This is to certify that the

thesis entitled

Individual Differences in the Ability

to Fake Across Non-Cognitive Measures

presented by

Lynn A. McFarland

has been accepted towards fulfillment
of the requirements for

Master degree in Arts

Major professor
Ann Marie Ryan

Date May 19, 1998

INDIVIDUAL DIFFERENCES IN THE ABILITY TO FAKE ACROSS NON-COGNITIVE MEASURES

By

Lynn A. McFarland

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

MASTERS OF ARTS

Department of Psychology

1998

ABSTRACT

INDIVIDUAL DIFFERENCES IN THE ABILTIY TO FAKE ACROSS NON-COGNTIVE MEASURES

By

Lynn A. McFarland

Applicant faking is often cited as a concern when using non-cognitive tests. This study demonstrates that there are individual differences in the extent to which individuals can increase their scores through faking three types of non-cognitive measures (i.e., personality test, biodata inventory, and an integrity test). Such differences in faking ability may change rank orders of applicants on these selection tests, and thus influence the usefulness of such measures or the interpretation of test scores. For all three non-cognitive tests the mean scores in the fake condition were significantly greater than in the honest condition. Additionally, self-efficacy for successfully faking each test was related to the extent of faking for the biodata form, integrity test, and two personality scales (extroversion and neuroticism). Neuroticism, conscientiousness, integrity, and self-monitoring were also related to the ability to fake some measures. Results also indicated that individuals faked fairly consistently across the measures. Thus, this study provides considerable evidence that there are individual differences in the ability to fake.

ACKNOWLEDGEMENTS

I would like to thank my committee chairperson, Dr. Ann Marie Ryan, for guidance through every step of this research process. She provided me with needed assistance, but also allowed me to determine the direction and scope of my thesis. I would also like to thank the other committee members, Drs. Rick Deshon and Neal Schmitt. Their helpful suggestions and valuable criticisms improved this thesis making it a better quality product.

This thesis, or any of my achievements, could not have been possible without the unwavering love and support of my family. They have helped me put things in perspective, and in doing so, have kept me sane throughout this process. I especially want to thank my parents, William and Rosanne McFarland. They have made me the person I am. Therefore, the completion of this thesis is as much a victory for them as it is for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

There are a number of benefits associated with using non-cognitive self-report measures such as biodata, personality inventories and integrity tests, for selection purposes. These tests generally yield high validities yet produce little adverse impact (Cunningham, Wong, & Barbee, 1994; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Kluger, Reilly, & Russell, 1991; Rothstein, Schmidt, Erwin, Owens, & Sparks, 1990; Sackett, Burris, & Callahan, 1989). However, there is concern that the validity of such self-reports may be attenuated by intentional distortion on the part of the applicant. For example, a number of large organizations do not use non-cognitive measures because they feel they are too easy to fake (Douglas, McDaniel, & Snell, 1996).

To quell concerns related to the fakability of non-cognitive measures, some researchers have found evidence that applicants do not fake on such selection tests and that even if faking occurs it does not negatively influence the validity of the test (Abrahams, Neumann, & Githens, 1971; Cunningham et al., 1994; Hough et al., 1990; McCrae & Costa, 1983; Ones, Viswesvaran, & Reiss, 1996). However, others disagree with these findings and provide their own evidence that faking does occur in selection settings and does attenuate the validity of non-cognitive measures (Douglas et al., 1996; Dunnette, McCartney, Carlson, & Kirchner, 1962; Kluger et al., 1991; Pannone, 1984).

The current study will contribute to the information on the fakability on noncognitive measures. The degree to which individuals can fake three types of non-cognitive measures (personality test, biodata inventory, and integrity test) will be

examined. Although faking has been discussed as an individual difference variable (Cohen & Lefkowitz, 1974; Fletcher, 1990; Steward & Carley, 1984), no one has looked at the ability to fake across types of selection tests.

This study will examine a number of individual difference variables that are potential correlates of ability to fake. A model of faking proposed by Douglas et al. (1996) will be used as the organizing framework for this investigation. This model is presented in Figure 1. Douglas et al. suggest that three factors, opportunity to fake, personal characteristics, and characteristics of the situation, have a direct effect on the extent to which applicants fake. First, some applicants may have more opportunity to fake because their true score on the test is relatively low as compared to other applicants. Therefore, some individuals have more room for improvement simply because their honest responses would not lead to high scores. The second factor which may influence the extent of faking are personal characteristics that may make some people better fakers. Risk taking tendencies, self-monitoring ability, and individuals' moral codes are examples of such differences. Third, situational characteristics may also determine faking. A person in a financial crisis may really need the job in question such that they would fake to do well on the selection test to ensure that they get the job.

Douglas et al. (1996) also suggest that their model explains why one should expect variance in the extent of faking between individuals. As will be discussed, interpreting test scores may be difficult if there is substantial variance in the extent to which applicants fake (i.e., some applicants improve their score a great deal through faking while others do not).

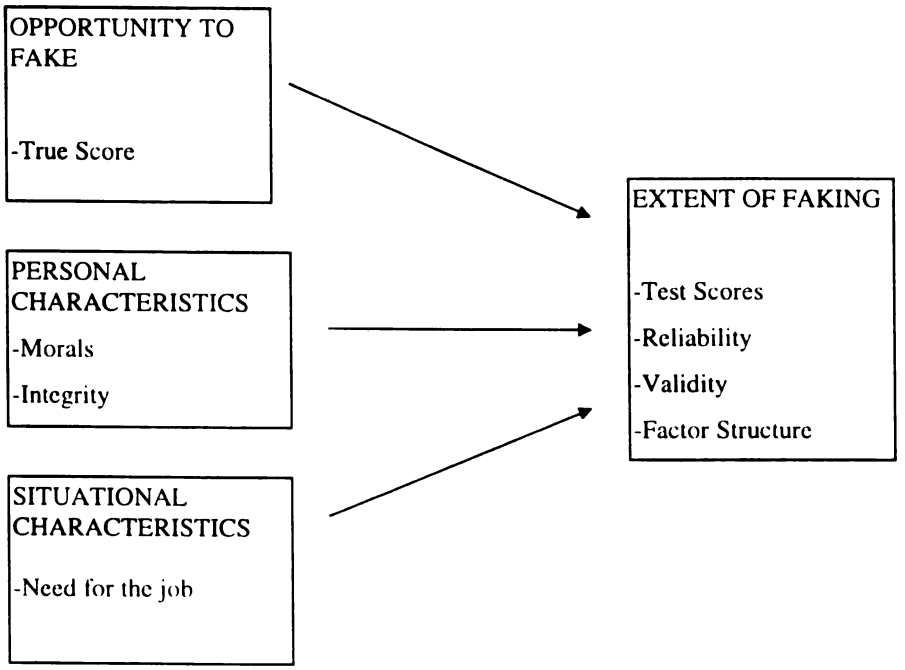

Figure 1. Douglas, McDaniel, and Snell Model of Faking.

Although the Douglas et al. (1996) framework is a good heuristic, it can be improved. One purpose of this study is to develop an expanded model of faking. Based on the results of this study the Douglas et al. model will be revised.

In the present study situational characteristics will be held constant; however, the extent to which personal factors and opportunity to fake influence the extent of faking on the three non-cognitive measures will be examined. First, however, faking will be defined and some reasons people may have for faking will be presented. Second, the research investigating whether applicants can and do fake will be examined. Third, the potential effects of faking on non-cognitive measures will be discussed and the literature providing evidence that faking biases results of such tests will be reviewed. Fourth, an alternative perspective that faking does not bias such measures will be presented. Fifth, some individual difference variables that are potential correlates of ability to fake will be discussed.

Definitions and Why People Fake

A number of labels have been given to describe the tendency of some individuals to select answers on tests that will result in others viewing them in the most favorable way. This has been referred to as response distortion, social desirability, faking, impression management, self-enhancement, and intentional distortion (Becker & Colquitt, 1992; Douglas, McDaniel, & Snell, 1996; Mumford & Stokes, 1992, Zerbe & Paulhus, 1987).

Paulhus (1984; 1986) suggests that socially desirable responding contains two components: self-deception and impression management. Self-deception refers to the

unconscious tendency to see one's self in a favorable light. Impression management is a conscious attempt to present false information to create a favorable impression on others. Therefore, when applicants taking a selection test consciously do not select the most accurate answer (the one that most closely describes them) but instead select the answer that they believe will make them look most favorable, they are using impression management. In the present study, this type of response bias, whereby an individual consciously distorts answers in order to be viewed favorably, will be referred to as faking. Paulhus (1984) suggests that unconscious self-deception is not affected by laboratory manipulations. The current study will ensure that only intentional faking is being measured as a within-subjects design will be used whereby subjects will respond to three non-cognitive measures honestly and then under a fake condition. Although self-deception will undoubtedly be present, it is assumed to be a constant.

There are a number of reasons to suspect that people may fake responses on a selection test. Leary and Kowalski (1990) suggest that people are motivated to manage their impressions when the impressions they make are relevant to the fulfillment of their goals and when these goals are greatly valued. Generally, when an applicant takes a selection test, he or she desires to get the job. Individuals may fake in order to obtain a higher score, thus increasing their chances of being hired. Pandey and Rastagi (1979) showed that attempts to manage impressions increased during a job interview when competition for the job became more intense. Therefore, it seems that as the stakes increase, attempts to look the best one possibly can also increase.

Other less obvious reasons individuals may fake are suggested by Schlenker and

Weigold (1992). They suggest that people regulate the impressions they make to maintain or enhance self-esteem. People generally act to maximize their self-esteem and will influence the content of self-presentation to do so. Answering a test in such a way that others will perceive one positively will help maintain self-esteem.

It is clear that there are reasons applicants may attempt to make themselves look favorable through faking on selection tests. The next questions to be addressed are: Can individuals fake such tests? If they can, do they? The following sections will review the literature suggesting that applicants can and do fake non-cognitive measures.

Evidence That People Can Fake

There is evidence that people *can* increase test scores by faking responses to non-cognitive measures. A number of studies have employed a within-subjects design, whereby subjects take the test in an honest condition and a faking condition, to examine whether people can increase their scores on non-cognitive measures through faking. Most of this research examines biodata inventories rather than personality and integrity tests. These studies provide evidence that applicants can distort their responses on non-cognitive measures.

Dunnette et al. (1962) had 62 salesmen take a forced choice self-description checklist twice in counterbalanced order. In one administration the subjects were asked to answer honestly, whereas in the other one they were told to fake (select the best answers). Subjects in the fake condition had mean scores more than half a standard deviation above those in the honest condition.

Doll (1971) had 300 college graduates receiving armed service training take a

biodata instrument under instructions to respond honestly. Some of the subjects were then told to fake to look good but to be prepared to defend their answers, another group was told to fake to look good but to be aware that a lie scale may be included to detect faking, and the last group was told to fake to look as good as possible. Results indicated that subjects could increase their scores by faking and that the largest score increases were in the "fake to look as good as possible" condition while the least amount of increase was seen by those in the condition which instructed subjects to be aware of a lie scale.

Schrader and Osburn (1977) had 148 volunteers take a biodata inventory under instructions to fake and two weeks later under instructions to answer honestly. Those instructed to fake to look as good as possible significantly increased their scores by as much as one standard deviation. The smallest increase in scores was for those who were told to fake but to be aware that a lie scale was included.

One criticism of this study and the one conducted by Doll (1971) is that in neither study did they counterbalance the conditions. In Doll's study all subjects took the biodata inventory under the honest condition first, while Schrader and Osburn (1977) administered the measures to subjects under fake conditions first. Klein and Owens (1965) have shown that order effects may exist. They found that those told to complete a biodata measure honestly and then told to complete it under instructions to fake received higher scores in the fake condition than those told to fake first and then respond honestly. The authors suggest that familiarity with the test may aid the ability to fake. Therefore, these studies should be considered in light of this weakness.

The fakability of an empirically developed biodata instrument was investigated by Thornton and Gierasch (1980). They had 94 male college students take the biodata inventory under both honest and fake conditions in counterbalanced order. They were told to respond in such a way as to increase their chances of being selected for a managerial job. Results revealed that there were significant score increases for seven of the ten biodata scales. Participants increased their scores by as much as one standard deviation. The authors conclude that even empirically keyed biodata inventories are fakable.

Other studies have used a between subjects design to examine if biodata inventories can be faked. Most of these studies find that the scores of those asked to fake the measure are significantly higher than those told to respond honestly. For example, Klein and Owens (1965) instructed half of their subjects to fake a scored life history blank and the other half to respond honestly. They found that the mean score of the fake condition was one standard deviation greater than the mean score in the honest condition.

Becker and Colquitt (1992) randomly placed 430 current employees into one of two conditions. Half of the employees were told to be honest when responding to a biodata form, while the remaining half were told to fake responses to look as good as possible so that they would most likely obtain the job they currently had, had they been a job candidate. Those in the fake condition had mean scores significantly higher (over one standard deviation) than those in the honest condition.

One criticism of this study is that the participants were told to fake so as to increase their chances of getting a job they desired. It is possible that the participants

were faking toward different jobs. Such responding may have actually resulted in underestimating the mean of the faking group. Thus, the true mean difference between the two groups may be even greater than this study suggests.

Although most research on faking has examined biodata inventories, a few studies have examined the extent to which individuals can fake personality and integrity tests. For instance, Douglas et al. (1996) administered a personality test and a biodata test, both of which contained an agreeableness and a conscientiousness scale, to 600 college students. They had a between subjects design in which half of the subjects were randomly placed into the honest condition, and the other half were placed into the faking condition (told to make themselves look as good as possible). Results revealed significant group differences of up to .70 standard deviations on the personality scale of agreeableness, .86 on conscientiousness, 1.30 on the biodata agreeableness scale, and 1.14 on the biodata conscientiousness scale, with all differences in favor of the faking condition.

Ryan and Sackett (1987) had subjects take an integrity test. Some subjects were told to take the test as if they were really job applicants, another group was told to answer truthfully, while a third group was asked to fake good. It was found that the subjects told to answer as applicants had higher scores than subjects asked to answer truthfully but less than those who were told to fake good.

Dwight and Alliger (1997) had 120 undergraduates take the Employee Integrity Index (EII) which is an overt integrity measure developed by Ryan and Sackett (1987). Subjects were placed in one of three conditions: fake good, honest, or a coaching

condition. Those in the coaching condition were told what integrity tests are designed to measure and taught how to respond to such items so as to obtain a high score. Results indicated that the mean score of the coaching condition was significantly higher (by almost .5 standard deviation units) than the mean score of those in the fake good and honest conditions. The mean score of the fake good group was also significantly greater (by over one standard deviation) than those in the honest condition.

Hurtz and Bohon (1997) examined the extent to which a personality-based integrity test could be faked. All subjects were told to play the role of an applicant for a retail sales job (they were all given a job description prior to taking the test). A between-subjects design was used in which applicants were assigned to either honest, fake good (fake so as to increase chances of getting the sales job), or an actual applicant condition (subjects told to answer as they would if they were really an applicant). The mean score of those instructed to fake good was over .50 standard deviation units higher than those in the honest condition.

These studies suggest that responses on non-cognitive measures can be faked. When asked to do so, subjects can increase their scores by as much as 1 standard deviation above honest respondents. It seems that knowledge of what the tests are designed to measure may also help applicants increase their scores. However, just because applicants have the ability to fake does not imply that they do fake. It is important to make the conceptual distinction between fakability and actual faking. If applicants can fake, but do not do so, then responses on a fakable measure may not be biased. Hence, it is crucial to examine the literature for any evidence indicating that

applicants actually do fake when responding to non-cognitive selection measures.

Evidence Applicants Do Fake

Dunnette et al. (1962) compared the scores of those in the fake and honest conditions to the scores of 64 applicants applying for a sales job. The applicants achieved a mean score somewhere between the mean scores of the subjects in the faking and honest conditions, indicating that the applicants faked to some extent. However, to determine whether all or only a few applicants tried to create a good impression on the checklist, a fake key for the checklist was developed. This was done by weighting those responses showing significant differences between the honest and fake conditions. By examining the fake score distributions it was determined that a cut score of 4 and above successfully identified 90% of the faked answer sheets at the expense of misclassifying only 9% of the honest ones. Through this investigation it was determined that only a few of the applicants actually faked. Therefore, it was the responses of these few that increased the mean for the applicant group. One must keep in mind that these conclusions are based on a small sample (n = 64) and thus must be viewed with caution.

Goldstein (1971) had 111 individuals applying for a position as a nurse's aide in a nursing home complete an application blank. The information individuals provided on the application blank was compared to the information provided by the applicant's last employer. Results indicated a number of discrepancies which included information as to the position which the applicant actually held, previous wages, and the duration of employment at that job. Moreover, 15% of the employers had said that the individual in question had never even worked for them. The authors concluded that these

discrepancies provided evidence that applicants faked the application blank. However, one must be aware that it may be the employer who provided inaccurate information. Therefore, the discrepancies found may have been a result of both applicant faking and the carelessness of responding on the part of the previous employer.

