Cornell University School of Hotel Administration
The Scholarly Commons

Articles and Chapters | School of Hotel Administration Collection

11-2009

The Utility of Integrity Testing for Controlling Workers' Compensation Costs

Michael C. Sturman
Cornell University, mcs5@cornell.edu

David Sherwyn
Cornell University, dss18@cornell.edu

Follow this and additional works at: http://scholarship.sha.cornell.edu/articles

 Part of the Hospitality Administration and Management Commons, and the Human Resources Management Commons

Recommended Citation

Sturman, M. C., & Sherwyn, D. (2009). The utility of integrity testing for controlling workers' compensation costs [Electronic version]. *Cornell Hospitality Quarterly, 50*(4), 432-445. Retrieved [insert date], from Cornell University, School of Hospitality Administration site: http://scholarship.sha.cornell.edu/articles/75/

This Article or Chapter is brought to you for free and open access by the School of Hotel Administration Collection at The Scholarly Commons. It has been accepted for inclusion in Articles and Chapters by an authorized administrator of The Scholarly Commons. For more information, please contact hlmdigital@cornell.edu.

The Utility of Integrity Testing for Controlling Workers' Compensation Costs

Abstract
Integrity tests can be a hiring tool to help employers determine which of their prospective employees will be more likely to engage in unproductive, dangerous, or otherwise risky actions on the job. Candidates are surprisingly candid when answering test questions about their workplace theft or drug use, but the tests also have control questions intended to indicate when an applicant provides false answers in an attempt to answer "correctly." Although tests represent an additional expense in the hiring process, a study of a large hotel chain found that the savings in screening out potentially expensive employees more than made up for the costs of conducting the tests, based on a substantial reduction in workers' compensation claims. A conservative estimate is that the company experienced a 50 percent one-year return on investment from the test. It is important to note that the tests do not violate U.S. employment laws, as data show that the tests create no adverse impact on protected groups.

Keywords
human resources, employment tests, integrity testing, workers' compensation

Disciplines
Hospitality Administration and Management | Human Resources Management

Comments
Required Publisher Statement
© Cornell University. Reprinted with permission. All rights reserved.

This article or chapter is available at The Scholarly Commons: http://scholarship.sha.cornell.edu/articles/75

© 2009 CORNELL UNIVERSITY
DOI: 10.1177/1938965509340989
Volume 50, Issue 4 432-445

The Utility of Integrity Testing for Controlling Workers' Compensation Costs

by MICHAEL C. STURMAN and DAVID SHERWYN

Integrity tests can be a hiring tool to help employers determine which of their prospective employees will be more likely to engage in unproductive, dangerous, or otherwise risky actions on the job. Candidates are surprisingly candid when answering test questions about their workplace theft or drug use, but the tests also have control questions intended to indicate when an applicant provides false answers in an attempt to answer "correctly." Although tests represent an additional expense in the hiring process, a study of a large hotel chain found that the savings in screening out potentially expensive employees more than made up for the costs of conducting the tests, based on a substantial reduction in workers' compensation claims. A conservative estimate is that the company experienced a 50 percent one-year return on investment from the test. It is important to note that the tests do not violate U.S. employment laws, as data show that the tests create no adverse impact on protected groups.

Keywords: human resources; employment tests; integrity testing; workers' compensation

Integrity tests are selection devices designed to measure an applicant's attitudes toward theft, dishonesty, drug use, and other counterproductive work behavior. Various forms of integrity tests have existed for well over three decades, and research into understanding whether, why, and how integrity tests work continues to accumulate (Berry, Sackett, and Wiemann 2007; Wanek, Sackett, and Ones 2003). Despite years of research demonstrating their validity,

practitioners still hesitate to use them—in the hospitality industry and other businesses, both in the United States and abroad (Ryan et al. 1999). In this article, we (1) describe integrity tests, (2) review the research that provides extensive evidence indicating the value of these tests to the industry, and (3) demonstrate some of the benefits of integrity tests using two separate examples of one such test that was implemented at two hotel companies. The results show that integrity tests can identify applicants' attitudes toward counterproductive work behavior, do not cause adverse impact, and can improve the profitability of companies choosing to implement them. As we show in this study, adding an integrity test to a selection system can prove to be a human resource intervention with a substantial return on investment (ROI).

Background on Integrity Tests

Employers have always been interested in avoiding unsafe or unethical employee behavior, most notably theft. Integrity tests began as a method of detecting dishonest applicants without having to use polygraph tests, which have been made illegal by federal law for most employment settings (Employee Polygraph Protection Act of 1988). With the polygraph unavailable, and with concerns increasing about employee theft and other unsafe behaviors in the workplace, employers have sought a selection device to help screen out prospective applicants who might engage in such activities (Ones and Viswesvaran 1998; Sackett 1994; Sackett and Wanek 1996).

