THE EFFECTS OF WARNINGS ON APPLICANT FAKING BEHAVIOR

Ruxandra Turcu
B.A., California State University, Sacramento, 2008

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF ARTS

in

PSYCHOLOGY
(Industrial/Organizational Psychology)

at

CALIFORNIA STATE UNIVERSITY, SACRAMENTO

SPRING
2011
THE EFFECTS OF WARNINGS ON APPLICANT FAKING BEHAVIOR

A Thesis

by

Ruxandra Turcu

Approved by:

__________________________________, Committee Chair
Gregory M. Hurtz, Ph.D.

__________________________________, Second Reader
Lawrence S. Meyers, Ph.D.

__________________________________, Third Reader
Jianjian Qin, Ph.D.

__________________________________
Date

ii
Student: Ruxandra Turcu

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

________________________________, Graduate Coordinator

Jianjian Qin, Ph.D.  Date

Department of Psychology
Abstract

of

THE EFFECTS OF WARNINGS ON APPLICANT FAKING BEHAVIOR

by

Ruxandra Turcu

The effectiveness of warnings at decreasing faking behavior on personality tests in the employment selection context along with test-taker reactions to those warnings were examined. CSUS undergrads were told to pretend they were job applicants, and were given written warnings along with the BIDR-IM and the NEO-FFI questionnaires, which were used to detect faking. Participants were most honest when they were first presented with a warning that informed them that items designed to detect faking exist in the personality test and if they are caught faking on the test, they will not be eligible for the prize offered for participation. Warnings which tap into test-takers’ subjective norms and beliefs about faking were also effective at reducing faking behavior, but to a lesser degree. Also, test-takers did not have either positive or negative reactions to the warnings. These findings support the use of warnings in the selection context.

_______________________, Committee Chair
Gregory M. Hurtz, Ph.D.

_______________________
Date
ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor and thesis chair, Dr. Gregory Hurtz and to my committee members, Dr. Lawrence Meyers and Dr. Jianjian Qin for their continuous support, guidance, motivation, mentorship, and knowledge throughout the writing of this thesis. I would also like to thank my family and friends for their support, patience, and understanding in any respect throughout the completion of this thesis.
TABLE OF CONTENTS

Acknowledgments........................................................................................................vi

List of Tables................................................................................................................x

Chapter

1. PERSONALITY TESTING IN THE EMPLOYMENT SELECTION PROCESS................................................................. 1
   Past Research................................................................................................................2
   Personality Assessment and Employee Selection.........................................................3
   Why Personality Tests Should not be Used in Selection.............................................6
   Why Personality Tests Should be Used in Selection..................................................8

2. PERSONALITY TESTS AND APPLICANT FAKING BEHAVIOR.............12
   Does Applicant Faking Occur?....................................................................................12
   Why do Applicants Fake?..........................................................................................15
   Faking Behavior Has a Negative Effect on Employee Selection..............................23
   Faking Behavior Has no Effect on the Selection Process.........................................27

3. THE USE OF WARNINGS AND APPLICANT FAKING BEHAVIOR........31
Appendix C  Control Groups’ Instructions and Experimental Groups’ Warnings ..83

References………………………………………………………………..………….86
LIST OF TABLES

1. Table 1 Correlation Matrix of All Measured Variables........................................58
2. Table 2 Means and Standard Deviations for the IM Scale and Personality Scale Variables across All Conditions .........................................................59
3. Table 3 Comparison of the Fake Good Control Group and the Honest Control Group.................................................................60
4. Table 4 Comparison of the Fake Good Control Group and the Warning Conditions.................................................................62
5. Table 5 Comparison of the Honest Control Group and the Warning Conditions.................................................................64
6. Table 6 Comparison of Warning Conditions.........................................................66
7. Table 7 Means, Standard Deviations, and Sample Size of Warning Conditions on Reactions to the Selection System Scale............................................68
The use of personality tests for making employment decisions has been and continues to be a controversial topic. One of the biggest reasons for this controversy is the fact that it is relatively easy for applicants to fake, or provide intentionally inaccurate responses in order to receive a more desirable score, on personality measures. Before discussing in depth the reasons for why faking has caused such controversy, and possible ways of reducing it, personality assessment in general must be discussed, as it relates to selection procedures. To start, personality refers to a person’s distinctive interpersonal characteristics and past behaviors, which are the best predictor of future behavior (Hogan, Hogan, & Roberts, 1996). People’s behaviors are a function of the kind of people they are, and when properly assessed, the patterns of behavior are consistent, and over time, people differ from one another in ways that are important to employers. It is for this reason that personality assessment was introduced to the employment selection process. Personality measurement is any procedure that systematically assigns numbers to the characteristic features of a person’s interpersonal style according to some explicit rules (Hogan et al.1996). Because the domains of personality and job performance are multifaceted, most performance criteria are best predicted by a combination of scales. Well-constructed measures of normal personality assess the characteristics that are
typically described as required for effective performance in a certain occupation, such as Conscientiousness (Hogan et al. 1996).

**Past Research**

For more than 25 years, the view of using personality tests in the personnel selection context was not endorsed by many employers due to a summary of 12 years of research that was published by Guion and Gottier in 1965. The authors concluded that, based on their summary, they do not advocate the use of personality tests in making employment decisions (Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). It was not until 1991 that it was concluded that personality measures should once again be used in selection contexts. This was due to the publication of two meta-analyses on the validity of personality tests for personnel selection by Barrick and Mount (1991) and by Tett, Jackson, and Rothstein (1991), which suggested that the corrected estimates of validity were actually meaningful, even though they were similar to those found in previous reviews (Morgeson et al. 2007). Since then, the use of personality testing in selection has been increasing. At the same time, the amount of research on the topic has grown substantially. However, this growing body of research also suggests changing opinions on the appropriateness of personality testing use. The bulk of the criticisms revolve around their criterion-related validity and their susceptibility to faking.
Personality Assessment and Employee Selection

In research by Rothstein and Goffin (2006), which reviewed and evaluated the main trends that have contributed to the increasing use of personality assessment in personnel selection, it was also stated that this growth stems from the 1991 meta-analytic studies. The authors explain that these meta-analytic studies provide a clearer understanding of the role of personality in job performance than did previous meta-analyses by examining the effects of personality on different criterion types and in different occupations. The authors also assert that although the meta-analyses have been conducted using the Five Factor Model of personality, there is more to personality than this model and the choice of personality measures to use in a selection context should consider a number of factors, such as the development of a predictive hypothesis on the relations expected between the personality measure and the performance criterion of interest.

Rothstein and Goffin (2006) state that personality-job performance relations are situation specific and reviews of research show that the assessment of applicant personality in a personnel selection context may provide organizations with predictive information on the likelihood that applicants may be involved in an accident, are likely to be satisfied with their job, will be motivated to perform, and will develop into leaders. This is because personality measures tap into many skills which are needed for specific jobs. For example, the authors explain that research by Barrick and Mount, (1993) found that Agreeableness and Openness to Experience are related to performance involving
interpersonal skills and that Conscientiousness and Extraversion predict managerial performance significantly better in jobs categorized as high in autonomy. Further, research by Judge and Ilies (2002) indicated that Neuroticism, Extraversion, and Conscientiousness correlated with performance motivation and research by McManus and Kelly (1999) found that the Five Factor Measure of personality provided incremental validity over biodata measures in predicting job performance and that personality data provided incremental validity over evaluations of managerial potential provided by an assessment center. Rothstein and Goffin (2006) also state that overall, Extraversion appears to be the best predictor of team-related behaviors and performance, with Conscientiousness and Emotional Stability also being good predictors. The authors conclude by holding that studies conducted in the 1990s repeatedly demonstrated that personality measures contribute to the prediction of job performance criteria and if used appropriately, may add value to personnel selection practices.

Further research has reported on the use of the Five Factor Model of personality regarding its use in selection procedures. Hurtz and Donovan (2000) have meta-analytically summarized the body of research that has developed in recent years where measures of the Five Factor Model were used as predictors of job performance. The authors suggested that perhaps, in previous research, the true predictive validity of personality was concealed by the lack of a common personality framework for organizing the traits being used as predictors. This is offered as an explanation for why the Five Factor Model of personality framework began to get adopted for selection procedures.
However, Hurtz and Donovan point out some flaws regarding the meta-analytic work by Barrick and Mount (1991) which led to the enthusiasm regarding the use of the Five Factor Model for selection purposes, and more specifically, its Conscientiousness dimension. The authors argue that these meta-analyses contain a potential threat to construct validity due to the methods the researchers used to derive their estimates of criterion-related validity. The authors explain that this is because the coefficients were based on studies that used measures that were not designed to measure the Five Factor Model personality measures.

In their examination of the Five Factor Model as it relates to personnel selection, Hurtz and Donovan’s findings indicate that the interpretation of the validity estimates for Conscientiousness provided by Mount and Barrick (1991) appear to be perhaps a bit too rosy. Despite lower validities than expected, Conscientiousness does appear to have the strongest relation to overall job performance. More specifically, those who describe themselves as hard-working, reliable, and organized appear to perform better than those who believe they are less strong in these characteristics. Emotional Stability also was found to have a small but stable influence on performance. More specifically, it appears that being calm, secure, well-adjusted, and low in anxiety has a consistent impact on job performance. Similarly, Agreeableness was found to have a small but consistent impact on job performance since being likable, cooperative, and good-natured is important in jobs that require interpersonal interactions. Also, Extraversion as well as Openness to Experience appears to have a small impact on job performance in customer service jobs.
Overall, the authors’ analyses suggest that the validities for the Big Five tend to be low to moderate in magnitude. They state that even the Conscientiousness dimension, which recently has been regarded as a valid predictor of job performance, is not as impressive as was expected. The authors explain that, with an estimated true validity of only .20, Conscientiousness adds a small portion of explained variance in job performance across jobs and across criterion dimensions which might potentially diminish if the relevant aspects of an applicant’s personality are already partially captured through other selection techniques. Lastly, the authors suggest that the magnitude of the Big Five correlations might be enhanced if the most relevant specific facets of these broad dimensions could be specified.

Why Personality Tests Should not be Used in Selection

Although numerous studies have been conducted on the use of personality measures as part of selection procedures, consensus on whether they should be or not be used in this context has not been reached. Morgeson et al. (2007) argue that there is very low validity of personality tests for predicting job performance, and therefore, they should not be used in this context. One of the main reasons for why the validity of personality tests is so low is because of applicant faking. This is because the “right” or most positive answer may be apparent to the candidates. The authors explain that the only thing preventing candidates from providing the positive answer when it is not true is their own honesty or lack of self insight, neither of which can be assumed when there is a desirable outcome at stake. The authors further state that faking casts doubt on the
soundness of personality inventories in selection situations where the test takers are often very motivated to do well on the tests. They also state that validity is likely to be greater when better personality tests are combined with cognitive tests. Although there is some potential for personality inventories to provide some incremental validity, since the inventories minimally overlap with measures of cognitive ability, they just do not tell us much about who will perform well or poorly on the job. Morgeson et al. conclude that common sense tells us that broad personality factors should be important, but the data suggests that they are not.

In a related article, Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007) question the usefulness of personality testing in selection procedures by asserting that the entire span of normal personality accounts for about 5% of the variance in job performance, and concluding that 95% of the variance in performance appears to have nothing to do with normal personality, as measured by currently available methods. The authors use this argument to conclude that personality should not be relied on too heavily in selection. They further explain that as low as the multiple correlations between personality and performance are, there are good reasons to believe that those figures and estimates of incremental validity are overestimates. This is because many meta analyses cited in papers rely on an unrealistically low estimate of the intercorrelations among the Big Five. Further, the authors assert that the actual observed, uncorrected values for validity cannot be ignored, and those values are even lower than the already low corrected values.
Why Personality Tests Should be Used in Selection

On the other side of the argument are those who believe that personality tests are useful and important tools for personnel selection. In research by Ones, Dilchert, Viswesvaran, and Judge (2007), meta-analytic research documenting the usefulness of personality variables for a broad spectrum of organizational attitudes, behavior, and outcomes was scrutinized and summarized. The authors found that the validities for the Five Factor Model of personality were on par with other frequently used assessment tools in personnel selection and assessment, such as situational judgment tests, assessment centers, structured interviews, and biodata. The researchers hold that if the amount of variance explained by personality is trivial, then those costly and time-consuming assessment tools do not fare better than traditional self-report measures of personality.

