.73**	.33**
(21) log CWB	228	1.4 (.41)	-.05	-.38**	.30**	.29**	.28**	.35**	-.48**	-.37**	-.11	-.52**	-.24**	-.18**	.17*

Variable	N	Mean (SD)	14	15	16	17	18	19	20	21
(14) MF (T2)	231	.42 (.48)	.89							
(15) SR-H	231	2.1 (.75)	-.13*	.86						
(16) SR-A	230	2.7 (.86)	.44**	-.01	.86					
(17) SR-F	231	3.5 (1.3)	.26**	-.31**	.28**	.93				
(18) IM	286	3.2 (.79)	-.22**	.02	-.16*	-.41**	.84			
(19) FA (T1)	286	.77 (.65)	.48**	-.15*	.19**	.22**	-.13*	.94		
(20) FA (T2)	231	.72 (.62)	.54**	-.13	.18**	.25**	-.11	.64**	.93	
(21) log CWB	218	1.4 (.41)	.27**	.11	.13*	.12	-.39**	.22**	.22**	.90

Note: CON = conscientiousness; ATT = attitudes towards faking; SN = social norms; PBC = perceived behavioral control; HON = honest integrity score; APP = applicant integrity score; FAKE = maximal faking integrity score; MF = motivated faking; SR = self-report faking behavior for honest (H), applicant (A), and maximal faking (F) conditions; IM = lie scale; FA = faking ability; CWB = counterproductive work behaviors.

* $p < .05$
** $p < .01$

Table 2
Self-reports of faking across Study 1 conditions

	Mean	*SD*	1	2	3	4
1. Intentions	2.75	.64				
2. Self-Report - Applicant	2.72	.86	.31**			
3. Self-Report - Honest	2.08	.75	.22**	-.01		
4. Self-Report - Fake	3.50	1.26	.05	.28**	-.31**	

* $p < .05$
** $p < .01$

Relationship between cognitions and intentions

As shown in Table 1, all three TPB variables were significantly correlated with intentions to fake (with respective correlations for attitudes, subjective norms, and perceived behavioral control of .55, .60, and .70, $p < .01$). When intentions to fake was regressed on the three TPB variables, the three variables explained 57% of the variance in intentions to fake (Table 3); this provides evidence that the TPB predictors are strong predictors of intentions to fake on a selection test. Attitudes toward faking were the strongest predictor of intentions to fake, followed by PBC and subjective norms (see Table 3 below). These results are consistent with previous studies that use the TPB to study intentions; for example, a review of TPB studies by Ajzen (1991) found an average multiple correlation of .71, similar to the results of the present study ($R = .75$).

In addition, the correlations between TPB variables and intentions were much weaker for applicant self-reports of faking behavior than for pre-selection test intentions to fake ($r = .21$, .16, and .25 for attitudes, social norms, and perceived behavioral control, $p > .05$). The set of these three variables produced a multiple correlation of .27

Table 3
Study 1 multiple regression analysis with intention to fake regressed onto attitudes, subjective norms, and PBC

Model	β	R	R^2	F
		.75	.57	125.43**
Attitude toward faking	.47**			
Subjective Norms toward faking	.15**			
PBC toward faking	.26**			

* $p < .05$
** $p < .01$

($R^2 = .07$) with self-reported faking behaviors in the applicant condition. This shows that cognitions towards faking are more strongly related to general pre-selection faking intentions than to specific post-test faking behaviors, which is a pattern to be expected based on the TPB.

Hypothesis tests

The first hypothesis was that intentions should mediate the effect of the three TPB variables on applicant faking behaviors. Hypothesis 1 was tested using a series of hierarchical regressions; separate models were tested for different operationalizations of applicant faking behavior (i.e., applicant minus honest difference scores, self-report faking, and lie scale scores). Results are shown in Table 4. In each regression model, two steps were used to enter predictors of faking behavior[4].

In the first model, the three theory of planned behavior variables (attitudes, social norms, and perceived behavioral control) were used to predict applicant faking behavior

[4] Although difference scores appeared to provide an acceptable operationalization of faking, tests of change may also be based on alternative methods, such as residualized gain score analysis. Results of the analysis of residualized gain scores for H1 were nonsignificant for Study 1 and 2.

Table 4
Study 1 hierarchical tests of mediation for different measures of faking

	SN	PBC	ATT	INT	R^2	*Adjusted* R^2	*F*	df	ΔR^2	$F(\Delta R^2)$
DS (model step 1)	.04	.14	.06	--	.04	.03	3.21	3	.04	3.21*
DS (model step 2)	.01	.07	-.03	.20*	.06	.04	3.56	4	.02	4.44*
DS (model step 3)	--	--	--	.24**	.06	.05	13.49	1	.06	13.49**
DS (model step 4)	.01	.07	-.03	.20*	.06	.04	3.56	4	.00	.29
SR-A (model step 1)	.01	.20*	.09	--	.07	.06	5.73	3	.07	5.73**
SR-A (model step 2)	-.03	.11	-.02	.27**	.10	.09	6.59	4	.03	8.63**
SR-A (model step 3)	--	--	--	.31**	.10	.09	24.67	1	.10	24.67**
SR-A (model step 4)	-.03	.11	-.02	.27**	.10	.09	6.59	4	.03	.61
IM (model step 1)	-.03	-.17*	.13	--	.08	.07	8.54	3	.08	8.54**
IM (model step 2)	-.00	-.11	-.03	-.23**	.10	.09	8.27	4	.02	6.95**
IM (model step 3)	--	--	--	-.31**	.10	.09	30.4	1	.10	30.4**
IM (model step 4)	-.00	-.11	-.03	-.23**	.10	.09	8.27	4	.01	.89

Note: DS = difference scores; SR-A = self-report applicant faking; IM = lie scale score (lower scores indicate greater lying).

* $p < .05$
** $p < .01$

(step 1); intentions to fake were then added to the model as a predictor (step 2). In the second model, the steps were reversed: in the first step, intentions to fake were used to predict applicant faking behaviors (step 3). In the next step, the three Theory of Planned Behavior variables were added to the model as predictors (step 4). For mediation between the set of predictors and the behavior to be supported, steps 1, 2, and 3 should explain significant levels of additional variance (ΔR^2), and step 4 should *not* be significant.

Hypothesis 1 was first tested using applicant faking behavior measured via applicant-honest difference scores at Time 2, as shown in Table 4. The model shows that the three Theory of Planned Behavior variables together predict significant levels of variance in Time 2 difference scores, and that adding intentions to fake as a predictor explains significant additional amounts of variation in Time 2 difference scores. Reversing the order of entry, the model shows that intentions to fake are a significant predictor of difference scores, but that once intentions to fake have been accounted for in the model, that the TPB variables no longer explain significant amounts of variance in applicant faking behavior, as measured using difference scores. These results support Hypothesis 1.

Hypothesis 1 was also tested using applicant self-reports of attempted faking behavior as a measure of applicant faking behavior. As shown in Table 4, the results supported Hypothesis 1: intentions to fake explained additional variance over and above the TPB variables, but the TPB variables did not predict significant levels of variance in faking behaviors after taking into account intentions.

Finally, a third operationalization of faking behavior – lie scale scores – were used to test Hypothesis 1. Results are shown in Table 4, and provide further support of Hypothesis 1. Intentions explained significant additional variance in lie scales, over and above the TPB variables, but the TPB variables did not explain significant additional variance over and above intentions to lie.

In addition to these tests of mediation, the indirect effects of the TPB variables on applicant faking behaviors were tested using a Sobel (1982) test (McKinnon, Lockwood, Hoffman, West, & Sheets, 2002). Results of these tests are shown in Table 5; there was a significant indirect effect of each TPB variable on all three types of faking measures. The p-value associated with direct effects indicate that once intentions are included in the model as a mediator, none of the TPB variables have a direct effect on faking. These results provide further support for Hypothesis 1.

Table 5
Sobel tests of Study 1 indirect effects

Relationship	Indirect effect (*SE*)	Z-value	Direct effect (*p*-value)
SN→INT→IM	-.14 (.04)	-3.31**	.97
PBC→INT→IM	-.09 (.04)	-2.42*	.17
ATT→INT→IM	-.18 (.07)	-2.69**	.62
SN→INT→SR-A	.19 (.05)	3.91**	.84
PBC→INT→SR-A	.13 (.04)	3.06**	.21
ATT→INT→SR-A	.27 (.08)	3.56**	.96
SN→INT→DS (T2)	.08 (.03)	2.81**	.79
PBC→INT→ DS (T2)	.06 (.02)	2.32*	.38
ATT→INT→ DS (T2)	.11 (.04)	2.67**	.98

* $p < .05$
** $p < .01$

Hypothesis 2 suggested that applicant faking behaviors should lead to an increase in observed scores on integrity tests taken as an applicant – that is, applicants who fake should get scores that make them look more honest than they actually are. This was tested by means of a repeated measures ANOVA, comparing the integrity test scores of individuals in the various conditions. If applicant faking leads to increased test scores, then participants who are told to fake should receive higher scores than those who are told not to fake.

The results of this analysis are presented in Tables 6 and 7, which show that individuals in the "respond as an applicant" condition received substantially higher scores than individuals told to respond honestly. The half-standard deviation difference in mean integrity scores between honest and applicant responses is similar to the results of studies based on real applicants (e.g., Hough, 1998; Rosse et al., 1998). This provides additional support that instructional sets to respond honestly or as an applicant were effective.

