

Predicting Correctional Officer Employment Success:

Criterion Validity for the Critical Hire - Screen

Anthony W. Tatman

Des Moines, Iowa

Author Note

Anthony W. Tatman is owner and founder of Critical Hire. Correspondence concerning this article should be addressed to Anthony Tatman at tonytatman@criticalhire.com.

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Abstract

Correctional employees, such as probation and parole officers, are justifiably given considerable power, authority, and trust by the courts, parole boards, and the community in general, making an officer's level of integrity and character an imperative factor to consider during the hiring process. This study examined the predictive validity of the Critical Hire-Screen (CH-S), a pre-employment integrity assessment, in predicting supervisory ratings for a sample of correctional officers. Analyses revealed a significant ability to identify mis-hires, as well as differentiate between mis-hires and non-mis-hires, demonstrating that the CH-S is an empirically valid and practically useful compliment to hiring processes for correctional officer applicants.

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Introduction

The use of integrity testing has increased substantially after the Employee Polygraph Protection Act was enacted in 1988 (Employee Polygraph Protection Act, 1988), which prohibited the use of pre-employment polygraph testing for all but a select few employment settings. Initially developed as a surrogate measure of honesty in place of the polygraph, integrity tests have since been classified into two main groups: overt and personality-based tests (Sackett, Burris, & Callahan, 1989). Overt integrity tests (also known as clear purpose tests) typically have two sections. The first is generally a self-report section soliciting admissions of past counterproductive workplace behaviors (CWBs), such as theft, taking bribes, assaulting people, or using illicit drugs. The second section in overt integrity tests often solicits information about an applicant's beliefs and values as it pertains to CWBs. Personality-based measures (also known as disguised purpose tests), on the other hand, measure normal-range, non-pathological, personality traits via composite measures of personality with the goal of predicting broad range CWBs such as disciplinary problems, violence on the job, and problems with authority. Subsequent research has shown that integrity tests have adequate validity in predicting both job performance and counterproductive work behaviors (CWBs) such as theft, drug use, tardiness, property damage, rule-breaking, absenteeism, and violence in the workplace (Berry, Sackett & Wiemann, 2007; Fine, 2013; Fine, Horowitz, Weigler, & Basis, 2010; Jones, Cunningham, & Dages, 2010; Marcus, Ashton, & Lee, 2013; Nicol & Paunonen, 2002; Ones, Viswesvaran, & Schmidt, 1993; Ones, Viswesvaran, & Schmidt, 2003; Schmidt & Hunter, 1998; Schmidt, Oh, Shaffer, 2016; Wanek, 1999). Ones et al.'s (1993) comprehensive meta-analysis explored the predictive validity of both overt and personality-based integrity measures, as well as possible moderators impacting validity estimates. Ones et al. revealed that, although both overt and

CRITICAL HIRE-SCREEN CRITERION VALIDITY

personality-based integrity measures significantly predicted overall job performance and CWB for low, medium and high-complexity jobs, as defined by Hunter, Schmidt, and Judiesch (1990), overt integrity tests were found to be slightly better predictors of CWBs than personality-based tests ($p = .55$ vs $p = .32$, respectively).

Integrity Testing with Correctional Applicants

An employee's integrity and character are critical in any occupational setting, particularly in correctional settings such as agencies providing supervision for individuals on probation and parole. Correctional officers (i.e., probation, parole, and residential officers) are expected to be models of character and integrity, closely monitor individuals under Court mandated supervision, protect the public, enforce the law, and serve as agents of change for their clients. With these expectations, correctional officers have been given considerable power and authority, as well as public and judicial trust, making a correctional officer's integrity and character essential. However, not all correctional officers have the level of integrity and character we as a community would expect and require, and unfortunately formal assessment of integrity and character has not been a standard component in correctional officer hiring processes (Tatman & Huss, 2019). Although correctional officer integrity and character are identified, and intuitively obvious, as being a critical component to consider during the hiring process, most of the instruction, guidance and "how to guides" for hiring correctional officers focus on interview techniques such as using behavioral questions, predetermined job competencies, structured interview guides, and behaviorally anchored interview scoring sheets (Wells, 2018). Tatman and Huss theorize that this disconnect may stem from a combination of the corrections field being unaware of the extensive research on the validity, and apparent value in, integrity testing in other job settings, and the lack of available integrity tests developed and normed on correctional