One such device is integrity tests, which are generally seen as coming in two forms: overt integrity tests and personality-based tests. This report examines the validity and utility of overt integrity tests, which are specifically designed to predict the predisposition of job applicants to engage in theft, drug taking, or dishonest or otherwise disruptive work behaviors. These tests ask specific questions about these behaviors, and thus the purpose of these tests is clear to those taking them. Making no attempt to disguise their purpose, these tests ask specific questions about attitudes toward theft, self-assessments of honesty, and admissions of unethical or illegal past behaviors (Ryan and Sackett 1987). Personality-based tests, by contrast, do not specifically try to detect theft or theft-related behaviors. Instead, they use measures of personality dimensions, such as reliability, conscientiousness, trustworthiness, and sociability, to predict a broad array of counterproductive behaviors. Consequently, the purpose of these tests is not immediately apparent to those taking the test, and many employees do not see personality-based tests as being job-related.

To gauge applicants' attitudes toward counterproductive work behaviors, overt integrity tests simply ask questions of applicants about their views and personal experiences with regard to topics such as stealing and taking drugs. For example, one question is, "What is the most you have ever stolen: (1) $0; (2) $1-$200; (3) $201-$500; or (4) more than $500." Another question frequently asked is, "Would you fail a drug test?" While it may seem obvious that the desired answer for the first question is "$0" and the desired answer for the second is "no," integrity tests rely on applicants' choosing other answers. Indeed, research on integrity tests has found substantial variance in how people respond when answering such questions. Applicants often give surprisingly candid answers (or at least, they do not always respond with what would be considered the "desired" answer) (Ones, Viswesvaran, and Schmidt 1993; Ryan and Sackett 1987; Sackett and Wanek 1996).

Although overt integrity tests are no longer seen as exact substitutes to polygraph tests, research has demonstrated

over the past thirty years that these tests can predict many of the employee behaviors that employers would like to screen out. Studies have often developed a scale of counterproductive behaviors. Such scales represent a broad spectrum of disruptive behaviors, such as actual theft, admitted theft, dismissals for actual theft, illegal activities, absenteeism, tardiness, and violence. Needless to say, employers would like to eliminate (or at least reduce) such behaviors in their workplace. An extensive meta-analysis by Ones and colleagues has shown integrity tests to be fairly strong predictors of counterproductive behaviors (Ones, Viswesvaran, and Schmidt 1993). It has also been shown that integrity tests predict overall job performance ratings with moderate validity.[1] However, when looking at only predictive studies with samples of job applicants, the estimated true correlation was even stronger.[2] The tests also have moderate validity for predicting theft. Again, research has showed that overt tests predicted theft with a fairly low correlation; however, because theft has a low base rate (i.e., it is a relatively rare event), correcting the correlation for this low base rate yields a much higher estimated validity for predicting theft.[3] Additionally, it has also been shown that integrity tests predict these behaviors at various levels of job complexity, and they work for both current employees and job applicants (Ones, Viswesvaran, and Schmidt 1993).

Given this fairly substantial body of research suggesting that integrity tests are useful, the question becomes, Why are so few companies employing them? Indeed, a study of 959 organizations in twenty countries found that integrity tests were rarely or never used (Ryan et al. 1999). One possible reason for this omission is that, despite the quantity of research related to integrity tests, the research failed to consider those tests from the perspective of an investment in human resources by evaluating ROI.

Instead, much of the research on integrity testing has considered the relationship of a particular test outcome with a follow-up questionnaire that gives a self-report of the item being measured. In this instance, when considering the relationship between the results of a particular integrity test and a follow-up questionnaire or other self-report measure of employee theft, it is unclear whether the results give evidence of the test's validity in identifying thieves or simply its reliability in encouraging consistent responses from the test taker. That is, if test takers are willing to report their unethical behavior on an integrity test, it is not surprising that they would then be willing to report the same activity on a follow-up questionnaire. The opposite is also true, in that if one is unwilling to admit to unethical behavior on a test, one may similarly not be willing to admit to such behavior on a subsequent questionnaire. Thus, the evidence of correlation based on self-report data of this kind has only marginal utility. Indeed, the relationship between integrity test results and admissions of counterproductive behavior are much higher than the relationship between test results and external measures of counterproductive behavior.

1. The Ones, Viswesvaran, and Schmidt (1993) study showed a relationship between integrity test score and overall performance ratings to have an average raw correlation of .18. Corrected for measurement error, the correlation was .30.
2. When looking at this subset of studies, which are arguably a better assessment of the true validity of integrity tests in an actual work setting, the estimated true correlation was .41.
3. Ones, Viswesvaran, and Schmidt (1993) show that the uncorrected correlation was .13. Corrected for measurement error and the low base rate, the estimated true correlation was .33.

Furthermore, since a comparatively small percentage of employees engage in counterproductive work behaviors (e.g., employee theft, workplace violence), these criteria are a poor means to evaluate integrity tests (Camara and Schneider 1994; Sackett and Wanek 1996).

While it is true that studies examining the validity of integrity tests have examined a number of criteria clearly of value to organizations (e.g., job performance, theft, turnover, absenteeism), the actual value (or cost) of these behaviors is notoriously difficult to quantify. Although methods exist to link such outcomes with financial consequences (Boudreau 1991; Cascio 2000; Cascio and Boudreau 2008), such methods can be mathematically complex (Sturman 2000) and are often difficult to explain to managers or leave managers unconvinced of the results (Latham and Whyte 1994; Whyte and Latham 1997). Thus, while these studies may present a strong case that integrity test results are associated with various desired outcomes, the research has so far not been able to provide empirical evidence of the specific value associated with investing in an integrity test as part of a selection system. We fill that void in this article by examining the actual ROI results of implementing a test at a large hotel chain. We examined the testing program's effect on workers' compensation claims to assess the validity and utility of implementing an integrity test in an actual work context.