As an additional test of Hypothesis 2, applicant integrity test scores at Time 2 were correlated with self-reports of attempted faking behaviors. The results were significant ($r = .16$, $p < .05$); this indicates that the more faking reported by a hypothetical "applicant," the greater their integrity test score. This provides additional evidence that faking does affect test scores, lending further support to previous studies that faking increases test scores (e.g., Ryan & Sackett, 1987).

Table 6
Time 2 Integrity scores by condition

Source	SS	df	MS	*F*	η^2
Instruction type	61.9	2	31.0	205.7**	.46
Error	71.7	476	.15		

Table 7
Pairwise comparison of Time 2 integrity scores (with Bonferonni corrections)

	Mean	*SD*	95%CI	Comparison to	Mean difference (*d* score)
Honest	3.57	.45	(3.51, 3.64)	-	-
Applicant	3.99	.42	(3.93, 4.04)	Honest	.42** (.96)
Fake	4.28	.59	(4.21, 4.36)	Honest	.72** (1.37)
				Applicant	.30** (.59)

* $p < .05$
** $p < .01$

According to Hypothesis 3, the main effect of applicant faking behavior on applicant test scores should be moderated by the level of faking ability possessed by the individual. As with previous hypotheses, separate models were run for different operationalizations of faking behavior, self-report faking and lie scale scores. The first model tested the interaction between applicant self-reports of attempted faking behavior and faking ability at Time 2 (Table 8). The model found support for Hypothesis 3. The interaction is displayed graphically in Figure 3: for increasing levels of faking ability, the effect of self-reported applicant faking behaviors on integrity test scores increases.

The same hypothesis was also tested using lie scale scores instead of self-report faking behaviors. The interaction was again significant (see Table 9), further supporting

Table 8
Interaction of Time 2 faking ability and applicant faking (SR) on applicant integrity scores (T2)

	β	B
Intercept		3.98**
Faking ability (FA)	-.32	-.21
Applicant faking (SR)	-.06	-.03
FA x Applicant faking (SR)	.55*	.12*

* $p < .05$
** $p < .01$

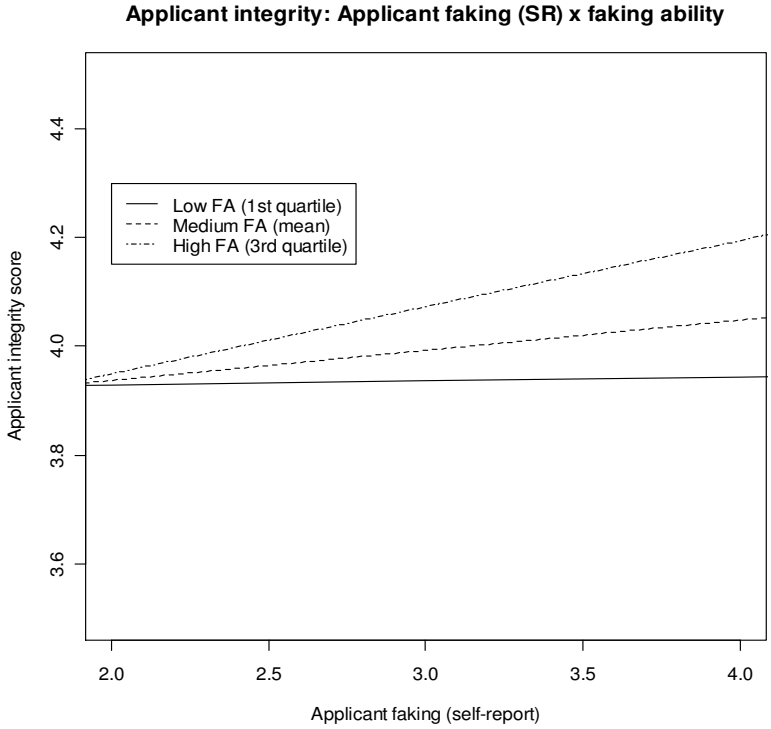

Figure 3. Applicant faking (SR) x faking ability[5]

[5] Possible values of the integrity and faking scores ranged from 1 to 5; values for CWB ranged from 0 to 7. The range of the X and Y axes in each figure (Figures 3, 4, 5, and 6) were chosen to reflect the range of the available data.

Table 9
Interaction of Time 2 faking ability and applicant faking (IM) on Time 2 applicant integrity scores

	β	B
Intercept		3.20
Faking ability (FA)	.41**	.22**
Applicant faking (IM)	.96**	.65**
FA x IM	-.81**	-.17**

* $p < .05$
** $p < .01$

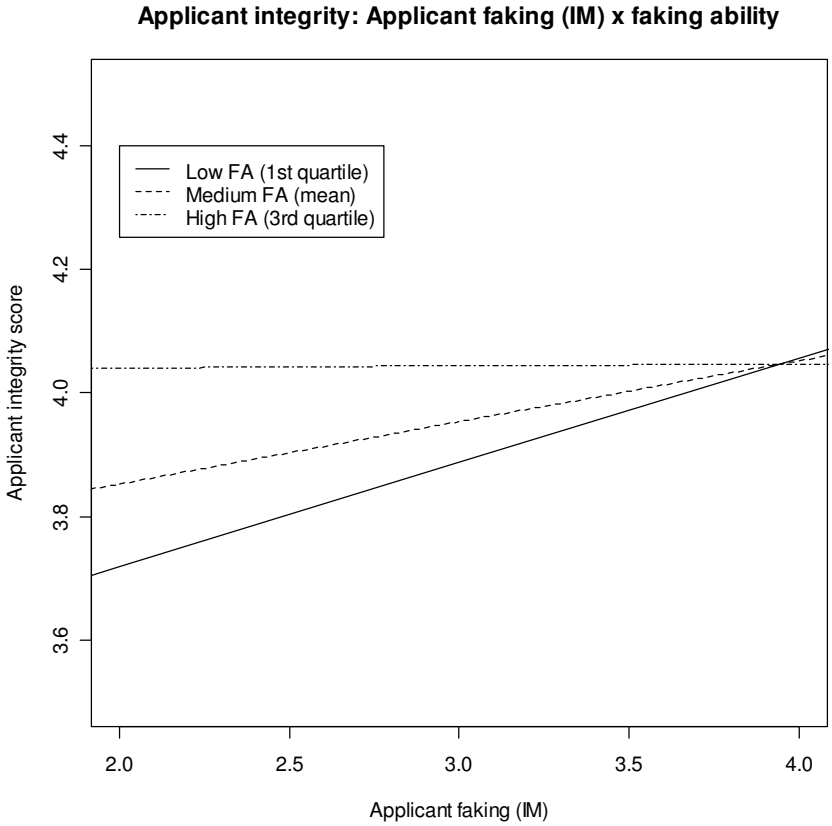

Figure 4. Applicant faking (IM) x faking ability

Hypothesis 3. Figure 4 highlights the interaction effect – keeping in mind that lie scale scores are such that a lower number represents more dishonest behavior, the figure

shows that individuals who are prone to lying *and* have high levels of faking ability tend to get significantly higher scores than individuals who are prone to lying but do not have high levels of ability.

Hypothesis 4 was tested by correlating individual levels of conscientiousness with self-reported attitudes towards faking behaviors. These correlated significantly (r = -.29, $p < .01$). This suggests that more conscientious participants tended to have significantly more negative attitudes towards faking behaviors, supporting Hypothesis 4. Furthermore, negative correlations between conscientiousness and intention to fake (r = -.28, $p < .01$) and self-report faking (r = -.18, $p <.01$) were observed. This suggests that people who are high on conscientiousness have lower intentions to fake and weaker applicant faking attempts.

Hypothesis 5 was tested by correlating SAT scores (combined math and verbal) with faking ability. SAT scores did predict faking ability at Time 1 (r = .18, $p < .01$), but not at Time 2 (r = .11, $p > .05$). Thus, there appears to be mixed support for Hypothesis 5. Furthermore, the correlation between SAT scores and perceived behavioral control was not significant (r = .02, $p > .05$); thus, Hypothesis 6 was not supported.

Hypothesis 7 considered the effect that faking behaviors had on applicant integrity test validity. This was tested by looking at the effect on CWBs of the interaction between applicant integrity scores and applicant faking behaviors. For

statistical reasons, CWB scores were transformed into an approximately normal variable by taking the logarithm[6] (Baba, 1990; Keene, 1995; Tabachnik & Fidell, 1996).

Support for Hypothesis 7 was somewhat mixed: there was a significant interaction between self-reported faking behaviors and applicant test scores (Table 10), but no interaction between lie scale scores and applicant test scores (Table 11). The

Table 10
Interaction between applicant faking (SR) and applicant integrity test (T2) validity for CWB

	β	B
Intercept		3.68**
Applicant integrity (AI)	-.66**	-.62**
Applicant faking (SR)	-.93	-.44
AI x Applicant faking (SR)	1.24*	.13*

* $p < .05$
** $p < .01$

Table 11
Interaction between applicant faking (IM) and Time 2 applicant integrity test validity for CWB

	β	B
Intercept		2.30**
Applicant integrity (AI)	-.08	-.07
Applicant faking (IM)	-.14	-.07
AI x IM	-.28	-.03

* $p < .05$
** $p < .01$

[6] Untransformed CWB scores were positively skewed and kurtotic in both Study 1 and 2 (skew = 1.1 and 1.3; kurtosis = 1.5 and 1.9, respectively); more importantly, the residuals of the standard linear regressions based on untransformed CWB scores were not normally distributed and had noticeable levels of heteroscedasticity, which could result in biased standard errors. After transformation, residuals were homoscedastic and normally distributed. In a recent study of CWBs that also used the Bennett and Robinson (2000) scale, Colbert, Mount, Harter, Witt, and Barrick (2004) suggested taking the logarithm of the scale.

interaction term from Table 10 is shown in Figure 5. The figure shows that high levels of faking behavior on the part of applicants reduce the magnitude of the relationship between applicant integrity scores and CWBs.