Pannone (1984) administered a rational biographical questionnaire to over 200 applicants for an electrician position. He included an item which made reference to a piece of electrical equipment that did not even exist. Those applicants who said they had used the piece of equipment were considered fakers. It was determined that over one-third of the applicants were fakers and there was a significant mean difference between the scores of fakers and non-fakers such that the mean score of the fakers was over .5 standard deviation units greater than those identified as non-fakers.

This study may be criticized in that it is possible that a number of the fakers were not identified, that is, some people who did fake throughout the test did not say that they had operated the fictional piece of equipment. Similarly, those who did not fake throughout the test may have been identified as fakers by their response to the fake item but actually were confused by the item or were just careless when responding to the item. Therefore, this one item measure of faking may also be tapping carelessness of responding rather than isolating intentional faking.

Kluger, Reilly, and Russell (1991) had subjects respond to two biodata inventories honestly. These same subjects also took the two inventories in a simulated applicant condition. Conditions were balanced. Results indicated that subjects simulating applicants were able to significantly increase their scores upwards of 1 standard deviation

when the biodata inventory was item-keyed (a procedure that assumes linear, monotonic relationships between item scores and the criterion and ignores possible nonlinear relationships) but not when it was option-keyed (a procedure whereby each item response option is analyzed separately and contributes to the score only if it correlates significantly with the criterion). These findings indicate that applicants do fake; however, faking may not always result in increasing scores. It may depend on the way in which the test is scored.

In the study described earlier, Becker and Colquitt (1992) also administered the biodata inventory to 250 real applicants. Results indicate that the scores of the applicants fell between the scores of those in the honest and the fake conditions (they scored lower than the fakers but higher than the honest respondents), indicating that applicants do fake responses.

The sample used by Becker and Colquitt (1992) may be criticized. The applicants in their sample had already gone through an extensive selection process. Therefore, only those who had passed several hurdles were administered the biodata inventory. It is likely that those who managed to pass all of these hurdles had high true scores on the biodata measure. It should not be surprising that those who just passed a multiple hurdle selection procedure would obtain scores greater than a voluntary sample of incumbents who were likely less motivated to take the test.

Stokes, Hogan, and Snell (1993) had both incumbents and applicants take an empirically developed biodata form. They computed a social desirability scale value (SDSV) for each of the 168 biodata items to determine if subjects were faking. They

developed this measure by averaging response desirability ratings obtained from 16 psychology graduate students. These raters were given information about the job and the organization for which the applicants were competing and told to rate all item responses on a seven point scale (+3 being extremely desirable and -3 being an extremely undesirable response) as to the extent to which such a response would make the respondent look good. Results indicated that applicants were much more likely to make socially desirable responses. In fact, 25% of the variance in applicant versus incumbent response mean difference scores was accounted for by social desirability.

The measure of faking used is one potential criticism of this study. Although it appears this scale was developed in a rational matter the authors provide no empirical support of their scale such that we do not know if it taps self-deception, as other social desirability scales tend to (e.g., the Marlow-Crowne scale).

There is considerable evidence to suggest that applicants do fake biodata inventories. However, only three studies could be located that found evidence that applicants fake their responses on personality and integrity tests. Schmit and Ryan (1993) administered the NEO-FFI, a personality test used to measure the traditional five factor model (extroversion, agreeableness, conscientiousness, emotional stability, and intellectance or openness to experience), to applicants and non-applicants (students). Using confirmatory factor analyses, they found the five factor model fit the data from the student sample but the applicant data contained six factors. The authors suggested this sixth factor (an "ideal-employee factor") that surfaced in the applicant sample may have resulted from applicant faking.

McDaniel, Douglas, and Snell (1997) had 192 subjects complete a conscientiousness and a social influence measure that were mailed to them. They also had respondents indicate their occupation and asked if they had misrepresented themselves in various ways when applying for jobs in the past. Respondents first completed the conscientiousness and social influence measures as they would if they were applying for a job that they really wanted. They were then instructed to respond to these measures a second time honestly. For the conscientiousness measure, 23% of the sample improved their score in the applicant condition and 26% improved their scores in the applicant condition for the social influence measure. However, over 65% of the sample obtained the same score across conditions for the social influence scale as compared to 72% for the conscientiousness measure. When responding as applicants, respondents increased their conscientiousness scale score by .26 standard deviations and their social influence scale score by .16 standard deviations. There were also differences in the extent to which people faked across occupations. For example, accountants reported lower levels of misrepresentation than those in sales and marketing jobs.

McDaniel et al.'s (1997) study may be criticized on the grounds that those who misrepresented themselves may have been less likely to respond to the mail survey for fear that their responses would be traced. However, if this is true then their results probably underestimate the true misrepresentation rate of the sample. This study may also be criticized because the conditions were not balanced; thus, order effects may be present.

In a study discussed earlier, Hurtz and Bohon (1997) instructed subjects to answer

a personality based integrity test as they would if they were applicants for a retail sales position. The mean score of those in this applicant condition was over .5 standard deviations higher than those who took the test under instructions to answer honestly.

In summary, these results suggest that at least some applicants do fake their responses on non-cognitive measures. There is also evidence that applicants are generally successful at increasing their score by as much as 1 standard deviation by doing so (e.g., Dunnette et al., 1962; Stokes et al., 1993). We also find that the way in which a biodata inventory is keyed may determine the ease with which applicants can increase their scores through faking (Kluger et al., 1991). Most of the research has been conducted on biodata inventories; more research needs to be done that examines if applicants fake personality and integrity tests.

Evidence Applicants Can Not and Do Not Fake

Few studies could be located which indicated that people could not fake to increase scores on a non-cognitive measure. However, Cunningham et al. (1994) conducted a study that indicated students asked to fake good on an integrity test may have actually decreased their scores. The authors suggested that their findings may have differed from those of Ryan and Sackett (1987) because unlike Ryan and Sackett their study used the Reid Report which is an empirically validated test. Such tests may be more resistant to faking than ones that have not been validated. However, Cunningham et al. point out that the students who were told to fake scored lower as a group than applicants taking the same test. It may be that the students were not successful at increasing scores through faking because they were not motivated enough. However,

applicants were motivated to fake as their answers influenced whether or not they got a job.

In a second study, Cunningham et al. (1994) provided some subjects with information about Reid Report scales. For example, the authors gave some subjects information about the punitive scale (what it was trying to measure and how it measured it), some about the projective scale, and some about both dimensions of the test while another group was given no information about any test dimensions. In an attempt to motivate subjects, they were told that those scoring within the top five percent would receive ten dollars. Those who scored lowest on the test were subjects who received no information on it. Those given information on a dimension of the test had significantly higher mean scores on that dimension as compared to those who did not receive this information. The groups receiving information on both dimensions had the highest mean total scores (.5 standard deviations greater than the control group). However, there were some generalized effects on the non-targeted dimensions such that information on one dimension increased the mean score on the other dimension slightly. These results suggest that applicants can improve scores through faking *if* they are provided with information about the test. These findings may be exaggerated given that students served as subjects. Real applicants may give more thought (due to greater motivation) to what the test is trying to measure and not need to get information on the nature of the test because they would be motivated to figure it out on their own. However, such information may have helped students who have less experience with selection tests and who are less motivated to increase scores through faking. Nonetheless, this study

indicates that the ability to figure out what the test is measuring may be a good predictor of faking. This implies that instruments or items that are more transparent (i.e., it is clear what the instrument or item is trying to measure) may be easier to fake.

Hough et al. (1990) administered a temperament inventory to military incumbents that included a social desirability scale to detect intentional distortion in a favorable direction. All respondents were told to answer the inventory honestly but some were later asked to answer so as to look good (fake condition). The inventory was also administered to 125 individuals going through the Military Entrance Processing Station (MEPS) who were told that the test would be used to make decisions about their careers. Results showed that incumbents were able to increase scores by faking when told to do so; however, the real applicants did not fake. The social desirability scores of the MEPS were similar to those of the incumbents in the honest condition.

One must keep in mind that the MEPS are not traditional applicants. That is, these are individuals who were enlisted in the military and they were told that this test may influence what jobs they would be doing in their military careers. It is possible that these individuals did not want the same types of jobs. Therefore, they may not have been faking in the same direction. The incumbents were asked to fake to look as good as possible so as to be selected by the army. The MEPS were already selected by the army. Therefore, they may not have been responding in the same manner as the incumbents who were asked to fake good, even if they were faking good, because they may have been targeting a specific job (i.e., they are faking for a classification rather than a selection goal). An applicant who fakes responses so as to look like the best candidate for a

position as a cook may have very different responses from someone who is faking because he desires to be an infantryman. Therefore, the MEPS may have been responding so as to appear desirable for a particular job while the incumbents were faking to appear desirable to the military in general.

With the exception of the study conducted by Cunningham et al. (1994), there is much evidence to suggest that respondents *can* fake non-cognitive measures when instructed to do so. These studies have demonstrated that applicants who are asked to fake may increase their scores through faking by as much as 1 standard deviation. Therefore, in the current study it is expected that subjects' mean scores will be significantly greater in the condition in which they are instructed to fake when compared to mean scores in the honest condition.

Hypothesis 1: All three tests (personality test, biodata inventory, and integrity test) will be fakable such that, for each test, mean scores in the faking condition will be significantly greater than in the honest condition.

Potential Impact of Faking on Non-Cognitive Measures

Most of the evidence leads to the conclusion that applicants can and do fake their responses on non-cognitive measures. Thus, a number of researchers have discussed how respondent faking may bias results. First, faking may be correlated with both the predictor and criterion, which may cause a spurious and misleading correlation between these variables (Nederhof, 1985; Ones et al., 1996). In such an instance there is an observed correlation between the two that may be due to their shared variance in faking and not due to shared variance in the constructs that the measures purport to tap. Second,

faking may hide relationships between variables by acting as a suppressor variable. In this case faking is uncorrelated with the criterion but correlated with the predictor (Nederhof, 1985; Ones et al., 1996). That is, a true correlation between a predictor and criterion may go undetected because of contamination by faking on the predictor. Partialling out faking would reveal the true relationship between the two variables. Third, faking may be a moderator. This can occur even if faking is not correlated with either the independent or dependent variables. However, there will be an interaction effect between the independent variable (the predictor) and faking (Ganster, Hennessey, & Luthans, 1983). The three ways in which faking may bias results all have the potential of clouding the relationship between the predictor and criterion and thus affecting the validity of non-cognitive selection tests.

Faking may also affect results by distorting scale reliabilities or changing factor structures. For example, faking may actually increases the reliability of scales because, when faking, people tend to respond in a consistently positive manner (e.g., Douglas et al., 1996). If people answer items honestly then their real inconsistencies may result in lower scale reliabilities. Others have suggested that faking may alter the factor structure of a measure such that a faking factor is created or that item loadings on factors change (e.g., Douglas et al., 1996; Schmit & Ryan, 1993).

Some authors take a broader perspective when considering faking's influence on tests. These authors suggest that faking can introduce bias, but this does not always occur. For example, Zerbe and Paulhus (1987) suggested that impression management or faking is only a problem when it is unrelated to the variable of interest. Therefore, faking

should not be controlled for when doing so may actually attenuate validity. Kluger et al. (1991) agreed, suggesting that the effect of faking on validity may depend upon the type of job performance being measured. Thus, if one wishes to determine applicants' experience with certain pieces of electrical equipment, faking may decrease validity (Pannone, 1984). However, if a test is designed to place applicants in a sales position, faking may serve to increase validity.

However, few researchers within this debate choose the middle ground and assume that faking is sometimes good and sometimes bad. Just as many people support the notion that faking should not be a concern when using non-cognitive measures as there are who suggest it ought to be. The literature representing both sides of this issue will be examined.

Evidence for the Influence of Faking on the Validity of Non-cognitive measures

In their study described earlier, Dunnette et al. (1962) found that faking exerted an obvious effect on the validity of their biodata checklist. Managerial ratings were provided for 45 of the salesmen who had answered the checklist under the honest and the fake conditions. Correlations between ratings and checklist scores under the honest condition were relatively large (reaching .38) and significant. However, when the ratings were correlated with scores of those in the fake condition, the correlations decreased (largest was -.25) and became non-significant. As described earlier, they determined whether all or only a few applicants tried to create a good impression by developing a fake key on the checklist and determined that only a few of the applicants actually faked. Therefore, the authors concluded that even the presence of a few fakers may decrease the

validity of such measures. Again, these results must be viewed with suspicion as all these analyses were conducted on a small sample ($\underline{n}$ = 64).

Pannone (1984) found that over one-third of applicants for an electrical job faked (at least one question) on a biodata measure. When fakers were excluded from the analysis the correlation between the biodata inventory and the criterion increased. Pannone concluded that faking introduced error variance that moderated the validity coefficients. However, one must view Pannone's results with caution. The criterion used by Pannone was a written test that was the selection test on which hiring decisions were based. The validity of this test was not provided. It is possible that the validity was very low, in which case Pannone's results are of little use.

Upon examination of a biodata inventory Stokes et al. (1993) found that, regardless of the keying procedure, there was very little overlap between keys developed on incumbent and applicant samples. In other words, very few items that were valid in the incumbent sample were valid in the applicant sample. One possibility is that applicant faking destroyed the validity of the keys because these keys were developed on incumbents who responded honestly.

Douglas et al. (1996) obtained job performance ratings from employers of subjects who worked and gave permission to contact their supervisor. Evidence that faking decayed the criterion-related validity of the scales (two conscientiousness and agreeableness scales) was found. The average validity of the scales was .26 for honest subjects and .04 for those who faked. Further analyses revealed results similar to those found by Dunnette et al. (1962). Douglas et al. randomly selected a few of the subjects

from the faking condition and added their data to the data of the 97 honest subjects so that the combined sample contained the data of 10 to 25 percent of the faking subjects. Through this procedure they found that as the percent of fakers in the sample increased, the number of fakers scoring within the top ten increased as well. For example, it was found that if 10% of the people faked, five of the top ten subjects were fakers, whereas, if 25% of the sample was faking, nine of the top ten subjects were fakers. The mean validity (for the personality test and the biodata inventory) decreased somewhat as the percentage of fakers in the sample increased. Moreover, it was found that as the number of fakers in the top 10 scorers increased, the mean validity of the measures dropped substantially for those who faked. For example, the mean validity of the full sample was about .20 when 25% of the sample was faking, but the validity for the fakers was near zero. They found that even a few fakers may actually distort results of non-cognitive measures.

A study by McDaniel et al. (1997) which was previously discussed had respondents rate how their supervisor would evaluate their overall job performance (on a four point scale) in addition to the measure of conscientiousness, social influence, and misrepresentation. Results indicated that the extent to which respondents misrepresented themselves was negatively correlated with self-reported job performance. They also found that the validities were higher in the honest condition for both the conscientiousness measure (.14 vs. .06) and the social influence measure (.23 vs. .19). A criticism of this study is that the two measures were validated against a self-reported performance measure. However, one would expect that if the self-reported performance

measure was not accurate it would be slanted in the positive direction. This would lead to an underestimation of the decline in validity. Therefore, these results may be conservative.

A study conducted by Graham, McDaniel, Douglas, and Snell (1997) concluded that biodata item attributes that predicted item validity for honest respondents were not the same as the item attributes that predicted item validity for the faking respondents. Two rationally keyed biodata inventories were administered to subjects. Half of the subjects were asked to fake good when filling out the biodata inventories while the other half were instructed to answer honestly. Performance ratings from the participants' supervisors were also obtained. The unit of analysis for this study was the item; thus, 162 biodata items were examined. Questions that assessed moralistic, sociability, and adjustment traits had significantly higher validities for the honest respondents (.22, -.22, and .19) as compared to the fakers (.02, .13, .01). Verifiable items were also more valid for honest respondents (.21) than for the fakers (-.04). On the other hand, items which assessed past behavior were significantly more valid for fakers (.15) than for the honest respondents (-.02). Overall, results indicated that predictors of item validity for honest respondents were almost always different from the predictors of item validity for the faking respondents, indicating that faking impacts the validity of biodata items. However, this study is limited in that it only examined rationally developed biodata inventories. Empirically validated biodata inventories may yield different results.

The evidence presented suggests that applicant faking can result in lower validities for biodata inventories. There is also evidence that the validity of both biodata

inventories and personality measures may be distorted even if only a few applicants fake. Because few studies have examined how faking influences integrity tests, no evidence suggesting that the validity of integrity tests are affected by applicant faking could be found.

