Applied H.R.M. Research, 1993, Volume 4, Number 1, pages 49-68.

Fakability of Paper and Pencil Honesty Tests

Johannes H. Snyman
New Mexico State University

The ability of pre-employment paper-and-pencil honesty tests to detect faking was examined. One hundred undergraduate students took three honesty tests in one of three conditions: Answer honestly, fake good, or simulate another person. Results indicated that honesty scales were fakable, but that the validity scales were extremely effective in detecting subjects' attempts at faking. Subjects in the personality simulation condition obtained scores similar to subjects in the answer honestly condition, indicating that the validity scales were not able to detect simulated responses.

With the increasing number of firms that use paper-and-pencil honesty tests comes the concern of their fakability: Can an applicant try to produce scores that will not be an indicator of his/her honesty and still be hired? Sackett and Harris (1984) identify eight tests that do not attempt to disguise their purpose, one test that has honesty items embedded in other items that are nonthreatening in nature, and five tests that utilize a social desirability scale to detect faking.

The question of fakability has been addressed by researchers using one of five research strategies: the contrasted groups approach, comparisons of attitude and admission scales with social desirability scales, partialling out the effects of social desirability from attitude and admission validity coefficients, test-retest reliability, and the effect of direct instructions to "fake good," or to "simulate honesty."

Contrasted Groups

This approach involves company test scores of groups known to be

honest with those of groups known to be dishonest. In one study, Ash (1974) compared the honesty scores of a group of job applicants and found that convicts scored significantly lower than employment candidates that were not recommended by the Reid Report. Based on this finding, he concluded that "...the RRI (Reid Report Inventory) is very resistant to 'faking good'" (p. 28). Morey (1981) also found a significant difference between tow criterion groups (an admission of theft and a convection for theft group) and their counterparts. Both groups' test scores were in the high risk category of the Wilkerson Pre-Employment Audit.

According to this strategy, the conclusion that honesty tests are not easily faked is based on the assumption that convicts would have tried to fake good to increase their chances of parole (Ryan & Sackett, 1987). According to Ash, although convicts have a tendency to try to maintain a crimeless record, their inner frame of reference is focused on ways "to beat the system" and " to commit the perfect crime." Without research into the motivations of convicts, this assumption is tenuous. Any number of motivations may be assumed. For example, Cunningham (1989a) hypothesized that dishonest individuals may try to outguess the test by admitting to criminal acts and attitudes; they may not take the honesty test seriously, believing that their scores will not be used in any decision; or, they may strongly believe that dishonest legitimizations are regarded as the norm in society. Sackett and Harris (1984) also mentioned that convicts could have elected to "fake bad" to please the researchers.

Comparison Of Honesty Scores With Social Desirability Scores

Using this approach, several researchers have correlated honesty scale scores with various social desirability or "lie" scale scores. Gough (1972) reported a mean correlation of .39 between the Personal Reaction Blank and the L, F, and K validity scales of the MMPI and the PDI Employment Inventory. A correlation of .30 was found between the Hogan Personnel Selection Series (Hogan & Hogan, 1986). Kochkin (1987) reported a correlation of .43 between the Reid Report and the Fake Good scale of the 16PF. Jones and Terris (1981) reported a correlation of .37 between the Personnel Selection Inventory and it's accompanying lie scale. Jones and Terris (1983a) also reported a correlation of .58 using the Personnel Selection Inventory and a response distortion scale. Martelli (1988) reported a correlation of .28 for males and a correlation of .27 for females using the Phase II Profile and the Marlowe-Crowne Social

Desirability Scale. Honesty tests have thus been found to correlate well with social desirability scales. This means that applicants who obtain passing scores on the theft attitude and admission scales tend to give socially desirable responses, that is, they tend not to give their honest attitudes toward theft, or to admit actual theft committed in the past. Sackett *et al.* (1989) therefore stated that there is a need for further research that can assess the effect of faking and social desirability on passing scores.

Partialling Out Social Desirability Effects

For studies using this approach, Sackett and Harris (1984) noted that subjects who conceal actual theft and who do not admit past theft can create a spurious relationship between theft admissions and honesty scores. The desire to appear socially desirable to an employer when a selection decision is made will inflate the correlation between theft admissions and honesty scores. Applicants with such a desire will not admit theft which will raise their honesty scores. Sackett and Harris (1984) called for research that would partial out the effects of social desirability. M.M. Harris (1980), cited in Sackett *et al.* (1989), did such a study and found a small reduction in correlations, indicating tat test validity is not established on social desirability influences alone.

