Personality and Absenteeism: A Meta-Analysis of Integrity Tests

DENIZ S. ONES1*, CHOCKALINGAM VISWESVARAN2 and FRANK L. SCHMIDT3

1Department of Psychology, University of Minnesota, USA
2Department of Psychology, Florida International University, USA
3Department of Management and Organizations, University of Iowa, USA

Abstract
Until recently, research focus has been on a variety of demographic, attitudinal, and organizational variables in predicting and explaining absenteeism. If personality traits predict absenteeism, then it may be possible to use measures of these traits to identify and select job applicants and thereby reduce absenteeism rates. In this research, our goal was to examine whether integrity tests could be used to predict absenteeism. Meta-analysis was applied to studies of the validity of pre-employment integrity tests for predicting voluntary absenteeism. Twenty-eight studies based on a total sample of 13,972 were meta-analysed. The estimated mean predictive validity of personality-based integrity tests was 0.33. This operational validity generalized across various predictor scales, organizations, settings, and jobs (SD = 0.00). Overt integrity tests, however, showed much lower predictive validity for absenteeism and greater variability than personality-based tests (ρ = 0.09; SD = 0.16). The results indicate that a personnel selection approach to reducing absenteeism in organizations may be a useful strategy, particularly if personality-based integrity tests are utilized. Potential explanations for differences between these results and those found for Big Five measures of personality are offered. Future research investigating models of absenteeism should incorporate the personality constructs assessed by integrity tests. Copyright © 2003 John Wiley & Sons, Ltd.

INTRODUCTION
Employee absences are a costly problem for employers (Hackett & Guion, 1985; Lyons, 1972; Muchinsky, 1977). Because of this, the correlates and antecedents of employee absenteeism have been researched extensively over the past 75 years (see e.g. Cooper & Payne, 1965; Evans, 1986; Hill & Trist, 1955; Kornhauser & Sharp, 1932; Naylor & Vincent, 1959; Noland, 1945; O’Hara, Johnson, & Beehr, 1985; Pierce & Newstrom, 1972; Muchinsky, 1977).

*Correspondence to: Deniz Ones, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455-0344, USA. E-mail: Deniz.S.Ones-1@tc.umn.edu

Copyright © 2003 John Wiley & Sons, Ltd. Accepted 19 December 2002
The focus has generally been on a variety of personal, attitudinal, and organizational variables both to predict and to explain absenteeism.

Personal factors empirically studied in conjunction with absenteeism include age (see e.g. Cooper & Payne, 1965; de la Mare & Sergeant, 1961; Hackett, 1990; Jackson, 1944; Naylor & Vincent, 1959; Schenet, 1945), gender and education (e.g., Steel & Rentsch, 1995), job tenure (see e.g. Baumgartel & Sobol, 1959; Hill & Trist, 1955; Noland, 1945), mood (see e.g. George, 1991), and other on-the-job behaviours and outcomes such as accident frequency. The most widely studied attitudinal correlate of absenteeism is job satisfaction. Hackett and Guion (1985), in a meta-analysis, found an average observed correlation of $-0.09$ between measures of absenteeism and job satisfaction. Reviews of organizational variables related to absenteeism are presented by Durand (1985), Muchinsky (1977), and Price and Mueller (1986). More recent seminal reviews of the literature (Johns, 1997, 1998, 2001; Harrison & Martocchio, 1998) note a greater variety of absenteeism correlates, including additional demographic variables, ability to attend, job characteristics, job satisfaction, and absence cultures. Multiple studies have shown that organizational support is negatively correlated with absenteeism (see e.g. Eisenberger, Fasolo, & Davis-LaMastro, 1990). Consistent with predictions using a social exchange paradigm (Chadwick-Jones, Nicholson, & Brown, 1982), employee perceptions of inequity are associated with higher levels of absenteeism (Geurts, Buunk, & Schaufeli, 1994; Schwarzwald, Koslowsky, & Shalit, 1992; van Dierendonck, Schaufeli, & Buunk, 1998; van Yperen, Hagedoorn, & Geurts, 1996).

