

Can Pre-employment Tests Identify White-collar Criminals and Reduce Fraud Risk in Your Organization?

*Connie O'Brien**

Introduction

Organizations struggle with effectively limiting the risk of white-collar crime. Personality tests and integrity tests are often used to screen new applicants for behaviors indicative of white-collar crime (Berry, Sackett, and Wiemann, 2007; Engleman and Kleiner, 1998; Sackett, Burris, and Callahan, 1989; Smith and Lilienfeld, 2013). Globally, annual white-collar crime losses were estimated to be \$3.7 trillion of the Gross World Product in 2013. Eighty-five percent of these losses are due to asset misappropriation (Association of Certified Fraud Examiners, 2014a). U.S. organizations lose approximately five percent of their annual revenues to white-collar crime (Association of Certified Fraud Examiners, 2014a). White-collar crime (WCC) is defined as an illegal act or series of illegal acts committed by non-physical means and concealment or guile, to obtain money or property, to avoid payment or loss of money or property, or to obtain personal advantage or business (Edelhertz, 1970). Many aspects of Edelhertz approach continues to be used in legal definitions of WCC (Podgor and Dervan, 2016). While WCC is the broadly defined term, fraud has come to be the popular term encompassing several illegal acts (Association of Certified Fraud Examiners, 2012). Fraud is defined as any intentional act committed to secure an unfair or unlawful gain (Forensic KPMG, 2006).

Despite the magnitude of this problem, it is estimated that only twenty percent of fraud is detected (Oliphant and Oliphant, 2001). Of that twenty percent, approximately only two-thirds are prosecuted (ACFE, 2014a). Due to the lack of detection and prosecution by organizations as a whole, employers have utilized a number of traditional techniques to pre-screen applicants and limit at-risk individuals from entering the organization (Brody, 2010; Henle and Gross, 2013). Brody (2010) reported that approximately ninety-six percent of all organizations use some technique to pre-screen dishonest job applicants, yet he also found traditional techniques (background investigations, reference checks, and resumes verifications) have had limited success. Another technique that is widely used by organizations is pre-employment tests.

Currently, U.S. organizations administer approximately three million pre-employment tests to applicants per year (Brody, 2010). Pre-employment tests are used to predict the likelihood of future deviant behaviors based on a number of criteria, such as personality traits, theft admissions, and integrity (Fine, Goldenberg, and Noam, 2015; Sackett et al., 1989; Sackett and Wanek, 1996). Despite their popularity, efficacy concerns have been raised regarding the ability of these tests to predict deviant behaviors, specifically fraud (Berry, Sackett, and Wiemann, 2007; Camara and Schneider, 1994; Coyne and Bartram, 2002; Lee, Ashton, and De Vries, 2005; Marcus, Ashton, and Lee, 2013; Martin, 1989; Murphy, 1993; Sackett, Burris, and Callahan, 1989; Sackett and Wanek, 1996; Van Iddekinge, Roth, Raymark, and Odle-Dusseau, 2012). These concerns may limit the reliability and effectiveness of pre-employment tests. Without additional research into the reliability and applicability of pre-employment tests to accurately identify at-risk hires, the benefits of these tests are unlikely to be fully realized (Camara and Schneider, 1994; Coyne and Bartram, 2002; Fine, Goldenberg, and Noam, 2015; Rieke and Guastello, 1995).

Taking these concerns into consideration, the goal of this study was to examine the following research question: How effective are pre-employment tests in identifying individuals with profiles indicative of a propensity for fraud?

Literature Review

Pre-employment tests are designed to pre-screen candidates for hire for things such as personality and deviant behaviors (Engleman and Kleiner, 1998; Sackett, et al., 1989). In the literature prior to 1990, integrity tests and personality test were often broadly grouped together in one category and referred to as integrity tests (Sackett, et al., 1989; Sackett and Wanek, 1996). After 1990, research further sub-divided integrity tests into two primary categories “overt” and “personality-based” integrity tests (Sackett et al., 1989). Overt integrity tests include direct theft omission questions, whereas personality-based integrity tests do not. Personality tests can include personality-based integrity tests and personality measures (Sackett et al., 1989; Sackett and Wanek, 1996). Overt-integrity tests are designed to assess admissions of historical theft and attitudes regarding theft or other deviant behaviors or illegal activities (Ones, 1993; Sackett, Burris, and Callahan, 1989). Overt integrity tests include questions that measure personality, but the primary constructs for analysis are behavior scenarios and overt omissions of theft such as, Have you ever stolen time from your employer? (Cullen and Sackett, 2004; Greitzer et al., 2010; Marcus, Höft, and Riediger, 2006; McCrae and John, 1992; Ones, Viswesvaran, and Schmidt, 2003).

A number of early studies attempted to analyze how integrity tests predict and explain the propensity for theft (Miners and Capps, 1996; Sackett, Burris, and Callahan, 1989; Sackett and Harris, 1984). They found that admissions of past wrongdoing correlated with integrity scores. They also discovered the design of integrity tests were evolving to incorporate a broader set of validation criteria. Consequently, the broader criteria actually detecting little actual theft; require larger samples to compensate for the low theft detection rates; had significant correlations with deviant behaviors instead of theft; and often had less than significant validity coefficients.

Integrity test providers continue to cite Ones (1993) and the operational validities within that study as foundational support for the effectiveness of integrity tests in practice (Arch Profiles, 2012). Ones (1993) examined over 600 validity coefficients from thirty-six available integrity tests in a comprehensive meta-analysis. Ones (1993) found the mean operational validity for both personality-based and overt integrity tests to be positive and substantial. They originally found integrity tests to be valid predictors ($p = 0.47$) of overall job performance for a composites of counterproductive behaviors on the job, which included theft and absenteeism, although these variables were not the primary focus of this study (Ones et al., 2003).