Test-Retest Reliability

Cunningham (1989a) cited a test-retest reliability study that he conducted using the Reid Report. Subjects were allowed to take the Report twice, but only an improvement of 3.2 points on a 70 point scale was found. They were not offered employment after the first sitting, but were allowed to take the Reid Report again for a different position. He speculated that "If someone were denied employment and felt they could improve their chances by faking, they probably would do so" (p. 3). Based on this assumption and the results, he concluded that faking would not improve one's score significantly. A problem with this interpretation is that subjects in the study were not instructed to fake good during the second administration. Thus, it is not known whether faking actually occurred.

Instructions To Fake Good

This last strategy uses a more direct approach. Subjects are assigned to several instructional sets and their test scores are compared. Ryan and Sackett (1987) created a test modeled after existing overt paper-and-pencil honesty tests and randomly assigned 148 college students to one of three instructional sets: answer honestly, fake good (i.e. attempt to appear honest), and respond as if you were applying for a job. The test consisted of a theft attitude and admission scale and a social desirability scale. Results indicated that it was easy to fake the theft attitude scale and to admit to lower dollar value theft or none at all. A difference was also found in the social desirability scale with students in the fake good condition obtaining lower scores than students in the other two conditions. Students asked to respond as job applicants had responses just slightly lower than students in the answer truthfully condition, but not low enough to conclude that they were able to fake the scales.

Cunningham (1989b) conducted two studies on the resistance of the Reid Report to faking. In the first study he compared scores of an honesty simulation condition group (i.e. subjects' attention was not focused on the dimension of honesty, only on completing the test) and a high seriousness fob candidate control group whose scores on the Reid Report were obtained from John E. Reid and Associates of Chicago. Subjects in the honesty simulation condition were 19% more successful than subjects in the low seriousness control group in obtaining higher scores, but 5% less successful in matching the scores of the employment candidates in the high seriousness control group. Although encouraging subjects to simulate a totally honest person produced higher scores, it did not increase the number passing the cutoff score of 50.

The second study employed four groups. Subjects in the control condition were asked to respond like a job condition similar to the honesty simulation condition of the first study, received more specific information about the nature of the test content. They were told that honest individuals were more punitive toward themselves and criminals and recommended punishment rather than forgiveness of criminal behavior. A projective information condition, the third group, received information that honest individuals tend to project an image of honesty and see others as honest as they are. The fourth group, a punitive and projective information condition, received the same instructions as the previous three groups combined. Subjects in all groups were also told tat if they obtained a score of 95% or above they would receive a bonus of $10. Re-

sults indicated that all the experimental groups obtained significantly higher scores than the control group, but that only 48.6 - 52.9% of subjects obtained a passing score of 50. This is still significantly lower than the employment candidates of the previous study. Based on these results, Cunningham concluded that the Reid Report is relatively resistant to faking.

Statement of Purpose

This study is designed to provide additional data on the fakability of honesty tests by using the direct approach. Building on the research of Ryan and Sackett (1987), Cunningham (1989b), and Streicher (1991), three direct instructional conditions were used: "answer honestly," "fake good," and "simulate an honest person you know," i.e. a personality simulation. The latter condition was utilized by Streicher (1991) to test the fakability of the Minnesota Multiphasic Personality Inventory (MMPI). She was interested in the ability of the validity scales to detect a personality simulation. Her findings indicated that the validity scales could not detect subject's simulation of their roommate's personality.

Hypotheses

It is hypothesized that:

1) Mean theft attitude and admission scale scores of subjects in the fake good condition will be significantly higher than scores of subjects in the answer honestly condition, indicating that successful faking has occurred on these scales.

2) Mean social desirability scale scores of subjects in the fake good condition will be significantly higher or lower (depending on how each testing company scores the scale) than scores of subjects in the answer honestly condition, indicating that the validity scales of honesty tests can detect faking.

3) Mean theft attitude and admission scale scores and scores on the social desirability scale for subjects in the personality simulation condition will not be significantly different from scores of subjects in the answer honestly condition, indicating that it is possible to fake honesty tests.