Evidence that Faking Influences Factor Structure

Schmit and Ryan's (1993) findings may provide further evidence that applicant faking biases the results of non-cognitive measures. They administered the NEO-FFI to applicants and non-applicants (students). They found that factor intercorrelations were higher for the applicant sample when compared to intercorrelations in the student sample. Moreover, the five factor model fit the data from the student sample but the applicant data contained six factors. The first factor for the job applicant sample was "a large, work-related personality-characteristic dimension" which they labeled the "ideal-employee factor." This factor contained most of the items from the Conscientiousness scale, but also included items from the other four scales. One possible explanation for this extra factor is that applicants, unlike students, wish to present themselves as a good employee. It may be that applicant faking created this first factor.

A similar study was conducted by Cellar, Miller, Doverspike, and Klawsky (1996). They examined the factor structure of the NEO-PI based on a sample of 423 flight attendant trainees. They also found that the six factor model fit better than the five factor model. However, the sixth factor in their study did not appear to be an ideal employee factor. Instead, this factor appeared to be a method factor. The differences between the findings of Schmit and Ryan (1993) and Cellar et al. may be attributed to the

samples used. Schmit and Ryan used actual applicants while Cellar et al.'s sample consisted of trainees who had already passed the selection procedure to become a flight attendant. Therefore, it seems that the sample used by Cellar et al. may have been less motivated to do well or to fake. Cellar et al. also examined applicants for a different type of job. These two influences could have resulted in different factor loadings on the sixth factor.

Douglas et al. (1996) found faking distorted the construct validity of their personality and biodata scales as measured by multi-trait and multi-method analyses and factor analyses. When the data of those in the honest condition were examined using an exploratory factor analysis, four factors were generated. All four were construct factors. The data from the faking applicants generated four factors but two were method factors and two were construct factors. Thus, a measure based on the responses of honest respondents may have a different structure when respondents fake. This is similar to what Schmit and Ryan (1993) found with the five factor model of personality.

Frie, Griffith, Snell, McDaniel, and Douglas (1997) conducted a confirmatory factor analysis on the biodata inventory data collected by Douglas et al. (1996). They tested for differences between the honest and fake good groups for the biodata inventory, using a four factor model (dependable, hard working, pleasant, and tolerant). A test for measurement invariance was conducted to determine whether the model had the same number of latent variables, the same factor loadings, the same measurement errors, and the same factor intercorrelations for the two samples. The results indicated that the GFI for the honest group was .86 and .69 for the fake good group, indicating that the specified

model fit the honest group better. This provides further evidence of factor structure differences between honest responders and fakers. Results also indicated that the relationships among the latent variables were significantly different across the honest and fake conditions.

These studies provide evidence that applicant faking may distort the factor structure of biodata inventories and personality tests, making it difficult to interpret results from these measures. Thus, if applicants fake we cannot be sure what our test is really measuring. No studies that examined if faking has the same effect on integrity tests could be located.

Evidence That Faking Impacts Scale Reliabilities

Douglas et al. (1996) found that internal consistency reliabilities for their four scales were higher for the faking condition than for the honest condition. The authors explain these findings by suggesting that honest responding hurts the homogeneity of the items. Subjects who fake make more of an effort to report consistently positive responses. In other words, faking creates an artificial inflation of consistency. Honest applicants report inconsistent behavior because their behavior truly is inconsistent. If it is true, and alphas do change as a result of faking, then these measures are tapping both conscious distortion and the construct which they were designed to measure. This is a violation of the unidimensionality assumption of alpha. This implies that alpha should not be used as a measure of reliability in settings where some respondents may be faking their answers.

This was the only study found that examined the influence of faking on scale

reliabilities. Obviously, more research needs to be done in this area across all three types of non-cognitive measures. If faking does artificially increase scale reliabilities, we again face a problem in which faking is making the interpretation of scores on non-cognitive measures difficult.

Evidence Faking Does Not Bias Results

Evidence has accumulated, particularly in the past few years, that indicates faking does not bias results of non-cognitive measures. However, no published studies could be found that provide explicit support for the assertion that faking does not distort factor structures or the reliability of scales in non-cognitive measures. Because most of these studies examine faking's impact on validity, the following section will be organized according to the type of test examined. First, evidence suggesting faking does not affect the validity of biodata inventories will be discussed, followed by a discussion of similar findings regarding integrity and personality tests.

Biodata. As previously discussed, Becker and Colquitt (1992) instructed some incumbents to fake and some to respond honestly to a biodata inventory. Over 250 real applicants were also administered the biodata. To determine faking's impact on selection decisions Becker and Colquitt (1992) assumed that every applicant distorted his or her score. They then adjusted the score distribution by subtracting three points from everyone's original score (this was the mean difference between the honest and applicant scores). The hiring decision that was made on the original score (hire or not hired) would stand for 83.1% of the applicants when the new distribution of scores was considered. The authors concluded by saying that this finding indicated that faking did not

significantly affect the selection decision.

However, the researchers based this analysis on the assumption that all applicants faked the same amount. As will be discussed later, there are several reasons to believe that applicants may differ in the extent to which they distort their scores. If the extent of faking was not uniform across subjects in Becker and Colquitt's (1992) study then these results may be questionable. This point will be discussed in more detail later.

Integrity Tests. Ones, Viswesvaran, and Schmidt (1993) conducted a meta-analysis of integrity test validities based on 665 validity coefficients across 576,460 data points. They suggested that faking is not a problem in integrity tests because they found substantial validities for studies conducted on applicants ($\underline{r} > .40$). Although they did find large validities across situations, they did not directly examine if faking lowered or affected these validities as they did not take it into account.

Comments made by Sackett and Wanek (1996) suggested that the results of Ones et al. (1993) should not be taken at face value. They pointed out that a number of their validity estimates were more like reliability estimates. For example, Ones et al. correlated overt integrity test scores with self-reports of past theft. A number of items in overt integrity tests ask respondents if they had committed theft. Therefore, it seems that Ones et al. determined test-retest reliability rather than estimates of validity. Thus, we cannot be sure if faking decreases the validity of integrity tests because if respondents lie or fake on the integrity test they are likely doing the same on the self-report measure of theft. Sackett and Wanek also pointed out that integrity tests have consistently been found to correlate with various faking or lie scales, yet little research has examined how

these correlations may influence hit rates and false positives when using integrity tests.

Lilienfeld, Andrews, Stone-Romero, and Stone (1994) administered the Reid Report Inventory and a personality measure with an impression management scale (similar to the MMPI Lie scale) to 184 undergraduates. They found no evidence that impression management moderated relationships between the integrity test and other measures (personality scale measures). However, the participants in this study were told their responses would be completely anonymous. Also, there was no fake condition included in this study. Therefore, the participants had little or no motivation to distort their responses. Impression management may not have moderated relationships in this instance because the impression management scores may have been lower or had less variance than would be found in a typical sample of job applicants (who are presumably more motivated to distort their responses).

Cunningham et al. (1994) set out to determine if social desirability acts as a moderator or a suppressor. The Reid Report and the Balanced Inventory of Desirable Responding (BIDR) was administered to over 100 students. The BIDR was designed by Paulhus (1989) to measure the two components of socially desirable responding (self-deception and impression management). Subjects were asked to imagine that they were taking the integrity test as applicants for a sales position. After the students completed the test they individually went to collect the five dollars which they were to receive for their participation in the study. All of the subjects were over paid (given ten dollars). The situation was such that it was fairly clear that the over payment would not be detected so that they would "get away" with taking the extra money.

Twenty-two percent of the participants kept the money they knew they should not have received. The measure of self-deception was not significantly associated with returning the money. However, there was a relationship between impression management and returning the money such that those who were higher on impression management were more likely to return it. However, it was determined that social desirability (self-deception and impression management) did not attenuate the predictive validity of the Reid report and did not act as a suppressor variable.

Sackett and Wanek (1996) suggested that the results of Cunningham et al. (1994) should not be over generalized. They pointed out that there may be large differences between passive acceptance of overpayment and active participation in theft. Most organizations who use such measures seek to determine which applicants will actively steal.

Personality Tests. The effect of faking on personality tests has also been explored. McCrae and Costa (1983) administered the NEO inventory and two social desirability scales (the Marlowe-Crowne and the Lie scale from Form A of the Eysenck Personality Inventory) to 215 volunteers. Six months later, the subjects' spouses were asked to rate their spouse's personality using the same personality inventory completed by the subject. All correlations between the self and spouse ratings were significant and ranged from .25 to .61. Partialling social desirability out led to significant correlation increases between self and spouse ratings for only two traits (activity and positive emotions). For most traits, the correlation between the self and spouse ratings dropped when social desirability was partialled out. The authors concluded that these findings

support the argument that socially desirable responding does not decrease the validity of personality measures. However, a number of individuals have incorrectly interpreted the findings of McCrae and Costa by saying that their findings provide evidence that faking does not reduce the validity of personality measures. McCrae and Costa only examined socially desirable responding, not intentional faking. The subjects in their study were volunteers so they were most likely not motivated to fake their responses. Also, the two social desirability scales used were not designed to detect respondents' attempts to consciously distort answers. These measures tap both self-deception and impression management. Therefore, this study says little about how faking may distort the validity of personality tests.

Hough et al. (1990) developed a temperament inventory called "Assessment of Background and Life Experiences" or ABLE. This inventory consisted of ten content scales that measured six temperament constructs and a social desirability scale that was included to detect intentional distortion in a favorable direction. Hough et al. administered the ABLE to military incumbents who took the test under both honest and fake conditions. All of these respondents were told to answer questions honestly, but half also took the test under instructions to fake good and the other half were told to fake bad (possibly so as to not be drafted). The conditions were counterbalanced. Results indicated that the response validity scales did detect the different types of distortion. Moreover, intentional distortion, in a positive direction, did not attenuate the criterion-related validity of the inventory.

The method Hough et al. (1990) used to obtain these results is questionable. The

data of the incumbents was skewed so that meaningful moderated regressions could not be performed to determine the impact of faking on the inventory. To deal with this the authors took the mean social desirability scale score of those in the fake good experimental condition and used this to divide the concurrent validity sample into two groups. Those scoring at or above the mean of the fake group were labeled "overly desirable," while those scoring below this mean were labeled "accurate." They then calculated the criterion-related validities of the ABLE for each of these groups separately.

Hough et al. (1990) found that the criterion-related validity did not change substantially for those in the "overly desirable" group; however, the incidence of faking in this group still may not have been very high. As Dunnette et al. (1962) showed, applicants can fake but few actually do. Therefore, it is likely that even those within the top half were not fakers. Instead of splitting the sample in half Hough et al. should have also examined the top one-third or top ten percent of scorers on the social desirability scale within the incumbent sample. Perhaps if this was done the criterion related validity would have dropped substantially for the "overly desirable" group. It may also be that the individuals in the "overly desirable" group had true high scores on the measure.

Recently, Ellingson, Sackett, and Hough (1998) reanalyzed the Hough et al. (1990) data. Specifically, they determined the effect of correcting for social desirability on applicant rankings and construct validity. When they corrected scores for socially desirable responding they found that the corrected mean score was lower than the mean score in the fake condition but still greater than the mean score of the honest condition. Also, the standard deviation of the corrected scores was still substantially larger than in

the honest condition.

Additionally, correlations between honest and faked scores were relatively low, confirming that responding in a socially desirable manner differentially affected individuals' scores (i.e., there was variance in the extent to which people faked such that rank orders changed). They then determined the percentage of correct top-down hiring decisions that would be made using the fake responses and then using the corrected responses (the "true" hiring decision was based on the applicant's standing in the honest condition). Results indicated that an increase in the selection ratio corresponded to an increase in the proportion of correct selection decisions in the faking condition. Also, holding everything else constant, as the percentage of individuals faking increased, the proportion of correct selection decisions decreased. For example, given a 10% selection ratio, if 30% of applicants fake only 34% of those who would have been hired under the honest condition would be hired. Results are similar for the corrected scores. That is, when faking is corrected, the rank orders are not similar to those in the honest condition.

To determine if faking influences the construct validity of their measure, Ellingson et al. (1998) factor analyzed the faked scores, the corrected scores, and the honest scores for the ten personality scales. Results indicated a one factor solution for both the fake and corrected scores, but a two factor (more complex) solution for the honest scores. The authors conclude that faking dissolves the complex, multidimensionality of their measure. Thus faking results in scores which no longer reflect personality traits but instead reflect faking behavior.

Therefore, using the same data as Hough et al. (1990), Ellingson et al. (1998)

concluded that applicants can fake a personality measure, faking will change rank orders, faking may change the nature of the construct being measured, and the effects of faking are not easily corrected.

Christiansen, Goffin, Johnston, and Rothstein (1994) examined how the criterion-related validity of a personality measure would be affected if corrected for faking. They administered the 16PF and a response distortion scale to 495 incumbents in a large factory products organization who were participating in an assessment center. Subjects were told that responses to these measures would be used for future operational purposes (e.g., selection, promotion). Of these subjects, 84 were selected for an upper level supervisory position on the basis of the assessment center evaluation. Using the data of the 84 selected applicants for whom criterion data were available, the effect of correction on the criterion-related validities of the traits relevant to the job was examined. Using faking-corrected scores based on the Krug (1978) formula and through using hierarchical regression analyses it was determined that there was not a significant increase in the criterion variance explained by corrected scores beyond that explained by uncorrected scores. The authors concluded that this demonstrated that faking does not influence the criterion related validity of the personality measure.

Christiansen et al. (1994), using the data from all 495 subjects, also examined the effect of correction on selection decisions for different selection ratios. It was determined that, depending on the selection ratio, correction of scores would have resulted in different hiring decisions than those that would have been made on the basis of uncorrected scores. For example, with a 15% selection ratio 12 candidates who would

have been selected on the basis of uncorrected 16PF scores would not have been selected on the basis of faking corrected scores. Results indicated that at lower selection ratios there was a greater percentage of discrepant hires. To account for the results, the authors noted that scores on the social desirability scale were positively correlated ($\underline{r}$ = .22) with predicted performance. That is, those who scored high enough on faking to warrant correction also tended to have high predicted performance scores. Thus, when the correction was made their standing decreased.

The social desirability scale used in Christiansen et al.'s (1994) study was developed by Windor, O'Dell, and Karson (1975). It is not clear if this scale is purely a measure of conscious attempts to fake or if it is also measuring the self-deception component of social desirability. If it is measuring both then the results of Christiansen et al. must be questioned as the authors suggest that they are examining how *intentional* faking biases a personality measure.

Ones et al. (1996) meta-analyzed the social desirability literature to determine if social desirability operates as a predictor of some criteria, as a mediator, or as a suppressor when personality tests are used as a predictor. They also set out to determine if social desirability is related to real individual differences in personality. They did this by comparing scores on social desirability scales to self and other reports on a personality inventory. The authors found social desirability correlated .37 with emotional stability, .20 with conscientiousness, .14 with agreeableness, and .06 with extroversion. There was no correlation between openness to experience and social desirability. To determine if these correlations represent the susceptibility of these scales to social desirable

responding or if the correlations represent the relation of social desirability to substantive personality constructs, other people's (e.g., friends, relatives) ratings of that individual on the personality measures were examined. There were similar correlations between social desirability scale scores and others' ratings of personality. Thus, there is evidence that social desirability is related to real individual differences in personality.

Ones et al. (1996) also found that social desirability did not predict job performance. They calculated the operational validity of social desirability's prediction of five frequently used performance criteria (school success, task performance, counterproductive behaviors, training success, and supervisory ratings of job performance). Only training success was predicted by social desirability scale scores.

To be a mediator social desirability must correlate with both job performance and personality variables. Because it does not correlate with most performance criteria, social desirability does not act as a mediator. In addition, by partialling social desirability from personality measures, Ones et al. (1996) were able to investigate if such responding impacts the criterion-related validities of the Big Five variables. When this was done, the validities did not change; thus, it was concluded that social desirability did not attenuate the criterion-related validity of the personality measure. It was also determined that social desirability did not act as a suppressor.

The authors concluded that the concern with the susceptibility of personality tests to social desirable responding is needless. Their results indicated that social desirability measures true variance in personality and did not attenuate criterion-related validity. Therefore, controlling for such responding is not necessary. However, Ones et al. (1996)

may have overstated their findings because the study contained a major flaw.

The scales used by Ones et al. (1996) to measure socially desirable responding not only measure respondent's conscious attempts to present a positive image (impression management) but also measure the self-deception component of social desirability. For example, it has been shown that the Marlow-Crowne scale and the MMPI K Scale, which were used by Ones et al., load on both the self-deception and impression management factors (Paulhus, 1984; 1986). Therefore, these scales are not only measuring conscious attempts to distort, but also the unconscious component.

It is not surprising that measures of self-deception would correlate with personality constructs and not attenuate the criterion-related validity of the personality measures. After all, it has been suggested that such measures, if controlled for, might decrease the usefulness of self-reports (Zerbe & Paulhus, 1987). Therefore, it may very well be that the self-deception component of social desirability was the cause of the correlation between desirability scale scores and personality dimensions. It would be interesting to see if this relationship still existed if self-deception was partialled out of the social desirability scores.