The Predicted Effects of Integrity Testing on Workers' Compensation Claims

To date, no academic research has looked at the effect of integrity testing on workers' compensation claims. While vendors of integrity tests have conducted their own research along this line (including the company whose integrity test was used in this study), no independent evaluation has been reported in the academic literature. Our purpose is not to question the vendors' studies but to make an independent evaluation of the tests (Camara and Schneider 1994).

In the United States, workers' compensation provides monetary benefits to employees injured on the job. Put generally, the workers' compensation system requires employers to give cash benefits and medical care for injuries or illnesses that arise out of and during the course of employment (Mathis and Jackson 2006). Each state has its own workers' compensation law, and thus benefits and liabilities vary, but the expense is considerable, ranging from 2 to 10 percent of payroll, depending on the employer and the state (Mathis and Jackson 2006). For any business, a reduction in workers' compensation claims means a direct benefit to the business. Workers' compensation claims directly reduce profitability, and thus a $1 savings in workers' compensation claims results in a $1 improvement in pretax profits.

The connection between integrity tests and workers' compensation claims may not immediately be apparent, so let us look at the reasons we would expect such a connection. First, because integrity tests ask individuals about their use of illegal drugs, the tests can be used to screen out those who are inclined to use drugs at work. Individuals under the influence of drugs are dangerous to themselves and to others. According to the U.S. Bureau of Labor Statistics, the hospitality industry has roughly five to six accidents per one hundred full-time workers per year. Between 10 and 20 percent of U.S. workers who die on the job test positive for alcohol or drugs (Weber and Cox 2001). The U.S. Bureau of Labor Statistics reported 5,320 deaths in 2006 for in all industries (U.S. Department of Labor 2006). On-the-job substance abuse is also a

significant predictor of the likelihood of workplace accidents (Frone 1998), and roughly 14 percent of U.S. employees report using illegal drugs at work (Frone 2006). Screening out those who admit to drug use at work can clearly lead to a safer work environment. Second, workers' compensation fraud is a growing concern, with some estimates that 25 percent of workers' compensation claims are fraudulent (Mathis and Jackson 2006). By screening out dishonest employees, an integrity test can reduce workers' compensation claims by reducing the number of fraudulent claims and decreasing the frequency of exaggerated claims. In sum, if companies could screen out those who would be more inclined to use drugs at work and engage in dishonest behaviors, this could help reduce the number of workplace accidents, reduce the number of falsely reported accidents, and decrease the instances of exaggerated claims.

Results from the Integrity Test

To test the validity and utility of implementing an integrity test, we examined integrity test data from a large hotel chain.[4] At the time the test was implemented, the company had 27,266 employees. Over a nearly one-year period, the test was given to 29,043 applicants, of whom 6,079 were hired.

To evaluate the performance of the integrity test, we consider two sets of results. First, we examine the following qualitative questions: (1) Does the test have adverse impact? (2) Do applicants actually indicate that they engage in high-risk behaviors? (3) Do individuals fake answers when taking the test? and (4) What does the test cost? Second, we calculate whether the test actually provides a positive ROI.

Tests for Adverse Impact

As with all preemployment screening tools, integrity tests invite potential legal concerns. Although tests may be able to predict valuable outcomes, it is important that they not violate relevant employment laws. In particular, the tests should not inadvertently discriminate against protected classes. Although the validity of job-related test results can be a defense against adverse impact claims, the costs of defending such claims are so great that it is vital that there not even be evidence suggesting a prima facie case of adverse impact (Sherwyn, Tracey, and Eigen 1999). Therefore, the first step in this analysis is to see whether there is any potential risk when implementing an integrity test.

Adverse impact occurs when there is a significant underrepresentation of a protected class in employment decisions (e.g., for hiring, placement, or promotion). For a selection device like an integrity test, a prima facie case for adverse impact is met if the passing rate for a given protected class (i.e., race, sex, age, national origin, color, religion, disability)[5] is less than 80 percent of the passing rate for the majority group (commonly called the 4/5ths rule). Contrary to what one might expect, the term "majority" for adverse impact purposes refers to the demographic group that performed the best on the policy or practice in question. Despite the often-held perception that majority refers to whites or men, the majority in this case is people of color (of

4. The company employed the Tescor Survey, which is a version of an overt integrity test. American Tescor provided the data for this study but provided no funding for writing this report. More information on American Tescor can be found at www.americantescor.com.
5. The protected classes listed above are based on U.S. federal law. It should be noted, though, that state and local laws can expand this list (e.g., sexual orientation and family status).