METHOD

Subjects

Subjects were 100 undergraduate students from a medium sized southeastern university who volunteered to take the three integrity tests described below. The majority of students were enrolled in a general education introductory psychology course and received extra credit for their participation. Their mean age was 18.94, with a range from 17-31. Sixty-one females and thirty-nine males participated in the study. Five percent of the subjects represented ethnic minorities(3 Blacks and 2 Asian Indians).

Test Materials

The selection of paper-and-pencil honesty tests for inclusion in this study was based on responses to a letter in which several companies interested in the study were asked to each supply 100 tests. Eleven companies reviewed by Sackett and Harris (1984) and Sackett et al. (1989) were mailed a copy of the letter. They were P.O.S. Corporation, London House, Inc., Milby Systems, Inc., Selection Research Publishing, Lousig-Nont & Associates, Inc., Psychological Security Corporation, Reid Psychological Systems, Psychometric Behavioral Group, Stanton Corporation, Wilkerson & Associates, Ltd., and Intergram, Inc. Only three, Lousig-Nont & Associates, Inc., Wilkerson & Associates, Ltd., and Reid Psychological Systems, agreed to participate in the study.

Phase II Profiles. The Phase II Profile, developed by Lousig-Nont & Associates, Inc., is a 117 item true/false and multiple choice overt integrity status inventory that assesses an individual's attitudes toward theft and his/her admissions of theft and other deviant behaviors. It is unique in that it measures six critical areas of a person's attitude towards honesty: the ability to rationalize dishonesty, how often a person thinks or plans dishonest acts, basic honest attitudes, basic dishonest attitudes, major admissions of dishonesty or a willingness to become involved in a dishonest act if given the chance, and minor admissions of dishonesty.

Combined with the attitude and admission items are 10 items that form a social desirability or "lie" scale. This scale detects individuals who

are not truthful in their responses. Although scores on the six attitude and admission sections are statistically combined with scores on the lie scale to give an overall score, separate scores for each of these sections are also given for closer scrutiny of an individual's attitude toward honesty.

Scoring the Profile can be accomplished by either using a personal computer and software provided by the company or by using the company's telephone services. The former method was used in this study.

Overall Profile scores are generated on a basis of 201 points. A score of 148 and above indicates an acceptable attitude regarding honesty. Applicants in this range can be considered as minimal security risks. A recommendation of "high desirability" is made provided no cautions are reported.

Cautions are questionable responses that do not affect the overall score. An applicant's confidence percentile is, however, lowered by any number of cautions. Lousig-Nont and Associates recommend that the applicant be interviewed so that clarification can be obtained regarding his/her questionable responses.

Scores between 136 and 147 with no cautions represent a marginal attitude toward honesty. This means the individual is recommended as "moderately desirable." Extreme caution should be exercised in considering the individual for employment because some answers reflect recent significant dishonest acts.

A score of 135 and below indicates a poor attitude toward honesty. This indicates that the individual has become involved in significant acts of dishonesty and theft. A recommendation of "low desirability" is given.

Three additional cautions may also generate a score of "low desirability." When the time taken to complete the Profile is less than 10 minutes or greater than 35 minutes a time caution is generated and a "low desirability" recommendation is given. If a Profile does not have at least six of the 10 validity (lie scale) questions, a "low desirability" recommendation is given. There are also many questions throughout the Profile that, if answered, are considered as inconsistent attitudes. If an individual answers enough of these questions, a "low desirability" recommendation is given.

For further interpretation of the overall scores, a confidence percentage is also generated. Individuals with passing scores should have a confidence percentage of at least 67. This indicates that the passing score is at least 67% of the time within accepted statistical norms.

Phase II Profile Addendum. Included with each Phase II Profile is a Phase II Profile Addendum, which is an alcohol and drug disclosure form. It is composed of five sections. Section one inquires as to how many times and how long ago a person has been involved in using, purchasing, and selling illegal drugs. Section two pertains to how many times and how long ago drugs interfered with the job or school. Section three pertains to how many times an applicant associated with people who use drugs and how long ago these associations took place. Section four places applicants in a managerial position and requests responses regarding their managerial attitudes toward drug use by people working for them. Section five pertains to alcohol use and asks questions about missing work or school due to alcohol, drinking on the job or at school, allowing the use of alcohol to interfere with work or school, and the number of times the person has been intoxicated in the last 12 months.