Subsequent studies by Ones and others revealed theft was less predictable than broad counterproductive behaviors and the validity was lower than initially reported. Ones and Viswesvaran (2001) focused on the incremental validity of personality measures used in the prediction of behaviors such as theft. After testing, the mean operational validity for prediction of theft was identified to be 0.13 for overt-integrity tests and 0.0 for personality-based integrity tests. This was considerably lower than the ($p = 0.47$) previously reported in Ones (1993).

Van Iddekinge et al., (2012) attempted to replicate Ones (1993), but were unable to. Van Iddekinge et al., (2012) found validity rates that were considerably lower than Ones (1993). They also found operational validity rates for deviant behaviors ranged from .06 to .27 instead of in the .30's as reported by reported in Ones (1993). Van Iddekinge et al., (2012) expressed concerns regarding some the methods and results of the Ones (1993) study. One concern, only ten percent of Ones (1993) data was from studies published in professional journals. Several studies were authored by the test publishers. Also, Van Iddekinge et al., (2012) were not able to obtain copies of the same tests in their study. The resulting analysis by Van Iddekinge et al., (2012) of the criterion-related validity revealed overt and personality-based integrity tests

have validity rates in all areas that are much lower than reported in Ones (1993) and Ones et al., (2003). The operational validity rates for deviant behaviors ranged from .06 to .27 instead of in the .30's.

Collin and Schmidt (1993) measured the personality differences between convicted white-collar criminals and non-offending, upper-level employees. They used a personality-based integrity test, the California Psychological Inventory, and a biodata scale to detect differences in personality traits and honesty. Based on a sample of 365 inmates and 344 employees, Collin and Schmidt (1993) found large psychological differences between white-collar inmates and upper-level employees. White-collar inmates were found to have lower levels of Conscientiousness, Responsibility, and Socialization (Collins and Schmidt, 1993), suggesting that white-collar inmates have a stronger inclination to be risk-takers, opportunistic, manipulative and unethical than the upper-level employees. Collin and Schmidt (1993) further described the white-collar inmates as self-reliant, irresponsible, and undependable. They found this combination to be indicative of serious future problems for the individual, including personal and financial difficulties (Collins and Schmidt, 1993).

Blickle, Schlegel, Fassbender, and Klein (2006) expanded on the Collin and Schmidt (1993) study. They measured the personality differences between convicted white-collar criminals and non-offending, upper-level employees. Blickle et al., (2006) found significant differences in personality traits of white-collar criminals and upper-level employees. Collins and Schmidt (1993) found white-collar inmates were often negatively associated with the dimensions of Socialization and Responsibility. This is characterized by the tendency to be undependable, self-centered, manipulative, opportunistic, and risk takers. Blickle et al., (2006) found white-collar inmates demonstrated similar characteristics: seeking higher amounts of pleasure, exhibiting lower degrees of self-control, and having greater difficulty in resisting temptation. These same traits have been associated with the high Emotional Stability and low Agreeableness (McCrae and Costa, 2010). In addition, they found low levels of Integrity and high levels of Narcissism in white-collar criminals. Unlike Collin and Schmidt (1993), Blickle et al., (2006) found white-collar criminals had high levels of Conscientiousness.

The goal of this study is to explore the differences between behavioral profiles of white-collar criminals and the general population. To further explore these differences, personality tests and overt integrity tests were selected for use in this study. Specifically, the FFM, NPI, and WINT overt integrity tests were selected.

Based on the literature review, the following behavioral relationships are proposed:

Proposition 1: Convicted white-collar criminals will demonstrate low levels of Integrity

Proposition 2: Overt-integrity tests will present a stronger indication for the propensity of deviant behaviors than personality-based tests

Research Model and Setting This study empirically tests the efficacy of pre-employment tests to elicit a profile of predictive indicators of white-collar deviant behaviors from within a population of white-collar inmates. The profiles were derived from a comparison to the general population (GP) as purported by the test publishers and research constructs and relationships as presented in the literature. The GP consists of a mixed group of male and female members, over the age of eighteen. They derive from a variety of backgrounds. None of the test publishers identified their GP as containing white-collar criminals or non-white-collar criminals.

The usefulness of this study is derived from studies such as Blickle et al., (2006); Collin and Schmidt (1993); Greitzer et al., (2010); Ones, (1993); Ones et al., (2003); and Sackett et al., (1989). They proposed: 1) that white-collar criminals will have personality and behavioral traits unique to WCC; 2) predictive personality and behavioral traits can be identified utilizing pre-employment tests; 3) predictive personality and behavioral traits can be utilized to limit at-risk hires; and 4) overt-integrity tests can be manipulated to fake honesty.

Method

This was an exploratory study, utilizing quantitative methods, to examine the efficacy of pre-employment tests as correlated and measured against a known deviant population of white-collar inmates. The study was conducted in two federal prison camps located in the eastern United States.

Sample Selection

Building off the studies of Blickle et al., (2006) and Collin and Schmidt (1993), this study will compare the profiles of a population of white-collar inmates within the federal prison system with the profiles of the GP as provided by the test publishers.

Recruitment

Arrangements, contacts, and permissions for this study were made directly through the chief psychologist and warden associated with each federal prison and the Federal Bureau of Prison Research. Official permission to commence with this project was granted by the Georgia State University Office of Human Research, the Federal Bureau of Prison Research Review Board, both wardens, both chief psychologists, and the regional director of prisons. Participants were contacted through a flyer at the participating federal prisons. Flyers were posted in common areas throughout the prisons explaining the general purpose, confidentiality, and date and time of the study. There were no personal inducements or rewards given for participation. Those volunteering and accepted for participation in this study had to: a) be currently serving time in the Federal Prison system for, b) a white-collar offense, and c) had to choose to voluntarily partake in this study.