Ones et al. (2007) also found that Conscientiousness was the only personality trait that predicts overall job performance with similarly consistent validities across different jobs. For different occupations, different combinations of the Big Five yield the best levels of validity. Overall, the authors assert that the Big Five personality variables as a set predict important organizational behaviors, such as job performance and leadership. In support of their stance on the matter, they state that the criterion-related validities of self-report personality measures are substantial, with effect sizes ranging from moderate to strong. The authors conclude that the accumulated evidence supports the use of self-report personality scales in organizational decision making, including personnel selection. A reason given for this conclusion was that personality measures are useful in
understanding, explaining, and predicting significant work attitudes and organizational behaviors. Further, personality measures can be expected to provide incremental validity, and thus, utility in applied settings.

Research by Fox and Dinur (1988) also supports the idea of using personality testing in a personnel selection context. Fox and Dinur’s study evaluated the validity of self-assessment in a natural, military setting. Although the authors acknowledge that people are naturally motivated to present themselves in a favorable light and because of this, self-assessment suffers from enhancement or inflation bias, their findings are promising. The authors found self-rating validities to be low, but significant, for predicting success over a two-year training period. Self-ratings were also significantly related to evaluations by two groups who had previously been used as raters. The researchers’ findings favor the notion that individuals possess the capability to reliably evaluate themselves in a manner similar to that of others and in a way that can predict subsequent performance. This study has substantial strength because the benefits of self-ratings were demonstrated in an actual selection setting.

The study also attempted to evaluate the possibility that instructional manipulation could improve the validity of self-assessment. The researchers wanted to test this possibility since when giving self-reports, people are influenced by the desire to present themselves accurately and the desire to present themselves favorably. The researchers tested this by directly manipulating beliefs that responses could be cross-checked. Their results showed that the manipulation did not lower self-enhancement and
that self-enhancement also did not lower the validity of self-assessment. From their study, Fox and Dinur (1988) conclude that self-assessment may serve as a valuable addition to personnel selection. The authors explain that self-assessment’s validity is gained at low cost and with relative ease and it can also provide a wide range of information which would be difficult to obtain from other sources.

Hogan, Hogan, and Roberts (1996), further offer support of personality testing in the personnel selection context. In an article that attempted to summarize the data needed to answer the most frequent questions about the use of personality measures in applied contexts, the authors reached important conclusions. They assert that well-constructed measures of normal personality are valid predictors of performance in virtually all occupations. This is because scores on well-developed measures of personality are stable over reasonably long periods of time and they predict important organizational outcomes. Further, they explain that well-constructed measures of normal personality do not result in adverse impact for job applicants from minority groups and that using personality measures for preemployment screening is a way to promote social justice and increase organizational productivity. However, when used for preemployment screening, personality measures should be used in conjunction with other information in regards to the applicant’s technical skills, job experience, and ability to learn. Lastly, it is also suggested by the authors that if jobs are classified by occupational type and then the Five Factor Model dimensional requirements and performance criteria relevant to that
occupational type are considered, the predicted relationships between personality and job performance will increase.
Chapter 2

PERSONALITY TESTS AND APPLICANT FAKING BEHAVIOR

It is apparent that the controversy surrounding the use of personality tests for selection purposes stems from whether applicant faking is believed to invalidate personality tests. However, some still question the notion of faking and whether it actually occurs, even though the research overwhelmingly indicates that it does occur. Others question the degree to which applicant faking occurs or the severity of it.

Does Applicant Faking Occur?

In their study, Griffith, Chmielowski, and Yoshita (2007) empirically tested whether applicants fake their responses to personality based employment inventories. The researchers concluded that applicants have the capability to fake non-cognitive measures, that at least some applicants fake in a selection setting, and that this faking may have an effect on the rank order of applicants, and perhaps the criterion validity of non-cognitive selection instruments. The researchers also explain that even when instructed to be honest, respondents may be unable to fully abide by these instructions because of unconscious biases. However, since the primary concern in faking behavior is not unconscious inflation, but the intentional distortion of personality, the authors have looked at how these unconscious biases might help in identifying intentional distortion. Hence, they further reveal that these unconscious biases may be able to aid in isolating the effects of intentional distortion by providing honest scores that would represent a
conservative estimate of the amount of applicant faking since the “honest” scores would be higher than the “true” scores. The overall results of Griffith et al.’s (2007) study demonstrate that applicant faking does occur, whether it be conscious or unconscious, and that it may have detrimental effects on subsequent employment decisions.

Similarly, in a meta-analytic study, Birkeland, Manson, Kisamore, Brannick, and Smith (2006) found that individuals applying for jobs distort their scores on measures of personality to portray themselves in a positive manner. However, the degree to which they distort their scores depends on the personality dimension being measured, type of job, and type of personality test being used. For example, the researchers found that applicants inflate their scores to a much larger degree on the Conscientiousness and Emotional Stability dimensions, which suggests that job applicants view these constructs as being particularly desirable by employers. Also, applicants applying for sales jobs inflated their scores on Extraversion scales and deflated their scores on Agreeableness scales, meaning that applicants are distorting their responses on personality dimensions they view as particularly relevant to the specific job they are applying for. These results support the assumption that job applicants may respond to personality measures in a way that increases their chances of getting a job. One of the strengths of Birkeland et al.’s (2006) study was that it examined intentional distortion that occurs under realistic, employee selection conditions. This is because they examined the degree of response distortion that occurs when participants are applying for a job and are not instructed to
fake responses, unlike several other studies conducted on the subject that induce faking. However, the researchers’ results are consistent with results from induced-faking studies.

Further support for the notion that applicant faking occurs comes from Donovan, Dwight, and Hurtz (2003). Using the randomized-response technique to estimate the base rate of entry-level job applicant faking during the application process, which protects the anonymity of survey respondents in a way that increases their willingness to admit faking behaviors, the researchers found higher base rates of the prevalence of faking than was previously found using other methods. More specifically, the results indicated that almost one-third of the recent applicants participating in the study said that they had engaged in faking behaviors. Further complicating the matter, these results might be underestimating the prevalence of faking on self-report selection measures. Although the participants were asked if they had exaggerated or de-emphasized their personality traits, the participants were not specifically asked if they had exaggerated or de-emphasized their responses on a personality test. That is, when answering to the researcher’s questions, participants might have been referring to faking behaviors regarding personality traits that might have occurred in an interview setting, and not on a personality test.

Donovan et al.’s (2003) results also indicate that applicant faking appears to be a relatively common occurrence and that the perceived degree of deceptiveness of the behavior is possibly a more important determinant of the prevalence of the behavior than the verifiability of the behavior. The researchers came to this conclusion after observing that individuals were more likely to engage in faking behaviors that were low in
perceived severity. Further, the researchers explain that since the most common forms of faking are low in severity, they may not have a large impact on the criterion-related validity or the hiring decisions made using a given selection measure. However, a large proportion of participants in this study indicated that the fake responses which they provided on an inventory when applying for a recent job were moderate to high in severity. The researchers describe how providing fake responses of moderate to high severity during a selection process is cause for concern by explaining that these more severely deceptive behaviors may be more likely to harm selection utility by increasing the rate of false positives due to hiring people who do not possess the qualities the prospective employer thinks they possess.

**Why Do Applicants Fake?**

After coming to the conclusion that applicant faking does occur, Robie, Brown, and Beaty (2007) sought to learn more about the severity of faking. In a study designed to examine the reasons why people fake on personality inventories, the researchers found that individuals vary in their motivation to fake and that individuals may be categorized based upon their faking motivation into one of three faking classes: Honest, Slight Fakers, and Extreme Fakers. The researchers observed that those classed as Honest responders made fewer corrections and took less time completing their personality inventories than those classed as either of the two types of Faking responders. These findings support earlier research by McFarland and Ryan (2000), which suggests that individuals vary considerably in their ability to fake. However, since this study used a
non-applicant sample, these findings were not in accordance with previous research by Robie et al. (2005), where a real applicant sample was used. The researchers explain that the difference between this study’s average personality scale scores and the previous study’s applicant sample is over one standard deviation. A reason for this might be that the induced motivation in this study is not as strong as an employment context in which an applicant is seeking a job.

Agreeing with the notion that faking is a real phenomenon, McFarland and Ryan (2006) sought to point out what some of the factors that predict intentions to fake are. Since intentions, in turn, predict behavior, knowing what those predictive factors are can help mitigate faking. The authors examined the influence of attitudes, subjective norms, and perceived behavioral control on the intention to fake and their results showed that those factors are indeed important predictors of the intention to fake. The authors explain that intentions to fake vary from person to person, since the predictive factors identified vary from person to person. For example, it has been found that positive attitudes toward a behavior lead to greater intentions to perform that behavior (Ajzen, 1991). Similarly, individuals vary in the subjective norms they possess. When individuals perceive that important people in their lives would approve or encourage the behavior in question, such as faking, they are more likely to intend to engage in that behavior. Lastly, individuals who perceive that they have control over a particular behavior have greater intention to perform that behavior (Ajzen, 1991). By examining factors that affect intentions to fake, McFarland and Ryan’s results demonstrate that faking behavior varies from person to
person and suggest that mitigating faking behavior might prove to be a much more
difficult task than one previously thought.

In an effort to also understand individual differences in faking, Snell, Snydell, and
Lueke (1999), proposed an interactional model of faking. The model suggests that the
main factors in faking are ability to fake and motivation to fake. The authors explain that
in order for an individual to fake, he or she must first have the ability to fake. Whether an
applicant has the ability to fake depends on several dispositional factors, experiential
factors, and test characteristics. For example, the authors describe that a dispositional
factor such as high emotional intelligence may give individuals an insight into given
behaviors, and this may allow them to choose more appropriate responses for a given
situation. Similarly, experiential factors such as experience in a customer service position
may make it easier to point out which characteristics are most important for a customer
service job. Also, test characteristics such as item type, item format and item scoring can
unintentionally make some tests easier to fake.

Motivation to fake is the second main factor identified by Snell et al. (1999) in
their model of applicant faking. However, whether an individual possesses the motivation
to fake can depend on several different factors, such as demographic factors, dispositional
factors, perceptual factors, and contextual factors. For example, it was found that age and
gender were both related to faking, with females and younger individuals reporting
higher levels of faking than males and older individuals (Newstead, Franklyn-Stokes, &
Armstead, 1996). Some dispositional factors which differ in individuals, and can
influence their motivation to fake, are integrity, Machiavellianism, manipulativeness, organizational delinquency, locus of control, and stage of cognitive development. Perceptual factors that can have the same influence on faking include others’ behavior, others’ attitudes, fairness, attitudes toward faking, expectations for success, and importance of the outcome. Motivation to fake can also be influenced by contextual factors, which include the general test-taking environment and the instructions given by the test administrator. Overall, the researchers’ model of applicant faking provides a framework for understanding both individual differences in successful faking and sample or situational differences in faking.

Agreeing that applicant faking is a real phenomenon, Levashina and Campion (2007) aimed to understand faking and the components that make it up. They believed that knowing more about the concept of faking will further the efforts of minimizing it. Although the researchers studied faking as it relates to the applicant interview process, and not the personality testing process, their findings are important because they are still applicable to faking in the selection process and employment decisions. Based on their results, Levashina and Campion suggest that the faking construct is represented by four factors: Slight Image Creation, Extensive Image Creation, Image Protection, and Ingratiation. More specifically, the researchers explain that when applicants engage in Slight Image Creation, they exaggerate, but they are still close to the truth. This differs from when they engage in Extensive Image Creating, during which they invent information, or lie. Further, when candidates engage in Image Protection, they
intentionally omit job-related information and finally, when they engage in Ingratiation, they are trying to make interviewers like them and give them a better score on the interview regardless of their performance. In the interview process, different faking factors can come at play in different ways. For example, the authors illustrate how undergraduate job candidates use significantly more Ingratiation than graduate job candidates.

Levashina and Campion (2007) further suggest that the faking construct is additionally made up of 11 subfactors: Embellishing, Tailoring, Fit Enhancing (subfactors of the Slight Image Creation factor), Constructing, Inventing, Borrowing (subfactors of the Extensive Image Creation factor), Masking, Distancing, Omitting (subfactors of the Image Protection factor), and Opinion conforming, Interviewer or Organization Enhancing (subfactors of the Ingratiation factor). In other words, faking is a multidimensional construct and it is likely that different variables would predict the likelihood of engaging in different faking behaviors. The authors explain that engaging in different faking behaviors, in turn, affects the interview outcome. For example, engaging in Extensive Image Creation increases the probability of getting another interview or job offer, whereas engaging in Image protection decreases the probability. Overall, Levashina and Campion’s study helps us understanding the different faking factors that make up the faking construct can help decrease faking in not only the interview process, but all selection processes.
In order to gain a better understanding of what faking-related behaviors applicants tend to engage in and why they engage in these behaviors, Hurtz and Hurtz (2005) asked recent job applicants in an open-ended approach about their faking behavior and their thoughts about the outcomes of faking. This research was also aimed at improving the methods of minimizing applicant faking. The researchers found that the most common dishonest applicant behaviors are lying about or exaggerating work experience, giving false or inappropriate references, lying about previous jobs, and lying about or exaggerating personal characteristics such as personality or attitudes. They also found that many individuals fake due to their positive beliefs regarding faking, which include that faking helps to get the job, helps to get a better starting salary or starting position, and helps to make them look better. On the other hand, beliefs that stop individuals from faking have to do with not being qualified or able to meet job expectations, getting caught, not getting the job, and the risk of getting fired later on. Further, individuals who possessed these positive beliefs regarding faking had stronger intentions to engage in deception in future application experiences than those who had stronger beliefs that negative outcomes would result from faking behavior.