Ones et al. (1996) *can* conclude that the aggregate social desirability scale scores (measuring both self-deception and impression management) used in their analyses correlated with personality dimensions and did not attenuate the validity of personality tests. However, they *cannot* conclude anything about the relationship between respondents' conscious attempts to distort their answers to present themselves positively and personality dimensions or validity. Therefore, Ones et al.'s suggestion that social

desirability should be of no concern to those using self-reports of personality are premature. Perhaps self-deception should be of no concern; however, there is still no conclusive evidence which indicates that applicant attempts to consciously distort responses does not bias personality test results.

Zickar, Rosse, and Levine (1996) have used a different approach to analyze the effect that faking has on non-cognitive measures. They conducted a Monte Carlo simulation using Item Response Theory (IRT) to model faking on a personality test. Results indicated that across different faking magnitudes (variance of faking between respondents) and faking prevalence rates (percentage of fakers in a validation sample), validities did not decrease a significant amount. Also, the mean observed scale difference between validation samples in which fakers were included and samples with no fakers was less than .25 standard deviations for all but the most extreme faking conditions. Thus, faking was not shown to moderate validity. However, their findings indicated that as the percentage of fakers in the sample increased, the percentage of fakers in the top end of the distribution increased as well. Such findings have implications for top-down selection procedures as this indicates that the fakers may be the ones selected.

Zickar (1997) sought to replicate the findings of Zickar et al. (1996) but based his analysis on a more complex and realistic model of faking. He examined a computer simulation of faking on the Work Orientation scale from the Army's Assessment of Background and Life Events (ABLE). In this simulation, Zickar considered three factors not considered by Zickar et al. First, he varied the variance of faking magnitude or the variance between subjects in the extent of faking. Second, the correlation between faking

magnitude and the latent trait was systematically manipulated. Finally, the percentage of items deemed fakable was varied. Results indicated that as the variance of faking increased, the validity decreased; however, this correlation was small and not significant. The percentage of items faked had a curvilinear effect on the validity correlation; that is, the most distortion occurred when 75% of the items were fakable and the least amount of distortion occurred when 25% and 100% of the items were fakable. The correlation between the squared percentage of items faked with observed validity was significant ($r = -.61$), while the correlation between faking magnitude and the latent trait and observed validity was not significant. Overall, validity was influenced the most when the variance in faking was large, there were negative correlations between faking and the latent trait, and 75% of the items were faked. Thus, it seems that in specific situations, faking may decrease the validity of non-cognitive measures.

Summary

Obviously we have reached no clear conclusion. A number of studies have found that individuals can fake non-cognitive measures (Becker & Colquitt, 1992; Dunnette et al., 1962; Douglas et al., 1996; Klein & Owens, 1965) and that applicants do fake (Cohen & Lefkowitz, 1974; Dunnette et al., 1962; Pannone, 1984; Stokes et al., 1993). However, Dunnette et al. also found that few applicants actually do fake, Cunningham et al. (1994) found faking may actually decrease individuals' scores on an integrity test, and Hough et al. (1990) found that applicants' scores are the same as the scores of honest respondents. There is also evidence to suggest that faking does bias results of non-cognitive measures. Dunnette et al., Pannone, Douglas et al., and Zickar (1997) found that faking does

decrease validity, while the findings of Schmit and Ryan (1993) and Douglas et al. suggest that faking may alter factor structure and scale reliabilities. Others refute these findings with their own evidence that faking does not attenuate the validity of non-cognitive measures (Becker & Colquitt, 1992; Cunningham et al., 1994; Hough et al., 1990; Ones et al., 1996; Zickar et al., 1996).

The debate has not yet been resolved. There are conflicting findings in the literature. However, there are a number of possible explanations for the discrepancies described here which will be addressed in this study. These include the operational definition of faking, consideration of job desirable responding, and variance in faking.

First, the measurement or operational definition of faking may account for these discrepancies. Some studies, such as the one conducted by Ones et al. (1996), based their analysis on scales which tap both self-deception and impression management. Other studies use mean differences between groups to determine who is faking, while others use one item (Pannone, 1984) to detect fakers. The inclusion of self-deception in a number of scales designed to measure faking may be clouding the true relationship between intentional attempts to fake good and non-cognitive self-report measures. The current study will ensure that only intentional faking is being measured as a within-subjects design will be used whereby subjects will respond to three non-cognitive measures honestly and then under a fake condition. Paulhus (1984) suggested that unconscious self-deception is not affected by laboratory manipulations. Therefore, differences between scores in each condition can be attributed to applicant attempts to *consciously* distort responses in a positive direction. It is understood that self-deception will still be

measured; however, because this is not influenced by lab manipulations and a within subjects design is being used, it should be constant across the two conditions.

Second, the majority of studies did not examine job desirability separately from social desirability. Social desirability is the intentional distortion of responses so as to present oneself as possessing more of the qualities deemed desirable in general, whereas job desirability refers to the intentional distortion of responses to present oneself as possessing more of the qualities deemed desirable for a specific job for which the applicant is applying (Kluger and Colella, 1993).

There are times at which social desirability and job desirability will be the same (elicit the same response). However, the two often diverge. For example, the socially desirable response to a personality test item stating "I do not mind seeing someone suffer" would most likely be "disagree." However, suppose someone was taking this personality test so that they can get a job as a nurse. A desirable characteristic for someone in such a position is that they do not mind seeing someone suffer, as nurses often have to deal with suffering patients and must be able to deal with this. Therefore, the most job desirable response to this item would be "agree."

Leary and Kowalski (1990) suggest that when individuals know they are being judged or are under intense scrutiny (as they are in a selection situation), they consciously search for cues regarding others' impressions and attend selectively to information that is relevant to making the right impressions. Therefore, individuals who are taking a selection test might try to tailor responses so as to appear most desirable for the specific job for which they are applying instead of responding in such a way as to appear as a

generally good person.

Studies examining the differences between job desirability and social desirability have been conducted. As discussed earlier, Schrader and Osburn (1977) had 148 volunteers take two empirically derived biodata inventories first under instructions to fake and two weeks later responding honestly. One of the biodata inventories was keyed for a sales job, while the other was keyed according to a managerial job. In the fake condition, half of the subjects were told to fake for a position as a salesman and were given information as to what the job would require them to do. The other group was told to fake more in general as they were only told that they were applying for a position as a manager and not provided with information about the tasks they would actually have to perform in such a job.

Subjects in both conditions increased their scores on both biodata inventories. Although it was expected that subjects who were given information for a particular job would increase their scores more for the corresponding biodata key, this did not happen. Those told they should imagine they were applying for a sales job did no better on the salesman keyed biodata inventory than those who were told they were applying for a managerial job and vice versa. These findings may be attributed to the fact that subjects in the sales job condition were provided unspecific information about the job (although it was more specific than those in the management condition) and the description was similar to what one would expect a manager to do. For example, those in the salesman condition were told that the job would require them to improve retail image and business practices through counseling dealers and that they would have to set plans and objectives

for each account. Thus, the failure to find expected results may be because the two jobs were very similar.

As previously discussed, Kluger et al. (1991) sought to determine whether item-keyed and option-keyed biodata scores would differ when subjects were given instructions to simulate applying for a non-specific job as compared to simulating that they were applying for a specific job (retail store manager) or answering the tests honestly. It was found that those given specific information about the job were able to inflate item-keyed biodata scores more than those who were only given general information.

In the study conducted by Becker and Colquitt (1992), items faked by applicants were more job relevant than those that were less job relevant. This finding is further supported by the work of Kluger and Colella (1993). They administered a 191 item biodata inventory to 429 job applicants for a nurse's assistant in a nursing home. Six graduate students rated the socially desirable direction of each item. Judges were first told to select the response that would be given by someone who was not very interested in the job but was very interested in being perceived as a good person. They were also asked to respond as someone wanting to maximize their chances of getting a nursing job in a geriatric nursing home. Judges also rated the transparency of each item (how obvious it was to select the most desirable answer). It was found that job desirability was a unique predictor, if not always the best predictor, of faking biases as indicated by the means and variances of biodata items. Kluger and Colella (1993) suggest that the superiority of job desirability over social desirability in predicting the bias suggests that

past studies may have somewhat underestimated the attenuating effects of faking because they measured only socially desirable responding.

With the exception of the findings of Schrader and Osburn (1977), the evidence seems to indicate that when applicants fake toward a specific job they are more successful in increasing their scores on a biodata inventory than those who fake in general (Becker & Colquitt, 1992; Bridgman & Hollenbeck, 1961; Kluger & Colella, 1993; Longstaff & Jurgensen, 1953). Therefore, the discrepancies in the literature that have examined faking and biodata inventories may be explained by the fact that few studies actually examined job desirability. The current study will address this issue. In this study half of the subjects will be told to fake good in order to obtain a specific job while the other half will be given instructions to fake so as to appear as a good applicant for a job they would desire. No studies could be located that examined the relationship of job desirability to integrity and personality tests. Also, the biodata inventory examined in this study is the only test that was validated for a specific job. Because of this, the second hypothesis will relate only to the biodata inventory.

Hypothesis 2: Those instructed to fake who are given specific job information will have higher biodata scores than those instructed to fake without job specific information.

A third explanation of the discrepancies in the literature is that often, when determining if faking introduces bias, authors assume that there is no variance in faking between individuals. For example, Becker and Colquitt (1992) assumed that all subjects faked the same amount so they subtracted three points from everyone's original score. If there was faking variance in Becker and Colquitt's sample then their results are

essentially uninterpretable. As Lautenschlager (1994) stated "One could argue that even if response shifts occur, as long as the shift is relatively uniform across subjects (i.e., the rank order of respondents is preserved), then the faking of the item is inconsequential" (p.409). In other words, if the extent of faking is not uniform across subjects then rank orders will change and such a change may result in lowered validity. However, is it reasonable to assume that applicants taking a non-cognitive selection test will all fake the same amount?

The Douglas et al. Model

As discussed earlier, Douglas et al. (1996) provides a model of faking that suggests faking is not uniform across subjects. Specifically, they suggest personal characteristics, situational characteristics, and the opportunity to fake all have a direct effect on the extent of faking. The following section will discuss literature that supports this model. The present study is not testing the entire Douglas et al. model, but is using it instead as an organizing framework.

Situational Characteristics. Sometimes the situation in which individuals are placed may determine the extent to which they fake responses to a non-cognitive measure. It has been suggested that people are more likely to present themselves positively when the expected benefits of doing so increase (e.g., getting a desirable job; Schlenker 1980). It follows that individuals who have more to gain from faking good on a selection test (if they believe they can successfully increase their score by doing so) and who see the job as very desirable or feel they have little to lose by attempting to fake, will be more likely to fake. Thus, a person who was just laid off and desperately needs a job

may fake more on a selection test than someone who already has a secure job.

Another situational characteristic that may influence the extent of faking is the possibility of getting caught. Applicants who are told that responses will be verified will be less likely to fake than those who are not. Also, those told that a lie scale may be included in the measure that can detect faking will be less likely to fake responses (Doll, 1971).

Opportunity to Fake. A number of factors may allow individuals more or less opportunity to fake selection tests. As stated earlier, those whose true score is already high have less opportunity to improve their scores through faking. However, those with low true scores are able to increase their scores a great deal if they have the ability to do so.

Those with more information about the construct the test is trying to measure will have a greater opportunity to fake. Cunningham et al. (1994) and Dwight and Alliger (1997) both showed that individuals who had greater knowledge of what the test was trying to measure and how it was measuring it were able to increase their scores more through faking when compared to those not given such information.

Personal Characteristics. There are a number of personal characteristics that may relate to differences in faking between individuals. Schlenker and Weigold (1992) suggested that an individual's self-efficacy may play a role in ability to fake. They point out that people who expect to perform poorly are less successful at attempts to manage their impressions. However, individuals who believe that they can perform a specific task well are generally more adept at managing the impressions they convey to others.

Therefore, it follows that those who believe they can successfully fake a biodata inventory, for example, may be more capable of doing so.

Other personal characteristics which may contribute to faking variance include personality traits (e.g., conscientiousness, neuroticism), integrity, and self-monitoring. Ones et al. (1996) found that social desirability correlated with conscientiousness (r = .15) and neuroticism (r = -.37). Faking as measured by the Eysenck Social Desirability Lie scale has also been shown to be negatively correlated with measures of neuroticism (r = -.37; Cunningham et al., 1994). It is not clear if the Eysenck scale measures self-deception, impression management, or both. However, it may be that the ability to fake can be predicted by scores on measures of both neuroticism and conscientiousness.

Cunningham et al. (1994) found that there was a high positive correlation (r = .55) between integrity test scores and impression management (a measure specifically designed to tap conscious deception). Lilienfeld et al. (1994) administered the Reid Report Inventory and an impression management scale (similar to the MMPI Lie scale which is a measure of conscious distortion) to 184 undergraduates. They found that impression management scores were significantly and positively correlated with scores on the integrity test (r = .23). Guastello and Rieke (1991) also found that impression management correlated positively with scores on an integrity test. None of these studies provided any incentive that may have motivated the participants to fake on these integrity tests. All of these studies examined the relationship between voluntary participation under instructions to be honest. Therefore, this positive correlation between integrity scores and impression management is difficult to explain, as it means that those who are

more honest fake more.

Self-monitoring is another individual characteristic that may be associated with variance in faking among individuals. There is reason to believe that, if instructed to do so, self-monitors would be able to increase their scores by faking to a greater extent than low self-monitors. Tobacyk, Driggers and Hourcade (1991) found evidence to suggest that self-monitors process information differently. Possibly because of this difference self-monitors tend to be better at controlling the impressions they make on others (Snyder, 1974). High self-monitors tend to be more sensitive to social cues and have a greater concern for social appropriateness.

The evidence seems to indicate that high self-monitors are more aware of their social surroundings, are more aware of what is socially appropriate, and are better at manipulating the impressions they make on others. All of this leads to the conclusion that high self-monitors may be more adept at faking to look good.

Several studies have been conducted that have investigated the relationship between socially desirable responding and self-monitoring. Snyder (1974) found that self-monitoring was not related to scores on the Marlow-Crowne, a measure of socially desirable responding. As discussed, this measure may tap self-deception; therefore, it is not clear whether or not Snyder's findings would hold if only intentional faking was examined.

Tunnell (1980) found that high self-monitors tended to rate themselves more highly than low self-monitors on items that were very desirable. However, Stewart and Carley (1984) found that high and low self-monitors did not differ in the extent of social

desirable responding. Their measure of socially desirable responding was determined by counting the number of positive adjectives respondents used in self descriptions. It is not clear whether this measure taps self-deception or impression management.

Obviously, the findings investigating the relationship between self-monitoring and faking are mixed. One must keep in mind that none of these studies asked respondents to fake. All of them administered a non-cognitive measure with a social desirability scale and a measure of self-monitoring to college students. The students were told to respond honestly.

Five personal characteristics that may influence the variance of faking will be examined: self-efficacy toward the ability to successfully (i.e., increase test score) fake each of the three non-cognitive measures, the personality constructs of conscientiousness and neuroticism, integrity, and self-monitoring. The following hypotheses will be examined:

Hypothesis 3: There will be a significant positive correlation between self-efficacy for successfully faking each test well and the extent of faking on that particular test.

Hypothesis 4: The extent to which individuals fake will be negatively related to neuroticism.

Hypothesis 5: There will be a significant positive correlation between self-monitoring and the extent to which individuals fake.

In addition to these hypotheses, the relationship between conscientiousness scores and the ability to fake will be examined. Studies have demonstrated that conscientiousness scores correlate with impression management scores. This study will

examine if there is a relationship between conscientiousness and the extent to which respondents can increase their scores (i.e., fake). Also, although studies have shown that integrity test scores under honest administrations correlate positively with scores on a measure of impression management, the relationship between integrity test scores and ability to increase scores through faking has not been investigated. I will examine how the integrity test score in the honest condition relates to the extent to which participants were able to increase their scores through faking. Because these issues have not been investigated, no hypotheses will be made but the issues will be explored.

The extent to which respondents fake across tests will also be examined. In other words, I will determine if those who successfully increase their scores on one type of test (e.g., integrity test) are generally successful in increasing their scores on the other two tests (biodata inventory and personality test).

Faking is often considered error; however, others view it as an individual difference variable. If respondents can successfully increase their scores through faking across tests (or generally fail to increase scores through faking across tests) this provides evidence that faking is an individual difference variable and may help to resolve some of the controversy surrounding faking. That is, this would demonstrate that some people are inherently better fakers than others. Because no published studies that have investigated this could be located, no predictions will be made but the issue will be explored.

One must keep in mind that just because individuals with certain characteristics have the ability to fake well does not mean that these are the individuals who actually do fake when taking such tests in applicant settings. However, the goal of this study is to

determine if certain individuals have the ability to fake better than others, to examine what characteristics may separate good fakers from the rest, and if this ability generalizes across the three non-cognitive tests.