Tests

Five instruments were used to discern the participants' opinions, personality, and integrity. The completion rate was 100% for the first four tests and eighty percent for the Fraud Motive Assessment (optional).

A demographic questionnaire was used to gather background information about the inmate, number of prior crimes committed, type of convictions, nature of convictions, and the violent or non-violent nature of those crimes, and to validate membership in the target population for this study. Although the population may appear to be a homogenous group of white-collar inmates, many factors may create variability within this population of interest. The personality inventory and the narcissism scale may enable identification of personality differences. The demographic questionnaire provided the ability to examine other differences in the inmates' personal background and the potential effect these factors may have had on the predictive indicators of WCC.

An overt integrity test, the Work Integrity Test (WINT) by ARCH Profiles (2012), was used to explore the relationship between WCC and pre-employment tests and measure theft. It combines five component scores: lenient attitude towards dishonest behavior, perceived frequency of dishonest behavior, rationalizing of dishonest behavior, self-reported dishonesty, and social desirability into a total Integrity score (Arch Profiles, n.d.). The standard WINT test scores range from zero to 100. The mean GP total integrity test score for the WINT is 69.65 ($SD = 16.51$) (Arch Profiles, 2012). According to test providers, scores greater-than forty generally indicate a potential for dishonest behavior, scores greater-than eighty generally indicate a strong potential for dishonest behavior (Arch Profiles, 2012). High scorers do not feel dishonest behavior should be punished, few people are honest, dishonesty can be justified, and are more likely to engage in dishonest behavior. The WINT contains transparent questions directly related to theft, which have been found to increase the probability of assessing the behavior of fraud.

The NEO-FFI-3 was used to explore the personalities of white-collar inmates. It is a sixty-question personality inventory that creates personality profile utilizing the five personality dimensions of the Five Factor Model (FFM) (Greitzer et al., 2010; Judge, Heller, and Mount, 2002; McCrae, Costa, Jr, and Martin, 2005; McCrae and John, 1992). FFM includes the traits of Neuroticism (N) or Emotional Stability (ES), Extraversion (E), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O) (McCrae, Costa, Jr, and Martin, 2005; Costa, Jr, and McCrae, 2013). Prior research has demonstrated that three of the five personality dimensions, Emotional Stability/ Neuroticism, Agreeableness, and Conscientiousness have a direct relationship to deviant work behaviors (Barrick et al., 2001; Berry et al., 2007; Mount, Ilies, and Johnson, 2006; Ones, 1993; Ones et al., 2003; Sackett and Devore, 2001; Salgado, 2002). Of these three, Conscientiousness has demonstrated the strongest relationship with deviant behaviors (Barrick et al., 2001; Berry et al., 2007; Ones, 1993; Ones et al., 2003).

Within the NEO-FFI-3 each trait is scored independently. Emotional stability/Neuroticism represent differences in an individual's experience with distress and the cognitive and behavioral styles that flow from distress (McCrae and John, 1992). High scorers have are less able to control impulses and cope with stress (McCrae and Costa, 2010). According to test providers, the mean GP score for Emotional stability/ Neuroticism is 20.8 ($SD = 7.7$) (McCrae and Costa, 2010). Individuals who score greater-than twenty-three in Emotional Stability/Neuroticism are considered sensitive and irrational (McCrae and Costa, 2010). Lower scorers, with less-than sixteen, are considered secure, emotionally more stable, and able to face stressful situations (McCrae and Costa, 2010).

Agreeableness is considered a reflection of an individual's willingness to fight for his or her own interests and eagerness to help others. According to test providers, the mean GP score for Agreeableness is 30.8 ($SD = 6.9$) (McCrae and Costa, 2010). Individuals who score greater-than thirty-three in Agreeableness are found to be more flexible, trusting, good-natured, cooperative, forgiving, and tolerant (Barrick and Mount, 1991) and are described as compassionate and sympathetic (McCrae and Costa, 2010). Lower scorers, less-than twenty-seven, are considered antagonistic, competitive, and proud (McCrae and Costa, 2010).

Conscientiousness and the development of it create the ability to resist temptations, manage desires, control impulses, and organize and carry out tasks (McCrae and Costa, 2010; McCrae and John, 1992). Individuals need conscientiousness to hold impulsive behavior in check. In addition, they need the will to achieve in order to direct and organize behavior. High scores in Conscientiousness reflect an individuals' dependability, also reflects an individual's tendency to be hard working, achievement-oriented and persevering (Barrick and Mount, 1991). According to test providers, the mean GP score for Conscientiousness is 32.6 ($SD = 7.3$) (McCrae and Costa, 2010). Individuals who score high in Conscientiousness, less-than thirty-five, are considered reliable and well organized (McCrae and Costa, 2010). Lower scorers, less than twenty-nine, are considered disorganized and easy-going (McCrae and Costa, 2010). Low levels of Conscientiousness have also been linked to low levels of integrity and unethical behavior (McCrae and Costa, 2010).

The FFM has consistently provided a "meaningful taxonomy for studying individual differences" in personality (Barrick and Mount, 1991) which, when evaluated, has been a valid predictor of counterproductive behaviors (Blickle, Schlegel, Fassbender, and Klein, 2006; Collin and Schmidt, 1993; Greitzer, Kangas, Noonan, and Dalton, 2010; Ones et al., 2003; Salgado, 2002) and compulsive behavior (Mowen, 2000) which is a focus of this study.