Yet, in the absenteeism literature few studies have been directed to explain absenteeism in terms of personality, despite the general interest in personality variables in work environments (Barrick & Mount, 1991; Hough, Eaton, Dunnette, Kamp, & McClay, 1990; Hough & Ones, 2001; Ones, Viswesvaran, & Schmidt, 1993; Salgado, 1997). The lack of interest in personality traits is very surprising since past absenteeism has been found to be one of the strongest predictors of future absenteeism (Breaugh, 1981; Farrell & Stamm, 1988; Keller, 1983; Rhodes & Steers, 1990, p. 102). Further, with empirical evidence suggesting that absence taking behaviour is more likely to be an individual differences phenomenon than a group level phenomenon (Yammarino & Markham, 1992), it becomes imperative to identify personality related variables that could be incorporated into models of absenteeism and used for predicting absence taking behaviour. If individual differences variables such as personality traits can be identified that predict absenteeism, then it may be possible to use measures of these traits to select job applicants and thereby reduce absenteeism rates. Considering personality dispositions in making staffing and placement decisions may become a feasible organizational intervention in combating the costly and disruptive problem of absenteeism.

### Personality and absenteeism

There are two lines of research that point to a dispositional basis for absenteeism. First, as already noted, prior absenteeism is the strongest predictor of future absenteeism (see e.g. Baguma, 2001; Garrison & Muchinsky, 1977; Landy, Vasey, & Smith, 1984). In a meta-analytic study, Farrell and Stamm (1988) found that absence history was correlated $0.65$ with current absence frequency and $0.71$ with time lost. In longitudinal research, Rentsch and Steel (1998) showed that frequency of absence predicted frequency of absence over 5 years ($r$ values between 0.74 and 0.53). Even when situations are drastically different, temporal stability of absenteeism persists. Brenner (1968) found that absenteeism in high
school was positively correlated with absenteeism in later employment. Moreover, past absenteeism has been found to predict subsequent absenteeism even in the presence of substantial job design changes (Ivancevich, 1985). Stability of absenteeism over time and across changing environments may in part be due to its enduring, personality based determinants (Ferris, Bergin, & Wayne, 1988; Froggatt, 1970a, 1970b).

Second, also in support of a dispositional basis for absenteeism, some empirical research has been suggestive of specific personality–absenteeism relationships (Furnham & Miller, 1997). In individual studies, neuroticism and conscientiousness related traits have been found to predict absences. Early research reported a positive relation between absenteeism and neuroticism (Bernardin, 1977; Cooper & Payne, 1965) as well as anxiety (Sinha, 1963). Negative correlations with conscientiousness related traits and absenteeism have been noted using the Ego Strength scale of the 16PF (closely related to self-discipline; Bernardin, 1977), overall conscientiousness (Conte & Jacobs, 1999; Hattrup, O’Connell, & Wingate, 1998), and need for achievement (Mowday & Spencer, 1981; Wegge & Kleinbeck, 1993).

In the only primary study to examine the relationships between the Big Five personality dimensions and absenteeism, Judge, Martocchio, and Thoresen (1997) found that extraversion and conscientiousness predicted absenteeism in a sample of 89 university employees. Salgado (2002) reported a meta-analysis of the Big Five personality traits and absenteeism. Eight to twelve studies contributed to the analyses; however, operational validities were disappointing (ranging between −0.08 and 0.06).

In industrial, work, and organizational (IWO) psychology applications, it may be important to recognize differences in personality scale origin and construction. To date, research on personality–absenteeism linkages has focused on measures of normal adult personality (e.g. NEO-PI, 16PF). The aim in the construction of these scales is the accurate description of individual differences in personality. The use of these inventories for personnel screening and selection is only one of their many applications. However, during the past decade, a number of measures of personality at work have come to the attention of the scientific community. Such measures are generically referred to as occupational personality scales (Ones & Viswesvaran, 2001a). The main aim of these measures is the accurate prediction of individual differences in work behaviours of interest. There are four defining characteristics of occupational personality scales: (i) they are inventories containing items similar to those found on traditional personality scales (traditionally these instruments were in paper-and-pencil format, but more recently computerized instruments have also been created (see e.g. Jones, Brasher, & Huff, 2002)), (ii) they were specifically developed to assess personality constructs of relevance for work environments, (iii) they were designed for use with job applicants (this is reflected in their normative data), and perhaps most importantly (iv) they were designed to predict work behaviours.