Every section is scored individually and is based on the individual risk factors associated with each section. A score of zero is an "acceptable score." "Marginally acceptable" scores fall between -1 and -5 and scores of -6 and below are "unacceptable." If an applicant marked all the "refuse to answer" choices, a "refuse" evaluation is given. This is to be interpreted as "disqualified or unacceptable." Also, if an applicant answered certain questions inconsistently, the question is invalidated and the section is marked "high risk," consequently making the entire Addendum "high risk."

An overall evaluation is also made for the entire Addendum. A "low risk" evaluation means that all sections are rated "low risk." A "moderate risk" means that all sections are rated "low" or "moderate risk." If any section is rated "high risk," the entire Profile is rated "high risk."

Although there is no scale embedded within the Addendum that may detect untruthful responses, applicants who are dishonest are thought to rarely pass the Phase II Profile Integrity Status Inventory (Lousig-Nont & Associates, 1989). It is therefore recommended that the two Profiles be used together. Lousig-Nont and Associates (1989) further caution that people have a tendency to minimize their admissions, making drug and alcohol use appear as if it took place long enough ago in order to score in the acceptable category.

Wilkerson Pre-employment Audit. The Wilkerson Pre-employment Audit is an overt paper-and-pencil honesty test that is divided into three sections. Section one is an extensive interview questionnaire that is

designed to cross-check statements made on a regular job application and to guide the face-to-face interview. It asks applicants for their name and address, job interests, education and training, physical handicaps, work experience, references, past residences, military experience, personal information such as the legality to work in the U.S. and their criminal record, driving record, and certification and understanding of section one. Wilkerson and Associates (1988) state that "The Audit is intended to be used as part of a hiring program, and not as the sole criteria for a decision to hire or not to hire,"(p. 9), hence the reason for the extensive interview questionnaire. Subjects in this study were instructed not to complete this section.

Section two consists of 100 multiple choice items that give scores related to social desirability (validity scale), attitudes toward and admissions of theft (integrity scale), and attitudes toward substance abuse (substance abuse scale). Items tapping this information are embedded among nonthreatening items (e.g., "Radio, television, and printed news is boring") to disguise the purpose of the Audit. These nonthreatening items are an attempt to reduce faking. Most items ask for a response that could be either agree, undecided, or disagree. Some items require a more specific response (e.g., outgoing and social, average, withdrawn and private). In most cases, the respondent has three choices.

Section three consists of questions about the applicant's use of drugs and alcohol. It also contains optional questions relating to theft behavior. Although subjects were asked to complete this section, their responses were not used in this study.

Scoring can be don two ways: telephonically and using computer software provided by the company. The latter was used in this study.

The validity scale provides information about the applicant's attempt to provide an inaccurate picture of him/herself. A score of 4 and below indicates little or no distortion, 5-7 a marginal distortion, and 8-10 a significant distortion. A score of 11 and above is invalid. This means that the Audit cannot be used as part of the selection process. On the integrity scale, a score of 30 and above places an applicant in the low risk range, 15-29 in the medium risk range, and 14 and below in the high risk range. A score below zero on the substance abuse scale means the applicant is at high risk. Zero-12 gives a medium risk evaluation and 12.01 and above is a low risk evaluation.

According to Wilkerson and Associates (1988), about 60% of all applicants score in the low risk range. If their validity score does not show

distortion, they should be considered for the job in question. About 22% of all applicants score in the medium risk range. These applicants should only be considered when there are not enough low risk applicants and when other screening measures, such as the polygraph (in industries allowed by law), or background checks, have been made. About 18% of all applicants score in the high risk range. This means that their scores are similar to the scores of convicted thieves and drug addicts. Thus, these applicants should not be hired.

In addition to the numeric scores on each of the three scales, any number of four standard caution paragraphs which point out areas in which the applicant's attitudes may need attention may appear. These paragraphs have nothing to do with the numeric scores or risk categories. They are to be discussed with the applicant once hiring has taken place.