The Narcissistic Personality Inventory (NPI) is a forty-item, forced-choice questionnaire designed to measure individual differences in Narcissism as a personality trait. The NPI combines the seven component scores for authority, exhibitionism, superiority, entitlement, exploitativeness, self-sufficiency, and vanity into one overall Narcissism score (Raskin and Terry, 1988). According to test providers, the mean Narcissism score for the NPI is 15.55 ($SD = 6.66$) (Raskin and Terry, 1988). While there is not a recognized cutoff point between "normal" and "excessive" or "subclinical" for trait Narcissism (Johnson,

Kuhn, Apostolou, and Hassell, 2012), the most widely used cutoffs for both men and women are (low ≤ 14 NPI; ≥ 20 = high NPI) (Emmons, 1984; Rhodewalt and Morf, 1998). Scores range from zero to forty. Responses are measured on a five-point scale (1 = *strongly disagree*, 5 = *strongly agree*).

Following the completion of the pre-employment instruments, participants were asked to complete a Fraud Motive Assessment. The Fraud Motive Assessment is a fourteen question, open-ended questionnaire designed to gather the participants' opinions and perspectives on their current white-collar crime and white-collar prevention.

The psychometric profiles created, along with individual responses to the demographic questionnaire and Fraud Motive Assessment, were used to assess the efficacy of the pre-employment tests and the relationships between deviant behavior and the proposed study variables.

Study Procedures

Participants were assigned an individual identification number to be used on the demographic questionnaire, the pre-employment instruments, and the Fraud Motive Assessment. This number was for internal tracking purposes only as no other identifying information was associated with these instruments. All instruments were administered via pencil and paper, on site, within each federal prison camp and supervised by the principal investigator. The testing time for the study took approximately sixty minutes. The total number of participants was twenty. No incentive was offered for participation. Participation was voluntary and confidential. No participants dropped out of the study. All tests and questionnaires were in English.

Results

The completed pre-employment instruments were analyzed. The means, standard deviations, and correlations of the study variables were analyzed in order to draw conclusions about the efficacy of these instruments' ability to identify indicators of WCC. The Fisher Exact Test method was employed to determine the statistical significance of data with regards to the research question. Fisher Exact Test is a nonparametric statistical test that measures the difference, or goodness of fit, between two or more samples to determine if the difference in the distributions of the samples is due to chance or limited sample sizes. T-tests and other nonparametric measures were also used to explore the relationships among the data. All of the data were compared to the GP on an individual participant level and as a whole, per instrument. This information was analyzed utilizing non-parametric tests within SPSS and Excel.

Demographics

The sample comprised twenty male white-collar inmates ($N=20$). Demographic information revealed that eighteen of the twenty participants were age forty or older and eleven were over the age of fifty-four, resulting in mean age of 52.2 years. At the time of their crime, fifteen were employed full time in a position of trust, such as owner (twelve) or upper management (three). The majority were college educated holding either a bachelor's degree (eight) or graduate degree (six). Fourteen (seventy percent) were considered "first time offenders." This percentage is slightly lower than the ACFE's purported eighty-five percent. Of the six prior offenses cited, five were white-collar offenses. The sixth was an unrelated non-violent offense. No participant reported convictions for violent offenses, past or present. Bank fraud, conspiracy, tax crimes, and wire fraud were the most common convictions. Note, while only twenty inmates participated in this study, some were convicted of multiple white-collar crimes. Therefore, there are more white-collar crimes reported than participants. In addition, two of twenty, or ten percent, of the participants were convicted of committing frauds while external to the organization and ninety percent were convicted of frauds while internal to the organization. [see Table I, pg 635]

Research Question Analysis

Non-parametric tests and other tests designed for small samples were used to examine the data as they relate to the proposed propositions since participation was low. Means, standard deviations, frequencies, *t*-tests, and correlations were used to evaluate the relationships.

An independent *t*-test compared the scores from each of the three pre-employment tests to the GP scores provided by the test providers. The scores from the sample were compared to the GP average. The *t*-scores, means, standard deviations, medians, range, and *N*'s are summarized in Table II. Based on the research, it was expected that white-collar inmates would present profiles that were significantly different from the GP in all five areas tested. Of the five variables measured, Emotional Stability/Neuroticism and Conscientiousness were significantly different from the GP at the $p < .05$ and $p < .10$ respectively. The white-collar population had lower than average level of Emotional Stability/Neuroticism as opposed to the high levels proposed in literature. In contrast, inmates presented higher than average levels of Conscientiousness as opposed to the low levels proposed in literature.

As for Integrity, this population was expected to have a lower level of integrity than the GP as represented by high total integrity scores. The average score was lower, although not significantly lower as expected. As a group, these results suggest that white-collar criminals may have behavioral profiles which are different from the GP and other predictors of deviant behavior. [see Table II, pg 636]

Proposition 1: Convicted white-collar criminals will demonstrate lower levels of Integrity than when compared to the Integrity scores for the GP

Individual Integrity scores as calculated by test publishers of equal to or greater than forty are indicative of potentially dishonest behavior. Integrity scores equal to or greater than eighty indicate a strong potential for dishonest behavior (Arch Profiles, 2012). Test providers make recommendations for hire based on the total Integrity score and other indicators. Low integrity scores should correlate with recommendations for hire and high scores should correlate with recommendations not to hire. Although, Sackett, Burris, and Callahan (1989) found integrity test recommendations for hire had significant correlations with deviant behaviors other than theft, which they were designed to detect.