These measures are further classified into: (i) Criterion-Focused Occupational Personality Scales (COPS), which include integrity tests, violence scales, and drug and alcohol scales, among others, and (ii) job-focused occupational personality scales (JOPS), which include sales potential scales and managerial potential scales, among others (Ones & Viswesvaran, 2001b). The former (i.e. COPS) have been constructed for the purpose of predicting particular criteria of interest. That is, integrity tests aim to predict dishonest behaviours at work; violence scales aim to predict violent behaviours at work; drug and alcohol avoidance scales aim to predict substance abuse at work; stress tolerance scales aim to predict handling work pressures well; customer service scales aim to predict serving customers well. The latter (i.e. JOPS) are geared to be predictive for particular occupational categories (sales people, managers etc).
Integrity tests, which were specifically developed to assess dependability, integrity, and honesty of applicants to help predict theft and future on-the-job deviant behaviours (Ones, 2002; Ones & Viswesvaran, in press-a; Sackett, 2002), are considered to be the prototypical criterion-focused occupational personality scales (Ones & Viswesvaran, 2001a, 2001b). Our primary purpose in this study is to present a meta-analysis of integrity test validities for predicting absenteeism.

**Integrity tests**

Ones and Viswesvaran (1998b) indicate that first integrity tests have existed since the late 1940s and that, in the US, there are over 40 off-the-shelf integrity tests available to organizations. Ones and Viswesvaran (1998b) write ‘Even by most conservative estimates, millions of people in the US have been tested using integrity tests’. As such, integrity tests are perhaps the most researched occupational scales in the literature. Examples of integrity tests include the London House Personnel Selection Inventory, Stanton Survey, Reid Report, PDI-Employment Inventory, Personnel Reaction Blank, and Hogan Personality Inventory–Reliability Scale (Ones et al., 1993; Ones & Viswesvaran, 1998a).

Ones et al. (1993), in a large-scale meta-analysis, found paper and pencil tests of ‘integrity’ to be valid predictors across jobs and organizations for overall job performance and also for composites of counterproductive behaviours on the job, such as tardiness, property damage, rule-breaking, and violence. In Ones et al. (1993), absenteeism was only one of several of the measures that contributed to the composite criterion of broad counterproductive behaviours on the job. In their meta-analysis, Ones et al. (1993) did not investigate the validity of integrity tests for predicting absenteeism per se.

Given the increased applications of pre-employment integrity testing for personnel selection, the question of whether these tests have predictive validity for the costly problem of absenteeism is certainly timely and important. If integrity tests can be shown to predict absenteeism, this may represent another benefit from use of integrity tests in selection (in addition to their validity in predicting employee theft, other counterproductive behaviours, and job performance).

There are several lines of evidence that hint at the potential validity of integrity tests for predicting absenteeism. The individual differences perspective on organizationally counterproductive behaviours posits that individuals who are irresponsible, untrustworthy, and dishonest will make overall poorer employees and will engage in more organizationally undesirable behaviours (Ones et al., 1993). Indeed, the study by Ones et al. (1993) established a relationship between the individual differences variable of integrity and a broad conceptualization of workplace counterproductivity. Given that integrity tests have been found to be predictive of counterproductive behaviours in general, one can postulate an integrity–absenteeism link.