METHOD

Sample

A sample of 224 undergraduates from a large university in the Midwest participated as subjects in this study. Of these, 161 (71.9%) were female and 56 (25%) were male. The sample consisted of 179 (79.9%) Whites, 14 (6.3%) Blacks, 15 (6.7%) Asians, and 9 (3.5%) indicating other. Seven participants did not indicate race or sex. The mean age was 20 with a standard deviation of 3.14.

Design

A 2 x 2 x 2 two-period mixed design was employed, with one within-subjects factor and two between-subjects factors. For the within-subjects factor, participants were asked to respond to three non-cognitive measures once honestly and once faking. Therefore, all subjects completed the three tests twice. Subjects were randomly assigned the order of presentation of the honest and fake instructions (one of the between-subjects factors). Two hundred and seventeen participants took the test under the honest condition. The second between-subjects factor was specificity of information. Under the general fake condition participants (n = 100) were asked to imagine that they were an applicant for a job that they desired and that they should fake responses on the tests to appear as the best applicant for the job they had in mind. Under the specific fake condition participants (n = 112) were instructed to fake to appear like the best applicant

for a specific job (they were provided a description of the job).

Table 1. Reliability of All Measures and Difference Scores

Measure	Alpha
Self-Monitoring	.69
Self-Efficacy toward Faking a Personality Test	.89
Self-Efficacy Toward Faking a Biodata Inventory	.94
Self-Efficacy Toward Faking an Integrity Test	.96
HONEST CONDITION	
Neuroticism	.88
Extroversion	.81
Openness	.79
Agreeableness	.73
Conscientiousness	.86
Biodata Inventory	.76
SPECIFIC FAKE CONDITION	
Neuroticism	.75
Extroversion	.69
Openness	.68
Agreeableness	.69
Conscientiousness	.80
Biodata Inventory	.62
GENERAL FAKE CONDITION	
Neuroticism	.87
Extroversion	.83
Openness	.63
Agreeableness	.76
Conscientiousness	.91
Biodata Inventory	.79
DIFFERENCE SCORES (across both fake conditions)	
Personality Test	
Neuroticism	.83
Extroversion	.75
Openness	.58
Agreeableness	.65
Conscientiousness	.81
Biodata Inventory	.66
Integrity Test	.77

Measures

Personality Test. The NEO-FFI (Costa & McCrae, 1989), a measure of the five factor model of personality, was used to assess personality. This measure is a short version of Costa and McCrae's (1985) NEO-PI. The five factors measured were neuroticism, extroversion, openness to experience, agreeableness, and conscientiousness. Each of the five factors were measured by 12 items with a likert response format ranging from 1 (strongly disagree) to 5 (strongly agree). See Table 1 for alpha coefficients for each of the five scales, across conditions.

Integrity Test. The London House PSI was used as the integrity measure. This is an overt integrity test. This test consists of two sections. One section measures attitudes toward theft while the other asks questions about beliefs of the frequency and extent of theft, perceived ease of theft, punitiveness toward theft, endorsement of common rationalization for theft, and assessments of one's own honesty. The typical alpha coefficient of this measure is .85 (Sackett et al., 1989).

Biodata Inventory. The biodata inventory that was used was designed and is used by a large organization with headquarters in the Midwest. This is a validated biodata inventory designed to select administrative and technical personnel. It is empirically keyed. See Table 1 for alpha coefficients.

Self-Efficacy. A measure of self-efficacy for the ability to fake was created for each of the three non-cognitive measures because no preexisting scales could be found in the literature. This scale was adapted from Brutus and Ryan (in press). For each of the three non-cognitive measures, a four item scale was used to measure self-efficacy for the

ability to increase one's score through faking. Each item had a likert-type response format ranging from 1 (strongly disagree) to 5 (strongly agree). The alpha coefficient of the self-efficacy for faking a personality test scale was .89; for the biodata form it was .94; and for the integrity test it was .96. See Appendix A for a reproduction of these scales.

Self-Monitoring. The shortened version of Snyder's (1974) Self-Monitoring Scale was used to assess self-monitoring. The scale contained 18 items with a Likert-type response format ranging from 1 (strongly disagree) to 5 (strongly agree). The reliability using coefficient alpha is .69. This scale is reproduced in Appendix B.

Demographic Questionnaire. A questionnaire assessing various subject demographics was also included and was used to assess such factors as race, sex, years in college, ACT and SAT scores, and GPA. See Appendix C for a reproduction of this questionnaire.

Procedure

Participants signed an informed consent form (see Appendix D). Participants took a personality test, an integrity test, and a biodata test under two experimental conditions. Half of the subjects were randomly assigned to take these tests under the honest condition first and the other half completed the tests under the fake good condition first. Participants who were administered the tests under the fake condition first were asked to respond to the self-monitoring and self-efficacy measures before this first administration. If participants received the honest condition first, these subjects were administered the self-efficacy measures after taking the test in the honest condition but

before taking the tests under the fake condition. This was so those who were administered the honest condition first would not be cued as to what they were going to be doing in the second administration of the tests. Subjects took these tests under the two conditions within the same testing session.

In the honest condition the participants were given the following instructions:

> Please answer the following questions on each of the three tests as honestly as possible. Your answers will remain completely confidential and anonymous. I have no way of connecting the test back to you. Your answers will be used for research purposes only, and will not be used to evaluate you in any way. The nature of the project requires that you answer the following questions as honestly as possible so please provide as accurate answers as you can.

Prior to the administration of the three non-cognitive measures in the faking condition, participants were asked to fill out the self-monitoring and self-efficacy measures. Before responding to the self-efficacy measures they were asked to look at three sample personality test items. After reviewing these sample items participants responded to the self-efficacy toward faking personality test measure. Participants then reviewed three sample biodata inventory items and then responded to the self-efficacy toward faking a biodata inventory measure. The same procedure was repeated with self-efficacy toward faking an integrity test. Participants sampled the items first because those who took the tests under the faking condition first would not have had a chance to see what kinds of items are on such tests. Therefore, unless they had an opportunity to review them they would have no information on which to base their self-efficacy toward faking.

In the faking condition participants were given one of two sets of instructions.

Half of the subjects were told to fake so as to look like a good applicant for a job they would desire. Those in this condition were given the following instructions:

> Please answer the following questions on each of the three tests in such a way as to make you look as good an applicant as possible for a job you would want. I am not interested in what your real answers for each question would be. Instead, for each item, select the response that you feel will make you look like the best applicant for a job you would like to have. Your answers will be kept completely confidential and anonymous.

The other half of the subjects were given a description of a job. This description provided very specific information such as what tasks one would have to perform and what skills one should have (e.g., interpersonal skills, communication skills) to perform the job effectively. A reproduction of this job description is in Appendix E. Subjects in the job specific faking condition were given the following instructions:

> When answering the questions on each of the three tests imagine that you are a job applicant for the job of an administrator as described by the job description you just read. Suppose this is a highly desirable position that you really want. Answer in the way that you think would most likely get you the position. I am not interested in what your real answers for each item would be. Instead, for each item select the response that you feel will make you look like the most favorable applicant for this job and will increase your chances of getting this job. Your answers will be kept completely confidential and anonymous.

In both fake conditions participants were told that their three test scores would be summed and compared to all of those participating in the experiment. Those who scored in the top 15% would receive $15.00 in a few weeks. This was done to motivate participants to fake good.

For a detailed description of the protocol that was followed see Appendix F.

RESULTS

Difference Scores

First, difference scores were calculated by taking the score obtained in the fake condition and subtracting the individual's score in the honest condition. This procedure was conducted for all participants and for all test measures and scales. Thus, each participant has seven difference scores (i.e., a difference score for neuroticism, extroversion, openness, agreeableness, extroversion, integrity, and for the biodata inventory). For all but the neuroticism difference score, positive scores indicate that a better score was obtained in the fake condition.

Difference scores are appropriate to use when one expects a participant by treatment interaction (Tisak & Smith, 1984). However, some have criticized the use of difference scores on the grounds that they are essentially meaningless, especially because they are generally unreliable (Edwards, 1993; Edwards & Harrison, 1993). However, others suggest that difference scores can obtain impressive reliability when the reliability of the tests are high, the correlation between them is small to moderate or negative, and the variance of each test is substantial (Rogosa, Brandt, & Zimowski, 1982). Most of the measures used in this study have high reliability. The correlation of each test across the two administrations is also moderate as will be seen shortly. There was also variance in the extent to which people faked the test. Most importantly, it was assumed that a treatment by participant interaction would be found as it was expected that some individuals would be able to increase their scores more than others in the fake condition. Therefore, the use of difference scores in this study is appropriate.

The reliability of the difference scores is listed in Table 1. The following formula was used to calculate the reliability of the difference scores: $r_{dd} = (\sigma_d^2 - \sigma_{ed}^2) \setminus \sigma_d^2$; where $\sigma_{ed}^2 = \sigma_h^2 (1 - r_{hh}) + \sigma_f^2 (1 - r_{ff})$; with h representing the measure in the honest condition, f for the fake condition, and σ_d^2 representing the variance of the difference score.

All difference scores have acceptable reliability with the exception of the personality scale of Openness which has a reliability coefficient of .58 and Agreeableness which has a reliability coefficient of .65. Thus, further analyses that include the difference scores for openness and agreeableness should be viewed with caution. Table 2 contains the descriptive statistics for the difference scores.

Descriptive Analyses

The reliability for each of the measures was examined before performing any analyses. Coefficient alpha was calculated for each a priori scale. The reliability of all scales for both the honest and the two fake conditions (i.e., general and specific) are presented in Table 1, as are the alphas for the self-efficacy and self-monitoring measures. For the integrity test, typical internal consistency estimates were provided by the publishers (they could not be calculated here because the publisher does not provide information on how to score the instrument). Therefore, for the integrity test, the reliability is only listed once and not for each condition. Almost all scales demonstrated adequate reliability (i.e., $\alpha \geq .70$).

Table 3 presents the means and standard deviations for each of the measures across the three conditions. Correlations between all of the test measures and demographics are presented in Table 4. There were moderate to substantial

intercorrelations between the different tests. For example, with the exception of openness, the five personality scales were correlated above .30 with each other. Furthermore, the biodata inventory and integrity measure correlated moderately with each other and with the personality measures (.17 to .44).

Of particular interest are the moderate to high correlations between difference scores. These correlations range from .16 to .61. Thus, it seems that those who increased their score through faking on one measure were generally successful at increasing their scores on the other measures.

A MANOVA was then performed to determine if there were order effects for relevant variables. A new variable was created whereby participants who took the test battery first under the honest condition followed by the fake condition were coded as 1 while those who took the tests under the fake condition first were coded as 2. This served as the independent variable. All scales and test scores in both the honest and fake conditions served as the dependent variables; thus, there were 14 dependent variables. Results indicated that order effects were not present as this analysis was not significant.

The equivalence of the two between subjects conditions (i.e., those selected to be in the faking toward a specific job manipulation vs. faking in general) was examined. This was done to ensure that differences between these groups were a result of the manipulation and not the result of demographic differences. Independent samples t-tests were conducted whereby the self-efficacy and demographic measures, and honest responses to all measures were compared across the two fake conditions.

Table 2. Descriptive Statistics for Difference Scores.

Test	n	Mean	S.D.
Difference Between Fake and Honest Responses Across Both Fake Conditions			
Neuroticism	212	-12.92	10.31
Extroversion	213	6.25	8.02
Openness	212	1.15	7.03
Agreeableness	211	6.35	7.26
Conscientiousness	209	11.09	7.51
Biodata Inventory	192	15.95	10.36
Integrity Test	200	41.35	26.63
Difference Between Fake and Honest Responses In Specific Fake Condition			
Neuroticism	102	-13.71	10.00
Extroversion	102	6.26	7.48
Openness	102	1.18	7.32
Agreeableness	102	6.92	7.09
Conscientiousness	102	11.71	7.19
Biodata Inventory	91	17.25	10.18
Integrity Test	98	40.88	26.22
Difference Between Fake and Honest Responses In the General Fake Condition			
Neuroticism	110	-12.20	10.58
Extroversion	111	6.23	8.52
Openness	110	1.12	6.77
Agreeableness	109	5.81	7.40
Conscientiousness	107	10.49	7.78
Biodata Inventory	95	15.01	10.53
Integrity Test	94	42.23	27.07

Note. All personality measures were assessed with a five point likert scale with high values indicating higher scores on that trait (high scores on Neuroticism indicate more neuroticism). Biodata inventory scores could range from –43 to 29 but scores in the honest condition ranged from –27 to 20 and scores in the fake condition ranged from –18 to 27. Integrity test scores ranged from 3 to 74 in the honest condition and from 3 to 99 in the fake condition. Difference scores were calculated by subtracting the honest score from the faking score.

Table 3. Descriptive Statistics For Measures and Non-Cognitive Tests in Honest and Fake Conditions.

Measure	n	Mean	S.D.
Self-Monitoring	216	3.13	.43
Self-Efficacy for Faking a Personality Test	216	3.06	.95
Self-Efficacy for Faking a Biodata Inventory	216	3.03	1.01
Self-Efficacy for Faking an Integrity Test	215	3.25	1.11
HONEST CONDITION			
Neuroticism	215	21.26	8.65
Extroversion	215	32.68	6.73
Openness	215	29.99	6.59
Agreeableness	214	32.17	6.10
Conscientiousness	212	33.08	6.77
Biodata Inventory	199	-.03	10.25
Integrity Test	202	24.56	19.69
GENERAL FAKE CONDITION			
Neuroticism	106	8.55	6.98
Extroversion	107	39.07	6.18
Openness	106	31.45	5.26
Agreeableness	106	39.00	5.71
Conscientiousness	106	43.93	4.92
Biodata Inventory	106	15.79	7.57
Integrity Test	94	67.77	28.32
SPECIFIC FAKE CONDITION			
Neuroticism	100	8.02	6.58
Extroversion	100	38.73	6.05
Openness	100	30.85	5.48
Agreeableness	100	38.10	6.20
Conscientiousness	100	44.39	5.79
Biodata Inventory	94	15.89	5.99
Integrity Test	98	65.51	25.43

Note. Self-monitoring, self-efficacy, and personality measures were all assessed with a five point likert scale with high values indicating higher scores on that trait. Biodata inventory scores could range from –43 to 29 but scores in the honest condition ranged from –27 to 20 and scores in the fake condition ranged from –18 to 27. Integrity test scores ranged from 3 to 74 in the honest condition and from 3 to 99 in the fake condition.

Table 4. Intercorrelations for All Scales and Test Measures.

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age													
2. Sex	-.09												
3. Race	.08	.10											
4. Self-Monitoring	.01	.12	.08										
5. Self-efficacy for Faking a Personality Test	.01	-.13	-.12	.05									
6. Self-Efficacy for Faking a Biodata Inventory	.11	-.06	.02	.09	**.72**								
7. Self-Efficacy for Faking an Integrity Test	.10	-.05	-.06	.07	**.67**	**.75**							
Raw Scale Scores (Honest Condition)													
8. Neuroticism	.00	.09	.02	-.09	**.15**	.03	.05						
9. Extroversion	-.06	**.15**	-.10	**.25**	**-.16**	-.06	-.08	**-.42**					
10. Openness	-.08	.03	-.04	.03	.03	.01	.04	.06	-.05				
11. Agreeableness	-.07	.14	-.07	-.09	-.11	-.09	-.05	**-.41**	**.43**	.03			
12. Conscientiousness	-.02	.09	-.04	-.08	-.13	-.06	-.11	**-.36**	**.33**	-.03	**.25**		
13. Biodata Inventory	.02	.02	-.13	.12	-.05	-.03	-.01	**-.36**	**.33**	.11	**.17**	**.44**	
14. Integrity Test	-.07	**.33**	.06	**-.29**	**-.14**	-.06	-.13	**-.30**	**.18**	-.02	**.37**	**.39**	**.25**

Table 4 (cont'd).