CRITICAL HIRE-SCREEN CRITERION VALIDITY

officer applicants. The Critical Hire - Screen (CH-S; Tatman, 2018) was developed to address this apparent disconnect by providing the corrections field a reliable and valid measure of integrity developed for, and with, correctional officers. The CH-S is the only known integrity test meeting the Uniform Guidelines on Employee Selection Procedure (UGESP; Uniform Guidelines on Employee Selection Procedure, 1978) guidelines for use with correctional officer applicants. In 1978, the Equal Employment Opportunity Commission (EEOC) adopted the UGESP, which provides a framework for determining the proper use of tests and other employee selection procedures. Based on the UGESP, selection procedures, such as integrity testing, must show empirical evidence for the instrument or process's validity, reliability, applicability for the specific job, and degree to which it has an adverse impact on protected populations. Research has found that the CH-S has appropriate internal consistency, test-retest reliability, convergent validity, content validity, and does not generate group differences based on gender, race, and age that would infringe on EEOC guidelines when used with correctional applicants (Tatman & Huss, 2019). A hole in the existing literature on the CH-S, however, is the lack of empirical evidence regarding its criterion validity or accuracy in predicting future correctional officer job performance. Therefore, the purpose of this study is to investigate the degree to which the CH-S can accurately predict supervisory ratings of correctional employee job performance. Specifically, this study will investigate the degree to which CH-S Critical Items, Critical Factors, and IMS accurately predict supervisor ratings. This study also sought out to replicate Tatman & Huss by measuring whether there are significant differences in CH-S scale scores based on race, gender, and age.

Methods

Participants

CRITICAL HIRE-SCREEN CRITERION VALIDITY

The CH-S has been used in community-based correction (CBC) agencies throughout the Midwest as part of a comprehensive hiring process for new and promotional candidates. As a result, archival data has been accumulated on new and seasoned correctional officer applicants. Using this archival data, 124 CBC employees (68 males and 56 females) completed the CH-S as they applied for promotions within their existing agency or employment with a different CBC agency. The sample had an average age of 38.16 ($SD = 9.16$) and ranged from 20 to 61 years of age. Racial composition consisted of 103 Caucasian, 14 African American, 5 Hispanic, and 2 Asian participants.

Measures

The Critical Hire - Screen (CH-S; Tatman, 2018) is a pre-employment, overt integrity assessment which can be used at a pre- or post-conditional offer phase in the hiring process. The CH-S measures applicant responses in three unique domains: Employment and Legal History, Personal Opinions and Beliefs, and impression management. The first domain measured by the CH-S is Employment and Legal History and utilizes direct admission questions inquiring about past work and legal experiences (e.g., “Have you ever been terminated”, “Have you ever been convicted of a crime”, or “How would you rate your last supervisor”). The Personal Opinions and Beliefs section measures five factors or constructs of integrity: Substances ($\alpha = .79 - .93$; use or selling of drugs in the workplace, and/or use of alcohol in the workplace), Theft ($\alpha = .73 - .82$; theft in the workplace), Authority ($\alpha = .61 - .78$; disparaging or conflictual opinions about management and supervisors), Rules & Deception ($\alpha = .70 - .82$; rule breaking, manipulating others, and deceptive behaviors), and Personal Responsibility ($\alpha = .71 - .82$; the degree to which an applicant places blame on victims for crimes committed against them) (Tatman & Huss, 2019). Personal Opinions and Beliefs questions are anchored with a five-point Likert scale