Recent construct validity work investigating what integrity tests measure has found that while some integrity tests focus on applicants’ attitudes toward theft, others attempt to measure poor impulse control, lack of conscientiousness, disregard of rules and regulations, and general organizational delinquency (Ones & Viswesvaran, 1998b). To systematically examine what personality constructs integrity tests tap into, Ones (1993) examined the correlations between integrity tests and the Big Five personality dimensions using both a large primary data set and meta-analytic cumulation. The highest three correlations were with conscientiousness, agreeableness, and emotional stability, in that rank order.
Ones and Viswesvaran (2001b) point out that the conglomeration of these three personality constructs corresponds to Digman’s (1997) factor alpha (i.e., the socialization higher order factor of personality—a higher order factor than the Big Five) and is particularly relevant in the prediction of behaviours at work. Indeed, factor alpha appears to be the construct responsible for the predictive validities for drug and alcohol scales, violence scales, stress tolerance scales, and even customer service scales (Ones & Viswesvaran, 2001a, 2001b). Therefore, it may be reasonable to expect that factor alpha, as measured by integrity tests, would also be predictive of absenteeism. Integrity tests may be particularly well suited to the task of predicting absenteeism: they incorporate conscientiousness elements of work ethic, dutifulness, rule following, impulse control, dependability, and reliability; agreeableness components of trust and non-hostility; and the emotional stability element of stress tolerance.

A focus on the criterion also suggests a link between integrity test scores and absenteeism. Following the conceptualization by Hogan and Hogan (1989), that organizational delinquency is a broad phenomenon that includes many types of disruptive behaviours on the job, we posit that unwarranted absenteeism is one of those behaviours. There is evidence in the article by Hogan and Hogan (1989) suggesting that voluntary absences are a part of the general phenomenon of counterproductivity. In support of this, a number of international studies have linked problem alcoholism and drug use (see e.g. Dash, 2000, in an Indian study; Upmark, Moeller, & Romelsjoe, 1999, in a Swedish study) and pathological gambling (see e.g. Martinez-Pina, Guirao de Parga, Fuste i Vallverdu, Serrat Planas et al., 1991, in a Spanish study) to absenteeism. Ones and Viswesvaran (in press-a) estimated that absenteeism correlated, on average, 0.44 (corrected $r = 0.58$) with other counterproductive behaviours, including aggression, alcohol use, antagonistic work behaviours, destruction of property, drug misuse, misuse of information, misuse of time and resources, substance abuse, theft, and unsafe behaviour. Different forms of counterproductive behaviours can be conceptualized as negative aspects of job performance (Kelloway, Loughlin, Barling, & Nault, 2002; Miles, Borman, Spector, & Fox, 2002; Viswesvaran & Ones, 2000). Viswesvaran (2002) found sizable true score correlations between absenteeism and (i) organizational records of productivity ($\rho = -0.21$), (ii) organizational records of quality ($\rho = -0.48$), (iii) supervisory ratings of effort ($\rho = -0.54$), and (iv) supervisory ratings of interpersonal behaviour ($\rho = -0.33$). These results suggest that absenteeism may be regarded as an opposite manifestation of work effort expanded by individuals. Perseverance of individuals at job tasks in the face of adversity is an important positive work behaviour. Together these findings suggest that a potential reason for the positive associations among different forms of work behaviours, including counterproductivity, is the presence of common individual difference antecedents such as integrity (Ones & Viswesvaran, in press-a).

This study is the first attempt of its kind to integrate the literature on how integrity measures predict and explain absenteeism. Our hypothesis is that excessive voluntary absenteeism may be one aspect of the overall phenomenon of counterproductivity on the job. Because voluntary absenteeism may be one facet of irresponsible behaviour at work, we hypothesize the following.