Individual Integrity scores within this sample range from thirty-nine to eighty-eight. A *t*-test failed to reveal a significant difference in the study mean for Integrity scores ($M=67.25$, $SD=15.427$, $t(19) = -.696$, $p = .495$) compared to the GP mean ($M=69.95$, $SD=16.51$, $N=1672$) (ARCH Profiles, 2012) (Table II). The median was also considered. The median of sixty-nine was not significantly different from the GP mean. However, by sub-dividing the sample into groups, categorized by recommended for hire and not recommended for hire and comparing the groups means, a significant difference is revealed. The recommended for hire group ($M=58.78$, $SD=12.337$, $t(8) = -3.745$, $p = .006$) compared to the not recommended for hire group ($M=74.18$, $SD=14.593$, $t(10) = 3.500$, $p = .006$). Comparing the groups to the GP, a significant difference in the means was revealed for the recommended for hire group ($M=58.78$, $SD=12.337$, $t(8) = -2.717$, $p = .026$). The study mean for Integrity and the mean Integrity score for the recommendations for hire were found to be significantly correlated ($r = .510$, $p = .022$). These findings lend support to integrity test recommendations for hire related to Integrity scores and to the supposition that white-collar inmates demonstrate lower levels of Integrity. Based on these findings, Proposition 1 is supported. The findings should be explored further with a larger population to determine if the results are consistent.

Proposition 2: Overt-integrity tests will present a stronger indication for the propensity of deviant behaviors than personality inventories

O'Bannon, Goldfinger, and Appleby (1989) and Sackett and Harris (1984) found integrity tests struggled with theft detection and theft admissions validation. Participants in this study were asked directly about the dollar value, commission, and age of most recent theft act. All members of this sample were convicted white-collar inmates. Of the twenty participants, only eleven admitted to committing thefts as

adults. Six admitted to additional prior convictions. One participant even claimed to have never committed any form of theft.

Coyne and Bartram (2002) argued that individuals who were more honest were more often penalized by test providers. On the contrary, this study found six of the nine who admitted to recent thefts of \$1000 or more were recommended for hire. A significant correlation was found between the mean Integrity scores and prior convictions ($r = .535$, $p = .015$). Nine participants were recommended for hire and eleven were not recommended for hire. Of the nine participants who were recommended for hire, four had prior white-collar convictions. A significant correlation was also found between the amount stolen and inclusion in the recommended for hire group even with prior convictions ($r = .770$, $p = .006$). The mean Integrity score was found to be significantly correlated with the amount of money stolen ($r = -.501$, $p = .024$). Of the eleven not recommended for hire, only five admitted to recently committing acts of theft, three of which were greater than \$1000. A significant correlation was found between recommendations for hire and Integrity scores and recency of theft acts ($r = .510$, $p = .022$). [see Table III, pg 636]

These findings suggest that test providers may take theft admissions into account when making recommendations for hire. These findings also lend support to Coyne and Bartram's (2002) claim that job applicants may modify responses to improve their chances of obtaining employment. Overt integrity tests directly ask about theft and theft behaviors, but they do not validate participants' answers. Participants are not obligated to reveal historical thefts or dishonest behaviors.

A frequency test revealed that the WINT overt integrity test determined that eleven of the twenty participants (fifty-five percent) should not be recommended for hire. Using a cross tabulation, the integrity test recommended hiring four participants who had additional prior white-collar convictions and five without out ($p < .640$, Fisher Exact Test). Using admission criteria, this indicates a failure rate of forty-five percent for the integrity test in this study raising concerns regarding the validity the recommendations in general.

Two participants presented profiles of high Emotional Stability/Neuroticism, low Conscientiousness, and trait Narcissism. This personality profile did not match the purported profile of white-collar inmates in previous research. Of these two participants, only one presented with low Integrity ($p < 1.0$, Fisher Exact Test) and was not recommended for hire (WINT score = fifty-eight). This inmate also admitted to a recent theft. The second inmate was recommended for hire (WINT score = seventy-eight) and did not admit to a recent theft. Personality tests in this study did not identify a profile of any participants as an individual likely to commit deviant behaviors.

Based on the sample population of white-collar inmates tested, it was expected that this sample would elicit "profiles" indicative of the deviant behaviors, which both personality tests and integrity tests are designed to identify and which organizations seek to identify in order to limit risk. As indicated previously, Murphy (1989) found failure rates of pre-employment tests ranged from thirty percent to sixty percent. Failure rates for personality tests in this study were found to be higher than for overt integrity test. While both tests failed to identify all members of the white-collar population as potential risks and the total integrity score proved to be an unreliable indicator of risk, overt integrity tests identified a significantly larger number inmates as potential risks than the personality indicators. Based on these findings, Proposition 2 is supported.

Discussion

Faking honesty and false positives continues to be an issue with integrity tests (Fine, S., J. Goldenberg, and Y. Noam, 2015; Sacket and Wanek, 1996). In order to improve their chances of obtaining employment, job applicants may modify responses to questions (Coyne and Bartram, 2002) especially if the applicants believed their dishonest answers could not be verified (Donovan, Dwight, and Hurtz, 2003). The WINT asked direct questions regarding the dollar value, commission, and age of the applicants most recent theft act (Table III). Only eleven of the twenty convicted white-collar inmates

(fifty-five percent), admitted to committing their current crime. For this population the false positives were higher than expected. It may be that the test itself is not sensitive enough to this type of population.

Organizations receive recommendations from the test providers to “hire” or “do not hire” based on the total integrity score. Eight of the nine inmates recommended for hire had integrity scores greater than forty, which according to Arch Profiles (2012), is indicative of dishonest behavior. The average Integrity score for the recommend for hire group was $M=58.78$. All eleven inmates not recommend for hire scored above 40, $M=74$. However, the distribution of scores for the two groups was similar: “hire” (thirty-nine to eighty-three) and “do not hire” (forty-four to eighty-eight). While the difference in the mean for the recommend for hire group and the GP average there was not a significant difference in the means of the do not recommend group or the overall integrity test score.