CRITICAL HIRE-SCREEN CRITERION VALIDITY

“Agree Strongly” to “Disagree Strongly.” Scores for the five Personal Opinions and Beliefs scales range from one (low concern) to seven (high concern), with four being average. When a Personal Opinions and Beliefs factor (e.g., Substances) reaches a score of five or above (i.e., elevated concern) it is identified as being a Critical Factor, prompting the hiring agency to follow up with additional, specific interview questions or inquiry into content pertaining to that particular Critical Factor. In addition to providing readers with scores for each of the five Personal Opinions and Beliefs factors, the CH-S provides users with a list of Critical Items. A Critical Item is generated when the applicant endorses a specific item on a Personal Opinions and Beliefs factor in a way that significantly deviates from the normative population, which then also triggers an area for further inquiry for the hiring agency. The CH-S also incorporates an impression management scale (IMS; $\alpha = .50 - .74$) which has been found to have strong reliability and concurrent validity in detecting an individual’s attempts to exaggerate their virtuousness on the CH-S (Tatman & Huss). In order to capture the intended purpose of the IMS in this study, IMS raw scores were converted into a dichotomous variable for this study indicating appropriate IMS scores versus elevated IMS scores suggesting significant social desirability.

Procedure

Supervisor Ratings of Job Performance

Supervisors for each participant who completed the CH-S were asked one question for this study: “Would you hire this employee again?” Supervisors answered this question without reference to, or knowledge of, the employee’s CH-S scores. Employees rated as being someone the supervisor would not hire again were termed “mis-hires,” while employees rated as being someone the supervisor would hire again was termed “non-mis-hires” in this study. The term

CRITICAL HIRE-SCREEN CRITERION VALIDITY

“mis-hire” is not meant here as a derogatory term, or necessary a term implying the mis-hired employee violated workplace rules or policies. A mis-hire can occur for a variety of reasons ranging from a poor job fit or personality and attitudes that conflict with company culture, to a disregard for authority and workplace policies or illegal conduct. This simple, dichotomous (“Mis-hire” or “Non-mis-hire”) outcome variable was chosen because it succinctly and concisely gets at the heart at a key concern shared by most hiring agencies and employers – “Will this applicant work out in my organization?” or “Will this applicant be a mis-hire?”

Statistical Analyses

Singh (2013) provides a review on the strengths and limitations of various statistical methods commonly used to measure predictive accuracy in risk assessment tools used in criminal justice and correctional settings. Conceptualizing the CH-S’s ability to predict the risk of mis-hires as being comparable to forensic risk assessment tools, Singh’s guidance on appropriate statistical methods to use to measure predictive accuracy was utilized in this study. Statistical methods used to measure predictive accuracy measure either calibration or discrimination, but not both. Calibration refers to the degree to which a tool’s prediction of risk agree with known, observed risk. Discrimination, on the other hand, refers to the degree to which an assessment is able to differentiate between two outcomes, which in this case is mis-hires and non-mis-hires. The type of predictive accuracy of interest is also important to consider when choosing the appropriate statistic. High risk accuracy, for example, is the tool’s ability to identify high risk groups, while low risk accuracy is the tool’s ability to identify low risk groups. Global accuracy, on the other hand, measures a tool’s ability to identify both high and low risk groups. Both calibration and discrimination, as well as all three types of predictive accuracy, are valuable to obtain a comprehensive understanding of a tool’s predictive accuracy and require unique

CRITICAL HIRE-SCREEN CRITERION VALIDITY

statistical methods. Therefore, in order to measure the CH-S's calibration and high risk accuracy positive predictive values (PPV) were conducted in this study. PPV answers the question "If you test positive for X, what is the likelihood you actually have X?" PPV, for this study, is the probability that an individual with an elevated score is classified as a mis-hire. In order to measure the CH-S's calibration and low risk accuracy negative predictive values (NPV) were conducted in this study. NPV answers the question "If you test negative for X, what is the likelihood that you do not have X?" NPV, for this study, is the probability that the individual without an elevated score is classified as a non-mis-hire. In order to measure the CH-S's discrimination and high risk accuracy sensitivity metrics were conducted. A tool's sensitivity is the degree to which the tool can correctly identify the issue or concern at hand. In this study, for example, sensitivity measures the degree to which the CH-S can correctly identify mis-hires from the sample population. Discrimination and low risk accuracy were measured with specificity metrics. A tool's specificity is the degree to which a tool can correctly identify the absence of the issue or concern in question. In this study specificity measures the degree to which the CH-S can correctly identify non-mis-hires out of the sample population. Global accuracy (calculated by true positive + true negative/sample size) was also conducted to identify the degree to which the CH-S can correctly identify the combination of people correctly identified as mis-hires and those who are non-mis-hires. Point-biserial correlation coefficients have also been identified as providing information about a risk tool's level of discrimination and global accuracy, as well as the direction and strength of association between dichotomous and continuous variables (Singh; Das Gupta, 1960). Therefore, point-biserial correlations, or when appropriate phi correlations (when measuring two dichotomous variables), were conducted in this study. Risk ratios, also known as relative risk ratios, were also calculated to identify the