**Hypothesis 1.** Integrity tests will be valid predictors of absenteeism.

Sackett and colleagues (Sackett, Burris, & Callahan, 1989; Sackett & Wanek, 1996) classified integrity tests into two categories: ‘overt integrity tests’ and ‘personality-based tests’. Overt integrity tests are designed to directly assess attitudes regarding dishonest
behaviours. Though intended to predict dishonest behaviours (particularly theft), these tests are valid predictors of a broad spectrum of counterproductive behaviours (Ones et al., 1993; Ones & Viswesvaran, 1998a; Schmidt, Viswesvaran, & Ones, 1997). Overt integrity tests include the London House Personnel Selection Inventory (PSI), the Employee Attitude Inventory (EAI), the Stanton Survey, the Reid Report, the Phase II Profile, the Milby Profile, the Trustworthiness Attitude Survey, and the Pre-employment Analysis Questionnaire. Ones (1993) found that there are sizable correlations among overt integrity tests. Personality-based measures, on the other hand, are intended to predict a broad range of counterproductive behaviours at work (e.g. violence on the job, absenteeism, tardiness, and drug abuse, in addition to theft) using personality scale items. Unlike overt tests, the initial intent in the construction of these measures has not been the prediction of theft or theft-related behaviours. Examples of personality-based measures that have been used in integrity testing include the Personal Outlook Inventory, the Personnel Reaction Blank, the Employment Inventory (Personnel Decisions, Inc.), and the Hogan Personality Inventory Reliability Scale. Compared to overt integrity tests, designed specifically to predict theft and dishonest behaviours, personality-based integrity tests were specifically developed to predict a variety of counterproductive behaviours and are therefore expected to be more valid predictors of absenteeism.

Thus, our second hypothesis is the following.

**Hypothesis 2.** Personality-based integrity tests will predict absenteeism better than overt integrity tests.

In personnel selection, an important question is whether concurrent validities can be used to estimate predictive validities. In the ability and aptitude domain, concurrent validities have been found to accurately estimate predictive validities (Bemis, 1968; Society for Industrial and Organizational Psychology, 1987). For personality measures in predicting job performance, Hough (1998) reported that concurrent validities were slightly higher than predictive validities. The meta-analysis by Ones et al. (1993) of the validity of integrity tests for predicting composites of counterproductive behaviours indicated that in the noncognitive domain of personality traits concurrent validity may overestimate predictive validities. However, results were varied depending on the criterion studied. For example, the validity of integrity tests for predicting supervisory ratings of overall job performance was found to be highest in predictive studies conducted using job applicants. Hence, in the present investigation, we sought to examine the impact of validation strategy (predictive versus concurrent) on criterion-related validities of integrity tests for predicting absenteeism.

**Research question.** For both overt and personality-based tests, do concurrent validities overestimate predictive validities?

**METHOD**

The hypotheses in this paper were tested using the interactive Hunter–Schmidt (1990a, p. 185) meta-analytic procedure. Meta-analysis is a statistical technique that seeks, among other goals, to determine (i) the mean correlation between two constructs unaffected by statistical artifacts and (ii) the extent to which the observed variance of findings across studies results from statistical artifacts. The interactive procedure uses artifact distributions.
to correct biases in observed validities caused by statistical artifacts. The artifacts operating across studies include sampling error, unreliability in the predictor and the criterion, range restriction, dichotomization of variables, and so on. If the validity is substantially dependent on the situation, statistical artifacts will not account for all or nearly all of the observed variation in the validities. In addition to estimating the portion of the observed variance that is due to statistical artifacts, meta-analysis also provides the most accurate obtainable estimate of the mean true (operational) validity. True or operational validity refers to the mean validity coefficient across studies corrected for unreliability in the criterion and range restriction, but not for unreliability in the predictor. As such, operational validity is the validity of a predictor expected to hold in job applicant samples.

If all or a major portion of the observed variance in validities is due to statistical artifacts, one can conclude that the validities are constant or nearly so. If the 90% credibility value is greater than zero, indicating that 90% of the estimates of true validity lie above that value, one can conclude that the presence of validity can be generalized to new situations (Hunter & Schmidt, 1990a). The lower credibility value is dependent on variance remaining after correction for statistical artifacts. In a meta-analysis, if the 90% credibility value is greater than zero, but there is variance in the validities after corrections, it can be concluded that integrity test validities are positive across situations, although the actual magnitude may vary somewhat across settings. However, the remaining variability may also be due to uncorrected statistical artifacts, other methodological differences, and unidentified moderators.