Regarding the mitigation of faking behavior, Hurtz and Hurtz (2005) explain that while means of lowering perceptions of control over the behavior could be implemented, such means would most likely have little impact in comparison to means that target beliefs about the outcomes of engaging in faking behaviors and beliefs about other people’s opinions of such behaviors. This is because applicants might simply just not
believe that they are unable to fake a good impression on a personality test, and if they still believe that they would be able to fake, they most likely would. Due to this, the researchers suggest the use of instructional interventions, in the form of warnings that are grounded in the beliefs about the outcomes of the behavior and the beliefs regarding the opinions of important others, as attempts at deterring faking behavior. Overall, Hurtz and Hurtz’s study offered insight into the different faking-related behaviors that applicants engage in, and more importantly, why they engage in those behaviors. Their results serve as the starting point for developing means of mitigating faking behavior.

Hurtz and Hurtz’s (2005) model of faking focuses on beliefs toward the outcomes of faking as the main influence that leads to faking behavior. However, Goffin and Boyd’s (2009) model of faking focuses on item-level responding because faking is likely to vary from person to person. Similarly to Snell et al. (1999), the researchers suggest that in order for an applicant to fake, he or she must have the perceived ability to fake and the motivation to fake. They also suggest that ability is likely to influence motivation since an individual is less likely to be motivated to fake if he or she does not have a self-perception of being able to fake. The authors further explain that the main influences that lead to an individual’s motivation to fake are self-monitoring, Machiavellianism, integrity, need for approval, conscientiousness, and emotional stability. They describe that those applicants high in self-monitoring would likely be conscious and aware of how they are presenting themselves in each item response and would tend to be motivated to fake to manage the impression they create. Applicants high in Machiavellianism have a
tendency toward lying and cheating and applicants high in a need for approval may be motivated to fake in order to present themselves in a more favorable light. Integrity, according to the researchers, has clear conceptual and empirical links to one’s motivation to fake, and narrow facets of conscientiousness and emotional stability would most likely be linked to a higher motivation to fake. Overall, the authors believe that certain personality traits have the potential to affect an individual’s motivation to fake.

In their model of faking, Goffin and Boyd (2009) also identified certain personality traits that have links to one’s ability to fake. Those personality traits are narcissism, social astuteness, innovation, self-esteem, and two narrow facets of conscientiousness-competence and achievement striving. The authors explain that those high in narcissism tend to engage in self-aggrandizement, and therefore, would most likely have a high perception of the ability to fake. Applicants high in social astuteness would likely perceive themselves as better able to fool testing professionals with their fake responses because they are manipulative of others. Applicants high in innovation are likely to perceive themselves as skilled at figuring out the most beneficial responses to provide on a personality test since they are creative and original. Applicants high in self-esteem would perceive themselves as generally more able to fake since they have a general sense of being capable. Finally, applicants high in achievement striving emphasize competition and ambition and would most likely have a higher perceive ability to fake and those high in competence would likely perceive themselves as more able to fake since they are very capable and effective. Goffin and Boyd also point out that other
factors, such as knowledge of the job for which the individual is applying and the nature of the test item make it easier for an applicant to fake. The researchers conclude by stating that faking has no absolute meaning because it reflects a complex process with multiple contextual and dispositional determinants and is itself, likely to be multidimensional.

**Faking Behavior Has a Negative Effect on Employee Selection**

It is clear that the vast majority of research acknowledges that applicant faking does occur. In order to get closer to finding the best ways of minimizing faking, the effects of faking must be understood. Understanding this is important in deciding whether faking really does matter or not in the selection process and whether the need to minimize faking behavior actually exists. In a study that explored the extent of response distortion on personality inventory scores in an actual applicant-testing environment and its potential effect on which applicants get hired, Rosse, Stecher, Miller, and Levin (1998) explain some of the harmful effects faking can have on employment decisions. By comparing job applicant personality scores to job incumbent personality scores, the researchers found substantial variance resulting from a large number of applicants exhibiting extreme levels of faking more than three standard deviations above the mean.

Rosse et al. (1998) explain that it is this variance in response distortion that changes the rank ordering of applicants and potentially lowers the validity and utility of personality scores. Because differences in faking translate into differences in personality scale scores, applicants had significantly more positive personality profiles than did job
incumbents, whose motivation to fake was presumed to be lower. More specifically, the differences in personality scale scores were particularly large for Neuroticism, Agreeableness, and Conscientiousness. The researchers hold that this is probably due to the socially desirable nature of many of the items in these facets, especially in an employment setting. The differences in the Conscientiousness facet are particularly important since scale scores on this facet are frequently recommended for use in making hiring decisions. The researchers conclude that since it is unlikely that the differences in scores reflect true differences in personality between job applicants and job incumbents, these differences raise concerns regarding the validity of job incumbent personality scores for selection purposes. They also suggest that the appropriate use of faking information hinges on the relevance of faking to job performance.

In a study intended to investigate the effects of situationally-induced response bias on the reliability and validity of non-cognitive measures, Douglas, McDaniel, and Snell (1996) found that how an applicant responds has implications for a variety of scale statistics. More specifically, the researchers found that when applicants desire to do so, they can substantially improve their scores on personality scales. Further, the researchers’ results showed that faking increases the internal consistency of scales. This is because truthful responding hurts the homogeneity of the items on a personality scale, since applicants truthfully report that their behavior is sometimes inconsistent. Therefore, internal consistency is increased with faking because faking responses are more homogeneous, since faking applicants report positive behaviors with consistency. Due to
this, the authors stated that faking decreases the construct validity of personality scales. This is because a personality test, when completed by applicants who fake, may have a different factor structure than when completed by honest applicants. Further, results showed that faking also decreases the criterion-related validity of personality scales. The researchers explain that scales which had useful levels of validity when administered to honest responders, had near zero validity when administered to faking responders. In other words, the criterion-related validity of non-cognitive measures will vary with the proportion of applicants who are faking. The authors conclude that since one does not know when applicants are faking, one can never be sure that the personality tests that were administered to them are valid and, therefore, personality tests should be subject to extensive review when used for personnel screening or selection purposes.

Similar results were found by Mueller-Hanson, Heggestad, and Thornton (2003) in a study that attempted to examine the effects of faking on the criterion-related validity of selection measures. In this study, the sample of participants was also placed in either an “honest” or “faking” condition. The results suggested that applicant faking may have a detrimental impact on the criterion-related validity of personality test scores and on the quality of decisions made using those scores. For participants in the “faking” condition, the criterion-related validity among those with scores in the lower third of the predictor-criterion distribution was significantly higher than the validity for those who scored in the upper third of the predictor-criterion distribution. In other words, less prediction error was found in the bottom third of the distribution than in the top third among those in the
“faking” condition. Results suggest that those in the “faking” condition quickly rose toward the top of the score distribution, creating serious implications for personality tests as they are used in selection contexts. Specifically, faking affects validity at the high end of the predictor distribution and can lead to hiring people whose performance is below what would be expected on the basis of their personality test scores.

In a study examining whether academic success can be predicted with a Big 5 personality questionnaire from different points of view, Ziegler, Danay, Scholmerich, and Buhner (2010) also looked at the impact of faking on the validity of the questionnaire, by using “faking” and “honest” conditions. The authors’ overall results showed that faking affected means and variances of almost all self-ratings but had no effect on the criterion validity of personality domains. More specifically, it was found that faking reduced differences between test-takers. The authors explained that in order for this to have happened, most individuals in the sample had to have had a similar idea of the ideal profile. Along this line of thought, it is reasonable to assume that a more heterogeneous sample might produce increased variance if the test-takers do not share a common stereotype.

Ziegler et al. (2010) also found that although there was no impact on criterion validity on the domain level, there was a significant change in criterion validity on the facet level of the questionnaire. More specifically, changes that occurred between the honest and faking conditions revealed that the correlation coefficients for some facets decreased while they increased for others. For example, for the facet “impulsiveness” the
validities obtained under the “faking” condition differed significantly from those obtained under the “honest” condition. Under the “honest” condition, “impulsiveness” did not affect performance but in the “faking” condition “impulsiveness” served as a negative performance predictor. The overall results of the study provide evidence for a differential effect of faking on the criterion-related validity of personality facets. However, the study possesses limited strength since a faking scenario was used, and not a real assessment situation in a real selection setting.

**Faking Behavior Has No Effect on the Selection Process**

In a study that attempted to offer support to the other side of the argument, which holds that faking behavior is not detrimental to personality testing, Ones and Viswesvaran (1998) found just that. After conducting a review of the extant literature and empirical research, the researchers came to the conclusion that social desirability is not a factor that destroys the criterion-related validity of personality measures and integrity tests. The authors claim that data from applicants and from large-scale meta-analyses indicate that faking does not matter in the prediction of personnel selection. A reason given is that partialling social desirability from personality measures does not have any impact on the criterion-related validities of the Big Five personality variables. Therefore, faking does not act as a mediator or a suppressor variable in the relation between personality and performance. The authors acknowledge that faking has been shown to explain variance in Conscientiousness and it is a predictor of a number of important work variables such as job satisfaction, organizational commitment, and supervisor ratings of
training success. However, they explain that faking is not a predictor of overall job performance and is only very weakly related to specific dimensions of job performance such as technical proficiency and personal discipline. In their conclusion, the authors hold that faking does not influence criterion-related validity and also state that personality and integrity tests have incremental validity over cognitive measures and this validity is increased while adverse impact is decreased when using those tests. A limitation of this study is that social desirability alone was used to measure faking. As Goffin and Boyd (2009) suggested in their model of faking, faking is multidimensional and social desirability does not capture the full breadth of it. Therefore, partialling out social desirability from a personality measure is not equal to partialling out all faking behavior from the measure.

Those affirming that personality tests should not be used for selection purposes hold that applicant faking is unavoidable and it is so detrimental to the results of the personality tests that they are rendered invalid. However, research by Hogan, Barret, and Hogan (2007) found that the majority of applicants who would be naturally motivated to fake did not fake. The researchers found that when applicants were given the same personality test a second time, after not being hired following the first personality test, only 7.3% had changed their scores. Additionally, of those 7.3% whose scores had changed, the scores were as likely to decrease as to increase the second time. When faced with the criticism that these data are hard to interpret because the applicants might
have been faking both times they took the personality tests, the researchers respond by stating that “faking doesn’t matter”.

Hogan et al. (2007) hold the view that faking does not matter because they believe that faking is part of the item response process. In other words, faking is normal and it does not adversely affect the validity of personality measures for employment decisions. For example, the authors explain that many experts wrongfully believe that responses to items on personality measures are self-reports, which are based on self-report theory. However, self-report theory possesses several problems and is inconsistent with the research regarding how memory works and with modern thinking about the nature of communication. For example, the nature of communication suggests that people construct their memories and use communication to manipulate others. In contrast, self-report theory assumes that when people report, they offer factual accounts of how an item matches their memory and that faking involves providing inaccurate reports about the match between the content of an item and the content of memory. The researchers’ view on faking falls in line with impression management theory, which maintains that during social interaction, most people try to maximize acceptance and status and minimize rejection and the loss of status. In support of their view, the authors conclude by adding that impression management theory has gained support through studies which have found that scores of people with good impression management skills are more consistent than scores of people with poor impression management skills.
Further support for the view that faking is not as detrimental to the selection process as some may believe is offered by Hough, Eaton, Dunnette, Kamp, and McCloy (1990) in a study that investigated the effects of faking on the validities of inventory scales. The researchers’ results suggested that faking of self-descriptions in a socially desirable way may not be the problem it has often been assumed to be. Specifically, results show that the criterion-related content scale validities for test-takers scoring high on Social Desirability differed by only very small amounts from the content-scale validities of test-takers scoring lower on Social Desirability. In their study, Hough et al. (1990) examined criterion-related validities for targeted predictor-criterion combinations of scales. For example, one of the targeted combinations consisted of Traditional Values, Nondelinquency, and Conscientiousness scales. The combination represented the Dependability construct, which was intended to predict the personal discipline criterion. Overall results show that validity results for the specific targeted predictor-criterion combinations demonstrated very small differences between test-takers who might have been trying to look good, as suggested by their higher Social Desirability scale scores, and those who were not. Applicants did not appear to distort their self-descriptions, and correlations with job performance were not weakened by such distortion.
Different Ways of Reducing Faking Behavior

Finding ways of reducing the hiring of applicants who fake can serve as an effective improvement in the selection process. However, ideally, employers would like to eliminate faking altogether and still be able to use the same personality measures they have used in the past. Since it would probably be impossible to completely eliminate applicant faking, employers, along with researchers, have been seeking ways of reducing faking. One of the ways of reducing faking that has been examined has been item placement. In a study conducted by McFarland, Ryan, and Ellis (2002), findings suggest that there may be an interaction between faking behavior and test format. More specifically, results showed that test-takers who were given a grouped item placement test tended to fake to a greater extent than those who were given a random item placement test. However, the significant effect was small and it was only present for the Neuroticism and Conscientiousness scales. Also, the grouped items have to be obvious as to what they are measuring as a group in order to facilitate faking. If the items that make up a personality scale are not obvious, item grouping will not aid the faking behavior. A possible reason for why faking is more likely to occur in a non-randomized, grouped item test than in a randomized test is that respondents are not placed under such a high cognitive load when items are randomized. When items are not randomized, applicants
could be influenced to not respond in an appropriate manner, and fake their responses. The overall results of this study suggest that randomizing test items should be continued because they appear to decrease faking on some measures and result in more sound psychometric properties of the measure.