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13
Raw Scale Scores (Fake Condition)													
15. Neuroticism	.10	-.03	.05	-.06	-.09	-.10	-.09	**.14**	-.04	-.02	-.11	**-.18**	**-.22**
16. Extroversion	-.10	.02	**-.15**	**.21**	.12	.11	.13	.07	**.22**	.02	.05	.09	**.18**
17. Openness	-.03	-.04	.03	**.20**	.04	.06	.07	.08	.03	**.32**	.09	.06	.05
18. Agreeableness	**-.14**	.11	-.10	.06	.03	-.01	.01	-.09	**.17**	.00	**.27**	**.21**	.07
19. Conscientiousness	**-.19**	.07	-.11	.02	.02	.07	.06	-.04	.13	.04	.13	**.25**	.12
20. Biodata Inventory	-.11	-.00	**-.12**	**.19**	**.19**	**.24**	**.27**	**-.15**	.09	.13	.05	.14	**.32**
21. Integrity Test	-.10	.04	-.07	-.07	.04	.06	.08	-.13	.06	-.02	**.15**	**.26**	**.16**
Difference Scores													
22. Neuroticism	.06	-.09	.01	.04	**-.16**	-.08	-.09	**-.75**	**.33**	-.07	**.27**	**.19**	**.17**
23. Extroversion	-.03	-.11	-.03	-.06	**.21**	.13	**.15**	**.40**	**-.67**	.06	**-.32**	**-.20**	-.15
24. Openness	.05	-.05	.06	.12	-.00	.04	.03	.00	.08	**-.70**	.04	.07	-.07
25. Agreeableness	-.05	-.05	-.02	.13	.13	.07	.05	**.27**	**-.23**	-.02	**-.63**	-.04	-.10
26. Conscientiousness	-.11	-.03	-.04	.08	.13	.10	.13	**.29**	**-.21**	.07	-.14	**-.72**	**-.31**
27. Biodata Inventory	-.11	-.03	.06	-.01	**.18**	**.17**	**.19**	**.24**	**-.28**	-.03	**-.15**	**-.32**	**-.77**
28. Integrity Test	-.05	**-.20**	-.12	**.15**	.13	.09	**.18**	.10	-.07	.00	-.11	-.03	-.02

Table 4 (cont'd).

Test	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Raw Scale Scores (Fake Condition)															
15. Neuroticism	-.07														
16. Extroversion	-.05	**-.58**													
17. Openness	-.14	**-.34**	**.53**												
18. Agreeableness	.12	**-.64**	**.64**	**.39**											
19. Conscientiousness	.12	**-.73**	**.63**	**.34**	**.61**										
20. Biodata Inventory	.06	**-.52**	**.38**	**.30**	**.35**	**.42**									
21. Integrity Test	**.37**	**-.57**	**.33**	.10	**.49**	**.45**	**.40**								
Difference Scores															
22. Neuroticism	**.21**	**.55**	**-.44**	**-.30**	**-.35**	**-.45**	**-.20**	**-.26**							
23. Extroversion	**-.19**	**-.41**	**.58**	**.38**	**.34**	**.36**	**.22**	**.20**	**-.61**						

Table 4 (cont'd).

Test	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
24. Openness	.09	**-.24**	**.39**	**.45**	**.30**	**.23**	.10	.10	**-.16**	**.23**					
25. Agreeableness	**-.22**	**-.42**	**.47**	**.25**	**.58**	**.40**	**.24**	**.27**	**-.51**	**.54**	**.21**				
26. Conscientiousness	**-.27**	**-.36**	**.35**	**.19**	**.25**	**.48**	**.19**	.10	**-.49**	**.44**	.08	**.32**			
27. Biodata Inventory	**-.20**	**-.19**	.13	**.17**	**.21**	**.18**	**.36**	.13	**-.32**	**.32**	**.15**	**.30**	**.43**		
28. Integrity Test	**-.36**	**-.51**	**.37**	**.22**	**.40**	**.36**	**.37**	**.73**	**-.41**	**.33**	**.17**	**.42**	**.29**	**.28**	

Note. Numbers in **Bold** indicates a significant correlation of $p \leq .05$.

Table 5. Paired Samples t-tests Comparing Means Across the Honest and Fake Conditions.

Measure	HONEST CONDITION		FAKE CONDITION		t-value	p	d
	Mean	SD	Mean	SD			
Neuroticism	21.26	8.68	8.34	6.85	-18.26	.001	-1.66
Extroversion	32.66	6.71	38.91	6.09	11.37	.001	.98
Openness	30.02	6.62	31.17	5.31	2.38	.05	.19
Agreeableness	32.19	6.13	38.54	5.87	12.71	.001	1.06
Conscientiousness	33.09	6.80	44.17	5.35	21.35	.001	1.82
Biodata Inventory	-.15	10.18	15.80	6.97	21.33	.001	1.86
Integrity Test	24.78	19.67	66.13	26.75	21.96	.001	1.78

Note. Effect sizes were calculated by subtracting the mean in the honest condition from the mean in the fake condition.

No significant differences were found indicating that the two groups were equivalent in terms of self-efficacy toward faking the three tests, demographics and on the personality dimensions, integrity, and background (as measured by the biodata inventory).

Hypothesis Tests

An alpha of .05 was used for all statistical tests. Hypothesis 1 predicted that all three tests would be fakable such that mean test scores in the faking condition would be significantly greater than in the honest condition. To test this hypothesis, paired-samples t-tests were conducted. Table 5 provides the result of these analyses. For all seven scales it was found that mean scores in the faking condition were significantly more positive than in the honest condition. In fact, some measures show over a 1 standard deviation increase in test score from the honest to the fake condition (e.g., conscientiousness and integrity). This result is consistent with the faking literature which finds that individuals can increase their scores up to 1 standard deviation through faking non-cognitive measures. Thus, Hypothesis 1 was supported.

Hypothesis 2 predicted that those who were instructed to fake and were given specific job relevant information would have higher biodata scores than those instructed to fake without job specific information. This was tested with an independent samples t-test. Results indicted no significant differences between those instructed to fake toward a specific job versus those told to fake in general (t (106, 94) = -.10, N. S.). Thus, Hypothesis 2 was not supported.

A mean difference between the general versus specific fake conditions was only predicted for the biodata inventory because this measure was keyed for the type of job described in the specific fake condition. Therefore, the specific job information was not

expected to influence the other measures. However, the effect of the specific information was also examined for the other six measures. There were no significant differences between the mean test scores of those in the general and specific fake conditions for the other six measures (i.e., neuroticism, extroversion, openness, agreeableness, conscientiousness, and integrity). Given that there were no significant mean differences between the specific and general fake conditions for any of the measures, the two fake conditions were combined for the remaining analyses.

Hypothesis 3 predicted a significant positive correlation between self-efficacy for successfully faking each test and the extent of faking on that particular test. To test this hypothesis, the self-efficacy measures were correlated with the difference scores for the seven measures (i.e., five personality scales, biodata inventory, and integrity test). As can be seen in Table 4, self-efficacy toward faking a biodata inventory was significantly and positively correlated with the biodata difference score ($\underline{r}$ = .17). Self-efficacy toward faking an integrity test was positively correlated with the integrity test difference score ($\underline{r}$ = .18). Thus, those who had high self-efficacy toward faking these tests actually did increase their scores to a greater extent through faking on these tests. However, self-efficacy toward faking a personality test was significantly correlated with the difference scores for only two measures, extroversion ($\underline{r}$ = .21) and neuroticism ($\underline{r}$ = -.16). This indicates that those who had high self-efficacy for faking a personality test tended to increase their extroversion scores through faking more than those low on self-efficacy. Those who had high self-efficacy toward faking a personality test tended to decrease their scores through faking on the neuroticism measure more than those with low self-efficacy. Keep in mind that for the neuroticism scale, a lower score is better (i.e., less neurotic).

Therefore, those with high self-efficacy for faking were able to get a better score on this measure. Thus, Hypothesis 3 was generally supported.

Hypothesis 4, that neuroticism would be negatively related to the extent of faking, was not supported. Table 4 shows that neuroticism was positively and significantly correlated with the difference scores for extroversion ($\underline{r}$ = .40), agreeableness ($\underline{r}$ = .27), conscientiousness ($\underline{r}$ = .30), and biodata inventory ($\underline{r}$ = .24). However, neuroticism was not significantly related to the openness and integrity test difference scores. This indicates that results are opposite to the predicted direction. That is, those high on neuroticism (as indicated by their neuroticism score in the honest condition) tend to increase their scores more through faking than those low on this trait for most of these measures.

Limited support was found for Hypothesis 5, which predicted that self-monitoring would relate to the extent of faking. As Table 4 shows, self-monitoring was only correlated with the integrity test difference score ($\underline{r}$ = .15). Therefore, high self-monitors were able to increase their scores through faking on the honesty test but not on the other measures.

Exploratory Analyses

The relationship between conscientiousness and faking was examined. To do this, conscientiousness scores from the honest condition were correlated with six of the difference scores (conscientiousness could not be correlated with the conscientiousness difference score as the score itself is imbedded within the difference score). Table 4 shows the honest conscientiousness score is positively correlated with the neuroticism difference score ($\underline{r}$ = .19) and negatively correlated with the difference scores for

extroversion (r = -.20) and the biodata form (r = -.32). These results indicate that those who were more conscientiousness were less successful at increasing their scores through faking.

Next, the integrity test score in the honest condition was correlated with six difference scores to determine if the extent to which a person was honest could predict faking (integrity could not be correlated with the integrity test difference score as the score itself is imbedded within the difference score). Results are presented in Table 4. Responses to the integrity test in the honest condition were positively correlated with the difference score for neuroticism (r = .21), and negatively correlated with the difference scores for extroversion (r = -.19), agreeableness (r = -.22), conscientiousness (r = -.27), and the biodata inventory (r = -.20). Thus, those with higher integrity scores were less likely to increase their scores through faking on the other measures.

Finally, Table 4 also shows moderate to high significant intercorrelations between the difference scores for the seven measures. These correlations are as high as .61. Only one correlation is not significant in the expected direction (i.e., the correlation between the difference scores for conscientiousness and openness). Therefore, those who were successful at increasing their scores through faking on one measure tended to do so on the other measures. This lends credence to the idea that the ability to fake is an individual difference variable.

DISCUSSION

The purpose of this study was to examine if there is variance between individuals in the extent to which they fake three types of non-cognitive selection tests and if individuals consistently faked across the three measures (i.e., if individuals who increased

their scores on one test were able to increase sores on the other two through faking). The results of this study provides general support for the hypotheses: (a) for all three tests the mean scores in the fake condition were significantly greater than in the honest condition, (b) participants given job specific information did not have higher mean scores on the biodata form than those given general information, (c) self-efficacy for successfully faking each test was positively correlated with the extent of faking for the biodata form, integrity test, and two personality scales (extroversion and neuroticism), (d) neuroticism was positively related to the extent of faking, and (e) self-monitoring was positively correlated with the extent of faking for the integrity test only. Furthermore, a number of exploratory analyses revealed interesting results: (f) individuals who were more conscientious were *less* successful at increasing their scores through faking, (g) those with greater integrity scores were less likely to increase their scores through faking, and (h) the difference scores for the seven measures were moderately to highly correlated. Each of these findings will now be discussed in the following sections.

Fakability of Non-Cognitive Measures

Across all seven scales, mean scores were greater in the condition in which participants were asked to fake when compared to the honest condition. Thus, this study clearly demonstrates that the three non-cognitive measures were fakable; that is, participants were able to increase their scores on these measures through faking. These results are consistent with the literature which has found that individuals can fake such tests.

Job Information and Faking

Hypothesis 2 predicted that individuals who were given specific job relevant

information (information on the job for which the biodata inventory was keyed) would be able to increase their biodata scores more through faking than those who were not given such information. This hypothesis was not supported as results indicated no significant mean differences between those instructed to fake toward a specific job versus those told to fake in general.

There are several reasons why the job relevant information may not have resulted in increasing the extent of faking on the biodata form. First, it may be that the information provided was not specific enough. The job description given to participants described the job of an administrator. As Appendix E shows, the tasks listed in the job description include such duties as clerical work and planning and organization. It also indicates that in order to perform well in the job one has to have initiative and be able to follow through on projects. It can be argued that this description could be given for almost any job, not just for a position as an administrator. Therefore, although job information was provided to some participants, the information may not have been specific enough for them to increase their scores.

Second, it may be that, instead of basing their answers on the job descriptions participants used their stereotype of the job to answer. Some participants may have felt they had a clear idea of what such a job entailed and how a person who would fit such a job would answer test questions; thus ignoring the description given to them and responding in accordance with their own stereotype. However, people are likely to have different ideas about what administrators actually do. It is possible that jobs vary in strengths of stereotypes. Stereotypes for some jobs may be very similar across people while others vary. The job of an administrator is not very well defined. Usually several

tasks fall under this category. It is unlikely that everyone shared the same stereotype. Therefore, some participant perceptions would be more accurate than others. Thus, this information failed to raise the group mean scores.

Third, the biodata inventory used in this study was empirically keyed. Therefore, even if one has a good idea of what the job of an administrator entails, it may still be very difficult to fake such a measure as it is not rational in nature.

Future research in this area should give very specific descriptions of the job and follow the procedures in this study to see if specific job relevant information can influence faking. Additionally, future research should investigate whether the keying of a biodata form influences the extent to which individuals can fake the measure. Specifically, research should address whether or not job specific information enhances the ability to fake on rationally or empirically keyed biodata forms differentially, if at all.

Self-Efficacy and Faking

As predicted, self-efficacy for successfully faking a test was positively correlated with the extent of faking on the measure. This was true of all scales except for the three personality scales of openness, conscientiousness and agreeableness. Therefore, overall, those who thought they could increase their scores through faking these measures were more likely to increase their scores by faking when compared to those with low self-efficacy.

It should be noted that the three measures of self-efficacy were highly correlated. That is, those who had high self-efficacy for faking one measure tended to have high self-efficacy toward faking the other two measures. This further supports the assertion that faking is an ability, as those who can fake in one situation believe they can fake in others.

Future research should examine why this relationship exists. It may be that individuals have high self-efficacy toward faking because they know they can fake such tests. That is, they have taken these types of tests before and failed to increase their scores through faking (e.g., they did not score high enough to get selected for a job). Therefore, their responses to this measure indicates their knowledge of their ability to fake such tests. However, it may be that those who believe they can fake were able to fake well just because they were more sure of themselves. In this instance, the self-efficacy measure is not tapping true knowledge of ones ability, rather a belief in themselves. Future research may try to determine which one of these explanations is most plausible.

Neuroticism and Faking

It was predicted that neuroticism would be negatively related to faking. However, exactly the opposite was found. Neuroticism was positively correlated with the difference scores for extroversion, agreeableness, conscientiousness, and the biodata inventory. Therefore, those high on neuroticism tended to increase their scores more through faking than those low on this trait for most of the measures.

The literature would seem to indicate that those higher in neuroticism would be able to fake less well (Ones et al., 1996). This is probably because such individuals are less emotionally stable and therefore, are likely to be less accurate in their perceptions of what the most desirable response would be. However, the research which has found neurotics fake less well have examined the two components of social desirability. It may be that neurotics are more concerned about how others view them (since the very nature of neurotics is an over concern with matters) and because of this they have become more

adept at presenting the image they want others to see. Future research should investigate this relationship to determine if these results hold, and if so, why those who are more neurotic would be better at increasing their scores through faking.

Self-Monitoring and Faking

It was predicted that self-monitoring would relate to the extent of faking, with those high in self-monitoring faking to a greater extent. However, faking on the integrity test was the only measure that was positively correlated with self-monitoring.

Future research should seek to determine why there are such discrepancies in the literature where self-monitoring is involved. It may be that self-monitoring does not measure the *ability* to determine what is socially appropriate. Perhaps high self-monitors are just more *concerned* about social appropriateness. Therefore, when a situation is weak (i.e., there is no need to monitor ones behavior) low self-monitors will not monitor their behavior but high self-monitors will (because they are always concerned about social appropriateness). However, if a situation requires one to monitor (i.e., job interview/selection test) both high and low self-monitors can behave in socially appropriate ways. Therefore, when asked to fake, high and low self-monitors can fake to equal extents. Further investigations may reveal the true nature of self-monitoring and its relationship, if one exists, to faking.

Conscientiousness, Integrity, and Faking

It was found that conscientiousness was positively correlated with the extent of faking on the measure of neuroticism, but negatively correlated with faking on the extroversion and biodata form. Therefore, those who were more conscientious were less successful at increasing their scores through faking while those low in conscientiousness

were more successful at increasing their scores.

These findings are inconsistent with those of Ones et al. (1996) who found faking was positively correlated with conscientiousness. However, as stated earlier, it is not clear whether or not the measures analyzed in their meta-analysis were tapping self-deception. If so, this may explain the different findings between this study and Ones et al.'s. Future research should be conducted to shed light as to why the two studies lead to different conclusions.

In the present study it is not clear whether conscientious individuals could not fake because they did not have the ability to or whether they just chose not to despite the instructions. That is, those who are high on conscientiousness may believe faking is inappropriate even when asked to do so. Therefore, future research should be done to determine if those who are more conscientious actually lack the ability to fake non-cognitive measures or if they just choose not to fake.

Like conscientiousness, those who scored high on a measure of integrity were less able to increase their scores through faking on other measures. Specifically, those with low integrity were able to increase their scores on neuroticism, extroversion, agreeableness, conscientiousness, and the biodata inventory measures.