The studies in the present meta-analyses were obtained from published reviews of the literature (McDaniel & Jones, 1986, 1988; O’Bannon, Goldinger, & Appleby, 1989; Sackett et al., 1989; Sackett & Harris, 1984) and from the test publishers/authors in the form of technical reports, personal communications, and raw data. A total of 28 studies were identified as relevant to this meta-analysis. These studies constituted a subset of the studies contained in the Ones et al. (1993) database. While Ones et al. (1993) did not investigate the validity of integrity tests for predicting absenteeism, data from the same studies were coded to investigate validities for predicting broad counterproductive behaviours. The first two authors of this study coded and created the database for meta-analysis. The intercoder agreement across all study results and characteristics coded was 93%. For this study, the degree of agreement for sample sizes and validity coefficients (input to the meta-analyses) was 100%. Disagreements in coding were easily resolved by checking the original articles coded and through discussions.

The validities coded for this particular meta-analysis represented nine different integrity measures. Of the 28 validities contributing to the analyses, seven were for the Accutrac integrity test, four were for the PDI-Employment Inventory, three used the Employee Reliability Inventory, and seven validities were reported for the Inwald Personality Inventory. There were two validity coefficients each on the Hogan Personality Inventory’s Reliability scale and the Honesty Scale of the Personnel Selection Inventory, whereas there was one validity coefficient reported for Rely and Hogan Personality Profile scales.

All 28 validity coefficients used non-self report measures of absenteeism; 24 used organizational records and four employed supervisory ratings. The time frame for absenteeism measurement varied from 90 to 1297 days with an average of 322 days. Further, all 28 validities were based on employee samples and had used the number of times absent as the specific absenteeism measure (i.e. a measure of voluntary absenteeism). Most of the validities (26 out of 28) were from the service industry. More specifically, there were eight validities on samples of hotel/restaurant employees, seven...
validities were based on samples of correction officers and security personnel, one validity coefficient each using insurance, lumber, home improvement, hospital employee samples, two validity coefficients using grocery store employees, and three validity coefficients on department store employees (the sample used for one validity coefficient was heterogeneous in terms of occupational composition). It may be important to note that we computed 14 of the 28 validity coefficients based on proprietary data provided by integrity test publishers. A summary of study characteristics is reported in Table 1.

In cases where dichotomized correlations were reported, they were corrected for dichotomization (Hunter & Schmidt, 1990b), and the corrected correlations were used in the meta-analysis. Some studies might have compared the pass/fail dichotomy on integrity test scores or used a dichotomized criterion of absenteeism (e.g. more than XX absences versus fewer than XX absences). Sample sizes for these corrected correlations were adjusted to avoid underestimating the sampling error variance. First, the uncorrected correlation and the study sample size were used to estimate the sampling error variance for the observed correlation. This value was then multiplied by the square of the dichotomization correction factor (the ratio of the corrected to uncorrected correlation), yielding the sampling error variance associated with the dichotomization-corrected correlation (Hunter & Schmidt, 1990b). This value was then used with the uncorrected correlation in the standard sampling error formula to solve for the adjusted sample size

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of validity coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry sector</td>
<td></td>
</tr>
<tr>
<td>Service</td>
<td>26</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>2</td>
</tr>
<tr>
<td>Specific industry</td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>1</td>
</tr>
<tr>
<td>Hotel/restaurant</td>
<td>7</td>
</tr>
<tr>
<td>Lumber</td>
<td>1</td>
</tr>
<tr>
<td>Department stores</td>
<td>3</td>
</tr>
<tr>
<td>Security/law enforcement</td>
<td>7</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>2</td>
</tr>
<tr>
<td>Hospital</td>
<td>1</td>
</tr>
<tr>
<td>Grocery stores</td>
<td>2</td>
</tr>
<tr>
<td>Home improvement</td>
<td>1</td>
</tr>
<tr>
<td>Absenteeism measures</td>
<td></td>
</tr>
<tr>
<td>Organizational records</td>
<td>24</td>
</tr>
<tr>
<td>Supervisory ratings</td>
<td>4</td>
</tr>
<tr>
<td>Validation strategy</td>
<td></td>
</tr>
<tr>
<td>Concurrent</td>
<td>22</td>
</tr>
<tr>
<td>Predictive</td>
<td>6</td>
</tr>
<tr>
<td>Integrity tests used</td>
<td></td>
</tr>
<tr>
<td>Accutrac</td>
<td>7</td>
</tr>
<tr>
<td>Hogan Profile</td>
<td>1</td>
</tr>
<tr>
<td>PDI-Employment Inventory</td>
<td>4</td>
</tr>
<tr>
<td>Personnel Reaction Blank</td>
<td>1</td>
</tr>
<tr>
<td>Employee Reliability Inventory</td>
<td>3</td>
</tr>
<tr>
<td>Inwald Personality Inventory</td>
<td>7</td>
</tr>
<tr>
<td>Hogan Reliability Scale</td>
<td>2</td>
</tr>
<tr>
<td>PSI-Honesty Scale</td>
<td>2</td>
</tr>
<tr>
<td>Rely</td>
<td>1</td>
</tr>
</tbody>
</table>
used in the meta-analyses. Entry of this sample size into the meta-analysis calculations results in the correct estimate of the sampling error variance of the corrected correlation in the meta-analysis.