Reid Report - modular version. The modular version of the Reid Report (copyrighted 1988) consists of four parts which may be administered as a single test or separately. Part one, entitled Attitude, contains 83 questions that reveal general attitudes about integrity or honesty and that predicts an applicant's likelihood of committing dishonest acts. Only 70 questions are scored. The remaining 13 questions are included for research purposes and for protection of the Report's scoring key. Part two, entitled Social Behavior, is designed to uncover significant admissions. It consists of 73 questions that inquire into an applicant's antisocial activities at work and elsewhere. Part three, entitled Substance Use, consists of 46 questions that measure an applicant's perception of drug and alcohol use within his/her peer group and his/her use of these substances on and off the job. Part four, entitled Personal Achievements, consists of biographical data covering previous job performance, attendance, safety record, circumstances of termination of employment, right to work in the U.S., and educational achievements. Parts one through three were administered as one test and part four was not used in this study.

Each part is scored as an individual unit. Results from part one yield three percentiles: a probability percentile, a percentile rank, and a punitive factor. The probability percentile indicates an applicant's likelihood of committing dishonest acts during the first year of employment. A score of less than 17% receives a Recommended evaluation. The percentile rank shows how the applicant's attitude toward honesty compares with all people that have ever taken the Report. The punitive factor measures an applicant's attitude towards punishment for theft. Information on

parts tow through four is evaluated according to job-relatedness and severity. Although parts two and three were answered, they were not used in this study.

An overall evaluation is based on the least favorable recommendation of any of the four parts. There are four possible overall evaluations: recommended, qualified recommended, not recommended, and no opinion. A recommended evaluation reflects positive attitudes toward honesty and positive information on the other three parts. Marginal attitudes toward honesty on part one, or attitudes and/or admissions of some likelihood of committing dishonest or counterproductive acts on the job receive a qualified recommendation. These applicants should be interviewed to clarify their responses. A Not Recommended evaluation reflects a significant likelihood of counterproductive or dishonest acts. Any one part that is evaluated as not recommended will result in an overall evaluation of not recommended. No Opinion means the applicant skipped some questions. Such applicants may be re-tested.

Three scoring methods are available: phone-in, mail-in, and personal computer scoring. For the present study, the mail-in method was used.

Procedure

Each subject was randomly assigned to one of three conditions: answer honestly (N=52), fake good (N=25), or simulate a very honest person you know (N=23). Subjects in the answer honestly condition were instructed to respond to the questions honestly. In the fake good condition, they were asked to make themselves look as good as possible on every question. Subjects in the third condition were asked to think of a very honest person they knew well, and to then respond to every question as if they were this person. To control for order effects, subjects completed each of the three tests in a counterbalanced order.

RESULTS

One-way analyses of variance (ANOVA) were conducted using the three instructional conditions as levels of the independent variable and scores on the variables presented in Table 1 as dependent variables.

As shown in Table 1, subjects in the fake good condition scored higher on the total theft attitude and admission score and confidence level of the Phase II Profile than subjects in the answer honestly and personali-

ty simulation conditions. Note that high total theft attitude and admission scores on the Phase II Profile do not necessarily mean that subjects have acceptable attitudes toward honesty or have committed less significant acts of theft in the past. Confidence levels of 67% and above, however, may be interpreted as acceptable.

Subjects in the personality simulation condition had the lowest total scores on the Phase II Profile, but slightly higher confidence scores than subjects in the answer honestly condition. In fact, they were seven percent more successful than subjects in the answer honestly condition, and nine percent more successful than subjects faking good in obtaining a high desirability recommendation (see table 2).

Subjects faking good were unable to pass the validity scale of the Phase II Profile, that is, they were unable to answer six out of ten questions correctly. In fact, the validity scale was extremely effective in detecting their attempt at faking good. No subject in this condition was able to obtain a high desirability recommendation. Although subjects in the fake good condition were able to obtain highly significant overall low risk recommendations (1 = low risk, 2 = moderate risk, 3 = high risk) on the Phase FII Profile Addendum, their scores would not be considered acceptable because they failed to pass the validity scale (Lousig-Nont & Associates, 1989). Subjects in the personality simulation condition passed the validity scale no more often than subjects in the answer honestly condition.

Overall scores of subjects in the personality simulation condition were slightly better (11 percent) than subject's overall scores in the answer honestly condition.

On the integrity scale of the Wilkerson Audit, subjects in the fake good condition were 53 percent more successful than subjects in the answer honestly condition and 66 percent more successful than subjects in the personality simulation in obtaining a low risk recommendation.