The final hypothesis suggested that faking by applicants should alter the rank order of integrity test scores. This was tested by use of the Spearman rank-order correlation (ρ), which provides a measure of the consistency of rank ordering between two variables. Perfect consistency in rank-ordering would be reflected in a perfect rank-order correlation (ρ = 1).

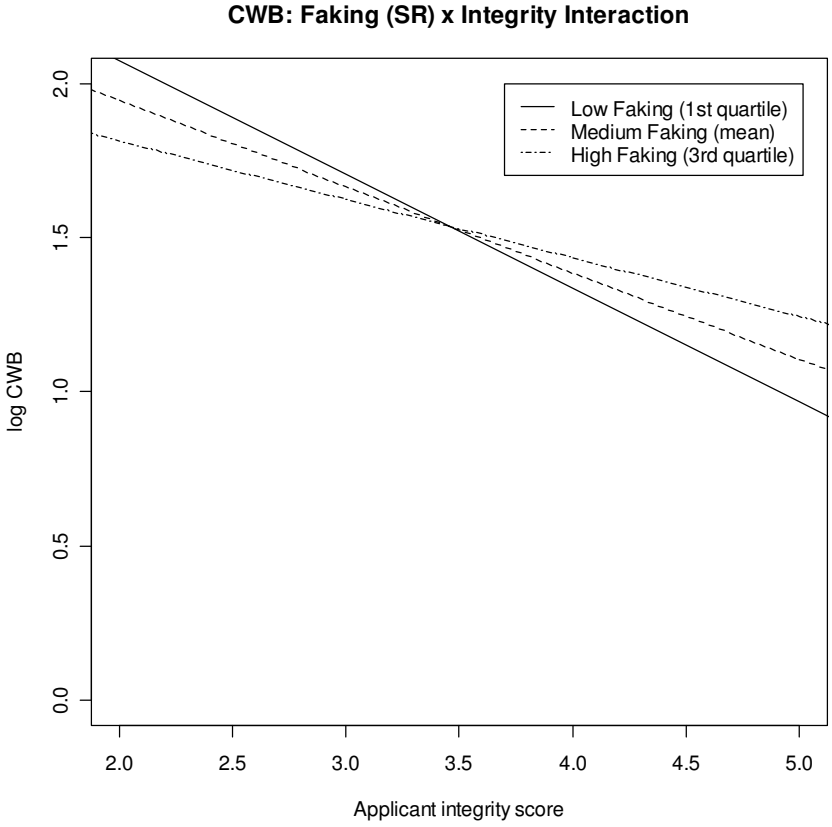

Figure 5. Faking behavior (SR) x applicant integrity test validity

To test the hypothesis, the method of Caruso and Cliff (1997) was used to construct 95% CI around ρ; the results indicate that applicant faking behavior altered the rank-order of integrity scores compared to honest scores at both Time 1 (rho = .66, 95%CI =[.62, .69]) and Time 2 (rho = .35, 95%CI = [.30, .40]). The CI did not include 1.0, indicating significant levels of deviation from perfect rank-order consistency between applicant and honest scores at both time periods, supporting this hypothesis.

The hypothesis was also tested by comparing changes in top scorers from the honest to the applicant condition. In Study 1, of the top 15% of scorers in the honest condition, only 39% remained in the top 15% of scorers in the applicant condition. In other words, 61% of the people who would have been in the top 15% of scorers under honest response conditions were no longer in the top 15% when responding as an applicant. This provides further evidence that changes in rank ordering were occurring.

STUDY 2 RESULTS

Descriptive statistics

Means, standard deviations, and correlations between variables used in Study 2 are shown in Table 12.

Manipulation checks

The efficacy of condition instructions was assessed in the same manner as in Study 1. A repeated measures ANOVA found significant differences in self-reported faking across conditions, $F(2,326) = 120.8$, $p < .01$; $\eta^2 = .43$, with the mean level of self-reported faking (Table 13) in each condition significantly different from the other conditions ($p < .01$). This suggests that participants understood the instruction sets and engaged in different degrees of faking effort across conditions. Furthermore, it can be seen that self-reports of faking in the applicant condition are not related to faking in the honest condition ($r = -.02$, $p > .05$) but are related to faking in the fake-good condition ($r = .30$, $p < .01$). This lends further support to the idea that the manipulated instructions elicited different faking efforts across conditions.

Relationship between cognitions and intentions

Support for the relationship between TPB variables and intentions were supported with significant ($p < .01$) correlations of attitudes toward faking ($r = .76$), subjective norms toward faking ($r = .62$), and PBC toward faking ($r = .66$) with intention to fake. A regression analysis (Table 14) showed that these three TPB variables explained 67% of the variance in the intention to fake. These results are also consistent

Table 12
Study 2 correlation matrix

Variable	*N*	Mean (SD)	1	2	3	4	5	6	7	8	9	10	11	12	13
(1) SAT	181	1019 (362)													
(2) GMA	180	25.4 (5.4)	.25**												
(3) CON	181	3.7 (.60)	.06	.07	.80										
(4) ATT	181	1.7 (.68)	.12	-.02	-.34**	.84									
(5) SN	181	2.0 (.80)	.03	-.13	-.32**	.63**	.73								
(6) PBC	181	2.5 (1.0)	.12	-.02	-.30**	.57**	.47**	.85							
(7) INTENT	181	2.6 (.66)	.08	-.01	-.33**	.76**	.62**	.66**	.85						
(8) HON	152	3.4 (.45)	-.18*	-.05	.22**	-.32**	-.25**	-.28**	-.40**	.93					
(9) APP	152	3.8 (.43)	-.04	.16*	.20*	-.33**	-.29**	-.29**	-.34**	.47**	.93				
(10) FAKE	152	4.3 (.53)	.02	.36**	.15	-.06	-.05	-.02	-.03	.22**	.37**	.96			
(11) MF	152	.32 (.46)	.15	.20**	-.03	.00	-.03	.00	.08	-.54**	.49**	.13	.87		
(12) SR-H	152	2.0 (.68)	-.02	-.31**	.02	.15	.15	.13	.23**	-.01	-.28**	-.26**	-.25**	.82	
(13) SR-A	152	2.4 (.85)	.04	.17**	-.04	.16*	.10	.16*	.25**	-.34**	.19*	.20*	.51**	-.02	.87
(14) SR-F	152	3.6 (1.3)	.14	.26**	-.04	-.04	-.09	.05	-.04	-.11	.19*	.34**	.29**	-.37**	.30**
(15) IM	181	3.2 (.75)	.04	.05	.28**	-.25**	-.24**	-.23**	-.32**	.46**	.34**	.19*	-.13	-.03	-.09
(16) FA	152	.80 (.62)	.14	.34**	-.03	.18*	.13	.19*	.27**	-.54**	-.03	.70**	.51**	-.21**	.42**
(17) log CWB	154	.79 (.40)	.02	.00	-.23**	.27**	.20**	.28**	.30**	-.55**	-.32**	-.17*	.25**	.01	.13

Variable	*N*	Mean (SD)	14	15	16	17
(14) SR-F	152	3.6 (1.3)	.93			
(15) IM	181	3.2 (.75)	-.09	.83		
(16) FA	152	.80 (.62)	.37**	-.17*	.93	
(17) log CWB	154	.79 (.40)	.14	-.49**	.26**	.91

Note: GMA = general mental ability; CON = conscientiousness; ATT = attitudes towards faking; SN = social norms; PBC = perceived behavioral control; HON = honest integrity score; APP = applicant integrity score; FAKE = maximal faking integrity score; MF = motivated faking; SR = self-report faking behavior for honest (H), applicant (A), and maximal faking (F) conditions; IM = lie scale; FA = faking ability; CWB = counterproductive work behaviors.

* $p < .05$
** $p < .01$

Table 13
Self-reports of faking across Study 2 conditions

	Mean	*SD*	1	2	3	4
1. Intentions	2.63	.66				
2. Self-report - Applicant	2.35	.85	.25**			
3. Self-report - Honest	2.01	.68	.23**	-.02		
4. Self-report - Fake	3.57	1.26	-.04	.30**	-.37**	

* $p < .05$
** $p < .01$

Table 14
Study 2 multiple regression analysis with intention to fake regressed onto attitudes, subjective norms, and PBC

Model	β	R	R^2	F
		.82	.67	122.07**
Attitude toward faking	.47**			
Subjective Norms toward faking	.17**			
PBC toward faking	.31**			

* $p < .05$
** $p < .01$

with typical studies on the TPB (Ajzen, 1991), although the multiple correlation between TPB cognitions and intentions in the current study ($R = .82$) were even higher than the average multiple correlation ($R = .71$) from other TPB studies reported by Ajzen (1991). Finally, attitudes, social norms and PBC were not nearly as strongly correlated with actual faking behaviors ($r = .16$, $p < .05$; $r = .10$, $p > .05$; $r = .16$, $p < .01$). These results show that, as expected, the three TPB cognitions are more closely related to general intentions to fake in the future than to specific reports of faking on a given test. These results also replicate the findings in Study 1.