The recommended for hire group had a larger percent of participants who admitted to recent, large thefts compared to the do not hire group (Table III). A correlation revealed a strong positive relationship between the total integrity test score and the amount of money stolen for the do not hire group, $r = .62, p < .05$. While the recommend for hire group did not reveal a significant correlation with the amount of money stolen $r = -.23, (p = n.s)$. It does raise the question, how are theft admission criteria used? Are the criteria applied consistently? Since all of the participants are convicted white-collar inmates serving federal prison sentences, it is presumed that all twenty should have elicited behavioral profiles that would have received do not hire recommendations. Ones et al., (2003) argued the narrower the criteria, the lower the predictive validity. This may have contributed to the inconsistent low integrity scores. It is also apparent that integrity tests do not measure honesty *per se* or make recommendation of integrity on easily identified or interpreted factors. Based on these findings, it can be derived that the total Integrity score alone is not a reliable indicator of integrity.

Despite the false positives, the WINT integrity test was able to identify eleven of the twenty white-collar inmates as lacking Integrity and not recommend them for hire. Failure rates and faking-honesty are still a concern, but in this study, the overt-integrity test was more effective than personality tests in identifying personalities with a propensity for deviant behaviors. Organizations need to be aware of the failure rates. Pre-employment tests, while able to identify some issues of deviant behavior, are not a fail-safe. They must be used in combination with other pre-employment screening tools. Based on the results of this study, an overt-integrity test may have a higher probability of detecting traits indicative of WCC than personality tests.

In summary, this study contributes to the body of knowledge through an examination of the relationships of personality traits, behavioral traits, and deviant behaviors of white-collar criminals. The findings highlight some traits and related issues which expand the understanding of white-collar criminals and pre-employment tests. This knowledge may encourage future research that may lead to more effective fraud risk mitigation.

Contribution to Practice

Integrity scores are based on the probability of an individual repeating similar behavioral traits and patterns in the future. An overt-integrity test is designed to identify those behaviors and make recommendations based on them. This study discovered there were inconsistencies with the theft admissions and recommendations for hire, total integrity score and recommendations for hire, and faking honesty. If these tests are to be affective in detecting and mitigating theft risks for organizations they need to improve the criteria used to identify and screen WC behaviors, as well as consistently apply the key indicators in the screening process. Additional research into WC related indicators might improve theft mitigation and hire recommendations. Organizations rely on these tests to provide candidate pools with a lower risk profile, not higher risk profiles.

Limitations

Methodologically, some limitations of the research must be noted. The most obvious limitation is sample size (N=20). Small sample sizes can result in false positives when conducting analysis. Also, the population sampled is a vulnerable population, therefore strict restrictions were imposed by the multiple Institutional Review Boards on this study regarding the questions, conversations, and interactions permitted with this population.

Conclusion

Fraud continues to be pervasive and expensive (ACFE, 2012; 2014a; Brody, 2010; Brody, Melendy, and Perri, 2012; Coyne and Bartram, 2002; Greitzer et al., 2010; Henle and Gross, 2013; MacLane and Walmsley, 2010; Perri, 2011; Perri and Brody, 2011; Van Iddekinge et al., 2012). As a result, organizations need to find ways to mitigate the risk of fraud. Test publishers continue to promote the abilities of pre-employment tests while ignoring the tests' shortcomings. Pre-screening job applicants with tools such as pre-employment tests may help organizations to reduce inaccuracies (Brody, 2010), but it may also expose the organization to additional risks (Lee et al., 2005). Due to this additional risk, it is important for organizations to keep these tests in perspective. They should not be the sole basis for employment decisions.

Despite these limitations, this research revealed significant insights into the relationship between pre-employment tests and WCC. WCC has a correlation with low-Integrity. Recommendations for hire are related to the total Integrity scores, amount of money stolen, and recency of theft. These relationships bring to question: What criterion are integrity tests measuring; theft, honesty, or integrity? Are they measuring the most reliable criterion for prediction of deviant behaviors? Although the failure rates for the overt-integrity test were found to be forty-five percent, it was found to be a better indicator of the propensity for deviant behavior than the personality measures, but is this reliable enough? This study adds empirical support to growing research on pre-employment tests and their ability to detect indicators of deviant behaviors, specifically fraud. This study also expands knowledge of the role of behavioral traits as predictive fraud indicators and their use in pre-employment tests.