Again, it is not clear whether those with high integrity are unable to fake or if they just choose not to despite being told because they never feel comfortable not telling the truth. However, it may also be that those low on integrity, because they are frequently dishonest, have no trouble embellishing responses. It could also be the case that these individuals are more able to fake because they tell untruths more often. It seems likely that the more one lies, the better one gets at it. Thus, those who have a tendency to be

dishonest are better at increasing their scores through faking on non-cognitive measures.

Faking Across Measures

One of the most interesting findings are the moderate to high significant correlations between the difference scores for the seven measures. These correlations indicate that those who were successful at increasing their scores through faking on one measure were generally able to do so on the other measures. This supports the notion that the ability to fake is an individual difference variable. It also provides evidence that the ability to fake is not test specific.

However, the correlations between the difference scores are not extremely high. Therefore, the extent of faking on one measure will not perfectly predict the extent of faking on the other two. There are several reasons the correlations are not perfect. Of course, the most obvious is measurement error. However, there are other explanations. For example, differential faking may be a function of item transparency. Transparent items are items in which the correct answer is obvious. Research has shown that transparent items are easier to fake (Lautenschlager, 1994). It may be that the transparency of the items differ across the three tests. That is, one measure may be more or less transparent than the others. Such differences would influence the extent of faking on these tests. Perhaps some people are better at faking both transparent and non-transparent items. This may also result in differential faking across the measures, assuming the number of transparent items was not consistent across these measures. It is not clear how the non-cognitive measures used in this study differ in terms of transparency. However, one could have experts rate the items for all three tests for the extent to which they are transparent. Future research should examine how non-cognitive

measures differ in transparency and how this may affect faking on these measures.

A Revised Model of Faking

The Douglas et al. (1996) model of faking provided a good organizing framework for the current investigation. However, this model assumes that personal characteristics, opportunity, and situational characteristics all have a direct effect on faking. The current research suggests the Douglas et al. model is inadequate as a number of variables that have been shown to be relevant are not in the model. Their model excludes relevant mediators and possible interactions. Therefore, this paper presents a expanded model of faking. This revised model is presented in Figure 2. This model is based on the research that has been conducted to date in this area, and also considers attitude-behavior links. This model may serve as a guide for future research on faking. In the following section each part of this model will be discussed.

Influences on Beliefs Toward Faking. This model proposes that individual differences directly influence the beliefs one has toward faking. For example, the present study found that conscientiousness was negatively related to faking; thus those low in conscientiousness were able to increase their scores more through faking. It may be that those who are very conscientious believe it is wrong to fake, for any reason. Therefore, these individuals did not attempt to distort their responses to a great extent, despite being told. One can only speculate if the extent of faking between those high and low on this measure would be even more pronounced in a selection situation (where individuals know they probably should not fake). One might expect the differences to be even larger than those found here.

Another individual difference variable that may influence one's beliefs toward

distortion may be integrity. The current study found that those high on integrity were less successful at increasing their scores through faking. Research has indicated that individuals who are inherently honest and have high moral standards are less likely to attempt to control the images they present to others (Leary & Kowalski, 1990; Schlenker, 1980; Schlenker & Weigold, 1992). In other words, these individuals seem to be less comfortable presenting a false impression of themselves. Therefore, it follows that applicants who have a strong belief in honesty and high moral standards are likely to believe that distortion is wrong; thus, they will not attempt to fake.

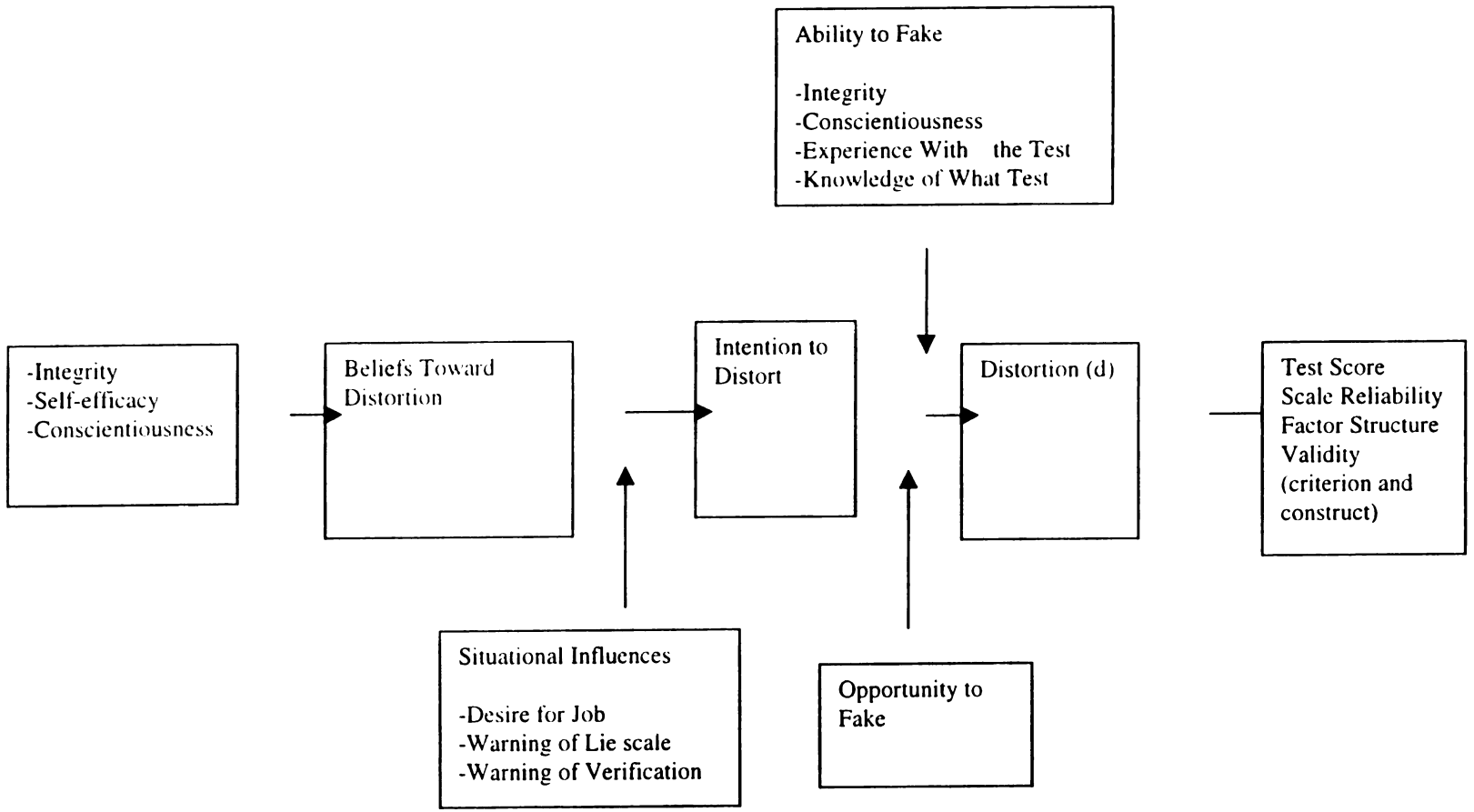

Figure 2. Model of Faking

Beliefs Toward Distortion and Intention to Distort. The model also proposes that those who believe distortion is acceptable are more likely to intend to fake their responses than those who believe such distortion is wrong. However, this effect is moderated by situational influences. It has been suggested that people are likely to manage the impressions they make on others when either the expected benefits of doing so increase or the expected costs of failing to do so increase (Schlenker, 1980).

It seems that if someone has a belief that faking is acceptable and believes that he or she will gain something from faking (i.e., increase their chances of getting the job), then they are likely to intend to distort their responses. However, even if someone holds a belief that faking is wrong, he or she may intend to fake if the situation is such that faking is seen as necessary. For example, a father of five who was just laid off from his job and has very few marketable skills is likely to be in great need of another job. Therefore, despite his belief that distortion is wrong, his need for the job may overcome this belief and he will distort his responses anyway in the hopes that this will get him the job.

Situations may also hinder faking. For example, even if individuals believe faking is acceptable, they may not intend to distort because it is believed that a lie scale is included in the test. For example, both Doll (1971) and Schrader and Osburn (1977) found that applicants told that there was a lie scale included in the non-cognitive measure scored lower than those who were not told this. Therefore, the concern with being caught faking will decrease the intention to fake, even when the belief toward faking is positive.

Intention to Distort and Distortion. The intention to distort has a direct effect on distortion. However, there are two potential moderators of this relationship: the ability to fake and the opportunity to fake. An individual may intend to distort and attempt to do so

but not have the resources available to successfully increase his score through distortion. That is, he may not have the ability to distort. For example, in the current study integrity was negatively related to extent of faking. Therefore, individuals who have high integrity tend not to increase their scores through faking as well as those low in integrity. It may be that individuals high in integrity do not have the *ability* to fake. The same may be true of those high in conscientiousness as these individuals also were less able to increase their scores through faking.

Experience with a test itself may lead to the ability to fake better. Klein and Owens (1965) have demonstrated that order effects exist in within-subject faking studies. Specifically, individuals who take the test under honest conditions first are able to increase their scores more in the fake condition than those who have not taken part in the honest administration. Therefore, exposure to a test may enable one to increase his score to a greater extent through faking. Cunningham et al. (1994) and Dwight and Alliger (1997) demonstrated that coaching individuals about the construct being measured increased individuals ability to increase their scores. Therefore, those who have knowledge of what the test is trying to measure may also help one fake.

Opportunity to fake refers to the extent to which individuals are able to increase their scores based on their true scores. As stated earlier, those whose true score is already high have less opportunity to improve their scores through faking. However, those with low true scores have the opportunity to increase their score a great deal through faking (assuming they have the intention to distort and the ability).

Effect of Distortion. This paper reviewed a number of studies that have shown how faking may distort test results. Although the vast majority of research has examined

the effect of faking on criterion-related validity, distortion has also been shown to influence test scores, scale reliabilities, factor structure, and construct validity (Cellar, Miller, Doverspike, & Klawsky, 1996; Douglas, McDaniel, & Snell, 1996; Dunnette, McCartney, Carlson, & Kirchner, 1962; Ellingson, Sackett, & Hough, 1998; Ryan & Sackett, 1987).

Alternatively, several studies have found that faking has little or no effect on test results (e.g., Christiansen et al., 1994; Cunningham et al., 1994; Hough et al., 1990; Ones et al., 1996). However, it is possible that these studies failed to find effects because they did not consider the rest of the model in their research. Specifically, they may fail to consider individual differences in faking.

Limitations

This study provides a number of important findings that extend our understanding of faking on non-cognitive measures. Although this research was carefully planned and thought out, a number of limitations still exist. First, instead of using "real" job applicants, college students served as participants. This may create doubt as to whether the findings here would generalize to applied settings. However, one must keep in mind that the goal of this study was to determine if there was variance in the extent to which individuals were able to fake non-cognitive measures, and if those who could fake one measure were also able to fake the other two. Therefore, there was no need to use applicants. The sample used adequately addressed these questions. However, the next step may be to conduct a similar study with an applicant sample.

Second, the job description used in this study may not have been specific enough, despite the fact that it was based on the actual job description used by the organization

that created the biodata form. Further studies should be conducted to determine if more specific job descriptions would lead to a greater ability to fake on the biodata form. Additionally, this study did not provide a manipulation check to ensure that participants were answering the biodata items based on the job description provided. They may have been relying on their own ideas of what a person who is successful in such a job is like. Thus, future research should include such a manipulation check.

Third, an empirically keyed biodata form was used in the current study. It is not clear if results would be different using a rationally keyed form. Therefore, further investigations should be conducted to determine possible differences in the extent of faking for the two types of keys.

Fourth, the measure of self-monitoring that was used in this study may not have been adequate. A number of studies have demonstrated that Snyder's (1974) self-monitoring scale may not be useful for measuring the construct it was intended to measure. This may explain why self-monitoring only related to faking on the integrity test. Perhaps a more well developed self-monitoring scale would reveal different relationships with faking.

Finally, the entire Douglas et al. (1996) model was not tested. Situational factors were held constant and only some personal characteristics that may relate to faking were examined. Future research should study the effects of faking on non-cognitive measures by testing the revised model.

Conclusions

This study has shown that the Douglas et al. (1996) model of faking has some merit for explaining variance in faking between individuals. The present research

suggests that there are individual differences in faking; thus, there is variance between individuals in the extent to which they fake. Additionally, individuals who fake well on one type of non-cognitive measure tend to be successful at increasing their scores on others. This further supports the notion of faking as an individual difference variable. Although the Douglas et al. (1996) model was a good start, the model introduced in this paper is more elaborate and includes a number of additional factors that may contribute to the extent of faking.

Understanding what contributes to the extent of faking is important to determine when we are likely to find effects due to distortion. Studies that have found faking does not have an effect on tests have failed to determine if individual differences in faking were present in that context. This study presents a model of faking that may be used as an organizing framework for future research in this area. By considering this model, future studies may be able to consider the relevant issues and influences on faking. Only then can we determine under what circumstances faking does matter.

APPENDICES

Appendix A

SELF-EFFICACY TOWARD THE ABILITY TO FAKE

Please read through the example personality test items below. These example items are similar to the actual items on the tests you are about to take (already taken and are about to take again). You need not answer the example items. Just read through the example items and answer all the reaction questions. Do not attempt to answer the example items, just read through them.

EXAMPLES OF THE TYPE OF ITEMS ON THE TEST

Example of Personality Test Items:

Due to the propriety nature of the NEO-FFI, the example items cannot be shown here. However five items will be given as examples.

Now that you have read the example personality test items please indicate the amount of experience you have had with such tests. Also, evaluate your ability to fake your responses on such items so as to obtain a high score on this measure. Evaluate how easy it would be for you to select the most desirable answer. That is, evaluate your ability to select the response that will likely give you the highest score on such a test by responding to the following items.

How many times have you taken this type of test?
1. Never
2. Once or twice
3. Three or four times
4. Many times

Use the scale below to answer items __ through __. Indicate the degree to which you agree with each of the following statements.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree

I am confident I can increase my score on the personality test through faking.

On the personality test I will be able to select the response that will result in increasing my score whether or not that response is true of me.

When taking the personality test I will be able to recognize which answer will result in increasing my score the most even if that response is not true of me.

If I fake, I am confident I can increase my score on the personality test.

Please read through the example biodata inventory items below. These example items are similar to the actual items on the tests you are about to take (already taken and are about to take again). You need not answer the example items. Just read through the example items and answer all the reaction questions. Do not attempt to answer the example items, just read through them.

EXAMPLES OF THE TYPE OF ITEMS ON THE TEST

Examples of Biodata Inventory Items:

Due to the proprietary nature of the biodata inventory sample items cannot be replicated here.

Now that you have read the example biodata inventory items please indicate the amount of experience you have had with such tests. Also, evaluate your ability to fake your responses on such items so as to obtain a high score on this measure. Evaluate how easy it would be for you to select the most desirable answer. That is, evaluate your ability to select the response that will likely give you the highest score on such a test by responding to the following items.

How many times have you taken this type of test?
1. Never
2. Once or twice
3. Three or four times
4. Many times

Use the scale below to answer items __ through __. Indicate the degree to which you agree with each of the following statements.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree

I am confident I can increase my score on the biodata inventory through faking.

On the biodata inventory I will be able to select the response that will result in increasing my score whether or not that response is true of me.

When taking the biodata inventory I will be able to recognize which answer will result in increasing my score the most even if that response is not true of me.

If I fake, I am certain I can increase my score on the biodata inventory.

Please read through the example integrity test items below. These example items are similar to the actual items on the tests you are about to take (already taken and are about to take again). You need not answer the example items. Just read through the example items and answer all the reaction questions. Do not attempt to answer the example items, just read through them.

EXAMPLES OF THE TYPE OF ITEMS ON THE TEST

Examples of Integrity Test Items:

Due to the proprietary nature of the London House PSI sample items cannot be replicated here.

Now that you have read the example integrity test items please indicate the amount of experience you have had with such tests. Also, evaluate your ability to fake your responses on such items so as to obtain a high score on this measure. Evaluate how easy it would be for you to select the most desirable answer. That is, evaluate your ability to select the response that will likely give you the highest score on such a test by responding to the following items.

How many times have you taken this type of test?
1. Never
2. Once or twice
3. Three or four times
4. Many times

Use the scale below to answer items __ through __. Indicate the degree to which you agree with each of the following statements.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree

I am confident I can increase my score on the integrity test through faking.

On the integrity test I will be able to select the response that will result in increasing my score whether or not that response is true of me.

When taking the integrity test I will be able to recognize which answer will result in increasing my score the most even if that response is not true of me.

If I fake, I am certain I can increase my score on the integrity test.