Hypothesis tests

Hypothesis 1 was tested using the same method described in Study 1 (Tables 15). However, compared to the results of Study 1, the results of the statistical tests of Hypothesis 1 were mixed: the mediation hypothesis was not supported using self-report attempted faking behaviors, nor using applicant minus honest difference scores, but was supported using lie scale scores.

In addition to the overall test of mediation, the Sobel tests of indirect effects associated with each of the three variables (self-report faking behaviors, difference scores, and lie scales) are shown in Table 16. The results indicate that each of the three TPB variables had significant indirect effects on faking behavior measured via lie scales and self-report, but not on faking measured using difference scores. These results provide partial support for Hypothesis 1.

A repeated measures ANOVA was used to test Hypothesis 2 (Tables 17 and 18). The results indicate that applicant faking behaviors led to increased integrity scores compared to honest responding, and that the magnitude of the increases was similar to past studies of actual applicants (Hough, 1998; Rosse et al., 1998). This provides further evidence of the effectiveness of the instructional manipulations. Furthermore, applicant integrity scores and self-reported faking behaviors correlated significantly at $r = .19$ ($p < .05$). These results support Hypothesis 2.

Tables 19 and 20 present the tests of the interaction between faking ability and applicant faking behavior. The interaction was not significant in either case – thus, Hypothesis 3 was not supported in this sample.

Table 15
Study 2 hierarchical tests of mediation for different measures of faking

	SN	PBC	ATT	INT	R^2	*Adjusted* R^2	*F*	df	ΔR^2	$F(\Delta R^2)$
DS (model step 1)	-.05	.01	.03	--	.00	.00	.08	3	.00	.08
DS (model step 2)	-.11	-.08	-.08	.26	.02	.00	.86	4	.02	3.19
DS (model step 3)	--	--	--	.08	.01	.00	.89	1	.01	.89
DS (model step 4)	-.11	-.08	-.08	.26	.02	.00	.86	4	.02	.85
SR-A (model step 1)	-.02	.10	.12	--	.03	.02	1.78	3	.03	1.78
SR-A (model step 2)	-.09	-.01	-.03	.35*	.07	.05	2.85	4	.04	5.89*
SR-A (model step 3)	--	--	--	.25**	.06	.06	10.44	1	.06	10.44**
SR-A (model step 4)	-.09	-.01	-.03	.35*	.07	.05	2.85	4	.01	.37
IM (model step 1)	-.11	-.11	-.12	--	.08	.07	5.45	3	.08	5.45**
IM (model step 2)	-.06	-.02	.00	-.27*	.11	.09	5.42	4	.02	4.97**
IM (model step 3)	--	--	--	-.32**	.10	.10	21.4	1	.10	21.4**
IM (model step 4)	-.06	-.02	.00	-.27*	.11	.09	5.42	4	.02	.19

Note: DS = difference scores; SR-A = self-report applicant faking; IM = lie scale score (lower scores indicate greater lying).

* $p < .05$
** $p < .01$

Table 16
Sobel tests of Study 2 indirect effects

Relationship	Indirect effect (*SE*)	Z-value	Direct effect (*p*-value)
SN→INT→IM	-.17 (.06)	-2.70**	.45
PBC→INT→IM	-.15 (.06)	-2.72**	.68
ATT→INT→IM	-.22 (.10)	-2.27*	.55
SN→INT→SR-A	.22 (.07)	2.99**	.31
PBC→INT→SR-A	.16 (.06)	2.42*	.82
ATT→INT→SR-A	.28 (.11)	2.46*	.58
SN→INT→DS	.06 (.04)	1.51	.21
PBC→INT→ DS	.04 (.04)	1.26	.40
ATT→INT→ DS	.09 (.06)	1.41	.29

* $p < .05$
** $p < .01$

Table 17
Integrity scores by condition

Source	SS	df	MS	*F*	η^2
Instruction type	53.68	2	26.84	179.24**	.52
Error	48.82	326	.15		

Table 18
Pairwise comparison of integrity scores (with Bonferonni corrections)

	Mean	*SD*	95%CI	Comparison to	Mean difference (*d* score)
Honest	3.45	.45	(3.38, 3.52)	-	-
Applicant	3.77	.43	(3.70, 3.83)	Honest	.32** (.73)
Fake	4.25	.53	(4.17, 4.33)	Honest	.80** (1.63)
				Applicant	.48** (.99)

* $p < .05$
** $p < .01$

Table 19
Interaction of faking ability and applicant faking (SR)

	β	B
Intercept		3.63
Faking ability (FA)	-.26	-.18
Applicant faking (SR)	.16	.08
FA x applicant faking (SR)	.19	.04

* $p < .05$
** $p < .01$

Table 20
Interaction of faking ability and applicant faking (IM)

	β	B
Intercept		3.00
Faking ability (FA)	.24	.17
Impression management (IM)	.42**	.24**
FA x IM	-.21	-.05

* $p < .05$
** $p < .01$

Hypothesis 4 was again supported, with conscientiousness correlating $r = -.34$ ($p < .01$) with attitudes towards faking. Furthermore, the correlation between conscientiousness and intention is -.33 ($p < .01$) and with self-report faking is -.04, ($p > .05$). This suggests that people who are high on conscientiousness have lower intentions and fewer actual faking attempts. Hypothesis 5 was partially supported: GMA scores correlated $r = .34$ ($p < .01$) with faking ability scores, but combined SAT scores did not ($r = .14$, $p > .05$). However, Hypothesis 6 was not supported; there was no correlation between perceived behavioral control of faking and either Wonderlic ($r = -.02$, $p = .78$) or SAT scores ($r = .12$, $p = .11$).

Tests of Hypothesis 7 indicate partial support for the proposed relationship. Specifically, CWB scores were predicted by the interaction between applicant integrity scores and self-report faking behaviors (Table 21), but not by the interaction between integrity scores and lie scale scores (Table 22). The interaction is visualized in Figure 6, which shows that individuals who reported greater levels of self-report faking behavior had a weaker relationship between applicant integrity scores and CWB scores.

Table 21
Interaction between applicant faking behavior (SR) and applicant integrity test validity for CWB

	β	B
Intercept		4.27**
Applicant Integrity (AI)	-.84**	-.79**
Applicant faking (SR)	-1.18*	-.57*
AI x SR	1.54*	0.17*

* $p < .05$
** $p < .01$

Table 22
Interaction between impression management and applicant integrity test validity for CWB

	β	B
Intercept		3.60**
Applicant integrity (AI)	-.48	-.35
Applicant faking (IM)	-1.15	-.54
AI x IM	.87	.08

* $p < .05$
** $p < .01$

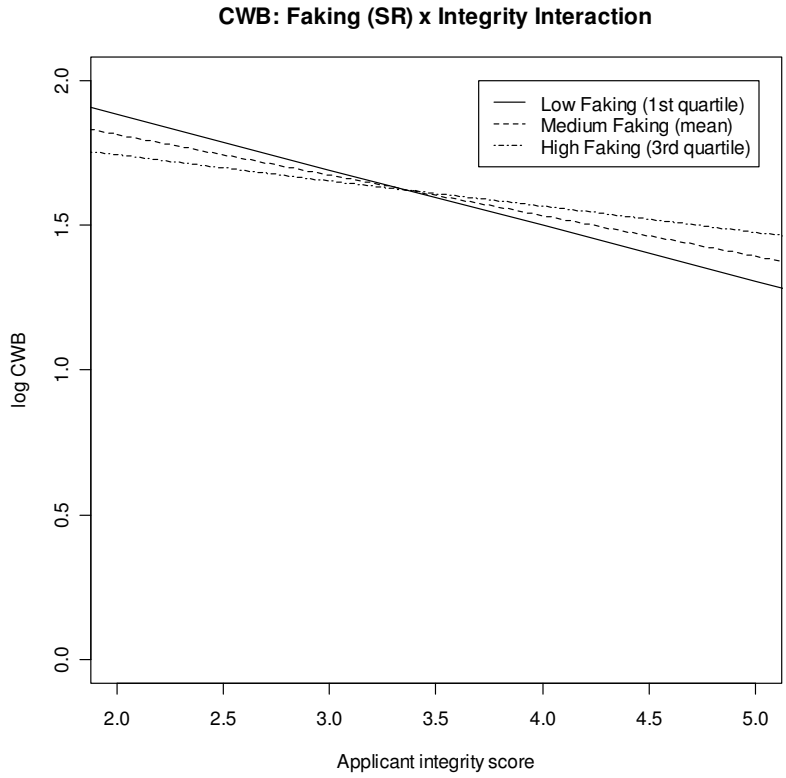

Figure 6. Effect of applicant integrity x self-report faking behaviors on log CWB scores

Finally, results from Study 2 support the last hypothesis. Spearman rank-order correlations between integrity scores from the applicant and honest conditions were r = .48 (p < .01), with a 95% confidence interval of [.35, .60]. Since this did not contain 1.0, the test indicated that there are significant differences in the consistency of the rank ordering of scores in the two conditions. As with Study 1, the change in rank ordering was also examined by identifying changes in the top scorers in the honest and applicant conditions. Of the top 15% of scorers in the honest condition, only 32% remained in the top 15% of scorers under applicant conditions.