Exhibit 1:
Adverse Impact Data on Test Results (in percentages)

Protected Class	*High Risk*	*Moderate Risk*	*Low Risk*	*Hirable*	*Percentage of Majority Group*
Total	31	19	50	69	—
Male	33	19	48	67	93
Female	28	19	53	72	—
White	33	22	45	67	92
Nonwhite	29	17	54	71	97
Black	30	19	51	70	96
Latino	27	16	58	73	—
Asian	33	17	50	67	92
Native American	32	16	52	68	93
Other	38	25	37	62	85
Over forty	31	17	52	69	99
Under forty	30	20	50	70	—

whom 71 percent were considered hirable, as moderate or low risk; and 54 percent were classified as low risk), as compared to whites (67 percent of whom were classified as hirable, and 45 percent of whom were classified as low risk).

The results from prior research on integrity tests suggest that adverse impact is not a significant risk. A study based on four large databases showed that there were no significant race differences for blacks, Hispanics, Asians, or American Indians compared to whites (Ones and Viswesvaran 1998). Individuals over forty actually scored somewhat better than those under forty. Test results did provide some evidence of adverse impact against men: women scored 0.16 standard deviations higher than men. Because of the large sample size, this difference is statistically significant; however, the magnitude of this effect is generally interpreted as being small (Cohen 1977). It is also not clear whether this effect would be large enough to violate the 4/5ths rule. Nonetheless, it is a point of potential concern.

The chain's results are similar to those reported in prior research. Tests for adverse impact for a variety of protected classes are shown in Exhibit 1. Women scored somewhat higher than men, with 72 percent being classified as hirable, compared to 67 percent of men. Men were thus classified as hirable at 93 percent of the rate of women, and thus the test did not violate the 4/5ths rule. Results also did not indicate adverse impact for race or age. All hiring ratios were at least 90 percent, with the exception of those classified as "Other," but this hiring ratio was still 85 percent of the majority group (in this case, Latinos) and 93 percent of the hiring rate of whites.

Overall, the results demonstrate that the integrity test is legally defensible in that it does not cause adverse impact. Next, we examine the outcomes associated with the integrity test to determine whether applicants actually do indicate inclinations to engage in high-risk behavior, whether the tests can detect faking, and whether the use of the integrity test can reduce workers' compensation claims.

Detecting High-Risk Behaviors

The chain's test data show that the integrity test can be used to distinguish

between applicants. Overall, of the 29,043 applicants who took the integrity test, 31 percent were classified as "high risk," because they reported that they engage in questionable behavior.

One dimension of the test is for test takers to report how much they have stolen. A total of 1,881 applicants reported stealing cash from previous employers. Of these, 698 (37 percent) reported stealing up to $25, 275 (15 percent) reported stealing between $25 and $500, and an astonishing 908 (48 percent) reported stealing more than $500.

Roughly 2,400 applicants admitted taking merchandise from previous employers, with 796 of these applicants (33 percent) reporting the theft of more than $500 in merchandise. Employees also responded positively to questions regarding whether they had shoplifted in the past year, would help a friend steal, would steal if they had low pay, and would fake time cards if never checked.

In addition to questions about theft, applicants admitted on the test that they used illegal drugs. The most common admissions were for cocaine (1,507), hashish (1,100), and hallucinogens (1,050). Additionally, 1,338 employees admitted regular drug use at work, 1,955 admitted drinking at work, and nearly 2,000 employees admitted that they would fail a urinalysis.

In all, the results clearly demonstrate that many job applicants openly admit to theft and drug use when completing an integrity test, and as a consequence this test screened out nearly one-third of applicants. Despite that observation, it is important to note that neither researchers nor vendors claim that these tests are completely accurate. Even though the tests typically contain questions designed to catch those who are faking answers, it is highly likely that some individuals do get away with faking answers on the drug or theft questions. Thus, employers should not think that integrity tests can perfectly eliminate all undesirable behaviors. Nonetheless, integrity tests clearly can be used to screen out a large number of individuals who freely admit that they engage in behaviors that employers simply do not want in an employee.

Detecting Faking

Our next analysis was to see whether the test's faking scale was able to identify those who were attempting to "game the test." To identify applicants who falsely answer questions about unsafe or illegal activities, integrity tests often include a number of items designed to detect applicants who are providing overly positive answers on the test. For example, tests include questions such as, "Have you ever told a lie?" and "Do you ever get angry?" Since an honest person will admit to having lied at some point or to having lost his or her temper, one can assume that a person who responds negatively to these questions is being dishonest on other questions.

The test's faking scale flagged roughly 9 percent of applicants as "high risk." An additional 20 percent were indicated to be "moderate risk" on the scale. Unfortunately, no research, including this study, can determine whether those who misrepresent themselves perform more poorly on the outcomes of interest to the organization. The chief reason for this lack of data is that those who are flagged as "high risk" for lying on the test were not hired by the organization.

In sum, it seems that integrity tests can elicit information that allows a hiring firm to differentiate between applicants. Although we have not yet tested how well the tool predicts outcomes of interest to the organization, the device seems to have face validity. To examine whether implementing such a test is a good business decision, let us turn

to considering the costs and benefits of implementing this integrity test.