The substance abuse scale yielded results similar to that of the integrity scale. Subjects faking good obtained significantly more low risk recommendations than subjects in the other two conditions. Subjects in the answer honestly condition also obtained more low risk recommendations than subjects in the personality simulation condition (see tables 1 and 2).

However, the Audit's validity scale (i.e., combining the significant distortion and invalid categories of table 2) was effective 76 percent of the time in detecting a subject's attempt at faking good. By comparison to

Table 1

Means and ANOVA Results by Instructional Condition

Test	Condition			F	p
	Honest	Fake Good	Simulate		
Phase II Profile					
Validity Scale	9.64 (.20)	1.92 (.28)	9.13 (.29)	271.01	.0001
Rationalization	4.65 (.42)	2.76 (.60)	6.13 (.63)	7.67	.0008
Bad Attitudes	11.06 (.76)	5.32 (1.09)	12.44 (1.14)	12.39	.0001
Minor Admissions	.98 (.11)	.08 (.16)	.61 (.17)	10.42	.0001
Major Admissions	5.67 (.75)	2.00 (1.09)	7.04 (1.13)	5.82	.0041
Good Attitudes	24.40 (1.25)	37.64 (1.80)	24.26 (1.87)	20.50	.0001
Total Score	45.23 (12.80)	108.52 (18.46)	26.22 (19.24)	5.60	.0050
Confidence	31.56 (2.67)	57.04 (3.85)	37.57 (4.01)	14.95	.0001

Phase II Profile Addendum

Drug Use	-9.00	-17.20	- 5.74	.69	.5053
	(4.96)	(7.15)	(7.45)		
Drugs & Work	-3.31	0.00	- 4.39	3.79	.0261
	(.82)	(1.19)	(1.24)		
Associations	-10.75	- .28	- 9.04	20.04	.0001
	(.96)	(1.38)	(1.44)		
Managerial	1.58	3.40	2.13	19.83	.0001
	(.17)	(.24)	(.25)		
Alcohol Use	-13.60	0.00	-14.17	8.63	.0004
	(1.99)	(2.87)	(3.00)		
Overall	2.83	1.12	2.61	83.99	.0001
	(.08)	(.11)	(.12)		
Wilkerson Audit					
Integrity Scale	14.63	56.28	13.52	29.12	.0001
	(3.30)	(4.77)	(4.97)		
Drug Scale	-.13	25.27	1.63	17.65	.0001
	(2.52)	(3.63)	(3.79)		
Validity Scale	-6.48	8.68	-4.39	77.24	.0001
	(.71)	(1.02)	(1.06)		
Reid Report					
Rank	14.65	74.16	19.65	52.26	.0001
	(3.42)	(4.93)	(5.14)		

Probability	58.71 (4.49)	15.60 (6.48)	60.61 (6.75)	17.09	.0001
Factor	28.81 (4.12)	66.44 (5.94)	32.96 (6.20)	14.19	.0001

Note. Means of Honest and Fake Good and means of Fake Good and Simulate differ significantly at $p < .05$ for all variables except Drug Use. Means of Honest and Simulate and means of Drug Use are not significantly different at $p < .05$.

the Phase II Profile, subjects in the fake good condition were more successful twenty percent of the time in obtaining passing scores on the validity scale of the Wilkerson Audit. Subjects in the answer honestly and personality simulation conditions showed no significant differences between them. They exceeded the scores of the subjects in the fake good condition, passing the validity scale by a margin of seventy-six percent.

The Reid Report had results on the theft attitude and admission scale similar to that of the Phase II Profile the Wilkerson Audit. Subjects in the fake good condition lowered their scores significantly, indicating that they were more desirable candidates. They were recommended by the test 92 percent of the time (see table 4). Subjects in the personality simulation condition were slightly less desirable as job candidates than subjects in the answer honestly condition.

DISCUSSION

Results of subjects faking good on the Phase II Profile indicate that they were able to obtain significantly higher overall scores and confidence levels than subjects in the answer honestly and personality simulation conditions. Though faking did occur, the validity scale prevented subjects in the fake good condition from obtaining scores that are high enough to pass the cutoff score of the test. In fact, no one in this condition was able to pass the cutoff score. Thus, the first hypothesis was only partially supported in that subjects in the fake good condition obtained significantly higher scores than subjects in the answer honestly or personality simulation conditions. The second hypothesis, however, was significantly supported by subjects faking good on the Phase II Profile.