DISCUSSION

The results of the two studies are summarized in Figure 7. The analysis provided initial support for the idea reflected in Hypothesis 1: that applicant faking may result – at least partly – from conscious intentions and decisions to fake, which is at least partly influenced by applicant cognitions regarding the appropriateness and utility of faking. However, these effects appeared to differ somewhat depending on the precise way in which faking behaviors were measured.

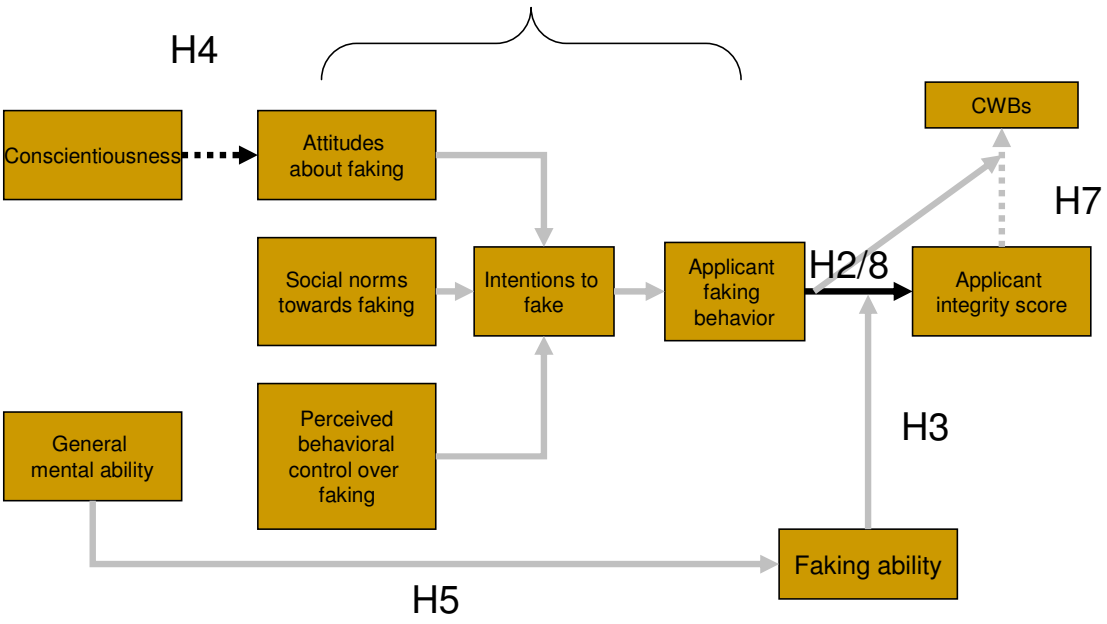

Figure 7. Graphical summary of results (Grey arrows indicate partial support. Black arrows indicate full support. Dashed arrows indicate negative relationships.)

Across both studies, tests for mediation based on changes in R^2 found initial evidence that intentions mediated the effect of the TPB variables on lie scale scores.

However, Study 1 also found some support for mediation associated with applicant-honest difference scores and self-report faking behaviors, relationships which were not supported in Study 2. These results were further confirmed by Sobel tests of the indirect effect associated with the TPB cognitions.

The second hypothesis was supported in both studies. This result indicates that applicant faking behaviors significantly increase applicant integrity scores, based on differences in integrity scores observed across response instruction categories, as well as differences within the applicant condition, based on self-reported faking behaviors.

The third hypothesis was supported in Study 1, but not Study 2. Specifically, Study 1 found a significant effect on applicant integrity test scores associated with the interaction between individual faking ability and faking behavior (measured using both self-reported faking behaviors and lie scale scores). This interaction was in the direction predicted, such that individuals who reported high levels of attempted faking (or who scored poorly on the lie scale) received higher integrity test scores when they also had high ability, compared to people with low ability to fake. However, no significant interactions were found in Study 2.

Hypothesis 4 was supported in both samples. Across both Study 1 and Study 2, conscientiousness was a significant predictor of attitudes towards faking. Specifically, individuals higher in conscientiousness tended to have significantly more negative attitudes towards faking, and reported lower intentions to fake and applicant faking attempts.

Hypothesis 5 was partially supported in both studies, based on SAT scores, but clearly supported in Study 2 on the basis of GMA scores. Because SAT scores are only an indirect measure of cognitive ability, the measure based on the Wonderlic arguably provides the clearer test of Hypothesis 5, and would suggest that Hypothesis 5 was supported.

Hypothesis 6 was not supported in either study. Results indicate that neither SAT scores nor GMA scores were a significant predictor of an individual's level of perceived behavioral control over faking. Thus, general mental ability did not appear to play any role in the extent to which individuals perceived themselves to be able to control faking behaviors.

Hypothesis 7 was partially supported in both studies. In both studies, applicant integrity score predictive validity with regards to CWBs were moderated by self-reported faking behaviors, but not by lie scale scores. This effect was in the direction initially hypothesized: individuals who reported more faking behaviors had a weaker relationship between applicant integrity test scores and CWBs than those who reported fewer faking behaviors.

Finally, the last hypothesis was supported in both samples. Results from both Study 1 and Study 2 found that applicant-condition and honest-condition integrity test scores displayed significant deviations from perfect consistency in rank-ordering – in other words, when individuals were asked to fake as applicants, it altered the rank order of integrity scores.

Overall, the study suggests that overt integrity test faking is at least partly driven by conscious, intentional processes, as well as by individual differences in cognitive ability and conscientiousness. Second, faking has substantial effects on both mean levels of integrity scores, relative levels of integrity scores (i.e., rank order), as well as on integrity test validity. Third, the effects of faking behaviors tend to be moderated by individual faking ability. Finally, cognitive ability does not appear to affect perceived behavioral control over faking.

Previous research (e.g., McFarland, 2000) has attempted to test a model of faking based on the Theory of Planned Behavior with mixed results, but the current study finds generally stronger evidence for the relationship between TPB cognitions, intentions, and actual faking behavior. Compared to the study of personality tests by McFarland (2000), the current study finds roughly equivalent predictive power of the TPB variables for faking intentions (R^2 = .57). However, the current study finds generally more robust relationships between intentions to fake and faking behavior than those found in McFarland (2000). This suggests that the TPB can be a useful framework for predicting applicant faking behavior on overt integrity tests.

The current study found that faking intentions had a mean correlation of .17 (sample-weighted) with applicant-honest difference scores, compared to an average correlation of r = .10 between difference scores found by McFarland (2000). In addition, intention-behavior correlations for overt integrity difference scores a week later were as high as .24 in the current study, compared to those found by McFarland for conscientiousness (r = .20, p < .05).

The variance between correlations found by McFarland and the current study supports the idea that an intentional, decision-based faking model may not be equally predictive across different non-cognitive selection measures. Specifically, the variance in intention-behavior relations suggests the possibility of important test-characteristic moderators such as test fakability or the social stigma associated with poor test performance. The existence of such moderators might determine the extent to which an intentional model of faking like the TPB is appropriate for a given test.

In the case of overt integrity tests, a TPB-based model may be more appropriate than for other non-cognitive selection tests such as personality measures. Overt integrity tests have easily-identified positive and negative directions; items which are specific, clear, and direct; and a potentially large stigma attached to poor test performance. The latter factor may be especially important for determining the relative appropriateness of decision-based faking models for selection tests. For instance, a low score on an overt integrity test would be associated with the ego-threatening stigma of “thief” or “liar.” In contrast, even low scores on very similar scales such as conscientiousness could be associated with very different stigmas, such as being perceived as lazy or disorganized.

In addition to evidence on the cognitive process of applicant faking, the current study also provides additional information on some of the individual differences that may play a role in how faking behaviors form and play out in applicant settings. Results support the idea that general levels of individual conscientiousness may be an antecedent of attitudes towards faking, and intelligence may play a role in the extent to which individuals are able to maximize their scores on integrity tests (i.e., faking ability).

Several authors (e.g., Tett et al., 2006) have recently called for models of faking that integrate individual differences, and the significant relations found in the current study provide empirical evidence of the potential utility offered by such an approach.

The evidence that skill or ability plays a role in determining faking outcomes is not new, and has been discussed under such terms as "testwiseness" (Edwards, 2003) and smart fakers or "test smarts" (Haas, Smith, & McDaniel, 1999; Ones et al., 1996). These constructs have often been operationalized using specific aspects of test knowledge, item identification accuracy, or the use of particular test-taking strategies. Our study finds that cognitive ability predicts faking behavior, but that ability to fake is not sufficient for response distortion to occur on integrity tests; having the ability to fake does not imply that an applicant will do so.

Another interesting finding of the current study is the way in which various operationalizations of faking behavior produced slightly different results. The difference between measures of generalized (i.e., not specific to a particular situation) faking behavior represented by lie scales and situation-specific measures of faking behavior (difference scores and self-reported behaviors) appears to be an important factor, especially for the relationship described by Hypothesis 7 regarding the effect of faking behaviors on integrity test validity.

In this case, situation-specific measures of faking were significant moderators of the integrity-CWB validity relationship, but a generalized, trait-like measure of intentional faking (i.e., lie scale) was not. This result supports findings reported by Griffith et al. (2006) that neither impression management (like the scale used in the

current study) nor social desirability (another trait-like measure representing unconscious faking) do a particularly good job of predicting difference scores occurring in specific testing situations. To some extent, this further calls into question the results of research into the effects of faking on test validity based on generalized measures of response distortion (Ones et al., 1996), and points to the potential importance of specific situational factors in determining the exact effects of faking on the selection process.