CRITICAL HIRE-SCREEN CRITERION VALIDITY

probability at which the CH-S can predict mis-hire in applicants with elevated scores as compared to applicants with non-elevated scores.

Results

Research Question 1 - *Can CH-S Critical Items accurately identify mis-hires?*

In order to answer this research question a point-biserial correlation (represented by r_{pb}) was conducted between supervisor ratings (i.e., mis-hire; dichotomous variable) and the number of Critical Items generated for each participant (i.e., number of specific CH-S items deviating from the normal population; continuous variable). Results indicated a significant correlation between the number of Critical Items generated on the CH-S and whether supervisors rated the participant as a mis-hire ($r_{pb} = .27, p = .003$). Comparative analyses were then conducted using sensitivity, specificity, PPVs, NPVs, and global accuracy rates on three options of Critical Item cut scores: 2, 3, and 4 Critical Items. Results from these comparative analyses revealed that all three options were rather indistinguishable in terms of specificity and global accuracy. However, the 3 Critical Items cut score option showed marked improvement in terms of PPV over the other two options, and therefore deemed superior to the alternative options. Using a cut score of 3 Critical Items produced a sensitivity rate of 19.05%, specificity rate of 99.03%, PPV of 80.00%, NPV of 85.71%, and a global accuracy rate of 86.18% for identifying mis-hires in this sample. The difference between the sensitivity rate and PPV found here is noteworthy. Sensitivity is a metric of the test itself, and provides information about the probability that a test result will be positive (i.e., 3 or more Critical Items) when the event (i.e., employee is identified as a mis-hire) is present. PPV, on the other hand, is a measurement of the population and provides the probability that a mis-hire would be identified if the applicant produces 3 or more Critical Items. In other words, PPV answers the question “What is the chance of a mis-hire if I get 3 or more

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Critical Items?”, while sensitivity answers the question “What is the CH-S’s ability, using a cut score of 3 or more Critical Items, to detect a mis-hire?” Therefore, based on these results it appears that when using a cut score of 3 Critical Items the CH-S detects mis-hires about 19% of the time (i.e., sensitivity). However, if a person does have 3 or more Critical Items there is an 80% likelihood that they would be identified as a mis-hire (i.e., PPV). In other words, and put quite simply, relatively few mis-hires generated 3 or more Critical Items. However, a considerable number of individuals generating 3 or more Critical Items were mis-hires.

Risk ratios were also conducted using 3 Critical Items and revealed a risk ratio score of 5.6 ($p < .0001$). This suggests that individuals who generate 3 or more Critical Items are over 5 times more likely to be classified as a mis-hire than employees generating 2 or fewer Critical Items.

Research Question 2 - *Can CH-S Critical Factors accurately identify mis-hires?*

Similar to Research Question 1, a point-biserial correlation was conducted between supervisor ratings and the number of Critical Factors generated for each participant (i.e., Personal Opinions and Beliefs factor scores over 5; continuous variable). Results indicated a significant correlation between the number of Critical Factors generated on the CH-S and supervisor ratings ($r_{pb} = .24, p = .008$). Comparative analyses were then conducted using sensitivity, specificity, PPVs, NPVs and global accuracy rates on cut scores using 1, 2, and 3 Critical Factors as comparison options. Results from these comparative analyses revealed that a cut score of 2 Critical Factors was superior in predicting mis-hires to using a cut score of 1 or 3 Critical Factors. Using a cut score of 2 Critical Factors produced a sensitivity rate of 42.86%, specificity rate of 88.35%, PPV of 42.86%, NPV of 88.35%, and a global accuracy rate of 81.30%. Risk ratios were also conducted using 2 or more Critical Factors, resulting in a ratio

CRITICAL HIRE-SCREEN CRITERION VALIDITY

score of 3.68 ($p = .0004$). This finding suggests that individuals who generate 2 or more Critical Factors are over 3 times more likely to be classified as mis-hires than employees generating 0 or 1 Critical Factors.