References

- ARCH Profiles. 2012. *WINT (Work Integrity Test) Psychometric Report*.
- ARCH Profiles. n.d. *WINT (Work Integrity Test) – Operational Definitions*.
- Association of Certified Fraud Examiners. 2012. *ACFE Report to the Nations on Occupational Fraud and Abuse, 2012 Global Fraud Study*. Available from http://www.acfe.com/uploadedfiles/acfe_website/content/rtt/2012-report-to-nations.pdf
- Association of Certified Fraud Examiners. 2014a. *ACFE Report to the Nations on Occupational Fraud and Abuse, 2014 Global Fraud Study*. Available from <http://www.acfe.com/rtt/docs/2014-report-to-nations.pdf>
- Barrick, M. R., and M. K. Mount. 1991. "The big five personality dimensions and job performance: a meta-analysis." *Personnel Psychology* 44, no. 1:1–26.
- Barrick, M. R., M. K. Mount, and T. A. Judge. 2001. "Personality and performance at the beginning of the new millennium: What do we know and where do we go next?" *International Journal of Selection and Assessment* 9, no. 1-2:9–30.
- Berry, C. M., D. S. Ones, and P.R. Sackett. 2007. "Interpersonal deviance, organizational deviance, and their common correlates: a review and meta-analysis." *Journal of Applied Psychology* 92, no. 2:410.
- Berry, C. M., P. R. Sackett, and S. Wiemann. 2007. "A review of recent developments in integrity test research." *Personnel Psychology* 60, no. 2: 271–301.
- Blickle, G., A. Schlegel, P. Fassbender, and U. Klein. 2006. "Some personality correlates of business white-collar crime." *Applied Psychology* 55, no. 2:220–233.
- Brody, R. G. 2010. "Beyond the basic background check: hiring the "right" employees." *Management Research Review* 33, no. 3:210–223.
- Brody, R. G., S. R. Melendy, and F. S. Perri. 2012. "Commentary from the American Accounting Association's 2011 annual meeting panel on emerging issues in fraud research." *Accounting Horizons* 26, no. 3:513–531.
- Camara, W. J., and D. L. Schneider. 1994. "Integrity tests: Facts and unresolved issues." *American Psychologist* 49, no. 2:112.
- Collins, J. M., and F. L. Schmidt. 1993. "Personality, integrity, and white collar crime: A construct validity study." *Personnel Psychology* 46, no. 2:295–311.
- Costa, P. T., Jr., and R. R. McCrae. 2013. "The NEO™ Inventories: NEO™ Five Factor Inventory-3 (NEO™-FFI-3) Publisher: PAR." Available from <https://shop.acer.edu.au/acer-shop/group/NEOFFI-3/X>
- Coyne, I., and D. Bartram. 2002. "Assessing the effectiveness of integrity tests: A review." *International Journal of Testing* 2, no. 1:15–34.
- Cullen, M. J., and P. R. Sackett. 2003. "Personality and counterproductive workplace behavior." *Personality and work: Reconsidering the role of personality in organizations* 14, no. 2:150–182.
- Donovan, J. J., Dwight, S. A., and G. M. Hurtz, 2003. "An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique." *Human Performance*, 16 no. 1:81–106.
- Edelhertz, H. 1970. *The Nature, Impact, and Prosecution of White-collar crime* (pp. 27). Washington, DC: National Institute of Law Enforcement and Criminal Justice.

- Emmons, R. A. 1984. "Factor analysis and construct validity of the Narcissistic Personality Inventory." *Journal of Personality Assessment* 48, no. 3:291.
- Engleman, D., and B. H. Kleiner. 1998. "Effective employment screening practices." *Career Development International* 3, no. 4:164–168.
- Fine, S., J. Goldenberg, and Y. Noam. 2015. "Integrity testing and the prediction of counterproductive behaviors in the military." *Journal of Occupational and Organizational Psychology*.
- Forensic, KPMG. 2006. *Fraud Risk Management: Developing a Strategy for Prevention, Detection, and Response*. Switzerland: KPMG International.
- Greitzer, F. L., L. J. Kangas, C. F. Noonan, and A. C. Dalton. 2010. *Identifying At-Risk Employees: A Behavioral Model for Predicting Potential Insider Threats*. Richland, Washington: Pacific Northwest National Laboratory.
- Henle, C. A., and M. A. Gross. 2013. "An examination of the relationship between workplace deviance and employee personality." In *Deviant and Criminal Behavior in the Workplace*, edited by Steven M. Elias. [book online] New York and London: New York University Press.
- Johnson, E. N., J. R. Kuhn Jr, B. A. Apostolou, and J. M. Hassell. 2012. "Auditor perceptions of client narcissism as a fraud attitude risk factor." *Auditing: A Journal of Practice and Theory* 32, no. 1:203–219.
- Judge, T. A., D. Heller, and M. K. Mount. 2002. "Five-factor model of personality and job satisfaction: A meta-analysis." *Journal of Applied Psychology* 87, no. 3:530.
- Lee, K., M. C. Ashton, and R. E. de Vries. 2005. "Predicting workplace delinquency and integrity with the HEXACO and five-factor models of personality structure." *Human Performance* 18, no. 2:179–197.
- MacLane, C. N., and P.T. Walmsley. 2010. "Reducing counterproductive work behavior through employee selection." *Human Resource Management Review* 20, no. 1:62–72.
- Marcus, B., M. C. Ashton, and K. Lee. 2013. "A note on the incremental validity of integrity tests beyond standard personality inventories for the criterion of counterproductive behavior." *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration* 30, no. 1: 18–25.
- Marcus, B., S. Höft, and M. Riediger. 2006. "Integrity tests and the five-factor model of personality: A review and empirical test of two alternative positions." *International Journal of Selection and Assessment* 14, no. 2:113–130.
- Martin, S. L. 1989. "Honesty testing: Estimating and reducing the false positive rate." *Journal of Business and Psychology* 3, no. 3:255–267.
- McCrae, R. R., and P.T. Costa. 2010. *NEO Inventories for the NEO Personality Inventory-3 (NEO PI-3), NEO Five-Factor Inventory-3 (NEO-FFI-3) and NEO Personality Inventory-revised (NEO PI-R): Professional Manual*. Psychological Assessment Resources.
- McCrae, R. R., P. T. Costa Jr, and T. A. Martin. 2005. "The NEO–PI–3: A more readable revised NEO personality inventory." *Journal of Personality Assessment* 84, no. 3:261–270.
- McCrae, R. R., O. P. John. 1992. "An introduction to the five-factor model and its applications." *Journal of Personality* 60, no. 2:175–215.
- Miners, J. B., and M. H. Capps. 1996. *How Honesty Testing Works*. Westport, CN: Greenwood Publishing Group.