The meta-analyses corrected the mean observed validity using the means of artifact distributions for criterion unreliability and for range restriction (Hunter & Schmidt, 1990a, p. 185). No correction for predictor unreliability (i.e. unreliability in integrity test scores) was applied to the mean validity, because our interest was in determining the operational validity of integrity tests for predicting absenteeism. However, the observed variance of validities was corrected for variation in predictor unreliabilities using the integrity test reliability distributions from Ones et al. (1993).

We used coefficient alphas or the equivalent in correcting the variation in validities for variability due to differences in predictor unreliabilities across studies. Test–retest reliability estimates over relatively short time periods also provided reasonably close approximations to alpha coefficients. This practice is consistent with previous meta-analyses on integrity tests (Ones et al., 1993). We could not use coefficients of equivalence and stability as these were not reported for integrity tests. We did not use generalizability coefficients in attenuation corrections as our interest was on measures (i.e. integrity tests) and not constructs used for personnel selection. The overall mean of the predictor reliability artifact distribution was 0.81 and the standard deviation was 0.11. The mean of the square roots of predictor reliabilities was 0.90 with a standard deviation of 0.06. Two other predictor reliability distributions were constructed: one for overt integrity tests and another for personality-based integrity tests. The mean of the overt test reliability artifact distribution was 0.83 and the standard deviation was 0.09. The mean of the square roots of overt test reliabilities was 0.91 with a standard deviation of 0.05. The mean of the personality-based test reliability artifact distribution was 0.72 and the standard deviation was 0.13. The mean of the square roots of the reliabilities was 0.85 with a standard deviation of 0.08. Each one of these predictor reliability distributions was used in analyses with corresponding predictor categories. That is, when validities of overt tests were being cumulated the predictor reliability distribution for overt tests was used, but when validities of personality-based tests were being meta-analysed the predictor reliability distribution for personality-based tests was used. Finally, when the analyses involved both overt and personality-based tests, the overall predictor reliability distribution was used.

The mean reliability values used in the corrections for unreliability of the absenteeism measures were obtained from the general absenteeism literature. There were 79 investigations found reporting test–retest reliability coefficients for absenteeism and also reporting the time period for which absenteeism records were kept. Using the Spearman–Brown formula, each of these 79 reliabilities was adjusted to a one-month period. The average reliability of absenteeism for one month was found to be 0.1680 (SD = 0.2002). Based on the Spearman–Brown formula, this mean reliability was adjusted to the time periods used in our integrity test studies; this was possible only for 20 studies out of 28; eight studies did not provide information on the time period over which absenteeism was measured. This criterion reliability distribution of 20 reliabilities had a mean of 0.70 and a standard deviation of 0.13. The meta-analysis used the square root of the absenteeism reliabilities, which had a mean of 0.83 (SD = 0.08).