Appendix B

SELF-MONITORING MEASURE

Using the scale below, indicate the degree to which you agree with each of the following statements.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree

1. I find it hard to imitate the behavior of other people.
2. At parties and social gatherings, I do not attempt to do or say things that others will like.
3. I can only argue for ides which I already believe.
4. I can make impromptu speeches even on topics about which I have almost no information.
5. I guess I put on a show to impress or entertain others.
6. I would probably make a good actor.
7. In a group of people I am rarely the center of attention.
8. In different situations and with different people, I often act like very different persons.
9. I am not particularly good at making other people like me.
10. I'm not always the person I appear to be.
11. I would not change my opinions (or the way I do things) in order to please someone or win their favor.
12. I have considered being an entertainer.
13. I have never been good at games like charades or improvisational acting.
14. I have trouble changing my behavior to suit different people and different situations.
15. At a party I let others keep the jokes and stories going.
16. I feel a bit awkward in public and do not show up quite as well as I should.
17. I can look anyone in the eye and tell a lie with a straight face (if for a right end).
18. I may deceive people by being friendly when I really dislike them.

Appendix C

DEMOGRAPHIC QUESTIONNAIRE

Please circle or fill in the appropriate response:

PID _____

1. **Year**:
   a. Freshman
   b. Sophomore
   c. Junior
   d. Senior
   e. Other

2. **Age**: _____

3. **Sex**:
   a. Male
   b. Female

4. **GPA**: _____

5. **Race**:
   a. White
   b. African American
   c. Hispanic
   d. Asian
   e. Native American
   f. Other

6. **SAT Scores**: (if you can only recall your total SAT score just indicate that score)
   Math _____ Verbal _____ Total SAT score ____

Appendix D

Selection Test Battery

Purpose of Study

The following study has two parts. You will complete three measures typically used in selection processes, two times. The purpose is to obtain a better understanding of how individuals respond to such tests. The entire experiment will take no longer than two and a half hours (2.5 hours).

Informed Consent

The nature and general purpose of this study have been explained to me and the attached statement has been read to me by Lynn McFarland, from the Department of Psychology. I understand that my participation is voluntary and that all information is confidential and my identity will not be revealed. I am free to withdraw consent and discontinue participation in the project at any time. Any questions I have about the project will be answered by the researcher.

On the basis of the above statements, I agree to participate in this project.

_____ _____

Participant's Signature Date

_____

Participant's Name (Printed)

Appendix E

Job Description - Administrative Job

People in this administrative job must process paper work and handle clerical tasks (e.g., they must process invoices). They must also work out and solve problems with both internal and external customers. They must display initiative and follow-through even if it means being able to overcome obstacles. They must be able to manage their time and resources effectively. Those in an administrative position must also display good communication skills by knowing and understanding the audience they are working with and by being a good listener. They must work well with others and be able to set priorities for others and for themselves. They must be able to handle multiple projects at once and be able to manage their time. They must be able to meet all of their deadlines.

Appendix F

PROTOCOL

Hello, my name is Lynn McFarland and I will by your experimenter today. The purpose of this study is to see how individuals respond to some tests typically used to select applicants.

In this experiment you will take three selection tests, a personality test, a biodata inventory and an integrity test. You will complete each of these tests twice, given different instructions. You will also fill out a survey that will ask for some demographic information and a second survey that will ask for other information about yourself. This entire experiment should take no longer than two and a half hours (2.5 hours). You will receive one credit for your class per half hour. However, even if this experiment takes less than two and a half hours you will receive the full five credits if you complete all the parts.

[hand out consent form]

The consent form I am handing out states that you are participating in this study voluntarily and that all of your responses will remain completely confidential and anonymous. You can leave this experiment any time you want. However, you will only receive credit for the amount of time you are here.

[wait until all participants have signed the consent form and answer any questions they might have. Then collect the forms.]

[Hand out demographic questionnaire]

The questionnaire I am handing out asks you some demographic information.

Please completely fill out this questionnaire. Please use your mother's birth date as your identification number and place that number in the section of your questionnaire that is marked PID. For example, if your mother's birth date is June 13, 1950, you should record "041745' in the section of your questionnaire that is marked PID. If your mother was born on September 6, 1955, you should record 090655 in the section of your questionnaire that is marked PID. If you do not know or want to use your mother's birth date, then use some other number that you will remember because you will have to use this PID number throughout the experiment.

[wait until all participants have filled out the questionnaire]

For both faking conditions

[hand out scantron answer sheets]

I am now handing out an answer sheet. In the section of your answer sheet that is marked PID write the same PID that you wrote on the demographic questionnaire. Please check to ensure that the PID you put on this answer sheet is the same one you placed on the demographic questionnaire.

[hand out self-monitoring/self-efficacy questionnaire]

Please open your test booklet to page ___. Please record your answers to the statements on pages __ and __ next to the corresponding number on your answer sheet. Please respond to these items as honestly as possible. Keep in mind that your answers are completely confidential and anonymous. Stop when you answer all of these questions until you are told to turn the page.

[wait for everyone to finish the self-monitoring and self-efficacy items]

This test booklet contains three types of tests typically used in selecting applicants

(if they already took the tests under the honest condition they will be told that they will be taking the same three selection tests again, in the same order as before).

For general faking condition

Please answer the following questions on each of the three tests in such a way as to make you look as good an applicant as possible for a job you would want. I am not interested in what your real answers for each question would be. Instead, for each item, select the response that you feel will make you look like the best applicant for a job you would like to have. Your answers will be kept completely confidential and anonymous.

For the specific faking condition

[hand out the job description]

I am now handing out a description of a job as an administrator in a large organization. Please read this job description carefully. [wait two minutes to allow participants to read the job description].

When answering the questions on each of the three tests imagine that you are a job applicant for the job of an administrator as described by the job description you just read. Suppose this is a highly desirable position that you really want. Answer in the way that you think would most likely get you the position. I am not interested in what your real answers for each item would be. Instead, for each item select the response that you feel will make you look like the most favorable applicant for this job and will increase your chances of getting this job. Your answers will be kept completely confidential and anonymous.

For the honest condition

[hand out scantron answer sheets]

I am now handing out an answer sheet. In the section of your answer sheet that is marked PID write the same PID that you wrote on the demographic questionnaire (or other scantron if they had already taken the tests under the faking condition). Please check to ensure that the PID you put on this answer sheet is the same one you placed on the demographic questionnaire.

[hand out test booklets]

Please do not open the test booklet until I instruct you to do so.

[after tests are all handed out]

Please answer the following questions for each of the three tests as honestly as possible. Your answers will remain completely confidential and anonymous. I have no way of connecting the test back to you. Your answers will be used for research purposes only, and will not be used to evaluate you in any way. The nature of the project requires that you answer the following questions as honestly as possible so please provide as accurate answers as you can.

All conditions

Please do not rush through these tests. Those who finish early will not be allowed to leave. Everyone will leave at the same time. The first test in this booklet is the personality test. This test contains 60 items. Please stop when you have answered all of the personality items. Stop when you come to page ___. Do not turn this page until you are told to do so by the experimenter.

[when it seems like everyone is done]

Has everyone completed the personality test? [if everyone is done] O.K., please turn to page ___ and begin answering the biodata inventory items. Please stop when you

complete all of the biodata inventory items. Stop on page ___. Do not turn this page until you are told to do so by the experimenter.

[when it seems like everyone is done]

Has everyone completed the biodata inventory? [if everyone is done] O.K., please turn to page ___ and begin answering the integrity test items. Please stop when you complete all of the integrity test items and wait for the experimenters instructions.

Has everyone completed the integrity test items? [if everyone is done] O.K. I am now handing out a debriefing form which will describe the purpose of this study. [I will read the debriefing form to them as they read it themselves]. Are there any questions? If anyone has any questions regarding this study feel free to call me at the number listed on the debriefing form. Thank you for your participation.

Appendix G

DEBRIEFING FORM

Research has indicated that the survey measures you just took (e.g., biodata inventories and integrity tests) can be faked, although there is considerable debate over the extent to which people do fake or the effects of faking. Only a few studies have examined which individual difference variables correlate with the ability to fake such measures. It has been suggested that many people are not adept at faking. Before determining how faking may influence the results provided by such measures, we must first understand what constructs are related to the ability to fake. For example, some suggest that an individual's self-efficacy may play a role in the ability to fake. Others have found that emotionality is significantly negatively correlated with scores on a measure of social desirable responding, while conscientiousness has been shown to be positively related. It has also been suggested that those high in self-monitoring may be better fakers than those who are low self-monitors.

This study attempts to examine a number of individual differences that are potential correlates of the ability to fake non-cognitive measures. For example, it is predicted that self-efficacy toward the ability to fake (i.e., the belief that one can successfully fake such tests) will relate to the extent that a person can increase his or her score through faking.

The extent to which individuals fake across three non-cognitive tests (biodata inventory, personality test, and an integrity test) will also be examined. That is, I will examine if those who can increase their scores through faking on one type of non-cognitive measure (personality test) can increase their score through faking on another (integrity test). Faking is often considered error; however, others view it as an individual difference variable. If respondents can successfully increase their scores through faking across tests (or consistently fail to increase scores through faking across tests) this provides evidence that faking is an individual difference variable and may help to resolve some of the controversy surrounding faking. That is, this would demonstrate that some people are inherently better fakers than others.

If you have any questions regarding this study please feel free to contact me by phone or email (355-2171 or lynn.mcfarland@ssc.msu.edu).

REFERENCES

REFERENCES

Abrahams, N. M., Neumann, I. & Githens, W. H. (1971). Faking vocational interests: Simulated versus real life motivation. Personnel Psychology, 24, 5-12.

Becker, T. E. & Colquitt, A. L. (1992). Potential versus actual faking of a biodata form: An analysis along several dimensions of item type. Personnel Psychology, 45, 389-406.

Bridgman, C. A., & Hollenbeck, G. P. (1961). Effect of simulated applicant stus on Kuder Form D occupational interest scores. Journal of Applied Psychology, 45, 237-239.

Brutus, S. & Ryan, A. M. (in press). Ambiguity and self-efficacy as components of preferential treatment. Journal of Business and Psychology.

Cellar, D. F., Miller, M. L., Doverspike, D. D., & Klawsky, J. D. (1996). Comparison of factor structures and criterion-related validity coefficients for two measures of personality based on the five factor model. Journal of Applied Psychology, 81, 694-704.

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

Cohen, J. & Lefkowitz, J. (1974). Development of a biographical inventory blank to predict faking on personality tests. Journal of Applied Psychology, 59, 404-405.

Costa, P. T., & McCrae, R. R. (1985). The NEO Personality Inventory manual. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1989). The NEO PI/FFI manual supplement. Odessa, FL: Psychological Assessment Resources.

Cotton, J. W. (1989). Interpreting data from two-period crossover design (also termed the replicated 2 x 2 Latin square design). Psychological Bulletin, 106, 503-515.

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, 643-658.

Doll, R. E. (1971). Item susceptibility to attempted faking as related to item characteristic and adopted fake set. Journal of Psychology, 77, 9-16.

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

Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behavior on a forced-choice self-description checklist. Personnel Psychology, 15, 13-24.

Dwight, S. A. & Alliger, G. M. (1997). Using response latencies to identify overt integrity test dissimulation. Paper presented at the 12th annual conference of the Society for Industrial and Organizational Psychology, St. Louis, April, 1997.

Edwards, J. R. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. Academy of Management Review, 36, 1577-1613.

Edwards, J. R., & Harrison, R. V. (1993). Job demands and worker health: Three dimensional reexamination of the relationship between person-environment fit and strain. Journal of Applied Psychology, 78, 628-648.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (unpublished manuscript). Social desirability and correcting for social desirability: Issues of applicant comparison and personality trait construct validity.

Fletcher, C. (1990). The relationships between candidate personality, self-presentation strategies, and interviewer assessments in selection interviews: An empirical study. Human Relations, 43, 739-749.

Frei, R. L., Griffith, R. L., Snell, A. F., McDaniel, M. A., & Douglas, E. F. (1997). Faking of non-cognitive measures: Factor invariance using multiple groups LISREL. Paper presented at the 12th annual conference of the Society of Industrial and Organizational Psychology, St. Louis, April, 1997.

Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social desirability response effects: Three alternative models. Academy of Management Journal, 26, 321-331.

Goldstein, I. L. (1971). The application blank. Journal of Applied Psychology, 55, 491-492.

Graham, K. E., McDaniel, M. A., Douglas, E. F., & Snell, A. F. (1997). Biodata validity decay and score inflation with faking: Do item attributes explain variance across items? paper presented at the 12th Annual Conference of the Society of Industrial and Organizational Psychology, St. Louis, April.

Guastello, S. J., & Rieke, M. L. (1991). A review and critique of honest test research. Behavioral Sciences and the Law, 9, 501-523.

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.

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

Klein, S. P. & Owens, W. A. (1965). Faking of a scored life history blank as a function of criterion objectivity. Journal of Applied Psychology, 49, 452-454.

Kluger, A. N. & Colella, A. (1993). Beyond the mean bias: The effect of warning against faking on biodata item variances. Personnel Psychology, 46, 763-780.

Kluger. A. N, Reilly, R. R., & Russell, C. J. (1991). Faking biodata tests: Are option-keyed instruments more resistant? Journal of Applied Psychology, 76, 889-896.

Krug, S. E. (1978). Further evidence on the 16PF distortion scales. Journal of Personality Assessment, 42, 515-518.

Lautenschlager, G. J. (1994). Accuracy and faking of background data. In G.S. Stokes, M. D. Mumford, & W.A. Owens (Eds.), Biodata Handbook (pp. 391-419). Palo Alto, California: Consulting Psychologists Press.

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

Lilienfeld, S. O., Andrews, B. P., Stone-Romero, E. F., & Stone, D. (1994). The relations between a self-report honest test and personality measures in prison and college samples. Journal of Research in Personality, 28, 154-169.

Longstaff, H. P. & Jurgensen, C. E. (1953). Fakability of the Jurgensen Classification Inventory. Journal of Applied Psychology, 37, 86-89.

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

McDaniel, M. A., Douglas, E. F., and 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, April.

Miller, M. L. & Thayer, J. F. (1989). On the existence of discrete classes in personality: Is self-monitoring the correct joint to carve? Journal of Personality and Social Psychology, 57, 143-155.

Mumford, M. D. & Stokes, G. S. (1992). Developmental determinants of individuals action: Theory and practice in applying background measures. Handbook of I/O Psychology.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. European Journal of Social Psychology, 15, 263-280.

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

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, 78, 679-703.

Pandey, J., & Rastagi, R. (1979). Machiavellianism and ingratiation. Journal of Social Psychology, 108, 221-225.

Pannone, R. D. (1984). Predicting test performance: A content valid approach to screening applicants. Personnel Psychology, 37, 507-514.

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

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleiner & J. S. Wiggins (Eds.), Personality Assessment Via Questionnaire New York: Springer.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 92, 726-748.

Rothstein, R. H., Schmidt, F. L., Erwin, F. W., Owens, W. A., & Sparks, C. P. (1990). Biographical data in employment selection: Can validities be made generalizable? Journal of Applied Psychology, 75, 175-184.

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

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

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

Schlenker, B. R. (1980). Impression Management. Monterey, CA: Brooks/ Cole.

Schlenker, B. R. & Weigold, M. F. (1992). Interpersonal processes involving impression regulation and management. Annual Review of Psychology, 43, 133-168.

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

Schrader, A. D. & Osburn, H. G. (1977). Biodata faking: Effects of induced subtlety and position specificity. Personnel Psychology, 30, 395-404.

Snyder, M. (1974). Self-monitoring of expressive behavior. Journal of Personality and Social Psychology, 4, 526-537.

Stewart, A. & Carley, L. (1984). Personality characteristics of extreme scorers on the self-monitoring scale. Perceptual and Motor Skills, 58, 199-205.

Stokes, G. S., Hogan, J. B., & Snell, A. F. (1993). Comparability of incumbent and applicant samples for the development of biodata keys: The influence of social desirability. Personnel Psychology, 46, 739-762.

Thornton, G. C., & Gierasch, P. F. (1980). Fakability of an empirically derived selection instrument. Journal of Personality Assessment, 44, 48-51.

Tisak, J. & Smith, C. S. (1994). Defending and extending difference score methods. Journal of Management, 20, 675-682.

Tobacyk, J. J., Driggers, E. C., & Hourcade, J. (1991). Self-monitoring and psychological type: A social cognitive information processing model. Journal of Psychological Type, 22, 33-38.

Tunnell, G. (1980). Intraindividual consistency in personality assessment: The effect of self-monitoring. Journal of Personality, 48, 220-232.

Windor, P., O'Dell, J. W., & Karson, S. (1975). New motivational distortion scales for the 16PF. Journal of Personality Assessment, 39, 532-537.

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

264.

Zickar, M. (1997). Computer simulation of faking on a personality test. Paper presented at the 12th Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, April.

Zickar, M., Rosse, J., & Levin, R. (1996). Modeling faking on a personality test. Paper presented at the 11th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, April.