Implementing the Test: Logistics, Costs, and Benefits

Integrity tests are available from a wide range of vendors. In addition to the test employed in this study developed by American Tescor, Inc., other integrity tests include the Personnel Selection Inventory, the Reid Report, and the Stanton Survey. Since designing integrity tests often requires psychometric testing, they are not simple tools to be assembled by nonexperts. However, tests can be purchased relatively easily. American Psychological Association (APA) guidelines for vendors of integrity tests are related to a particular test's construct validity, the development and reporting of scores, marketing and promotional materials, independent evaluation of a test, and the proper use of the test and training of test users (APA 1991; Camara and Schneider 1994). The Tescor integrity test complies with the recommendations issued by the APA, and Tescor is listed as a reviewer of the original APA report.

When considering the purchase and use of integrity tests, companies should seek information on both the design of the instrument and evidence of its validity. Ideally, tests should be evaluated independently. Companies should also keep their own data on the results of integrity tests and validate their own results (as occurs in this study). Ideally, a company should have a clear idea of what it wants an integrity test to accomplish: boosting job performance, reducing turnover and absenteeism, or cutting workers' compensation claims. Companies should collect their own data and verify these results for their own situation.

Integrity tests may be paper-and-pencil, electronically scored, or web-based. The test examined in this article can be either electronically scored or web-based. The test generally takes fewer than twenty minutes to complete. Additionally, integrity tests are available in many different languages[6] and are generally written at the fifth-grade reading level.

Test Cost

The cost of purchasing a test is often a function of multiple factors, including how many tests will be purchased. Although the numbers may sound large, the cost of an integrity test is relatively low, considering the potential cost of workers' counterproductive behaviors. The Tescor screening test costs at most $20 per test, with discounts for volume. In the example of this hotel chain, the cost per test ended up at roughly $9.30. Thus, the total cost of implementing the test in our sample was $270,100.

The Predictability of Workers' Compensation Claims

After the integrity test had been implemented for a full year, we compared the workers' compensation claims of screened employees with those of unscreened employees (i.e., those hired before the integrity test was implemented). Using the existing set of unscreened employees as a control group has a number of advantages, but the primary disadvantage is that it is not a true experiment given that employees are not randomly assigned into the experimental and control groups. Hence, this study proceeds using a quasi-experimental design (Shadish, Cook, and Campbell

6. The Tescor Survey is available in Albanian, Bosnian, Chinese (Mandarin), Creole (Haitian), English, French, German, Hmong, Oromo, Polish, Portuguese-Brazilian, Portuguese-Continental, Romanian, Russian, Samoan, Somalian, Serbo-Croatian, Spanish, Tagalog, Tongan, and Vietnamese.

Exhibit 2:
Workers Compensation Claims for Screened versus Unscreened Employees

Screening Status	*Number of Employees*	*Claim Frequency*	*Number of Claims*	*Total Claims per Group*	*Average Cost per Claim Filed*	*Average Cost per Employee in Group*
Unscreened	27,265	2.82%	769	$2,665,712	$3,466	$97.77
Screened	6,079	1.46%	89	$188,589	$2,119	$31.02
Total	33,344	2.57%	858	$2,854,301	$3,327	$85.60

Note: The data above are reported after the elimination of one extreme outlier. One employee from the unscreened group was involved in an accident with more than $600,000 in claims. Because this was so much greater than all claims from either group, it was eliminated so as to not unduly bias the result.

2002). This procedure also faces the risk of maturation, since those in the unscreened group could experience natural changes that could cause differences with the screened group (i.e., they are older than when they were employed and have gained experience). Thus, we need to carefully evaluate any differences between the two groups and be careful to rule out alternative explanations for those findings.

The advantages this procedure has over other studies on integrity tests include the fact that it uses actual outcomes (i.e., workers' compensation claims) and it investigates an objective outcome that has clear business implications. So while an experiment may provide better control when manipulating subjects, a field study such as this one benefits from a realistic and generalizable setting (McGrath 1982). The other advantage of this approach is that the control group provides a basis for a realistic comparison, because (1) it records actual employees' job behavior, (2) the applicants who took the integrity test were doing so in an actual employment experience, and (3) the outcome is objectively measured and has a direct effect on organizational profitability. Thus, while no methodology is perfect (McGrath 1982), this quasi-experimental design, even with noted limitations, provides a valuable opportunity to assess the utility of an overt integrity test.

Comparisons of workers' compensation claims, including frequency and costs, are provided in Exhibit 2. As the table shows, the percentage of workers' compensation claims for those screened was 1.46 percent, whereas employees who were unscreened filed at a rate of 2.82 percent. These percentages are small, reflecting the typical frequency of workers' compensation claims. Nevertheless, those who were employed without being screened had workers' compensation claim frequencies nearly double those of employees screened by the integrity test. Statistical testing showed that the observed difference was greater than we might expect to observe due to chance (statistically significant at $p < .0001$).

Not only was the screened employees' claim frequency lower, but the average size of a claim from a screened applicant was smaller than that of an unscreened incumbent. As shown in Exhibit 2, the average claim filed by a screened candidate was $2,119, whereas the average claim for an unscreened incumbent was $3,466. To compare these quantities, we looked at two potential differences: the variances of the distributions and the means of the distributions.