Past research has shown that removing fakers from applicant samples has little positive impact on the mean performance levels of the sample that was selected, and in some cases, can actually result in lower mean-level performance (Schmitt & Oswald, 2006). Although this type of research has led many to believe that faking does not reduce the validity of personality test scores for selection purposes, faking is still generally seen as detrimental to the process of making employment decisions. Because of this, some researchers have sought ways of eliminating fakers from getting hired. For example, in a recent study, Peterson, Griffith, and Converse (2009) examined the impact of applicant faking on personnel selection outcomes across single predictor (conscientiousness alone) and multiple predictor (combinations of conscientiousness and cognitive ability) selection methods. The researchers’ findings showed that hiring decisions based on a combination of cognitive ability and conscientiousness measures generally did not result in the selection of a significantly smaller percentage of fakers than those based on a conscientiousness measure alone. The authors explain that while most of the reductions in the percentage of fakers in the hired sample were not significant, some ability-conscientiousness composites resulted in as much as a 13.5% reduction. In other words, the results of this study showed that the percentage of hiring applicants who fake could
be greatly reduced by using the composite predictors. Since faking that substantially affects hiring decisions is a concern not only for the effective prediction of performance, but also for the fairness of selection practices involving personality assessments, reductions of this size could offer considerable improvements to selection practices.

The Use of Warnings to Reduce Faking Behavior

Although removing fakers from applicant samples can mitigate faking behavior, other ways of reducing faking have shown more promising results. One of these more promising ways of reducing faking has been the use of warnings. For example, in a study by Kluger and Colella (1993), which aimed to demonstrate that a warning against faking a biodata test affects job applicant responses, applicants were randomly chosen to be shown a warning before answering questions regarding themselves and their experiences which may be used to determine their fit to the job. Those chosen to be shown a warning received a warning which alerted them that a special scoring system is used such that dishonest responses may reduce their total score and will not increase their chances of getting the job. Kluger and Colella found that warning against faking can mitigate faking behavior. Their findings show that warnings reduce the extremeness of the item means and increase item variability for scales composed of mostly obvious or transparent items in regard to job desirability, which was defined by the authors as presenting oneself as possessing qualities that are perceived to be important for the particular job. For non-obvious items, warnings did not have an item means and reduced item variances. The overall results suggest that warnings work for specific test items and that, depending on
how obvious an item is in regard to job desirability, the effect of faking may or may not be detected in item means and item variances.

In a similar study, Dwight and Donovan (2003) sought to find whether warnings can reduce faking on noncognitive selection measures using scales specifically designed to measure applicant faking. The researchers also chose to experiment with different types of warnings to see which work better at reducing faking. In their study, participants were assigned to one of four conditions. There was the Unwarned condition (UW), where the applicants did not receive a warning; the Warning of potential identification condition (WI), where participants received a warning that cautioned them that there are test questions designed to identify those who attempt to fake their responses; the Warning about consequences of faking condition (WC), where applicants received a warning about the potential results or consequences of faking but did not receive information about how they might be identified; and the Warning of both identification and consequences condition (WB), where applicants received a warning which both warned them that there are questions designed to identify faking and about the consequences of faking. The participants were also randomly assigned to either an “honest” condition or a “faking” condition, in which they were advised to try to beat the test by making themselves look as good as possible.

Dwight and Donovan (2003) found that those who were in all three of the Warned groups (WI, WC, and WB) scored lower on the personality scales than those who did not receive a warning (UW). Their results also indicated that those in the Warning of both
identification and consequences of faking condition were the only participants whose scores were statistically significantly lower than personality scores found in the Unwarned condition. Additionally, extreme fakers were only found in the Unwarned condition and more potential fakers were found in the Unwarned group, than in any of the Warned conditions. Dwight and Donovan’s findings suggest the importance of developing effective ways of minimizing applicant faking since, in their study, some fakers did rise to the top (were ranked in the top 10 participants). It would be a great detriment to a company if these scores were used to hire individuals under a top-down selection strategy. This study provided a more direct look at the impact of warnings by assessing the impact of the warnings on three measures of faking. Because of these measures, the overall results of the study suggest all three warnings impacted faking by reducing it, with warnings that contain identification and consequence information having provided the largest reduction. This reduction in faking could impact the quality of selection decisions and it is suggested by the authors that personnel selection specialists should consider the use of warnings when utilizing non-cognitive measures in a selection context.

In an extension of the research by Dwight and Donovan (2003), Robson, Jones, and Abraham (2008) sought to examine the effects of applicant warnings against faking on the convergent validity of self-observer ratings. In their study, they only had one warning group, the Warning of both identification and consequences group, previously identified by Dwight and Donovan. Although this warning condition was the only
condition identified by Dwight and Donovan that produced statistically significantly lower scores than the scores produced by the Unwarned group, Robson et al. (2008) found that using a warning statement with consequences does not improve convergent validity and it might actually negatively impact it. The researchers explain that this might have happened because introducing a warning statement might have created a systematic response bias where respondents overcompensated to ensure that they are not caught faking.

Another possible reason for the negative impact on convergent validity might have been that warning statements effectively removed relevant trait variance, which suggests that the reduced convergent validity could be because of the effects of warning statements on the variance related to socially desirable responding. The researchers add that the idea that social desirability measures relevant trait variance rather than error variance has long been supported. Due to this, practitioners should use caution when using warnings in selection contexts since mean scores are reduced by warning statements. Despite this drawback, the warning statements did demonstrate significant mean differences consistent with the research by Dwight and Donovan (2003). Robson et al. (2008) conclude by suggesting that warning statements were effective in reducing faking on some scales, but, in doing so, may have also introduced a systematic response set from applicants and may have a negative impact on convergent validity where applicants are motivated to do well.
Further research on warnings has been conducted by Landers, Tuzinski, and Sackett (2010). The authors examined whether an interactive warning displayed on computer screens for online applicants had any effects on reducing whether the test-takers faked their responses by using blatant extreme responding, which refers to the use of only extreme responses (all 1s and 5s). Blatant extreme responding was identified by the researchers as a common way of “beating the test” used by applicants. This is because if candidates could determine which pole of the scale (either 1 or 5) was the desirable one for the job in question, then this response strategy would be effective in producing a very high score on the test. In order to test whether warnings deterred applicants from using blatant extreme responding, the researchers implemented warnings for a portion of the applicant sample. For that portion of applicants, a general warning was shown prior to the start of the test and a pop-up warning, which they called a real-time warning, was also shown if the applicants completed the first third of personality items in the test using a response pattern that the system identified as 100% 1s and 5s.

The results of Landers et al.’s (2010) study showed that the real-time warning reduced the rate of blatant extreme responding. However, the authors explain that although the real-time warning is effective at deterring blatant extreme responding, it may not have the same effect on other forms of socially desirable responding, because socially desirable responding as it is typically studied would be more difficult to detect. Also, ways of detecting faking, by detecting certain response patterns, have to be developed continuously because as test takers become aware that their technique no
longer works, or is easily detected, they develop new techniques of faking. This, in turn, causes the warnings which were developed for those response patterns, such as the real-time warnings in this study, to be ineffective. The authors conclude by stating that correlates and determinants of deliberate faking must be discovered, or this pattern will repeat indefinitely.

**Test-Taker Reactions to Warnings**

Research has demonstrated that warnings against faking do indeed reduce applicant faking behavior. However, this also comes with a price. Research has shown that certain warnings have negative effects on test-takers. For example, Converse, Oswald, Imus, Hendricks, Roy, and Butera (2008) analyzed how criterion-related validity and test-taker reactions are affected across conditions by comparing an explicit warning against faking vs. no warning. The researchers used two types of warning conditions, a positively framed warning and a negatively framed warning. More specifically, the positively framed warning informed the applicants that they get a one in four chance to earn some money if they answer as accurately and honestly as possible. The negatively framed warning informed participants that if they answer dishonestly, they will lose the one in four chance to earn some money if they answer inaccurately or dishonestly. In order to measure test-taker reactions to the warnings, the participants were administered two measures examining the level of difficulty experienced on the personality inventory, two measures potentially related to perceptions of the opportunity to convey job-related personality characteristics, and two measures of affect. These measures were
administered after the participants had completed the personality inventory and had been exposed to the warning conditions.

Regarding the overall effects of warnings, Converse et al. (2008) did not find strong evidence that warnings enhance criterion-related validity or that they undermine incremental validity for personality measures. However, their results indicated that warnings against faking are another way to reduce response distortion. Regarding test-taker reactions, the researchers’ results indicated that those test-takers given a negative warning reported higher test-taking anxiety than those given a positive or no warning. At the same time, negative warnings were not found to be more effective than positive warnings in deterring faking. Negative warnings may result in negative reactions because applicants may feel that the employer is distrusting or limiting their ability to present themselves as they would like (McFarland, 2003). These results point to a potential drawback to the use of some warnings. Converse et al. conclude by suggesting that, given that the negative warning was not clearly more effective than the positive warning, researchers and practitioners should consider exploring other methods of strengthening warnings, as opposed to framing them negatively, that may be less likely to produce negative reactions.

Also investigating the effects of warnings on test-taker reactions, McFarland (2003) examined the effects of warnings on procedural justice perceptions. The participants were placed in one of two conditions, Warned or Unwarned, and were asked to take a personality test as part of a simulated selection process. The participants in the
Warned condition were alerted that a social desirability scale was included in the personality measure and that such scales are used to identify individuals who give inaccurate or false responses. They were further alerted that those detected as faking will be removed from the selection process. The participants in the Unwarned condition were not given this information. In contrast to results obtained by Converse et al. (2008), Mcfarland (2003) found that warnings had little effect on procedural justice perceptions. More specifically, it was found that individuals who were warned had significantly more positive perceptions of selection information, and unexpectedly, perceived predictive validity, opportunity to perform, and question impropriety were more favorable in the warned conditions. McFarland’s overall results suggest that warnings do not generally have a negative effect on test-taker reactions. In other words, it seems that warnings do not affect how individuals view the testing process and the test administrators. For example, individuals who perceived the testing process as fair were found to be less likely to fake on the test, regardless of whether or not they were given a warning.

The Current Study

Past research demonstrates how contradicting findings have been in regard to not only whether personality tests should be used as part of the selection process, but whether faking has a detrimental effect on their validity and whether means of reducing faking are actually effective and should be used. Acknowledging that all of these questions are critical in determining the role of personality tests in the employee selection context, the current study does not attempt to answer all of these questions, since some of them are
beyond the scope of the study. The current study also does not attempt to take any sides on the matters which are beyond the scope of the study. The purpose of the current study is to determine what types of warnings reduce applicant faking behavior and what types of test-taker reactions those warnings produce, if any. It is clear that the research regarding these matters is quite limited and that the few studies that have been conducted are in disagreement with each other. Nevertheless, the importance of possessing effective means of mitigating faking behavior is apparent, as they can be the decision-making factor in deciding whether to use a personality test during the selection process. Deciding whether to use a personality test when selecting employees can, in turn, greatly affect the quality of employees selected, whether this be a positive or a negative effect.

In addition to the important information regarding warnings and applicant reactions that will aid the overall selection process, the experimental simulation in the current study will also add to the literature a comparison of warnings that has not been studied before. These warnings have not been previously used in the same study and test-taker reactions have not been previously examined for this specific set of warnings. The current study will further knowledge regarding ways of deterring applicant deception and will give insight into whether those ways, even if highly effective, are beneficial for an organization to use.