- Mowen, J. C. 2000. *The 3M Model of Motivation and Personality: Theory and Empirical Applications to Consumer Behavior*. Berlin: Springer Science and Business Media.
- Mount, M., R. Ilies, and E. Johnson. 2006. "Relationship of personality traits and counterproductive work behaviors: The mediating effects of job satisfaction." *Personnel Psychology* 59, no. 3:591–622.
- Murphy, K.R. 1993. *Honesty in the workplace*. Belmont, CA: Brooks/Cole.
- O'Bannon, R. M., L. A. Goldfinger, and G. S. Appleby. 1989. *Honesty and Integrity Testing: A Practical Guide*. Atlanta, GA: Applied Information Resources.
- Oliphant, B. J., and G. C. Oliphant. 2001. "Using a behavior-based method to identify and reduce employee theft." *International Journal of Retail and Distribution Management* 29, no. 10:442–451.
- Ones, D. S. 1993. "The Construct Validity of Integrity Tests." (Unpublished doctoral dissertation). University of Iowa, Iowa City.
- Ones, D. S., C. Viswesvaran, and F. L. Schmidt. 2003. "Personality and absenteeism: a meta-analysis of integrity tests." *European Journal of Personality* 17, no. S1:S19–S38.
- Perri, F. S. 2011. "White-collar criminals: The 'kinder, gentler' offender?" *Journal of Investigative Psychology and Offender Profiling* 8, no. 3:217–241.
- Perri, F. S., and R. G. Brody. 2011. "The Sallie Rohrbach story: Lessons for auditors and fraud examiners." *Journal of Financial Crime* 18, no. 1:93–104.
- Podgor, E. S., and L. E. Dervan. 2016 "White Collar Crime': Still Hazy After All These Years." *Georgia Law Review* 50, no. 3.
- Raskin, R., and H. Terry. 1988. "A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity." *Journal of Personality and Social Psychology* 54, no. 5:890.
- Rhodewalt, F., and C. C. Morf. 1998. "On self-aggrandizement and anger: A temporal analysis of narcissism and affective reactions to success and failure." *Journal of Personality and Social Psychology* 74, no. 3:672.
- Rieke, M. L., and S. J. Guastello. 1995. "Unresolved issues in honesty and integrity testing." *American Psychologist* 50, no. 6:458–459.
- Sackett, P. R. 2002. "The structure of counterproductive work behaviors: Dimensionality and relationships with facets of job performance." *International Journal of Selection and Assessment* 10, no. 1-2:5–11.
- Sackett, P. R., L. R. Burris, and C. Callahan. 1989. "Integrity testing for personnel selection: An update." *Personnel Psychology* 42, no. 3:491–529.
- Sackett, P.R., and C. J. Devore. 2001. "Counterproductive behaviors at work." In N. Anderson, D. Ones, C. Sinangil, and C. Viswesvaran (Eds.), *International Handbook of Work Psychology* 1:145-164. London: Sage.
- Sackett, P.R., and M. M. Harris. 1984. "Honesty testing for personnel selection: A review and critique." *Journal of Personnel Psychology* 37:221–244.
- Sackett, P. R., and J. E. Wanek. 1996. "New developments in the use of measures of honesty integrity, conscientiousness, dependability trustworthiness, and reliability for personnel selection." *Personnel Psychology* 49, no. 4:787–829.
- Salgado, J. F. 2002. "The Big Five personality dimensions and counterproductive behaviors." *International Journal of Selection and Assessment* 10, no. 1–2:117–125.

Smith, S. F., and S. O. Lilienfeld. 2013. "Psychopathy in the workplace: The knowns and unknowns." *Aggression and Violent Behavior* 18, no. 2: 204–218.

Van Iddekinge, C. H., P. L. Roth, P. H. Raymark, and H. N. Odle-Dusseau. 2012. "The criterion-related validity of integrity tests: An updated meta-analysis." *Journal of Applied Psychology* 97, no. 3:499.

Table I: Demographic Summary

Demographic Characteristics of White-Collar Participants				
	Demographic Characteristic		Frequency (N = 20)	Percent
1	Gender	Male	20	100
2	Age	18–25	0	0
		26–32	1	5
		33–39	1	5
		40–46	5	25
		47–53	2	10
		Over 54	11	55
3	Position	Employee	0	0
		Management	0	0
		Upper Management	3	15
		CEO	2	10
		Owner	12	60
		Board Member/External to Company	3	15
4	Inmates With Prior White-Collar Convictions	Employee	0	0
		Management	0	0
		Upper Management	2	10
		CEO	0	0
		Owner	2	10
		Board Member/External to Company	0	0
5	White-Collar Related Convictions	Bank Fraud	5	19.2
		Bribery	1	3.8
		Conspiracy	5	19.2
		Tax Fraud	4	15.4
		Mail Fraud	3	11.5
		Wire Fraud	5	19.2
		SEC/Anti-trust	3	11.5

Table II: White-Collar Inmates vs. GP Statistics

Trait Measured	White-Collar Inmates					General Population			
	T-scores	M	SD	Median	Range	Sig	N	M	SD
NEO-FFI 3									
Emotional Stability	-2.295	16.65	8.087	17	3–34	.033	635	20.8	7.7
Agreeableness	.294	32.55	6.840	33	16–44	.772	635	32.1	6.0
Conscientiousness	1.848	35.95	8.351	37	17–47	.080	635	32.5	6.3
NPI (Narcissism)	-.539	14.65	7.050	11	6–29	.596	1018	15.5	6.66
WINT (Integrity)	-.783	67.25	15.427	69	39–88	.443	1672	69.95	16.51

Table III: Summary Theft Admissions

Theft Admissions (N=20)							
	Admitted Theft Amounts	Integrity Score	Under 18 Years Old	Over 18 Years Old	Never Stolen	Prior Convictions	
Recommend to Hire	\$1–5	83	X				
	\$5–20	54	X			Priors	
		58	X				
	\$1000–\$10,000	58			X		Priors
		39			X		
		70			X		
	Over \$10,000	50			X		
53				X		Priors	
61				X		Priors	
Do Not Recommend	\$0	74			X		
	\$1–5	68	X				
		45	X				
	\$5–20	88			X		
		84		X			
	\$20–100	85			X		Priors
		86		X			
	\$100–500	84			X		
	Over \$10,000	78			X		
44				X		Priors	
50			X				