Before comparing the means of the distributions, we tested to determine whether the two distributions' variances were equal. Results clearly show that claims in the

unscreened group were far more variable than those in the screened group. In the unscreened group, the standard deviation of claims was $9,543, whereas it was $4,452 for the screened group. This difference was statistically significant ($p < .0001$), thus indicating that a test comparing the means should be based on an analysis assuming unequal variances for the two groups. More practically, given that claims were bounded by a minimum value of $0, these results show that the distribution of claims for unscreened employees had far more cases of high values than did the screened employees' distribution.

The comparison of the means showed that the two groups are statistically significantly different ($p < .05$). That is, the unscreened group had higher average claims ($3,466) than did the screened group ($2,119), and this difference was greater than we would expect to observe due to chance fluctuations.

Another way of looking at the same data is to examine the percentage of "large" claims filed by employees. Only 20 percent of employee claims were greater than $2,000, but these claims accounted for more than 89 percent of total claim amounts. Of that group, only 17 were from screened employees (0.28 percent of the screened group), whereas 158 came from unscreened employees (0.58 percent of the unscreened group). Furthermore, the average claim size within this group was $9,564 for screened employees and $15,070 for unscreened employees. Both the probability of being in this large claim group and the average size of the claims were statistically significantly different for those screened as compared to those not screened. Specifically, a screened employee was statistically significantly less likely to file a large claim ($p < .01$), and the average size of a large claim for screened employees was also significantly smaller than for unscreened employees ($p < .01$). This indicates that risky employee behaviors (such as drug use) are related to the probability of filing a claim and to the size of a claim that is filed.

Looking at the hotel chain's overall situation, the 6,079 screened employees produced a total of 89 claims, totaling $188,589. The average size of those 89 claims was $2,119. Dividing the total claim amount by the total number of screened employees, the overall expected average cost of a screened employee was $31.02. The 27,265 unscreened employees were responsible for a total of 769 claims, totaling $2,665,712, for an average claim size of $3,466. Using the same overall calculation as above, the average cost per unscreened employee was $97.77. On average, a screened employee was expected to cost $66.75 less (per year) than an unscreened employee, and this is only in terms of workers' compensation claims. We did not examine the benefits of avoiding theft, reducing absences, or having better-performing employees (since we had no data on these outcomes).

Having made those calculations, we cannot conclusively state that the difference is solely attributable to the use of the integrity test, because this methodology provides for no controls that could provide strong conclusions of causality. Nevertheless, it is reasonable to believe that the results are truly attributable to the integrity test and not to other possible explanations.

First, the company made no other major changes to its selection system beyond the inclusion of the integrity test during this time period. Applicants who failed the integrity test were excluded from further consideration, but the rest of the selection system proceeded without any significant changes. Other, nonsystemic changes—such as different recruiters being involved in the process, different experience levels of those conducting interviews, different

positions being open, different applicant pools over different years—should not be sufficiently strong to affect the results; and in any event, we cannot control for all such factors in this analysis. Given that there was no strategic change to the process, we suggest that it is highly unlikely that the magnitude of differences observed in this study could be due to the sort of random fluctuations that affect all recruitment systems over time. With the decision processes in the selection system largely otherwise unchanged, we submit that the evidence is compelling that the introduction to the system of the integrity tests led to this substantial decrease in workers' compensation claims.

Second, one could argue that the differences between the newer (screened) employees and the unscreened employees may have caused the decrease in claims. For example, perhaps the type of people who applied for jobs when the company was screening applicants with the integrity test just happen to be less accident-prone than the current employees of the organization. Research on accidents, though, suggests that if anything, the opposite would occur. Specifically, prior research on workplace accidents shows that both age and tenure are negatively related to the likelihood of workplace accidents (Cellier, Eyrolle, and Annick 1995). In general, the addition of a year of organizational tenure and job experience are both related to increases in job performance (Sturman 2003b). The unscreened group by definition has more organizational tenure than the screened group (because they were hired before the screened group); thus, the comparison of the screened group to job incumbents is actually a relatively conservative test of the selection device's performance. Note also that screened employees were employees present after at least one year of employment, and thus this group had undergone a year of self-selection (employees quitting) and termination. Given that prior research has shown drug use to be correlated with firing (Mathis and Jackson 2006), this again suggests that the potential true benefit of the integrity test is even greater than the comparison against the unscreened group already suggests. Thus, despite our method requiring us to compare cohorts from different years, the strong effects observed for the screened group further emphasized the potential effect that integrity testing has on reducing workplace accidents.

Estimating the Return on Investment

The return for this hotel chain's investment in the test seems to be quite favorable. Using the example from the large hotel chain, at the cost of $9.30 per applicant, an average hire screened by the integrity test is expected to cost $66.75 less than one unscreened. Given that there were 29,043 applicants screened, and 6,079 employees hired, the test had a benefit of $405,773 at a cost of $270,100. In other words, there was cost reduction of $135,673, which goes directly to the bottom line. Thus, the investment in the integrity test yielded an ROI of 50 percent.

We have stated this return conservatively. If employees remain with the company for multiple years, the value of this benefit should increase, because the average benefit of reduced workers' compensation claims should last for each year the employee remains with the firm. Thus, the benefits of integrity tests can compound as screened employees remain in their jobs for multiple years. Again, the value of any other benefits of integrity tests, including improved employee performance and reduced counterproductive behaviors, provide additional financial return, though that value cannot be quantified in this study.