Research Question 3 - *Can CH-S IMS scores accurately identify mis-hires?*

Analyses on the Impression Management Scale (IMS) were also conducted to measure the degree to which the IMS can independently identify mis-hires in this sample. This analysis was conducted based on existing research finding that an applicant's propensity to endorse psychological assessment items in a socially desirable manner significantly correlates with subsequent counterproductive work behaviors (Sellbom, Fischler, & Ben-Porath, 2007; Weiss, Davis, Rostow, & Kinsman, 2003; Weiss, Vivian, Weiss, Davis, Rostow, 2013). A Phi correlation measuring the degree to which IMS scores (dichotomous variable; appropriate vs significantly elevated) was correlated with supervisor ratings (dichotomous variable; mis-hire vs non-mis-hire) was calculated. Results showed a significant relationship between IMS scores and supervisor ratings ($\Phi = .41, p > .0001$). The degree to which the IMS uniquely identified mis-hires was then measured by calculating sensitivity, specificity, PPVs, NPVs accuracy, and risk ratios. The sample population used in this study contained eight (6.45%) CH-Ss with significantly elevated IMS scores, indicating significant social desirability while completing the CH-S. Results of these analyses revealed a sensitivity rate of 28.57%, specificity rate of 98.06%, PPV of 75.00%, NPV of 87.07%, and an overall accuracy rate of 86.99%. Risk ratios were also conducted, resulting in a ratio score of 5.80 ($p > .0001$). This finding suggests that individuals who generate elevated scores on the IMS are nearly 6 times more likely to be classified as mis-hires than employees generating non-elevated scores on the IMS.

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Research Question 4. *Are there significant differences in CH-S scale scores based on race, gender, and age?*

A series of independent samples t-tests were conducted to investigate the degree to which the various CH-S scale scores showed significant differences based on race. Due to the relatively small number of minority participants in the various racial groups found in this sample, it was decided to combine all minorities into a single “minority” group for this analysis. Consistent with prior research (Tatman & Huss, 2019), results revealed no significant differences between racial minorities (N = 21) and Caucasian (N = 103) participants on the various CH-S scales (Table 1).

A series of independent samples t-tests were also conducted to investigate the degree to which the various CH-S factors showed significant differences between male and female participants. Results revealed no significant differences between male and female participants on the various CH-S factors (Table 2), suggesting that the CH-S does not generate significant group differences based on gender.

A series of independent samples t-tests were also conducted to investigate the degree to which the five CH-S factors showed significant differences based on age. The Age Discrimination in Employment Act (ADEA; Age Discrimination in Employment Act, 1967) prohibits age discrimination against people who are age 40 or older. Therefore, to be consistent with this ADEA guideline, the following t-test comparisons were based on participants who were 40 or older (N = 56) against participants under 40 (N = 68). Results of these analyses revealed that the two age groups did not significantly differ on the various CH-S factors (Table 3), suggesting that the CH-S does not appear to generate group differences based on age of the applicant, as defined by the ADEA.

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Summary and Discussion

This study sought out to explore the degree to which the Critical Hire-Screen (CH-S) accurately identifies correctional employee mis-hires. Results obtained from this study revealed that the CH-S adequately identified mis-hires and differentiated between mis-hires and non-mis-hires. This study also provides readers with three separate indices in which to evaluate CH-S scores obtained from their applicants. All three indices measured in this study (i.e., two or more Critical Factors, three or more Critical Items, and a significantly elevated IMS score) showed promise in being able to predict correctional applicant mis-hires. Results from this study also provide supporting evidence for the application of the CH-S with protected populations. The Uniform Guidelines on Employee Selection Procedures (UGESP) and Equal Employment Opportunity Commission (EEOC) Guidelines require that employment selection tools and processes do not discriminate against racial minorities, females, and employees over the age of 40. Consistent with existing research (Tatman & Huss, 2019), the present findings indicate that CH-S results do not generate group differences based on age, gender, or minority status, providing additional evidence that the results generated from the CH-S does not contribute to an adverse effect on these protected populations. This is an important finding, not only because it replicates prior findings (Tatman & Huss), but also for the evidence it provides correctional organizations and human resource personnel that the CH-S complies with UGESP and EEOC guidelines.