Another issue that our results suggest may be important is the issue of sample characteristics. An analysis of sample characteristics from the two studies showed that they differed in a number of sample characteristics, including sex ($\chi^2(1) = 8.3$, $p < .01$), school classification ($\chi^2(1) = 52.2$, $p < .01$), hours worked per week ($\chi^2(1) = 27.9$, $p < .01$), and whether or not they responded to the test with a particular type of job in mind ($\chi^2(1) = 35.6$, $p < .01$). Specifically, Study 1 appeared to have a greater proportion of females, a broader representation of different student classifications (freshmen, sophomores, etc.), more students who worked for more than 10 hours per week, and more students who had a specific job in mind when responding. The samples did not appear to differ in terms of ethnicity ($\chi^2(1) = 4.1$, $p > .05$) or GPA ($\chi^2(1) = 4.0$, $p > .05$). Thus, the different results observed in the two studies may be partly the result of differences in sample characteristics, with stronger findings in the sample that reported greater job experience and specific job applicant roles used to guide responding. The lack of clear applicant schemas and job experience on the part of applicants in Study 2 may have affected the results.

One of the major questions in faking research is the conditions under which faking results in non-optimal selection decisions. Simulation-based studies of selection decisions (Schmitt & Oswald, 2006) have suggested that changes in rank ordering may reduce the efficiency and fairness of selection decisions, and the present study finds that in the case of integrity tests, shifts in rank ordering associated with applicant faking behavior occur.

Other authors (Tett et al., 2006) have argued that faking may not be a serious problem to the extent that other criterion-related variables are also related to faking; for example, an individual who fakes on an integrity test has a lower 'true score' on the construct than their observed test results would indicate. In some cases, this lower 'true score' may be compensated by other variables related to faking. For example, intelligence is related to job performance, and this study finds that it is also related to faking ability. If the integrity test was being used to predict job performance, then higher cognitive ability would produce greater levels of faking, but also increased levels of job performance, thus tending to reduce any adverse effect of faking behavior on integrity test validity.

This study identified two major individual differences that were indirectly or directly related to faking behaviors: conscientiousness and cognitive ability. Conscientiousness was a significant negative predictor of CWBs, but had null or negative relationships with specific faking behaviors. Intelligence was positively related to at least one aspect of faking (faking ability), but not related to the criterion (CWBs). Thus, the current study indicates that high levels of conscientiousness and cognitive

ability (two important variables in a selection context) may have different effects on the relationship between faking and criterion-related validity, at least in the specific case of integrity tests and CWBs.

Implications

Overall, these results have a number of implications for research on integrity test faking. First, it strongly reinforces recent psychometric-based studies on the significance of faking for test validity and outcomes. This is in direct contrast to suggestions by some authors (see Morgeson et al., 2007; Ones & Viswesvaran, 1998a; Ones et al., 1996) that the effects of faking are minimal and pose no threat to test validity. The study clearly indicates that intentions to fake are a significant predictor of individual faking behavior, and that this behavior affects applicant test scores, rank on the test, and the predictive validity of the test with regards to self-reported counter-productive work behaviors. Therefore, instead of calling for a moratorium on faking research as some authors have done (e.g., Morgeson et al., 2007), this study suggests that a better understanding of applicant faking behavior can help ensure the validity and efficiency of integrity tests currently in use by organizations. More specifically, the current model provides general support for the important role of conscious, intentional processes on the part of applicants.

Second, the current study provides initial support for the distinction between the typical applicant faking behavior (i.e., “will do” faking) and applicant faking ability (i.e., “can do” faking) on overt integrity tests. Such a distinction is not just of theoretical interest, because it suggests the possibility of an important moderator of faking

behaviors (i.e., faking ability), which has not yet been fully integrated into the applicant faking literature. The majority of this literature has concerned itself with the causes and consequences of faking behaviors, not the separate causes or consequences of faking ability (or quality of faking). However, instead of focusing purely on *whether* people faked, researchers may need to pay more attention to *how well* applicants fake. Applicants do not fake uniformly, there are clear alterations to applicant rank-ordering based on test scores, and there is some initial support for the relationship between cognitive ability and faking ability.

This raises the specter of an interesting dilemma for organizations: applicants with high cognitive ability and strong motivations to get the job may be the most likely to successfully fake. However, these are the applicants most wanted by organizations. The current model suggests that hiring based on cognitive ability could lead to an increase in the average level of applicant faking. This could shed light on the findings that fakers are disproportionately ranked among the top of the applicant pool (e.g., Douglas et al., 1996; Rosse et al., 1998). Computer simulated studies that varied the number of fakers and amount of faking also found that fakers rise to the top of the rank-order (Zickar et al., 1996, 2000). Since faking ability increases the effect of applicant faking, this could reduce the usefulness of a measure of integrity, especially given the use of top-down selection. Organizations interested in both cognitive ability and integrity should therefore consider the consequences of their selection system carefully, and base their decision on a cost-benefit analysis that takes into account the effect of faking ability on the predictive validity of integrity tests.

The relationship between cognitive ability and faking ability also suggests one potential explanation for the slightly inconsistent (and sometimes null) results on the effects of faking on predictive validity, especially for job performance. Individuals with high cognitive ability will tend to receive higher test scores when they fake. Normally, this might result in a decrease in predictive validity for outcomes like job performance; however, the effects of any increase in faking associated with increasing ability could be hidden by the substantial relationship between cognitive ability and performance. This is consistent with one of the possible explanations by Tett et al. (2006) for why faking may not always result in a reduction in predictive validity.

Third, the study provides increased support for the role of conscious intentions in integrity test faking behavior. Applicant test-taking behaviors may well be affected by unconscious biases such as self-deception, but individuals who report a greater intention to fake also tend to report greater levels of faking behavior. Previous research has found partial support for the role of intentions on faking behavior for personality tests (e.g., McFarland & Ryan, 2006), but the current study suggests that overt integrity tests, with their clear-purpose nature, could be affected even more strongly by conscious faking decisions than are other non-cognitive measures.

The significant role played by conscious decision-making processes on overt integrity tests suggests directions for research on interventions designed to reduce faking. For example, warnings are occasionally used to reduce faking (Dwight & Donovan, 2003; Hurtz & Bohon, 1997), and have shown some promise. The current research suggests that other interventions designed to increase awareness about the

negative consequences of faking, produce social pressure to avoid faking, or prime anti-faking attitudes could prove especially efficacious, since they would directly influence applicant intentions to fake. Organizations interested in reducing or avoiding faking on the part of applicants could adopt policies directed towards removing situational or cognitive determinants associated with applicant faking cognitions.

Limitations and future directions

While the current study found a number of interesting results, it also had a number of limitations that need to be addressed in future research. Although participant responses indicated that most students took the test seriously and tried to follow instructions, one potential limitation is the reliance on a student (non-applicant) sample. Concerns regarding the use of student participants should be limited because the jobs held by college students (e.g., finance and sales) make them a fair representation of a large group of test takers who have had prior experiences with integrity tests (Ryan & Sackett, 1987; Sackett et al., 1989).

Another way in which the results of the current study are limited is that the proposed model and hypothesis tests do not provide an explicit test of the moderators which may influence the appropriateness of a TPB-based model of faking for different types of non-cognitive selection tests. Results of the current study and McFarland (2000) suggest some potential moderators of the faking intention-faking behavior relation, but ultimately inconclusive. The identification of specific test characteristics such as item type (i.e., attitudinal versus admittance items), test type (e.g., overt versus covert), or test

publisher – that might moderate the relationship between intention and faking behavior – could provide a basis for future work on faking interventions.

A third limitation is the way in which integrity was measured. Recent reviews (Berry et al., 2007) have suggested that integrity is actually a multifaceted, multidimensional construct. Research on recently developed multidimensional integrity scales suggests that different facets may be faked to different extents (Slaughter, Payne, & Yu, 2006). However, the integrity test used in the current study (while widely used and well-validated) does not allow for easy testing of facet-specific hypotheses. However, the use of such a scale would potentially provide information on the extent to which specific integrity test item categories are faked, how item content might moderate the effect of faking intentions on test outcomes, or how the various combinations of item types on commercial tests can be used to help organizations predict which tests will be more or less susceptible to intention-based faking.

Beyond addressing specific limitations of the current study, results suggest a number of future directions for research on faking and integrity tests. One such extension would be to identify other correlates (e.g., ethicality or psychopathy) of the TPB variables relevant to organizations. The current model focuses on trying to explain the source and effects of faking behavior on integrity tests, but variables like social norms towards faking, perceived behavioral control over faking, and attitudes towards faking may be useful predictors of other individual behaviors. For example, in the current study, faking-related TPB cognitions predicted almost as much variance in

CWBs as conscientiousness and applicant integrity scores, two variables that are often used to predict CWBs.

Related specifically to process models of integrity test faking, the current model could be expanded by identifying variables that affect the cognitive accessibility and direction of applicant faking cognitions. For example, individuals may participate in different types of social networks (work, friendship, etc.) which may or may not overlap. Each network represents a possibly different social setting; thus, cognitions of social norms accessed by an applicant may be partly dependent on which social setting or network is driving an applicant's perceptions. Social norms of applicant faking that might be acceptable in one setting (with strangers) might not be acceptable in another (with friends).