Research into applicant faking as it relates to the employee selection process is ideally performed in a real life job applicant setting. This is because the motivation to do well on a personality test in a real selection setting is difficult to mimic in an
experimental setting. Consequently, the motivation to fake is substantially different in the two settings. However, although having its own downfalls, the experimental design can also prove to be very useful in such a situation. For example, having the opportunity to manipulate variables in a controlled setting is a valuable attribute of the traditional experimental design. This allows for the different types of warnings in the current study to be manipulated, while also controlling for other possible reasons for the results obtained. Although the setting in the current study is not natural, measures have been taken in order to ensure that the participants would be as closely motivated to respond in the same way as real life applicants as possible, increasing the external validity of the study. These measures are described in greater detail in the following chapters of this paper.

Since the main purpose of the study is to examine what type of warning, if any, will have an impact on faking behavior, several warning conditions will be compared. Specifically, the study used two unwarned groups and three warned, experimental groups. The first unwarned (control) group was told to fake good on their responses. Specifically, the participants were instructed to present themselves in the best possible light possible and answer the items in a way which would increase their chances of getting hired. This control group will be referred to as the Fake Good or FG control group for the remainder of this paper. In contrast, the second unwarned (control) group was told to answer as honestly as possible since the personality test is being used solely for research purposes and frank, open, and honest responses are needed. This group was also alerted that their
responses would not be shared with anyone else. This control group will be referred to as the Honest or HC control group for the remainder of this paper. The two unwarned groups were used as control groups, in order to compare their results to the results of the warned groups. This served to indicate whether different types of warnings do have an effect on applicant faking behavior. The two unwarned groups were also compared to each other. This served to indicate whether applicants followed instructions properly.

The first warned group was the Identification and Consequences of Faking, warning condition used by Dwight and Donovan (2003). This warning condition alerted participants that the personality test they were about to take can detect whether they were intentionally providing inaccurate answers. The warning also indicated that if the participants were caught deliberately providing inaccurate answers, they would not be eligible to be entered in the raffle for the $25 prize. This warning will be referred to as the Identification and Consequences or IC condition for the remainder of this paper. The second warned condition tapped into participants’ attitudes toward faking. This warning condition informed the participants that faking on these types of personality tests often leads to applicants getting caught and not getting the job. Participants were also informed that faking behavior is discouraged for a number of reasons, such as getting a bad reputation and if hired, it would only increase their chances of getting fired later on. This warning will be referred to as the Attitudes toward Faking or AF condition for the remainder of this paper. The third warned condition tapped into the subjective norms associated with applicant faking. This warning notified the participants that the
organization for which they were applying condones honesty, and that honesty is greatly needed in order to help provide a better match between applicants and jobs. This warning will be referred to as the Subjective Norms or SN condition for the remainder of this paper.

**Hypotheses**

Based on previous warning literature, the questionnaire scores of participants in the group who is told to fake their responses in a desirable way (FG control group), are expected to reflect personality traits typically endorsed by employers to a higher degree than the group that was told that the results of this study will only be used for research purposes (HC control group). This is because it is expected that the participants who are told that their scores will only be used for research will answer more truthfully and their scores will therefore reflect a broader range of personality traits, including personality traits not endorsed by employers. In other words, the FG control group’s scores will indicate more faking than the HC control group’s. A reason for this is because participants in the HC control group have no motivation to fake and make themselves look like they possess more desirable traits. For this reason, it is hypothesized:

**Hypothesis 1:** There will be a greater tendency to fake in the FG control group than in the HC control group.

The FG control group participants’ scores are also expected to show higher levels of faking than all three experimental, warned groups (IC condition, AF condition, SN condition). This is because even though the participants in the warned groups might fake
their answers to a certain degree, it is expected that the warnings presented in those conditions will have a reducing effect on faking behavior. Even if there were no warnings present, it is expected that the group who was instructed to fake good will show higher levels of faking behavior than groups who were not told to fake. For this reason, it is hypothesized:

Hypothesis 2: There will be a greater tendency to fake in the FG control group than in the SN condition group.

Hypothesis 3: There will be a greater tendency to fake in the FG control group than in the AF condition group.

Hypothesis 4: There will be a greater tendency to fake in the FG control group than in the IC condition group.

The three warned groups will be compared to each other in order to determine which type of warning has the stronger reducing effect on faking behavior. Due to past research by Dwight and Donovan (2003), it is expected that the IC condition will have a very good reducing effect on faking. It is expected that this effect will be stronger than the effects produced by the SN warning and the AF warning since the IC condition directly and negatively affects the respondent. Participants are expected to be deterred from faking by the IC condition because the warning states that they will not be eligible to be entered in the raffle for the $25 prize if they fake. For this reason, it is hypothesized:

Hypothesis 5: There will be a greater tendency to fake in the SN condition group than in the IC condition group.
Hypothesis 6: There will be a greater tendency to fake in the AF condition group than in the IC condition group.

Although the AF warning is not expected to have as strong of an effect on reducing faking behavior as the IC condition, it is expected to have a stronger effect than the SN warning. This is because, in previous research, positive attitudes toward faking behavior were identified as one of the main causes of faking (Hurtz & Hurtz, 2005). Therefore, it is expected that warnings which directly disconfirm those attitudes and beliefs will have a quite strong effect in reducing faking behavior. This effect is expected to be stronger than the effect the SN warning will have since the SN warning simply tells participants that faking is not condoned by the organization for which they are applying and that honesty is needed in order to provide a better match between the applicants and jobs. This warning does not directly affect the participant and does not challenge their way of thinking, in contrast to the AF warning. For this reason, it is hypothesized:

Hypothesis 7: There will be a greater tendency to fake in the SN condition group than in the AF condition group.

Although determining which type of warning is the most effective at reducing applicant faking behavior is the main purpose of the proposed study, it is also important to identify what kind of test-taker reactions those warnings produce. This is because negative reactions to warnings might result in turning away a qualified applicant by producing negative impressions of the organization and its management. It is hypothesized that the IC condition will produce more negative test-taker reactions than
the other two warnings (AF and SN) because its negatively worded nature might threaten participants and cause them to feel defensive toward the selection process since it implies that they are not trusted. For this reason, it is hypothesized:

Hypothesis 8: There will be a greater tendency for participants to have more negative reactions to the selection test when the IC condition is given as compared to the SN or AF warnings.
Chapter 4

METHOD

Participants

The participants in this study were 218 undergraduate Psychology students (50
Freshmen, 60 Sophomores, 48 Juniors, and 33 Seniors) from California State University,
Sacramento. There were 155 females, 38 males, and 25 who did not report their gender.
The average age of the participants was 21 years old, with ages ranging from 18 to 46.
There were 91 Caucasians, nine African Americans, 35 Hispanics, 33 Asians, one Native
American, 23 who reported “Other” ethnicity, and 26 who did not report their ethnicity.
There were 125 participants currently employed. The average number of hours worked
per week reported by the employed participants was 23. The three most frequently
reported jobs held by the employed participants were sales associate, restaurant worker
(cashier, host, server), and customer service representative. There were 68 participants
who were currently not employed. On average, the non-employed participants reported
that the last time they had held a job was approximately six and a half months prior to
their participation in the study. The three most frequently reported jobs that the non-
employed participants last held were sales associate, restaurant worker, and customer
service representative. The average number of previous jobs held by the participants
(including current job) was three.
Measures

Participants were administered three measures: The BIDR-IM (Impression Management Scale from the Balanced Inventory of Desirable Responding), the NEO-FFI (Five Factor Inventory), and a Reactions to the Selection System Scale. See Appendices A and B for the Balanced Inventory of Desirable Responding - Impression Management Scale and the Reactions to the Selection System Scale, respectively.

The BIDR-IM was administered in the current study in order to measure intentional response distortion, or faking. The BIDR-IM was chosen for use in this study because it does not appear to possess the same issues as other intentional response distortion scales possess. Specifically, a large number of previous studies have used the Marlowe-Crowne scale, which does not appear to be a pure measure of intentional response distortion (Rosse et al.1998). Other scales possessing the same problems are the Edwards Social Desirability Scale, the Desirability scale from the Personality Research Form, and the K scale from the MMPI (Rosse et al.1998). The BIDR-IM is a relatively pure measure of intentional response distortion and has been found to measure conscious faking not related to substantive dimensions of personality that may be related to the broader construct of impression management (Rosse et al.1998). Additionally, the BIDR-IM appears to have a stable factor structure, with internal consistency coefficients alpha ranging from .75 to .86. The BIDR-IM scale can be scored both on a 5 point or a 7 point scale. For the current study, it was scored on the 5 point scale, with a score of 1 indicating that the respondent *strongly disagrees* with the statement, a score of 2
indicating that the respondent *disagrees* with the statement, a score of 3 indicating that the respondent *neither disagrees nor agrees* with the statement, a score of 4 indicating that the respondent *agrees* with the statement, and a score of 5 indicating that the respondent *strongly agrees* with the statement. The scoring procedure starts with reverse coding the items 21,23,25,27,29,31,33,35,37,39 on the scale. Then, the scale is dichotomized by replacing every 4 or 5 rating with a 1 and all of the other ratings with a 0. The minimum score that can be obtained on the BIDR-IM is a 0 and the maximum score a 20, with 20 indicating complete deceptiveness. The BIDR-IM items were randomly interspersed with the NEO-FFI items and the same 5 point response scale was used for both.

The NEO-FFI (Five Factor Inventory) was primarily administered as an additional measure of detecting faking behavior. Since it is not a measure specifically designed to detect faking behavior like the BIDR-IM is, the way faking behavior was measured has to be explained. First, mean differences between conditions on the personality dimensions were used as indicators of faking behavior. A statistically significant mean difference between two conditions on the same dependent variable indicated an inflation of scores on one of the conditions. Second, the participants were asked to pretend they were applying for the occupation of “psychologist”, it was thought that the Neuroticism dimension of the NEO-FFI was an undesirable trait and the Extraversion, Openness, Agreeableness, and Conscientiousness were desirable traits to have for that specific occupation. Therefore, a significant mean difference between two conditions on the
Neuroticism scale indicated faking had occurred for one of those conditions, and since Neuroticism is an undesirable trait, the condition that displayed a less neurotic image in Neuroticism scores was the condition for which participants displayed more faking behavior. Similarly, a significant mean difference between two conditions on the Extraversion, Openness, Agreeableness, and Conscientiousness scales indicated faking had occurred for one of those conditions, and since these personality traits are desirable traits for the job, the condition that displayed more extraversion, openness, agreeableness, or conscientiousness was the condition for which participants displayed more faking behavior.

The NEO-FFI was also administered in order to mimic a real-life selection testing situation where a personality test is administered as part of the selection process. This version of the NEO consists of only 60 items as opposed to the 240 items of the NEO-PI-R. The shortened version of the NEO is ideal for use on college students, which is what the sample of participants used was, since it provides college-age norms. The NEO-FFI was selected for this study because it measures a number of the major personality dimensions and because the constructors of the test have provided a set of observer rating scales that parallel the self-report dimensions (Topping & Gorman, 1997). These provide a ready source of information on the validity of the self-report dimensions. The NEO-FFI can also help understand an individual's basic emotional, interpersonal, experiential, attitudinal, and motivational styles. This personality test was selected for this study also because previous research on the NEO-FFI has found that participants could create
specific impressions in a manner consistent with instructional sets provided (Scandell & Wlazelek, 1996). Since the current study is aimed at investigating the types of effects specific instructional sets can have on personality tests, this feature of the NEO-FFI was thought to serve as helpful. Lastly, the NEO-FFI was chosen for the current study because it shows good reliability values, with internal consistency coefficients ranging from .74 to .89. Similarly, the NEO-FFI scales show correlations of .75 to .89 with the NEO-PI validimax factors.

A Reactions to the Selection System Scale was administered following the NEO-FFI and BIDR-IM scales in order to measure whether the warnings had any effects on participants’ reactions to the selection system. The Reactions to the Selection System Scale that was used was previously used by Rosse, Miller, and Stecher (1994) in a study that looked at job applicants’ reactions to personality and cognitive ability testing. The Reactions Scale is made up of two different scales, one of which measures perceived privacy, and the other which measures perceived appropriateness. The scale that measures privacy has been shown to have an internal consistency of .84 and the scale that measures appropriateness has been shown to have an internal consistency of .94. Items for both scales are measured using five-option Likert-type questions that was averaged to form scale scores. Because the two scales are highly intercorrelated ($r = .86$), they were combined into a single scale with an alpha of .94, the Reactions to the Selection System Scale. The scoring procedure for the Reactions to the Selection System Scale starts by
reverse coding the items 2, 4, and 5. Items on the Reactions Scale are scored so that high scores indicate more positive reactions to the selection procedure and faking warnings.