We must note that the price this company paid for the integrity test was reduced due to bulk purchase. Since another company's return will depend on the actual price paid for the test and the ratio of the number of hires to number of applicants, let us run the ROI calculations using the "full price" of $20 per applicant. We again use the figures of 69 percent of applicants classified as employable and the average benefit of $66.75 per hire, as experienced by this hotel company. With those assumptions, the average return of the integrity test per hire is $37.75 per hire, for an ROI of 130 percent. This is calculated as follows: If 69 percent of applicants are classified as employable, then the company needs to test 1.45 applicants per vacant position. This yields an expected cost of $29 per hire (1.45 applicants × $20 per applicant). With an average benefit of $66.75 per hire, the expected return from administering the integrity test is $37.75 (or $66.75 – $29). Thus, the ROI is $37.75/$29, or 130 percent.

This analysis pertains only if the integrity test is the only selection device used in the hiring decision. Most companies use other selection devices (as they ought to), and so the return on the test investment would decrease as more individuals need to be tested. Some applicants classified as employable by the integrity test will be rejected based on the results of interviews, background checks, or other selection devices. Despite clear benefits from the tests, companies should consider the role that integrity tests will play in their overall selection process.

A Follow-Up Replication

As noted at several points earlier, this study has limitations that make us temper our conclusions. In particular, we had to compare two cohorts of employees that could have been so different as to render any comparisons moot. While this allowed an examination of the implementation of the integrity tests and the business outcomes associated with the two cohorts, it is impossible to rule out all alternative explanations for the effects that we observed. Similarly, while there were no major changes to the selection system (aside from the introduction of the integrity test), we could not examine the actual decision making or decision makers in the selection process to ensure consistency across the two years. It was also infeasible for us to collect additional information from subsequent years to assess whether the effects associated with the integrity test continued past the first year. Thus, to provide some additional evidence as to the utility of the integrity test in a hospitality context, we conducted the same analysis with a different implementation of the test. In this replication, we examined data from a different hotel management company that had implemented the same integrity test in its properties. Note that all the data in this replication sample are completely independent of the data from the first sample.

The replication confirms the results from the original hotel chain. As shown in Exhibit 3, the probability of a claim and the average claim size for this latter company are both significantly lower in the screened group than in the unscreened group (probability of a claim, $p < .0001$; average claim, $p < .05$). The selection device cost the company $7.14 per test, and a total of 77 percent of employees passed, of whom 1,464 were ultimately hired. The total cost of adding the integrity test to the selection method was $43,086. As shown in Exhibit 3, the average cost of a screened employee is expected to be $278.46 less

Exhibit 3:
Replication: Workers Compensation Claims for Screened versus Unscreened Employees in Second Sample

Screening Status	*Number of Employees*	*Claim Frequency*	*Number of Claims*	*Total Claims per Group*	*Average Cost per Claim Filed*	*Average Cost per Employee in Group*
Unscreened	2,788	6.35%	177	$958,372	$5,415	$343.75
Screened	1,464	2.94%	43	$95,580	$2,223	$65.29
Total	4,252	5.17%	220	$1,053,952	$4,791	$247.87

than an unscreened employees.[7] Thus, the total benefit of the integrity test for hiring the 1,464 employees was $407,665: a return of $364,579 ($407,665 – $43,086), or an ROI of 846 percent.

The results of this replication lend confidence to our conclusion that administering an integrity test to job candidates can help control workers' compensation costs. The results of the second analysis are even stronger than in the original example. Again, we acknowledge the limitations of this quasi-experimental design and accept that we cannot rule out all possible alternative explanations. Nonetheless, the finding of the same pattern of results in two independent situations where integrity tests were added to an existing selection system lends fairly high confidence that the effects we observed in each case are at least partially the result of the new screening tool.

Conclusion

While these results are interesting and provide evidence that integrity tests can be a valuable component of a selection system, no study is perfect. This study contributes to research on integrity testing by examining actual results (i.e., workers' compensation claims) that have a direct effect on the profitability of the organization; however, the nature of these data prevents us from being able to conduct a controlled experiment. Unlikely though it may be, it is possible that changes occurred across years that increased the likelihood and severity of workers' compensation claims among the existing workers. Certainly, the inclusion of the independent replication provides some additional confidence in the original findings, but further study on integrity tests using controlled experimental methods would provide a useful alternative perspective on the sort of effects integrity tests should yield.

A longitudinal study would also be useful. The study described here is essentially a snapshot in time, since one year is a brief time to evaluate a selection device. Although our finding of the same pattern of results in the independent replication provides support for our conclusion that the changes we observed were indeed attributable to the integrity tests, a longitudinal study would provide greater confidence in our findings.