The integrity and substance abuse scales of the Wilkerson Audit were successfully faked 88 percent of the time by students in the fake good condition. However, the validity scale was 76 percent effective in detecting subjects' attempts at faking good. Hypotheses one and two were supported by the results of subjects faking good on the Wilkerson Audit.

The Reid Report was more susceptible to faking than the Wilkerson Audit or the Phase II Profile. Subjects in the fake good condition obtained acceptable scores 92 percent of the time. Although these results support the first hypothesis, it is not consistent with the research done on the fakability of the Reid Report. Cunningham (1986), cited in Sackett et al. (1989), compared the results of college students instructed to fake good with the results of job applicants and found that subjects instructed to fake good were more prone to theft. A lack of randomization may have caused differences that may be attributed to factors other than faking (Sackett *et al.*, 1989). In another study (1989b), Cunningham found that students instructed to present themselves as being honest could not pass the Reid Report by comparison to job applicants. Again, the scores of the job applicant group were obtained from John E. Reid and Associates of Chicago, causing a lack of randomization. Since subjects in the present study were randomly assigned to conditions, no interpretational difficulties can be attributed to methodology. The conclusion that the Reid report can be faked under direct instruction may be evident in the present study, but should not be interpreted as indicating that a flaw exists in the Report's ability to detect faking. Future studies need to look carefully at the attitudes towards honesty of people in their teens and twenties, since the failure rate is so high (Sackett *et al.*, 1989), before any conclusive statements can be made regarding the fakability of the Reid Report. It may be necessary to use a social desirability scale with the Reid Report until the high failure rate can be explained scientifically.

There was no significant difference for the validity scales and the integrity and substance abuse scales between the answer honestly and personality simulation conditions. The validity scales were not able to detect the subjects' attempt at presenting the theft attitudes of persons they regarded as very honest. The third hypothesis was thus supported. However, the subjects' test recommendations were not different from the answer honestly subjects' test recommendations. High desirability or low risk recommendations were at best only twenty-six percent. This means that although subjects in the personality simulation condition were able to pass the validity scales, they were unable to pass the cutoff scores on each

of the three tests.

A new question emerges: how accurately did subjects in the personality simulation condition simulate an honest person they knew? They may have responded to the items according to their own honesty or they may have chosen someone close to them in age, whose responses may have been similar to their responses. Since no measures were obtained from the "very honest group," no correlation could be calculated to determine the degree of accuracy. Streicher (1989) paired subjects in her study on the fakability of the MMPI and found that subjects were able to simulate their roommate's personalities very accurately. There is therefore a need for further research that could pair subjects in the personality simulation condition with their chosen counterparts. This need becomes even more evident in an increasingly competitive job market that accelerates the use of honesty tests. Applicants for employment may regard their own honesty below average and may attempt to distort their responses by simulating the honesty of someone they may perceive as very or more honest them themselves.

According to the results of the present study, subjects in the answer honestly condition had a passing rate ranging from two percent on the Phase II Profile to twenty-five percent on the integrity scale of the Wilkerson Audit. The average of twenty percent across all three tests is consistent with previous research (Ash, 1972; Morey, 1981; Lousig-Nont & Associates, 1982a). Sackett et al. (1989) reported a range from 40 to 70 percent, inclusive of all ages. No published research to date, however, has addressed the high failure rate amongst people in their teens and early twenties.

A possible explanation for the low passing rate may be immaturity. Kochkin (1987) administered the 16PF and the Reid Report to 179 applicants (mean age = 30.7 years) for employment with an airline. He found that applicants recommended for employment by the Reid Report were more emotionally mature and stable; exhibited higher ego strength; were more conscientious about moral standards; were less self-indulgent; exhibited higher super-ego strength; exuded more self-confidence; were less guilt prone; were more controlled, compulsive, and apt to behave in socially desirable manners; were more relaxed; were less driven by nervous excitement; exhibited less poorly directed id impulses and anxiety; and were more venturesome and inhibited. Young people are still immature with regard to attitudes to theft. However, as they are the prime candidates for jobs that need some measure of honesty (Ryan &