Although the results obtained from this study are promising for the accuracy and ultimate utility of the CH-S, it should be stressed that these are initial findings. The CH-S was not developed or intended to be used as the single determinant for hiring decisions (Tatman, 2018). However, cut scores could help identify thresholds that could trigger additional interview

CRITICAL HIRE-SCREEN CRITERION VALIDITY

questions, collateral contacts or more extensive criminal or employment background checks to inquire further into the particular CH-S area of concern. Since these remain initial findings, additional, replicative studies are recommended before using data generated from this paper as cut scores for this purpose during correctional officer hiring processes. As with any hiring process, it should also be stressed that hiring decisions using the CH-S should not rest on one single piece of information, but rather should come from multiples sources of credible information.

In addition to identifying applicants with a higher risk for future workplace problems, results obtained in this study could also have a financial impact on hiring agencies. For example, one business saw a 50% reduction in terminations historically caused by employee misconduct such as theft, illegal drug use, and violence over a five-year period after implementing integrity tests into their application and hiring process (Brown, Jones, Terris, & Steffy, 1987). This reduction in turnover through terminations has considerable cost savings for correctional agencies. Boushey and Glynn (2012) report that employees making \$30,000 or less a year cost at least 16% of that employee's annual salary to replace, while employees earning between \$30,000 and \$75,000 a year can cost approximately 21% of the employee's salary. Replacing employees earning \$100,000 or more, however, can cost agencies upwards of 213% of that employee's annual salary, making senior management / officers, and specialized personnel, extremely expensive to replace. The U.S. Department of Labor share similar figures, estimating that the cost of a mis-hire can equal 30% of the employee's potential first-year earnings (Fatemi, 2016). With correctional officers earning an average annual salary of between \$35,578.00 and \$43,550 (Correctional Officer Salary, 2018; Glassdoor, 2019 April; Glassdoor, 2019 May), correctional departments could expect to pay between \$7,000 to \$9,000 to replace one officer. These costs,

CRITICAL HIRE-SCREEN CRITERION VALIDITY

however, do not include many indirect costs such as legal fees and settlement costs, lost productivity, strained workplace moral, negative publicity, or fractures in public trust that can co-occur with mis-hires. Therefore, reducing the odds of hiring mis-hires could have a significant financial impact on correctional agencies.

There are a couple limitations and recommendations for future research that should be noted. The sample population used in this study was comprised solely of correctional agencies within the Midwest. Additional research should be conducted in different geographic regions to provide empirical evidence for the generalizability of these findings across more diverse settings. For example, the sample used in this study consisted of 103 (83%) Caucasian, 14 (11%) African American, 5 (4%) Hispanic, and 2 (2%) Asian participants. The population density from which the participants for this study originated came from largely rural towns and relatively small (210,000 citizens) cities (World Population Review, 2019). Additional research is recommended with samples consisting of more racial diverse participants, as well as geographically diverse settings, to help enable and support generalizations made to these various groups of individuals.

Also, the original, normative sample population for the CH-S included certified police officers (Tatman & Huss, 2019), making the CH-S applicable for this subset of the law enforcement community. However, criterion validity, as it pertains to predictive accuracy, for the CH-S has not yet been established with police officer candidates. Therefore, replicating the present study with police officer applicants would be a valuable and informative extension of this research.