Cognitions may also be affected by various facets of the organization's selection process. The use of other selection tests that are not as easily subject to faking or cheating (e.g., cognitive ability tests or reference checks) may reduce the anticipated gains of faking perceived by applicants. The use of computer adaptive testing procedures to select items and administer tests may reduce an applicant's level of perceived control over faking (Richman, Kiesler, Weisband, & Drasgow, 1999). Framing applicant instructions regarding honesty and faking could affect attitudes towards faking.

CONCLUSION

Applicant response distortion is a process that is not fully understood, but it is clear that intentional faking plays an important role. The present study provides information on the cognitions that influence the level of faking applicants display on overt integrity tests, and some of the individual differences that influence these factors. The results indicate that applicant faking behaviors are directly related to faking intentions and indirectly related to other antecedent cognitions. It provides empirical evidence for the utility of differentiating between baseline levels of applicant faking and individual differences in faking ability, demonstrating that faking ability can stem from general cognitive ability. Additionally, it provides evidence for the negative effect of faking on the criterion-related validity of overt integrity tests, and for changes in the rank-order of applicants associated with faking.

REFERENCES

Alliger, G. M. & Dwight, S. A. (2000). A meta-analytic investigation of the susceptibility of integrity tests to faking and coaching. *Educational and Psychological Measurement, 60*, 59-72.

Alliger, G. M., Lilienfeld, S. O, & Mitchell, K. E. (1996). The susceptibility of overt and covert integrity tests to coaching and faking. *Psychological Science, 7*, 32-39.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: From cognition to behavior*, (pp. 11-39). Heidelberg, Germany: Springer.

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

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology, 22*, 453-474.

Arthur, W., Jr., Woehr, D. J., & Graziano, W. G. (2001). Personality testing in employment settings: Problems and issues in the application of typical selection practices. *Personnel Review, 30*(6), 657-676.

Baba, V. V. (1990). Methodological issues in modeling absence: A comparison of least-squares and tobit analyses. *Journal of Applied Psychology, 75*, 428-432.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287.

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.

Beck, L., & Ajzen, I. (1991). Predicting dishonest actions using the theory or planned behavior. *Journal of Research in Personality, 25*, 285-301.

Bell, B. S., & Kozlowski, W. J. (2002). Goal orientation and ability: Interactive effects on self-efficacy, performance, and knowledge. *Journal of Applied Psychology, 87*(3), 497-505.

Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology, 85*(3), 349-360.

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.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*, 425-440.

Brown, R. D., & Cothern, C. M. (2002). Individual differences in faking integrity tests. *Psychological Reports, 91*, 691-702.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge, UK: Cambridge University Press.

Caruso, J. C., & Cliff, N. (1997). Empirical size, coverage, and power of confidence intervals for Spearman's rho. *Education and Psychological Measurement, 57*, 637, 654.

Christiansen, N. D. (1998). Sensitive or senseless? *On the use of social desirability measures to identify applicant response distortion*. Paper presented at the 13th Annual Conference of the Society of Industrial and Organizational Psychology, Dallas, TX.

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.

Colbert, A. E., Mount, M. K., Harter, J. K., Witt, L. A., & Barrick, M. R. (2004). Interactive effects of personality and perceptions of the work situation on workplace deviance. *Journal of Applied Psychology, 49*, 599-609.

Cullen, M. J., & Sackett, P. R. (2004). Integrity testing in the workplace. In J. C. Thomas (Ed.), *Comprehensive handbook of psychological assessment: Industrial and Organizational Assessment* (Vol.4, pp. 149-165). Hoboken, NJ: John Wiley & Sons Inc.

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.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford, NY: Oxford University Press.

Donovan, 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, August). The validity of non-cognitive measures decays when applicants fake. *Academy of Management Proceedings*, 747-772.

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

Edens, P. L., & Arthur, W., Jr. (2000). *A meta-analysis investigating the susceptibility of self- report inventories to distortion*. Paper presented at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Edwards, B. D. (2003). *An examination of factors contributing to a reduction in Race–based subgroup differences on a constructed response Paper–and–pencil test of achievement*. Unpublished doctoral dissertation, Texas A&M University, College Station.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*(3), 265-287.

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.

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology, 73*(3), 421-435.

Fishbein, M., & Ajzen, I. (1975). *Beliefs, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesely.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability.

Psychological Science, 15, 373-378.

Griffith, R. L. (1998). Faking of noncognitive selection devices: Red herring is hard to swallow. Unpublished doctoral dissertation, The University of Akron, Akron, OH.

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.

Griffith, R. L., Malm, T., English, A., Yoshita, Y., & Gujar, A. (2006). Applicant faking behavior: Teasing apart the influence of situational variance, cognitive biases, and individual differences. In R. L. Griffith & M, H, Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 151-178). Greenwich, CT: IAP.

Griffith, R. L, & McDaniel, M. (2006). The nature of deception and applicant faking behavior. In R. L. Griffith & M, H, Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 1-19). Greenwich: IAP.

Goldberg, L. R. (1999). A broad-bandwidth, public-domain personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Guion, R. M. (1965). Synthetic validity in a small company: A demonstration. *Personnel Psychology, 18*, 49-63.

Haas, A. Smith, K., & McDaniel, M. (1999). *Faking strategies: Effects on a situational judgment test. "Smart fakers" factor (a test-taking strategy) was not related to improved test scores.* Paper presented at the 14th Annual Conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.

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

Hogan, J., & Brinkmeyer, K. (1997). Bridging the gap between overt and personality-based integrity tests. *Personnel Psychology, 50*, 587-599.

Hough, L. M. (1998). The millennium for personality psychology: New horizons or good ole daze. *Applied Psychology: An International Review, 47*(2), 233-261.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response

distortion on those validities. *Journal of Applied Psychology, 75*, 581-595.

Hough, L. M., & Oswald, F. L. (2007). Personality testing and I-O Psychology: Asking questions, offering answers, discussing unknowns, and providing direction. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*(3), 1-51.

Hurtz, G. M., & Alliger, G. M. (2002). Influence of coaching on integrity test performance and unlikely virtues scale scores. *Human Performance, 15*(3), 255-273.

Hurtz, G. M., & Bohon, L. M. (1997). *Impression management on personality-based employee selection tests.* Paper presented at the 12th Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Jackson, D. N., Wroblewski, V. R., & Ashton, M. C (2000). The impact of faking on employment tests: Does forced choice offer a solution? *Human Performance, 13*, 371-388.

Keene, O. N. (1995). The log transform is special. *Statistics in Medicine, 14*, 811-819.

McDaniel, M. A., Douglas, E. F., & Snell, A. F. (1997). *A survey of deception among job seekers*. Paper presented at the 12th annual conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.

McFarland, L. A. (2000). *Toward an integrated model of applicant faking.* Unpublished doctoral dissertation, Michigan State University, East Lansing.

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

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*(4), 979-1016.

McKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30*(5), 525-312.

Millsap, R. E. (1997). Invariance in measurement and prediction: Their relationship in the single-factor case. *Psychological Methods, 2*, 248-260.

Millsap, R. E. (2007). Invariance in measurement and prediction revisited. *Psychometrika, 72*(4), 461-473.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*, 683-729.

Mueller-Hanson, R., Heggestad, E. D., & Thornton III, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select out perspectives. *Journal of Applied Psychology, 88*(2), 348-355.

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

Murphy, K. R. (1993). *Honesty in the workplace*. Pacific Grove, CA: Brooks/Cole.

Murphy, K. R. (2000). What constructs underlie measures of honesty or integrity? *Problems and solutions in human assessment*; honoring Douglas N. Jackson at seventy, (pp. 265-283). Boston, MA: Kluwer Academic Publishers.

O'Bannon, R. M., Goldinger, L. A., & Appleby, G. S. (1989). *Honesty and integrity testing: A practical guide.* Atlanta, GA: Applied Information Resources.

Ones, D. S. (1994). *The construct validity of integrity tests*. Unpublished doctoral dissertation, University of Iowa, Iowa City.

Ones, D. S., & Viswesvaran, C. (1998a). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11*(2/3), 245-269.

Ones, D. S., & Viswesvaran, C. (1998b). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology, 83*(1), 35-42.

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

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1995). Integrity tests: Overlooked facts, resolved issues, and remaining questions. *American Psychologist, 50*(6), 456-457.

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*(4), 679-703.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2003). Personality and absenteeism: A meta-analysis of integrity tests. *European Journal of Personality, 17*, S19-S38.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*(2), 329-335.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 589-609.

Paulhus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego: Academic Press.

Rees, C. J., & Metcalfe, B. (2003). The faking of personality questionnaire results: Who's kidding whom? *Journal of Managerial Psychology, 18*(2), 15-165.

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

Robie, C., Brown, D. J., & Beaty, J. C. (2007). Do people fake on personality inventories? A verbal protocol analysis. *Journal of Business and Psychology, 21*(4), 489-509

Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal, 38*(2), 555-572.

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83-92.

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.

Ryan, A.M., & Sackett, P. R. (1987). Pre-employment honesty testing: Fakability, reactions of test takers, and company image. *Journal of Business and Psychology*, *1*, 248-256.

Rynes, S. L. (1993). Who's selecting whom? Effects of selection practices on applicant attitudes and behavior. In N. Schmitt & W.C. Borman (Eds.), *Personnel selection in organizations*. San Francisco: Jossey Bass Publishers.