Table 2

Percentages of Test Recommendation

Test	Condition		
	Honest	Fake Good	Simulate
Phase II Profile			
High Desirability	2	0	9
Marginal Desirability	17	4	0
Low Desirability	81	96	91
Phase II Profile Addendum			
Low Risk	2	88	13
Moderate Risk	13	0	13
High Risk	85	12	74
Wilkerson Audit Validity Scale			
Valid Scores	96	20	96
Marginal Scores	2	4	4
Significantly Distorted Scores	2	24	0
Invalid Scores	0	52	0
Wilkerson Audit Integrity Scale			
Low Risk	25	88	22
Medium Risk	23	4	35
High Risk	52	8	43
Wilkerson Audit Substance Abuse Scale			
Low Risk	33	88	26
Medium Risk	19	4	35
High Risk	48	8	39
Reid Report			
Recommended	19	92	26
Qualified Recommended	2	0	0
Not Recommended	79	8	74

Sackett, 1987), more research needs to be conducted to establish different norms for the young job applicant, without increasing the false positive rate.

Another explanation for the low passing rate is the consequences associated with risk taking behavior. Stewart and Hemsley (1984) stated that when consequences of risk taking behavior are severe, the willingness to engage in these behaviors will be low. Martelli (1988) found that students who perceived the consequences of cheating on an academic exam or stealing from work as severe scored higher on the Phase II Profile. Since no severe consequences existed with the present study, subjects may not have felt the need to present an honest profile of themselves. Furthermore, subjects in the personality simulation condition may have chosen models with honesty profiles closer to them in age, resulting in lower passing rates. If the consequences of not obtaining a passing score were severe, however, they may have elected to simulate the honesty of someone more mature and honest than themselves. Future research may present more meaningful results pertaining to young people's attitudes toward theft and the models they choose to simulate, if the ethical problems involved in creating circumstances in which the consequences of risk taking behavior are severe can be overcome.

REFERENCES

Ash, P. (1974). Convicted felon's attitudes toward theft. *Criminal Justice and Behavior, 1*, 21-29.

Cunningham, M. R. (1989a). Test-taking motivations and outcomes on a measure of on-the-job integrity: analyses of responses to the Reid Report. *Journal of Business and Psychology, 4*, 119-127.

Cunningham, M. R. (1989b). *Resistance to faking in integrity tests*. Unpublished manuscript, University of Louisville, Department of Psychology, Kentucky.

Gough, H. G. (1972). *Manual for the Personnel Reaction Blank*. Palo Alto, CA: Consulting Psychologists Press.

Harris, M. M. (1980). *A research statement on the Reid Report*. Unpublished manuscript, University of Illinois at Chicago.

Hogan, J., & Hogan, R. (1986). *Hogan Personnel Selection Series manual*. Tulsa, OK: authors.

Jones, J.W., & Terris, W. (1983a). Personality correlates of theft and drug abuse among job applicants. *Proceedings of the Third International Conference of the 16PF Test*, 85-94.

Kochkin, S. (1987). Personality correlates of a measure of honesty. *Journal of Business and Psychology, 1*, 236-247.

Lousig-Nont & Associates. (1989). *Phase II Profile Profile Administration Guide*. Las Vegas, NV: Lousig-Nont & Associates, Inc.

Martelli, T. A. (1988). *Preemployment screening for honesty: The construct validity, criterion-related validity, and test-retest reliability of a written integrity test*. Unpublished doctoral dissertation, Ohio University, Athens.

Morey, L. C. (1981). *Statistical properties of the Wilkerson Pre-employment Audit*. Houston, TX: Wilkerson & Associates.

Paajanen, G. E. (1987). *The prediction of counterproductive behavior by individual and organizational variables*. Unpublished doctoral dissertation, University of Minnesota.

Ryan, A. M. & Sackett, P. R. (1987). Pre-employment honesty testing: Fakability, reactions of test takers, and company image. *Journal of Business and Psychology, 11*, 248-256.

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

Streicher, P. J. (1991). Personality simulation as a response set on the M.M.P.I.. *Applied H.R.M. Research, 2*(1), 68-83.

Wilkerson, O. M. (1980). *The Wilkerson Pre-employment Audit*. Houston, TX: Wilkerson & Associates.

Wilkerson & Associates. (1988). *Audit Instruction Manual*. Houston, TX: author.