CRITICAL HIRE-SCREEN CRITERION VALIDITY

References

- Age Discrimination in Employment Act of 1967, Pub. L. No. 90-202, § 621, 29 (1967).
- Berry, C. M., Sackett, P. R., & Wiemann, S. (2007). A review of recent developments in integrity test research. *Personnel Psychology*, 60(2), 271-301. doi: 10.1111/j.1744-6570.2007.00074.x
- Boushey, H. & Glynn, S. J. (2012). There are significant business costs to replacing employees. Center for American Progress. Retrieved from <https://www.americanprogress.org/wp-content/uploads/2012/11/CostofTurnover.pdf>
- Brown, T. S., Jones, W. J., Terris, W., & Steffy, B. D. (1987). The impact of pre-employment integrity testing on employee turnover and inventory shrinkage losses. *Journal of Business and Psychology*, 2(2), 136-149. doi.org/10.1007/BF01014208
- Correctional Officer Salary. (2018). CorrectionalOfficerEDU.org. Retrieved from <https://www.correctionalofficeredu.org/salaries/>
- Das Gupta, S. (1960). Point biserial correlation coefficient and its generalization. *Psychometrika*, 25, 393-408. doi: 10.1007/BF02289756
- Employee Polygraph Protection Act, 29 U.S.C. § 2001 (1988).
- Fatemi, F. (2016, September 28). The true cost of a bad hire – It's more than you think. *Forbes*. Retrieve from <https://www.forbes.com/sites/falonfatemi/2016/09/28/the-true-cost-of-a-bad-hire-its-more-than-you-think/#38e97b494aa4>
- Fine, S. (2013). A look at cross-cultural integrity testing in three banks. *Personnel Review*, 42(3), 266-280. doi: 10.1108/00483481311320408
- Fine, S., Horowitz, I., Weigler, H., & Basis, L. (2010). Is good character good enough? The effects of situational variables on the relationship between integrity and

CRITICAL HIRE-SCREEN CRITERION VALIDITY

counterproductive work behaviors. *Human Resource Management Review*, 20(1), 73-84.

doi: 10.1016/j.hrmr.2009.03.010

Glassdoor (2019, April). Probation Officer Salaries. Retrieved from

https://www.glassdoor.com/Salaries/probation-officer-salary-SRCH_KO0,17.htm

Glassdoor (2019, May 30). Parole Officer Salaries. Retrieved from

https://www.glassdoor.com/Salaries/parole-officer-salary-SRCH_KO0,14.htm

Jones, J. W., Cunningham, M. R., & Dages, K. D. (2010). Pre-offer police integrity testing:

Scientific foundations and professional issues. In P. A. Weiss (Ed.), *Personality*

Assessment in Police Psychology: A 21st Century Perspective (pp. 159-187). Springfield,

IL: Charles C Thomas Publisher.

Marcus, B., Ashton, M. C., & Lee, K. (2013) A note on the incremental validity of integrity tests

beyond standard personality inventories for the criterion of counterproductive behaviour.

Canadian Journal of Administrative Sciences, 30(1), 18-25. doi 10.1002/cjas.1235

Nicol, A. M., & Paunonen, S. V. (2002). Overt honesty measures predicting admissions: An

index of validity or reliability. *Psychological Reports*, 90(1), 105-115.

Ones, D. S., Viswesvaran, C., & Schmidt, F. (1993). Comprehensive meta-analysis of integrity

test validities: Findings and implications for personnel selection and theories of job

performance. *Journal of Applied Psychology Monograph*, 78(4), 679-703.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2003). Personality and absenteeism: A meta-

analysis of integrity tests. *European Journal of Personality*, 17(SPEC. 1), S19-S38. doi:

10.1002/per.487

Sacket, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An

update. *Personnel Psychology*, 42(3), 491-529. doi: 10.1111/j.1744-6570.1989.tb00666.x

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings.

Psychological Bulletin, 124(2), 262-274. doi: 10.1037/0033-2909.124.2.262

Schmidt, F. L., Oh, I-S., Shaffer, J. A. (2016). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 100 Years of Research Findings. Retrieved from <https://www.testingtalent.net/wp-content/uploads/2017/04/2016-100-Yrs-Working-Paper-on-Selection-Methods-Schmidt-Mar-17.pdf>

Sellbom, M., Fischler, G., & Ben-Porath, Y. (2007). Identifying MMPI-2 predictors of police officer integrity and misconduct. *Criminal Justice and Behavior*, 34(8), 985-1004.