Yet even with these limitations, we argue that these findings indicate that an integrity test can have a positive return for companies in the hospitality industry. Prior research evidence and the examples from the two hotel companies in this article show that integrity tests (1) can predict outcomes of importance to organizations, (2) do elicit responses from applicants that

7. An unscreened employee has an "expected" cost of $343.75, computed as the total cost for the unscreened employees divided by the number of unscreened employees. Similarly, the expected cost of the screened employees is $65.29. Hence, a screened employee is expected to cost $278.46 less than an unscreened employee.

allow one to select out potentially high-risk employees, (3) do not create adverse impact, and (4) produce returns that even when conservatively estimated are greater than the cost of administering the test. We calculated an immediate 50 percent ROI for one company and as much as an 800 percent ROI for the other. It thus appears that integrity tests are a potentially highly useful selection tool, one that hospitality companies should consider implementing.

References

American Psychological Association. 1991. *Questionnaires used in the prediction of trustworthiness in pre-employment selection decisions: An APA task force report*. Washington, DC: American Psychological Association.

Berry, C. M., P. R. Sackett, and S. Wiemann. 2007. A review of recent developments in integrity test research. *Personnel Psychology* 60:271-301.

Boudreau, J. W. 1991. Utility analysis for decisions in human resource management. In *Handbook of industrial and organizational psychology*, vol. 1, ed. M. D. Dunnette and L. M. Hough, 621-745. Palo Alto, CA: Consulting Psychologists Press.

Camara, W. J., and D. L. Schneider. 1994. Integrity tests: Facts and unresolved issues. *American Psychologist* 49:112-19.

Cascio, W. F. 2000. *Costing human resources: The financial impact of behavior in organizations*. 14th ed. Boston:Kent.

Cascio, W., and J. Boudreau. 2008. *Investing in people: Financial impact of human resource initiatives*. Upper Saddle River, NJ: Pearson Education.

Cellier, J. M., H. Eyrolle, and B. Annick. 1995. Effects of age and level of work experience on occurrence of accidents. *Perceptual and Motor Skills* 80:931-40.

Cohen, J. 1977. *Statistical power analysis for the behavioral sciences*. San Diego, CA: Academic Press.

Frone, M. R. (2006). Prevalence and distribution of illicit drug use in the workforce and in the workplace: Findings and implications from a U.S. National Survey. *Journal of Applied Psychology, 91*, 856-869.

Fronem M. R. (1998). Predictors of work injuries among employed adolescents. *Journal of Applied Psychology, 83*, 565-576.

Latham, G. P., and G. Whyte. 1994. The futility of utility analysis. *Personnel Psychology* 47:31-46.

Mathis, R. L., and J. H. Jackson. 2006. *Human resource management*. 11th ed. Mason, OH: South-Western.

McGrath, J. E. 1982. Dilemmatics: The study of research choices and dilemmas (pp. 69-102). In *Judgment calls in research*, ed. J. E. McGrath, J. Martin, and R. A. Kulka. Beverly Hills, CA: Sage.

Ones, D. S., and C. Viswesvaran. 1998. Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology* 83:245-69.

Ones, D. S., C. Viswesvaran, and F. L. Schmidt. 1993. Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology* 78:679-703.

Ryan, A. M., L. McFarland, H. Baron, and R. Page. 1999. An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology* 52:359-91.

Ryan, A. M., and P. R. Sackett. 1987. Pre-employment honesty testing: Fakability, reactions of test takers, and company image. *Journal of Business and Psychology* 1:248-54.

Sackett, P. R. 1994. Integrity testing for personnel selection. *Current Directions in Psychological Science* 3:73-76.

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* 49:787-829.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin.

Sherwyn, D., J. B. Tracey, and Z. Eigen. 1999. In defense of mandatory arbitration of employment disputes: Saving the baby, tossing out the bath water, and constructing a new sink in the process. *University of Pennsylvania Journal of Labor and Employment Law* 2:73-150.

Sturman, M. C. 2000. Implications of utility analysis adjustments for estimates of human resource intervention value. *Journal of Management* 26:281-99.

———. 2003a. Introducing utility analysis for hospitality: A tool for quantifying the value of human resource interventions. *Cornell Hotel and Restaurant Administration Quarterly* 44 (2): 106-16.

———. 2003b. Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management* 29:609-40.

U.S. Department of Labor. 2006. *Census of fatal occupational injuries, 2006*. Washington, DC: U.S. Bureau of Labor Statistics, U.S. Department of Labor.

Wanek, J. E., P. R. Sackett, and D. S. Ones. 2003. Towards an understanding of integrity test similarities and differences: An item-level analysis of seven tests. *Personnel Psychology* 56:873-94.

Weber, W., and Cox, C. (2001). Work-related fatal injuries in 1998. *Compensation and Working Conditions*, vol. 6, Iss. 1 (Spring), pp. 27-29.

Weber, W., and C. Cox. 2001. Work-related fatal injuries in 1998. *Compensation and Working Conditions*, Spring.

Whyte, G., and G. P. Latham. 1997. The futility of utility analysis revisited: When even an expert fails. *Personnel Psychology* 50:601-10.

Michael C. Sturman, PhD, Cornell University, is an Associate Professor at Cornell University's School of Hotel Administration, where he is also the Kenneth and Marjorie Blanchard Professor of Human Resources. **David S. Sherwyn,** J.D., Cornell University is an Assistant Professor of Cornell University's School of Hotel Administration.