Control/Unwarned Groups

Fake Good Control Group

Participants randomly assigned to this condition were presented with instructions that told them to respond to the questionnaire in a way that makes them look best in order to increase their chances of getting hired. These instructions told the participants they could lie about themselves in order to appear as if they possessed all of the qualities that they believe an employer would look for in a very impressive applicant. In other words, these participants were instructed to “beat the tests” and create a very good impression of themselves. These participants were expected to self-report in an overly positive manner.

Honest Control Group

Participants randomly assigned to this condition were presented with instructions that asked them to answer the items on the measure as frankly, openly, and honestly as they possibly can. These participants were also informed that their answers to the questionnaire would not be used to make any psychological evaluations about them, would not be shared with anyone else, and would only be used for research purposes.

Experimental/Warned Conditions

Identification and Consequences of Faking Condition

Participants randomly assigned to this condition were presented with a warning that alerted them that there are questions in the questionnaires designed to identify faking.
These participants were also warned that there are consequences if they are caught faking (they would not be eligible to be entered into the drawing for the $25 prize).

**Subjective Norms of Faking Condition**

Participants randomly assigned to this condition were presented with a warning that alerted them that the organization for which they are applying, including all potential coworkers, supervisors, and managers, values honesty and openness. Therefore, the organization seeks honesty and openness in anyone seeking to join the team as well, in order to help provide a better match between applicants and jobs.

**Attitudes toward Faking Condition**

Participants randomly assigned to this condition were presented with a warning that informed them that many job seekers choose to exaggerate applicant information in order to help them get the job and increase their chances of getting a higher starting salary or a better starting position. The warning further alerted the participants that, since providing inaccurate answers often leads to the applicant getting caught and not getting the job, they are discouraged from this type of behavior. Participants were also informed that this type of behavior leads to applicants receiving positions for which they are not qualified, and for which they would not be able to meet expectations. Therefore, faking will often only give them a bad reputation, will make them feel bad about themselves, and if hired, would increase the risk of them getting fired later on. (See Appendix C for all conditions)
Procedure

Participants were first given a consent form to sign and a sheet of paper to sign into. The participants were then randomly assigned a packet that contained a randomly picked warning condition or control condition, the personality inventory composed of the NEO-FFI and the BIDR-IM items, Reactions to the Selection System Scale, an answer sheet, and a demographics page. The packets also had a sheet of paper attached to the outside of them which instructed the participants to pretend that they are applying for a job as a psychologist and the questionnaire inside of the packet would be used to make hiring decisions. Before opening the packet, the participants were instructed to carefully read all instructions before proceeding to fill out the contents of the packet. The participants were allowed one-half hour to read all instructions and complete the items in the packet. It should be noted that participants were not instructed to stop pretending they were applying for a job and to stop taking the warning and control group instructions into consideration before responding to the Reactions to the Selection System Scale. Once the participants were done with the packets, they were debriefed by informing them of the purpose of the study and providing them with the contact information of the researcher. Despite the warning which indicated that if participants are caught lying they would not be eligible to be entered into the raffle to win the $25 prize, all participants were entered into the raffle upon completion of data collection, and the winner was notified by e-mail.
Chapter 5

ANALYSIS AND RESULTS

The primary purpose of this study was to examine whether different warning conditions have a minimizing effect on faking behavior on a personality test as part of an employment selection process. The purpose of Hypotheses 1 through 7 was to examine the expected patterns of differences in faking behavior across the various experimental groups. These analyses were started by first conducting a correlation of all of the personality scale variables along with the IM scale and the Reactions scale variables. The correlations of the measured variables are presented in Table 1.

Next, a one-way omnibus multivariate analysis of variance (MANOVA) was conducted using the IM variable and all of the personality scale variables. The reason for this analysis was to test for differences between the warning conditions on a linear combination of the Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness personality variables, and the IM variable. The means and standard deviations are presented in Table 2. The MANOVA analysis showed a significant effect at the .05 level, Wilks’ Lambda = .79, $F(24, 712.88) = 2.05$, $p = .002$, $\eta^2 = .06$, indicating significantly different scores on the dependent variables between participants based on warning condition. Due to the significant effect found, simple contrast MANOVAs were conducted to further examine which dependent variables were significantly different for the different warning conditions in accordance with the hypothesized effects.
Six simple contrast analyses were conducted in order to compare specific conditions. The contrasts were evaluated using a Bonferroni corrected alpha of .008. The first simple contrast MANOVA compared the two control groups (FG) and (HC). This was done in order to test Hypothesis 1. The overall MANOVA was significant at the .008 level, Wilks’ Lambda = .79, \( F(6, 78) = 3.52, p = .004, \eta^2 = .21 \). The nature of the multivariate effect was explored with \( t \) tests for contrast coefficients on each of the six personality scores, and again a corrected alpha of .008 was used for these tests. As seen in Table 3, the comparison of the two control groups was significant at the .008 level for the Neuroticism and Openness personality scores. These significant effects indicate that participants randomly assigned to the two control groups had significantly different scores on the Neuroticism and Openness personality scores. This contrast suggests that participants assigned to the FG control group presented a less neurotic image and a more open image than those assigned to the HC control group. While not all contrasts deemed statistically significant, a consistent effect was found in that all of the contrasts were in the expected direction. This comparison served as a manipulation check, showing that the two control groups had the desired effects on the participants.
Table 1

Correlation Matrix of all Measured Variables

<table>
<thead>
<tr>
<th></th>
<th>IM</th>
<th>N</th>
<th>E</th>
<th>O</th>
<th>A</th>
<th>C</th>
<th>RS</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM</td>
<td>(.49)</td>
<td>.02</td>
<td>-.03</td>
<td>-.04</td>
<td>-.03</td>
<td>.05</td>
<td>.02</td>
<td>8.01</td>
<td>2.63</td>
</tr>
<tr>
<td>N</td>
<td>(.85)</td>
<td>-.45**</td>
<td>-.41**</td>
<td>-.57**</td>
<td>-.62**</td>
<td>-.34**</td>
<td>15.40</td>
<td>7.60</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>(.70)</td>
<td>.31*</td>
<td>.45**</td>
<td>.45**</td>
<td>.32**</td>
<td>33.83</td>
<td>5.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>(.71)</td>
<td>.36**</td>
<td>.27**</td>
<td>.27**</td>
<td>29.05</td>
<td>6.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>(.78)</td>
<td>.54**</td>
<td>.37**</td>
<td>35.46</td>
<td>5.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>(.85)</td>
<td>.35**</td>
<td>39.18</td>
<td>6.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RS</td>
<td>(.84)</td>
<td>44.95</td>
<td>5.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * Correlation is significant at the .05 level; ** Correlation is significant at the .01 level. Values in parentheses are coefficient alphas. IM = Impression Management Scale; N = Neuroticism, E = Extraversion, O = Openness, A = Agreeableness, C = Conscientiousness, RS = Reactions to the Selection System Scale
Table 2

Means and Standard Deviations for the IM Scale and Personality Scale Variables across all Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IM</td>
<td>N</td>
</tr>
<tr>
<td>FG</td>
<td>45</td>
<td>8.13(3.19)</td>
</tr>
<tr>
<td>HC</td>
<td>39</td>
<td>7.71 (1.62)</td>
</tr>
<tr>
<td>SN</td>
<td>41</td>
<td>7.88(2.75)</td>
</tr>
<tr>
<td>AF</td>
<td>39</td>
<td>7.88(2.54)</td>
</tr>
<tr>
<td>IC</td>
<td>49</td>
<td>8.37(2.73)</td>
</tr>
<tr>
<td>Grand Mean</td>
<td>218</td>
<td>8.01(2.63)</td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses are standard deviations. n is the valid sample size. IM = Impression Management Scale; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness. FG = Fake Good; HC = Honest; SN = Subjective Norms; AF = Attitudes toward Faking; IC = Identification and Consequences of Faking.
Table 3

Comparison of the Fake Good Control Group and the Honest Control Group

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>HC - FG</th>
<th>t test</th>
<th>p value</th>
<th>d statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM score</td>
<td>-0.41</td>
<td>-0.73</td>
<td>.465</td>
<td>-0.16</td>
</tr>
<tr>
<td>N score</td>
<td>6.71</td>
<td>4.42</td>
<td>.000</td>
<td>0.97</td>
</tr>
<tr>
<td>E score</td>
<td>-1.74</td>
<td>-1.54</td>
<td>.128</td>
<td>-0.34</td>
</tr>
<tr>
<td>O score</td>
<td>-3.66</td>
<td>-2.83</td>
<td>.006</td>
<td>-0.62</td>
</tr>
<tr>
<td>A score</td>
<td>-3.12</td>
<td>-2.37</td>
<td>.020</td>
<td>-0.52</td>
</tr>
<tr>
<td>C score</td>
<td>-2.96</td>
<td>-2.18</td>
<td>.032</td>
<td>-0.48</td>
</tr>
</tbody>
</table>

*Note.* p value is set at .008. d statistic = Cohen’s d effect size. HC = Honest Control group; FG = Fake Good; IM = Impression Management Scale; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness.

Next, a simple contrast MANOVA compared the first control group (FG) with warning conditions SN, AF, and IC for each dependent variable. This was done in order to test Hypotheses 2, 3 and 4. The overall MANOVA was significant at the .008 level, Wilks’ Lambda = .79, $F(18, 467.18) = 2.25$, $p = .002$, $\eta^2 = .08$. Hypothesis 2 stated that there would be a greater tendency to fake in the FG control group than in the SN condition. As seen in Table 4, the comparison of the FG control group and the SN condition was significant for the Neuroticism and Openness personality scales. This contrast indicates that on the Neuroticism scale, participants assigned to the FG control group presented a less neurotic image than participants assigned to the SN condition and
on the Openness scale, participants presented a more open image than those assigned to the SN condition, with effect sizes being large in magnitude.

Hypothesis 3 stated that there would be a greater tendency to fake in the FG control group than in the AF condition. As seen in Table 4, the comparison of the FG control group and AF condition was significant at the .008 level for the Neuroticism personality scale. This contrast indicates that on the Neuroticism scale, participants assigned to the FG control group presented a less neurotic image than those assigned to the AF condition with the effect size being large in magnitude.

Hypothesis 4 stated that there would be a greater tendency to fake in the FG control group than in the IC condition. As seen in Table 4, the comparison of the FG control group and IC condition was significant at the .008 level for all of the personality variables: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. This contrast indicates that on the Neuroticism scale, participants assigned to the FG control group presented a less neurotic image than those assigned to the IC condition and that on the Extraversion, Openness, Agreeableness, and Conscientiousness scales, participants presented more extraverted, open, agreeable, and conscientious images than those assigned to the IC condition, with effect sizes being large in magnitude.
Table 4

Comparison of the Fake Good Control Group and the Warning Conditions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>SN - FG</th>
<th></th>
<th></th>
<th>AF - FG</th>
<th></th>
<th></th>
<th>IC - FG</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
<td>d statistic</td>
<td>p value</td>
<td>D</td>
<td>d statistic</td>
<td>p value</td>
<td>D</td>
<td>d statistic</td>
<td>p value</td>
</tr>
<tr>
<td>IM score</td>
<td>-0.19</td>
<td>-0.05</td>
<td>.762</td>
<td>-0.22</td>
<td>-0.05</td>
<td>.728</td>
<td>0.28</td>
<td>0.07</td>
<td>.636</td>
</tr>
<tr>
<td>N score</td>
<td>5.77</td>
<td>0.57</td>
<td>.000</td>
<td>4.82</td>
<td>0.47</td>
<td>.003</td>
<td>6.27</td>
<td>0.64</td>
<td>.000</td>
</tr>
<tr>
<td>E score</td>
<td>-2.10</td>
<td>-0.29</td>
<td>.058</td>
<td>-0.77</td>
<td>-0.11</td>
<td>.492</td>
<td>-3.45</td>
<td>-0.50</td>
<td>.001</td>
</tr>
<tr>
<td>O score</td>
<td>-3.42</td>
<td>-0.41</td>
<td>.008</td>
<td>-3.07</td>
<td>-0.36</td>
<td>.019</td>
<td>-3.79</td>
<td>-0.48</td>
<td>.002</td>
</tr>
<tr>
<td>A score</td>
<td>-1.43</td>
<td>-0.18</td>
<td>.239</td>
<td>-0.07</td>
<td>-0.01</td>
<td>.957</td>
<td>-3.17</td>
<td>-0.42</td>
<td>.007</td>
</tr>
<tr>
<td>C score</td>
<td>-2.47</td>
<td>-0.29</td>
<td>.063</td>
<td>-1.84</td>
<td>-0.21</td>
<td>.170</td>
<td>-5.04</td>
<td>-0.61</td>
<td>.000</td>
</tr>
</tbody>
</table>

Note. p value is set at .008. D = mean difference. d statistic = Cohen’s d effect size. IM = Impression Management Scale; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness; SN = Subjective Norms; FG = Fake Good; AF = Attitudes toward Faking; IC = Identification and Consequences of Faking.
Next, a MANOVA was conducted in order to compare the second control group (HC) with warning conditions SN, AF, and IC for each dependent variable. The overall MANOVA was not significant at the .008 level, Wilks’ Lambda = .87, $F(18, 453.03) = 1.31, p = .180, \eta^2 = .05$. As seen in Table 5, the comparison of HC and SN condition was not significant at the .008 level for any of the dependent variables measured. This contrast indicates that participants assigned to the SN condition did not differ in their responses from those assigned to the HC group. The comparison of the HC control group and AF condition was not significant at the .008 level for any of the personality scales. This contrast indicates that participants assigned to the AF condition did not differ in their responses from those assigned to the HC group. The comparison of HC and the IC condition was not significant at the .008 level for any of the dependent variables measured. This contrast indicates that participants assigned to IC condition did not differ in their responses from those assigned to the HC group.