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences and the Law*, 31, 8-22. doi: 10.1002/bsl.2052

Tatman, A. W. (2018). Critical Hire-Screen: Professional Manual. Des Moines, Iowa. Self-published.

Tatman, A. W., & Huss, M. (2019). Initial reliability and validity for the Critical Hire - Screen. *Personnel Assessment and Decisions*. In press.

Uniform Guidelines on Employee Selection Procedure (1978); 43 FR 38295 (August 25, 1978). Retrieved from <http://uniformguidelines.com/uniformguidelines.html#20>

Wanek, J. E. (1999). Integrity and honesty testing: What do we know? How do we use it? *International Journal of Selection & Assessment*, 7(4), 183-195. doi: 10.1111/1468-2389.00118

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Weiss, P. A., Vivian, J. E., Weiss, W., U., Davis, R. D., & Rostow, C. D. (2013). The MMPI-2L scale, reporting uncommon virtue, and predicting police performance. *Psychological Services, 10*(1), 123-130. doi: 10.1037/a0029062

Weiss, W. U., Davis, R. D., Rostow, C. D., & Kinsman, S. (2003). The MMPI-2 L scale as a tool in policy selection. *Journal of Police and Criminal Psychology, 18*(1), 57-60. doi: 10.1007/BF02802608

Wells, L. (2018, December 17). Reducing recidivism through the hiring process. Retrieved from <http://blogs.iu.edu/speaindy/2018/12/17/recidivism-probation-officers-hiring-hr/>

World Population Review (2019). Iowa Population 2019. Retrieved from <http://worldpopulationreview.com/states/iowa-population/>

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Table 1

Scale Means, Standard Deviations, and t-tests by Racial Status

	Race	<i>M</i>	<i>SD</i>	<i>t</i> (df)	<i>p</i>
IMS	Caucasian	1.04	.19	-1.69(21.93)	.11
	Minority	1.19	.40		
Substances	Caucasian	1.46	1.07	.39(23.06)	.40
	Minority	1.81	1.78		
Theft	Caucasian	3.14	1.21	.11(122)	.22
	Minority	3.10	1.45		
Authority	Caucasian	2.66	1.30	-1.23(122)	.22
	Minority	3.05	1.32		
Rules & Deception	Caucasian	3.91	1.07	-1.71(122)	.09
	Minority	4.33	.97		
Responsibility	Caucasian	3.72	1.01	-1.84(122)	.07
	Minority	4.14	.65		

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Table 2

Scale Means, Standard Deviations, t-tests by Gender

	Gender	<i>M</i>	<i>SD</i>	<i>t</i> (df)	<i>p</i>
IMS	Male	1.28	.29	1.23(116.50)	.22
	Female	1.20	.19		
Substances	Male	1.41	1.07	-1.05(122)	.30
	Female	1.64	1.38		
Theft	Male	3.03	1.25	-.98(122)	.33
	Female	3.25	1.25		
Authority	Male	2.69	1.31	-.32(122)	.75
	Female	2.77	1.32		
Rules & Deception	Male	4.01	1.07	.45(122)	.66
	Female	3.93	1.06		
Responsibility	Male	3.79	1.15	.05(122)	.96
	Female	3.79	.71		

CRITICAL HIRE-SCREEN CRITERION VALIDITY

Table 3

Scale Means, Standard Deviations, and t-tests by Age

	Age	<i>M</i>	<i>SD</i>	<i>t</i> (df)	<i>p</i>
IMS	Under 40	1.07	.26	.45(122)	.66
	40 +	1.05	.23		
Substances	Under 40	1.62	1.39	1.06(118.77)	.29
	40 +	1.40	.97		
Theft	Under 40	3.28	1.26	1.48(122)	.14
	40 +	2.95	1.23		
Authority	Under 40	2.60	1.29	-1.15(122)	.25
	40 +	2.88	1.32		
Rules & Deception	Under 40	4.07	1.03	1.13(122)	.26
	40 +	3.86	1.10		
Responsibility	Under 40	3.85	.97	.79(122)	.43
	40 +	3.71	.99		