Sackett, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology, 42*, 491-529.

Sackett, P. R., & Harris, M. M. (1984). Honesty testing for personnel selection: A review and critique. *Personnel Psychology, 37*, 221-245.

Sackett, P. R. & Wanek, J. E. (1996). New developments in the use of measures of honest, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel Psychology, 49*, 787-829.

Schifter, D. B., & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behavior. *Journal of Personality and Social Psychology, 49*, 843-851.

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. *Psychological Bulletin, 124*(2), 262-274.

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.

Slaughter, A. J., Payne, S. C., & Yu, J. (2006). *Integrity tests and impression management: Differences across item categories*. Paper presented at the 21st Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

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

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

Sobel, M. F. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt, (Ed.), *Social methodology* (pp. 290-312). Washington, DC: American Sociological Association.

Stark, S., Chernyshenko, O. S., Chan, K., Lee, W. C., & Drasgow, F. (2001). Effects of

the testing situation on item responding: Cause for concern. *Journal of Applied Psychology, 86*(5), 943-953.

Tabachnik, B. G. & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: Harper Collins Publishers.

Tett, R. P., Anderson, M. G., Ho, C., Yang, T. S., Huang, L., & Hanvongse, A. (2006). Seven nested questions about faking on personality tests: An overview and interactionist model of item-level response distortion. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 43-83). Greenwich, CT: IAP.

Van Iddekinge, C. H., Raymark, P. H., Eidson, C. E., & Putka, D. J. (2003). *Applicant-incumbent differences on personality, integrity, and customer service measures*. Poster presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Wanek, J. E., Sackett, P. R., & Ones, D. S. (2003). Towards an understanding of integrity test similarities and differences: An item-level analysis of seven tests. *Personnel Psychology, 56*, 873-894.

Whitney, D. J., Diaz, J., Mineghino, M. E., & Powers, K. (1999). Perceptions of overt and personality-based integrity tests. *Overt and Personality-Based Integrity Tests, 7*, 35-45.

Wonderlic, E. F. (1992). *Wonderlic Personnel Test User's Manual*. Libertyville, IL: E. F. Wonderlic.

Wooley, R. M., & Hakstian, A. R. (1992). An examination of the construct validity of personality-based and overt measures of integrity. *Educational and Psychological Measurement, 52*, 475-489.

Zickar, M. J. (2000). Modeling faking on personality tests. In D. Ilgen & C. L. Hulin (Eds.), *Computational modeling of behavior in organizations* (pp. 95-108). Washington, DC: American Psychological Association.

Zickar, M. J. & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement, 20*(1), 71-87.

APPENDIX A: INSTRUCTIONS

Introduction:

Howdy! My name is Janie Yu. I will be your experimenter today.
The purpose of this study is to see how individuals respond to a selection test that is commonly used to choose applicants into a variety of jobs. This study requires you to take the same test 3 times. Before you take the selection test, you will be asked to fill out a survey that will ask you for some questions related to the selection test.

The consent form I am handing out states the purpose of the study. It also states that your responses will remain confidential, your participation is voluntary, and you may omit any items and leave at any point of the study. [Pass out and sign consent form and answer questions.] I will now hand out the survey packets then read the instructions.

Respond as an Applicant

The test you are about to take is one that is frequently used by employers to select employers for various job types (e.g., sales, managerial positions, finance, etc.). When responding to the questions on this test, TRY TO PICTURE YOURSELF AS THE JOB APPLICANT. Try to answer the questions as if you were the person who is trying to get a job. Please keep in mind that your answers will be kept completely anonymous. To make the situation more like a job applicant situation, you will be given an incentive of $20 if you score among the top 15% on this test.
[Source: See McFarland (2000)]

Respond Honestly

Please answer the following questions on this test as honestly as possible. Please note that your answers will remain completely anonymous (i.e., no one, not even the researcher will be able to link your answers to your name). Keep in mind that the score you obtain on this test will not be used outside the context of this experiment. That is, your answers will be used for researcher purposes only, and will not be used to evaluate you in any way, so please ANSWER AS TRUTHFULLY AS POSSIBLE.
[Source: See McFarland (2000)]

Fake good to "beat the test"

When responding to the questions on this test, select responses that you feel would help you OBTAIN THE BEST POSSIBLE SCORE, regardless of how much you have to lie. Select answers that you believe will ensure that you will get the highest score possible. If you score among the top 15% on this test, you will receive $20.
[Source: See Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Ryan & Sackett, 1987]

APPENDIX B: MEASURES

Previously published measures used in the study are described below, along with information on the instructions and response anchors used. Copies of full measures may be obtained directly from the source provided. In addition, original measures used in the study are reproduced entirely.

Conscientiousness Measure (Goldberg, 1999)
Conscientiousness was assessed using a 10-item scale from Goldberg (1999). Respondents are asked to describe how accurately each statement describes them. Response anchor ranged from 1 (very inaccurate) to 5 (very accurate).

BIDR-IM subscale version 6 (Paulhus, 1991)
The Impression Management subscale of the BIDR measure consists of 12 items. A 5-point response anchor ranging from 1 (not true) to very true (5) was used.

Interpersonal and Organizational Deviance Scale (Bennett & Robinson, 2000)
This scale provides a 19-item measure of workplace deviance. Respondents are asked to estimate the frequency with which they engage in certain types of behaviors. Response anchors include: Never, once a year, twice a year, several times a year, monthly, weekly, and daily.

TPB Questionnaire (McFarland, 2000)
These items asked about respondent's feelings toward lying on employment tests. *Attitudes towards faking* items assess attitudes toward faking by responding to a 5-item measure based on five different semantic differential-type response scales: good bad, pleasant-unpleasant, foolish-wise, useful-useless, and unattractive-attractive. *Social norms for faking* were assessed using a 4-item measure on a 5-point scale response anchor ranging from strongly agree to strongly disagree. *Perceptions of behavioral control over faking* were measured using 3 items on a 5-point response anchor ranging from strongly agree to strongly disagree. *Intentions to fake* on future tests were measured using ten items on a 5-point scale, ranging from strongly agree to strongly disagree.

Self-Report Faking Measure (McFarland, 2000)
Items for the self-report faking measure were obtained from McFarland (2000), who used the items as a manipulation check for an applicant test-taking condition. This scale provides a 7-item measure of attempted faking. Respondents are asked to report the extent to which they faked on the test they completed. Response anchors ranged from strongly disagree to strongly agree.

Employee Reliability Index (Ryan & Sackett, 1987)
The ERI is a research tool designed to closely mirror existing overt integrity measures.

There are 63 items used to measure overall integrity. Respondents are asked to either report the extent to which they agree with a statement (on a scale of 1 to 5, strongly disagree to strongly agree), or for certain items, to choose a specific value presented in a 5-option multiple-choice format.

Demographic Questionnaire

1. What is your gender? [(a)Male, (b) Female]

2. With which of the following ethnic groups do you most closely identify?

    a) African American (Black)
    b) Asian or Pacific Islander
    c) Caucasian (White, Non-Hispanic)
    d) Latino/Latina
    e) Other

3. What is your class year?

    a) Freshman
    b) Sophomore
    c) Junior
    d) Senior

4. On average, approximately how many hours per week have you worked for pay during the last year?
    a) Less than 10 hours per week
    b) 11 to 19 hours per week
    c) 20 to 95 hours per week
    d) 30 to 40 hours per week
    e) More than 40 hours per week

5. When responding to items on the employment selection test, I had a particular job in mind. [(a) Yes, (b) No]

6. When responding to items on the employment selection test, I had in mind a job that is similar to:
    a) Retail or sales
    b) Food industry
    c) Banking industry
    d) Administrative position
    e) None of the above

7. I was completely truthful when I was instructed to respond honestly:
    a) Strongly agree
    b) Agree
    c) Neither agree nor disagree
    d) Disagree
    e) Strongly disagree

8. I believe that these types of test should be used to select job applicants.
    [(a) True, (b) False]

9. The $20 cash prize was an adequate amount to motivate me.
    [(a) True, (b) False]

10. The extra credit was an adequate amount to motivate me.
    [(a) True, (b) False]

11. I have completed test(s) that are similar to the employment selection test in this study.
    [(a) Yes, (b) No]

APPENDIX C: ORDER OF ADMINISTRATIONS FOR STUDY 1 AND STUDY 2

Table C1: Order of measurements

Measures Administered		
	Study 1	Study 2
Time 1	SAT TPB predictors (SN, PBC, ATT) Intention to fake Conscientiousness Impression Management Integrity-Applicant Integrity-Honest Integrity-Fake Demographics	SAT TPB predictors (SN, PBC, ATT) Intention to fake Conscientiousness Impression Management Wonderlic Personnel Test Demographics
Time 2 (about 1 week later)	Integrity-Applicant Self-Report Faking Integrity-Honest Self-Report Faking Integrity-Fake Self-Report Faking Impression Management CWB	Integrity-Applicant Self-Report Faking Integrity-Honest Self-Report Faking Integrity-Fake Self-Report Faking Impression Management CWB

VITA

Name: Janie Yu

Address: Department of Psychology
c/o Dr. Daniel Newman
Texas A&M University
College Station, TX 77843-4235

Email: JanieYu@tamu.edu

Education: Ph.D., Psychology, Texas A&M University
M.A., Psychology. California State University, Long Beach
B.A., Psychology, University of California, Los Angeles