Hypothesis 5 stated that there would be a greater tendency to fake in the SN condition than in the IC condition. A simple contrast MANOVA was conducted comparing the SN condition to the IC condition with NEO-FFI scores as the dependent variables. The overall MANOVA was not significant at the .008 level, Wilks’ lambda = .91, $F(6, 83) = 1.44, p = .210, \eta^2 = .09$. As seen in Table 6, the contrast was not significant at the .008 level for any of the personality scales. This contrast suggests that participants assigned to the SN condition were not presenting more favorable
Table 5

*Comparison of the Honest Control Group and the Warning Conditions*

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>SN - HC</th>
<th>AF - HC</th>
<th>IC - HC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
<td>d statistic</td>
<td>p value</td>
</tr>
<tr>
<td>IM score</td>
<td>0.23</td>
<td>0.06</td>
<td>.681</td>
</tr>
<tr>
<td>N score</td>
<td>-0.94</td>
<td>-0.09</td>
<td>.559</td>
</tr>
<tr>
<td>E score</td>
<td>-0.36</td>
<td>-0.05</td>
<td>.735</td>
</tr>
<tr>
<td>O score</td>
<td>0.24</td>
<td>0.03</td>
<td>.857</td>
</tr>
<tr>
<td>A score</td>
<td>1.68</td>
<td>0.22</td>
<td>.161</td>
</tr>
<tr>
<td>C score</td>
<td>0.50</td>
<td>0.06</td>
<td>.684</td>
</tr>
</tbody>
</table>

*Note.* p value is set at .008. D = mean difference. d statistic = Cohen’s d effect size. IM = Impression Management Scale; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness; SN = Subjective Norms; HC = Honest Control group; AF = Attitudes toward Faking; IC = Identification and Consequences of Faking.
images than participants assigned to the IC condition, and vice versa.

Hypothesis 6 stated that there would be a greater tendency to fake in the AF condition than in the IC condition. A simple contrast MANOVA was conducted comparing the AF condition to the IC condition with NEO-FFI scores as the dependent variables. The overall MANOVA was not significant at the .008 level, Wilks’ lambda = .83, $F(6, 81) = 2.60$, $p = .024$, $\eta^2 = .16$. As seen in Table 6, the contrast was not significant at the .008 level for any of the personality scales. This contrast suggests that participants assigned to the AF warning condition were not presenting more favorable images than participants assigned to the IC group, and vice versa.

Finally, Hypothesis 7 stated that there would be a greater tendency to fake in the SN condition than in the AF condition. A simple contrast MANOVA was conducted comparing the SN condition to the AF condition with NEO-FFI scores as the dependent variables. The overall MANOVA was not significant at the .008 level, Wilks’ lambda = .97, $F(6, 73) = .97$, $p = .877$, $\eta^2 = .03$. As seen in Table 6, the contrast was not significant at the .008 level for any of the personality scale variables. The non-significant effects observed for the personality variables between conditions SN and AF indicate that there are no statistically significant differences in scores on the NEO-FFI personality scores between participants who were in the SN condition and those who were in the AF condition. This contrast suggests that participants assigned to the SN group were not presenting more favorable images on the personality test than participants assigned to the AF group, and vice versa.
Table 6

Comparison of Warning Conditions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>IC - SN</th>
<th>IC - AF</th>
<th>AF - SN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
<td>d statistic</td>
<td>p value</td>
</tr>
<tr>
<td>IM score</td>
<td>0.47</td>
<td>0.17</td>
<td>.427</td>
</tr>
<tr>
<td>N score</td>
<td>0.50</td>
<td>0.07</td>
<td>.734</td>
</tr>
<tr>
<td>E score</td>
<td>-1.35</td>
<td>-0.27</td>
<td>.210</td>
</tr>
<tr>
<td>O score</td>
<td>-0.37</td>
<td>0.06</td>
<td>.764</td>
</tr>
<tr>
<td>A score</td>
<td>-1.74</td>
<td>-0.34</td>
<td>.117</td>
</tr>
<tr>
<td>C score</td>
<td>-2.57</td>
<td>-0.47</td>
<td>.029</td>
</tr>
</tbody>
</table>

Note. p value is set at .008. D = mean difference. d statistic = Cohen’s d effect size. IM = Impression Management Scale; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness; SN = Subjective Norms; AF = Attitudes toward Faking; IC = Identification and Consequences of Faking.
The present study also sought to examine whether certain warnings evoked more negative reactions in participants than other warnings. Hypothesis 8 stated that there would be more participants who have negative reactions to the IC condition than to the SN warning and the AF warning. In order to analyze this, a one-way analysis of variance (ANOVA) was conducted. The main effect of Reactions to the Selection System was significant at the .05 level, $F(4, 211) = 3.30, p = .012, \eta^2 = .06$, indicating that scores on the Reactions to the Selection System variable were significantly different across warning conditions. Means and standard deviations of measured variables on the Reactions to the Selection System scale are presented in Table 7. In order to determine if the hypothesized condition combinations were the cause for the significant effect, planned comparisons were analyzed. Specifically, the comparison between the IC condition and the SN warning was examined as well as the comparison between the IC condition and the AF warning. The planned comparison indicated that there was no significant effect between the IC and the SN warnings, $p = 1.000$, as well as no significant effect between the IC and the AF warnings, $p = .724$.

Although the apriori comparisons were not significant, the remaining comparisons were examined under the Bonferroni framework to determine which group difference was the source of the significant ANOVA. The comparisons showed that there was a significant mean difference at the .05 level in scores on the Reactions to the Selection System scale between the Fake Good control group and the Honest control group, $p = .038$. The comparisons also showed that there was a significant mean difference at the .05
level in scores on the Reactions to the Selection System scale between the FG control
group and the IC condition, p = .021. Since the participants were not instructed to stop
pretending they were applying for a job and to stop taking the warnings or control
instructions into consideration before responding to the Reactions to the Selection System
questionnaire, they might have still been “faking good” on this dependent variable. This
limitation in the study could serve as an explanation for the significant effects observed
on the Reactions to the Selection System Scale.

Table 7

Means, Standard Deviations, and Sample Size of Warning Conditions on Reactions to the
Selection System Scale

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG condition</td>
<td>46</td>
<td>47.11</td>
</tr>
<tr>
<td>HC condition</td>
<td>39</td>
<td>43.46</td>
</tr>
<tr>
<td>SN condition</td>
<td>42</td>
<td>45.07</td>
</tr>
<tr>
<td>AF condition</td>
<td>40</td>
<td>45.65</td>
</tr>
<tr>
<td>IC condition</td>
<td>49</td>
<td>43.45</td>
</tr>
</tbody>
</table>

Note. FG = Fake Good; HC = Honest Control group; SN = Subjective Norms; AF = Attitudes toward Faking; IC = Identification and Consequences of Faking.
Chapter 6
DISCUSSION

The results of the present study offer important evidence which can be used in support of the usefulness of warnings against faking behavior on personality tests. The main effect of warning condition was non-significant when looking at Impression Management scores alone, since scores on the Impression Management scales were not statistically lower for participants who were shown warnings than for participants who were in the unwarned, or control conditions. However, when determining faking behavior by examining scores on the NEO-FFI personality scale, significant differences were found between personality scores of participants who were randomly assigned to specific conditions and scores of participants who were randomly assigned to the control conditions. However, significance was not found to indicate that participants had more negative reactions when presented with certain warnings than with others.

The main analyses began by first examining whether an overall significant effect on the IM scale and personality scale dependent variables was present after random assignment to the different warning conditions. A significant main effect of warning condition was found, indicating participants who had been assigned to the different conditions had significantly different scores on the dependent variables. These results indicate the presence of faking. This overall main effect of warning supports prior research by Griffith et al. (2007) which concluded that applicants have the capability to fake non-cognitive measures and that at least some applicants fake in a selection setting.
Since the overall analyses only indicated that there was a main effect of warning condition present, follow-up analyses were conducted to examine specifically which warning conditions had significant effects on which dependent variables. First, the two control groups, Fake Good and Honest, were compared in order to determine if the participants had been following the instructions and manipulations when filling out the questionnaires. The comparison was significant for two of the personality variables, Neuroticism and Openness. The comparison was not significant for the Impression Management scale, indicating that the manipulations were not effective for this scale. The significant effects indicate that participants who were assigned to the Honest control group were being more honest in their responses by presenting a more neurotic image than those assigned to the Fake Good control group and that those assigned to the Fake Good control group were faking their answers by presenting a more open image than those assigned to the Honest group. These results were expected since Neuroticism is regarded as a “bad” personality trait, and Openness as a “good” personality trait. However, it was expected that those who were assigned to the Fake Good group would fake their answers on the other “good” personality traits as well. This is because previous research has found that applicants inflate their scores to a much larger degree on the constructs which they view as being particularly desirable by employers, such as Conscientiousness (Birkeland et al.2006). Even though not all differences were significant, all differences were in the expected direction. Hypothesis 1 was supported since there was a greater tendency to fake in the FG group.
The Fake Good control group was compared next with each of the three experimental warning conditions, Subjective Norms, Attitudes toward Faking, and Identification and Consequences of Faking, in order to further examine if the overall significant main effect of warning condition can also be partly attributed to any differences on dependent variable scores between the control group and the experimental groups. There were significant differences in scores on the Neuroticism and Openness personality scales between the Fake Good control group and the Subjective Norms condition, indicating that participants who were shown the Subjective Norms warning were being more honest in their responses since they indicated a higher level of neuroticism and a lower level of openness than those assigned to the Fake Good condition. A greater tendency to fake was shown by participants assigned to the Fake Good control group, therefore, Hypothesis 2 was supported. These results support previous research by McFarland and Ryan (2006) which shows that faking is directly related to an individual’s subjective norms regarding faking. The authors explain that when individuals perceive that important people in their lives would disapprove or discourage the behavior in question, such as faking, they are more likely not to engage in that behavior.

The comparison also detected significant differences in scores on the Neuroticism personality scale between the Fake Good control group and the Attitudes toward Faking condition, indicating that participants who were shown the Attitudes toward Faking warning were being more honest in their responses since they indicated a higher level of neuroticism than those assigned to the Fake Good control group. Since a greater tendency
to fake was demonstrated by those assigned to the Fake Good group, Hypothesis 3 was supported. These results support previous research by Hurtz and Hurtz (2005). The authors found that many individuals who have negative beliefs about faking, such as not being qualified for the job and not being able to meet expectations, getting caught, not getting the job, and having the risk of getting fired later on, do not engage in faking behavior as much as those who possess positive attitudes toward faking.

Also detected by the comparison were significant differences in scores on all of the personality variables, Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, between the Fake Good control group and the Identification and Consequences of Faking condition. These results indicate that participants who were shown the Identification and Consequences of Faking warning were being more honest in their responses since they indicated a higher level of neuroticism and a lower level of extraversion, openness, agreeableness, and conscientiousness than those assigned to the Fake Good control group. These results support Hypothesis 4. These results support past research by Dwight and Donovan (2003) that found scores which indicated faking behavior were statistically lower for participants who were shown the Identification and Consequences of Faking warning than for unwarned participants.

Comparisons were conducted to further examine the effectiveness of the experimental conditions on deterring faking behavior. Specifically, the experimental conditions were compared with the Honest control group. No hypotheses were tested here but it was expected that no significant differences would be found since the experimental conditions were predicted to discourage participants from faking. Therefore, participants
assigned to the experimental conditions were expected to provide similar results as those assigned to the Honest control group. As anticipated, there were no significant differences in scores on any of the dependent variables between participants assigned to any of the experimental conditions and those assigned to the Honest control group.

Specific comparisons between the experimental conditions and the control groups have shown that the presentation of a warning, as opposed to no presentation, does reduce the tendency for test-takers to fake their behavior on personality measures. Further comparisons were made in order to determine which warning is the most effective at deterring faking. Therefore, the warning conditions were compared to each other. First, the Subjective Norms condition was compared to the Identification and Consequences condition. There were no significant differences, indicating that neither one of the warnings is more effective at reducing faking behavior on personality tests than the other. These results do not support Hypothesis 5, which predicted that the Identification and Consequences warning would be more effective than the Subjective Norms warning at deterring faking behavior.

The Attitudes toward Faking warning and the Identification and Consequences of Faking warnings were compared next. Again, the results were not significant for any of the dependent variables, indicating that neither one of the warnings is more effective at reducing faking behavior than the other. These results do not support Hypothesis 6, which predicted that the Identification and Consequences warning would be more effective at reducing faking behavior than the Attitudes toward Faking warning. The comparisons between the Identification and Consequences condition and the Subjective
Norms and Attitudes toward Faking conditions do not support previous research by Dwight and Donovan (2003) which found that, out of all the warnings they tested, the Identification and Consequences warning was significantly more effective than the others at reducing faking behavior.

Another experimental condition comparison was made, this time between the Subjective Norms warning and the Attitudes toward Faking warning. Similarly to the other experimental condition comparisons, results were not statistically different, indicating that neither one of the warnings is more effective at reducing faking behavior than the other. These results do not support Hypothesis 7, which predicted that the Attitudes toward Faking warning would be more effective at reducing faking behavior than the Subjective Norms warning.

Although the results of the comparisons between the experimental conditions indicated that all three experimental warnings are equally effective at reducing faking behavior, when comparing all three experimental warnings to the Fake Good control group, it is clear that the Identification and Consequences of Faking warning effectively reduces faking on more dependent variables than the Subjective Norms and the Attitudes toward Faking warnings. The reason for this discrepancy among the comparisons is the nature of the analyses. The comparisons between the experimental conditions and the Fake Good control group yielded higher mean differences because the control group was set up as a condition that indicated a high level of faking. This was done in order for the effects obtained through the comparisons between this condition and conditions in question to be clearer. None of the experimental conditions were set up in such an
extreme manner and that is why no significant mean differences between them were detected.

Past research has shown some concern for the use of warnings. The main concern regards test-taker reactions to the warnings and whether the presentation of a warning would instill the feeling of not being trusted by the organization, which in turn, would turn applicants away from wanting to be a part of the organization. In order to test this in the present study, mean differences between all of the conditions on the Reactions to the Selection System dependent variable were examined. Since the “selection system” includes the warnings, differences on this measure were thought to indicate differences in reactions to the warnings in addition to reactions to the personality scale. Results showed that there was a significant main effect of warning condition on the dependent variable, indicating that participants assigned to different conditions reported significantly different reactions to the Selection System. Upon further examination of the results, it was determined that the experimental warning conditions were not the cause for the main effect, therefore, not causing significant test-taker reactions. Due to this, Hypothesis 8, which predicted that the Identification and Consequences warning would cause more negative test-taker reactions that the Subjective Norms and Attitudes toward Faking warnings, was not supported. These results do not support past research by Converse et al.(2008), which found that test-takers who were given negative warnings (such as the Identification and Consequences warning) reported higher test-taking anxiety than those given a positive or no warning. The results are more in line with research by McFarland
(2003), whose findings suggest that warnings do not generally have a negative effect on test-taker reactions.

Implications and Future Research

The present study makes several contributions to our understanding of the use of warnings in personnel selection contexts. First, it was found that certain warnings do have an effect on the way participants respond to personality questionnaires and they do reduce faking behavior. This information is valuable for any company or agency that utilizes personality tests as part of their employee selection process. This is because faking behavior could be a manifestation of low integrity and therefore detrimental to future performance. However, further studies have to be conducted with a real job applicant sample in order to confirm the present study’s findings and whether they are generalizable to real life settings.

The present study adds to the literature a combination of warnings which has not been previously examined, along with test-taker reactions to those warnings. The current study also compares the experimental warnings to two different control groups in order to have two sources of data confirming the results. This has not been previously done with the same variables used in the current study. The present study also had two means of detecting faking behavior: the BIDR-IM and the NEO-FFI. This is a strength of the study because faking behaviors that were not detected by the BIDR-IM were detected by the NEO-FFI. It is possible that the presence of faking behavior has gone undetected in studies that used BIDR-IM scores only to determine the presence of faking behavior.
Further studies have to be conducted to examine whether warnings used to deter faking behavior actually do have a negative impact on the selection process since existing research appears to be split on the topic; some supports this notion and some does not. It is unclear whether the non-significant findings regarding test-taker reactions in the present study are due to experimental error or limitations such as the sample of participants used, since it was expected that the Identification and Consequences warning would cause negative reactions in test-takers. This is further discussed in the limitations section. More research should explore the possibility of possible detrimental effects of warning. This is important because even if a warning is effective at reducing faking behavior, an employer might decide it is not worth using it if it turns away potentially good job candidates. However, further research should utilize a real job applicant sample in order for the results to be generalizable to the real life applicant selection process.

Further studies should also focus on finding the best measure of detecting faking behavior. The analyses in the current study indicated that the BIDR-IM scale, which was developed to identify whether responders are being dishonest in their responses, did not detect any faking behavior in the current study. This might have been due to the sample used in the study or to experimental error. However, it might also be the case that the items on the BIDR-IM are too transparent or obvious as to what they are measuring. Also, they might not have blended in well enough with the NEO-FFI items. Further research should explore these possibilities and alternative faking detecting measures should be explored.
Further, faking behavior was detected on some of the personality scales more than others. For example, almost all of the comparisons conducted identified faking behavior on the Neuroticism scale, whereas only a couple of the contrasts identified faking behavior on the Extraversion scale. Further research should examine whether certain scales on the NEO-FFI are easier to fake than others. Such information could be important when revisions are made to these personality scales. Detecting which items are too easy to fake could result in their revision or in them being removed from the scale altogether in the future.

**Limitations**

The present study had considerable limitations. One of the limitations was the sample of participants used. Undergraduate psychology students were used due to their immediate availability. This is limiting to the study because these participants did not have the same motivation driving them to respond to the personality questionnaire as real job applicants would have had. The participants were only instructed to pretend they were applying for a job they really wanted. The motivating factor for real job applicants would have been to obtain the job they were applying for, where in the present study the motivating factor was possibly winning a $25 prize. Also, it would be expected that real job applicants would have taken the warnings more seriously since they would have directly affected their chances of obtaining the job they were applying for. Further, using a real job applicant sample would have increased the external validity of the present study. It is uncertain that real job applicants would have had the same reactions to the
warnings since the environments and circumstances of them receiving the warnings during the present study were different than for a real job applicant.

Another limitation of the study has to do with the rating the participants provided on the Reactions to the Selection System scale. The participants were not instructed to stop pretending that they are applying for a job and to stop following the warning instructions before they were to begin responding to the Reactions to the Selection System questionnaire. Due to this, their responses on the Reactions to the Selection System questionnaire might have been influenced by the previous instructions and warnings, which were only intended for the NEO-FFI and IM items. This might have been the reason why no significant effects were found when analyzing whether participants had more negative reactions when exposed to certain warnings, as was found in previous research.

The current study first and foremost alerts us of the problems of using self-report measures in the personnel selection process. The results indicated that participants were, to a certain degree, faking on the personality measure administered to them and although the current study provides evidence for the usefulness of warnings, faking behavior was not completely eliminated by utilizing warnings. When making employment decisions, employers have to be familiar with the potential consequences of using self-report measures and base their employment decisions on the results of personality measures only when used in conjunction with other selection methods.
### APPENDIX A

**BIDR Version 6 - Form 40A**

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

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly Disagree</td>
<td>Disagree</td>
<td>Neutral</td>
<td>Agree</td>
<td>Strongly Agree</td>
</tr>
</tbody>
</table>

___ 21. I sometimes tell lies if I have to.
___ 22. I never cover up my mistakes.
___ 23. There have been occasions when I have taken advantage of someone.
___ 24. I never swear.
___ 25. I sometimes try to get even rather than forgive and forget
___ 26. I always obey laws, even if I'm unlikely to get caught.
___ 27. I have said something bad about a friend behind his/her back.
___ 28. When I hear people talking privately, I avoid listening.
___ 29. I have received too much change from a salesperson without telling him or her.
___ 30. I always declare everything at customs.
___ 31. When I was young I sometimes stole things.
___ 32. I have never dropped litter on the street
___ 33. I sometimes drive faster than the speed limit.
___ 34. I never read sexy books or magazines.
___ 35. I have done things that I don't tell other people about.
36. I never take things that don't belong to me.
37. I have taken sick-leave from work or school even though I wasn't really sick.
38. I have never damaged a library book or store merchandise without reporting it.
39. I have some pretty awful habits.
40. I don't gossip about other people's business.

Scoring key for BIDR Version 6 - Form 40A

Impression Management (IM): Items 21 - 40

Dichotomous Scoring procedure

1. Reverse the Likert ratings for the items indicated above.
2. Add one point for every '4' or '5' on the scale.

The minimum score is 0; the maximum is 20.

Reliability: Typical alphas are .77-.85.

Norms: Means and standard deviations for 177 UBC undergraduates under two scale formats and two instructional sets.

<table>
<thead>
<tr>
<th>5-point scale</th>
<th>Males (122)</th>
<th>Females (248)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respond Honestly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>5.5 (3.5)</td>
<td>6.1 (3.6)</td>
</tr>
<tr>
<td>Play Up Your Good Points</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>11.8 (4.3)</td>
<td>12.5 (4.5)</td>
</tr>
</tbody>
</table>
APPENDIX B

Reactions to the Selection System Scale

1. I was treated respectfully
2. They asked too much personal information
3. My rights as an applicant were respected
4. It was an invasion of my privacy
5. The process was too subjective and allows [the company] to justify any hiring decision that is made
6. It seemed fair
7. It made me feel that [the company] is concerned about hiring the most qualified applicants
8. [The company] seems concerned about hiring employees who will enjoy their work
9. Most of the information requested seemed relevant for the job
10. It was an appropriate way for [the company] to make employment decisions
11. The care [the company] takes in hiring employees makes me want to work there.
APPENDIX C

Control Groups’ Instructions and Experimental Groups’ Warnings

Fake Good (FG) control group instructions:

We are interested in your capability to present yourself in the best possible light on this measure. We want you to respond to the items in a way that makes you look best in order to increase your chances of getting hired. While completing the NEO-FFI, please keep in mind that the more impressive you look on this measure, the better your chances of getting the job will be. What we would like you to do is make yourself look as if you possess all of the qualities that you believe we look for in the applicant we would like to hire. Although you may not always be sure how to respond to these items in order to appear as impressive as possible, just do your best. So remember, respond to each item as you think we believe an extremely impressive applicant would respond.

Honest (HC) control group instructions:

While completing the NEO-FFI, please keep on mind that it is being used solely for research purposes, and all we want you to do is answer the items as frankly, openly, and honestly as you possibly can. Your answers to the questionnaire will not be used to make any psychological evaluations about you and will not be shared with anyone else. Please give as truthful and accurate a description of your attitudes, opinions, preferences, and behaviors as possible and remember to answer as honestly as you can.
Subjective Norms of Faking (SN) condition warning:

At our organization, we believe your potential coworkers, supervisors, and other individuals (that you may have the pleasure to work with) value honesty and openness. Therefore, we look for honesty and openness in anyone seeking to join the team as well, in order to help provide a better match between applicants and jobs.

Attitudes toward Faking (AF) condition warning:

Since providing inaccurate answers or lying when answering the items on the measure often leads to getting caught or not getting the job, we discourage applicants from this type of behavior. This type of behavior leads to applicants receiving positions for which they are not qualified and for which they will not be able to meet requirements or expectations. Providing inaccurate answers on the measure could give you a bad reputation, and if hired, will only increase your chances of getting fired later on.

Identification and Consequences of Faking (IC) condition warning:

In order to minimize the possibility that applicants who provide inaccurate answers or lie when answering the items on the measure are hired, we use a special scoring system such that dishonest responses may reduce your total score and will not increase your chances of getting the job. More specifically, some items in this measure are designed to detect if you attempt to lie in your responses. Therefore, faking behavior is highly discouraged. Additionally, if our scoring system indicates that you have
provided dishonest answers, you will automatically be disqualified from the raffle to win the $25 prize.