Because integrity tests are often used to select job applicants, the validity calculated using an employee sample may be affected by restriction in range. Further, additional direct and indirect processes (e.g. attraction, selection, attrition) may result in reduced variability in predictor scores among employees. In the present analysis, range restriction
corrections were applied to the mean validities to correct for direct range restriction on the predictor variable. A distribution of range restriction values was constructed from the studies contributing to the database of Ones et al. (1993). The range restriction ratio was calculated as the ratio of study to reference group standard deviation \( (s / S) \). In most cases, reference groups were job applicants. In a few cases where applicant data were unavailable, we relied on technical manual norms. The latter practice has been shown to yield similar results to those of applicant samples (Ones & Viswesvaran, in press-b). The mean ratio of the restricted sample’s standard deviation to the unrestricted sample’s standard deviation used was 0.81 and the standard deviation was 0.19. The range restriction artifact distributions for overt and personality-based integrity tests were not any more different for the two types of test than would be expected by sampling error. Therefore, we used the more robust combined distribution in our corrections. We should note that the range restriction values for integrity tests are quite similar to those found for other personality measures (e.g. Big Five; see Salgado, 2002). However, the degree of range restriction in this research domain is less than for cognitive ability (Alexander, Carson, Alliger, & Cronshaw, 1989). Thus, range restriction corrections were much smaller in the present research than in meta-analyses in the ability domain. A summary of statistical artifact distributions used in this meta-analysis can be found in Table 2.

### META-ANALYSES AND RESULTS

Results of the overall meta-analyses are presented in Table 3.

The first meta-analysis estimated the validity of all integrity tests, overt and personality based, for predicting absenteeism. The total sample size across 28 studies reporting such a correlation was 13,972. This meta-analysis indicated that only 22% of the variance observed in validities was due to statistical artifacts we corrected for (i.e. sampling error, unreliability in the measures, and range restriction). The best estimate of the operational validity of integrity tests with the criterion of absenteeism is 0.20. Relative to the mean, however, the standard deviation of the true validity was substantial, 0.19. The lower bound for the 90% credibility interval was 0.02. Therefore, the first hypothesis that all integrity tests will be valid predictors of absenteeism across organizations and jobs was not confirmed.

To test the second hypothesis, overt versus personality-based integrity tests were meta-analysed separately. Personality-based integrity tests were found to have a mean operational validity of 0.36 in predicting absenteeism. All of the observed variance was accounted for by statistical artifacts. SD\(_p\) for personality-based integrity tests was zero, and the lower credibility value of 0.36 indicated that validity of personality-based integrity

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>( K )</th>
<th>Mean ( s / S )</th>
<th>Mean ( s / S )</th>
<th>Mean ( s / S )</th>
<th>Mean ( s / S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrity test reliabilities</td>
<td>124</td>
<td>0.81 0.11</td>
<td>0.81 0.11</td>
<td>0.81 0.11</td>
<td>0.81 0.11</td>
</tr>
<tr>
<td>Overt integrity test reliabilities</td>
<td>97</td>
<td>0.83 0.09</td>
<td>0.83 0.09</td>
<td>0.83 0.09</td>
<td>0.83 0.09</td>
</tr>
<tr>
<td>Personality-based test reliabilities</td>
<td>27</td>
<td>0.72 0.13</td>
<td>0.72 0.13</td>
<td>0.72 0.13</td>
<td>0.72 0.13</td>
</tr>
<tr>
<td>Absenteeism reliabilities(a)</td>
<td>20</td>
<td>0.70 0.13</td>
<td>0.70 0.13</td>
<td>0.70 0.13</td>
<td>0.70 0.13</td>
</tr>
<tr>
<td>Range restriction ( s / S ) values</td>
<td>79</td>
<td>0.81 0.19</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

\(a\) Based on the average reliability of absenteeism for one month of 0.1680 (SD = 0.2002) across 79 absenteeism reliabilities. See the text for explanation.
<table>
<thead>
<tr>
<th>Analyses</th>
<th>Total N</th>
<th>K</th>
<th>$r_{\text{mean}}$</th>
<th>$SD_r$</th>
<th>$SD_{\text{res}}$</th>
<th>$\rho$</th>
<th>$SD_\rho$</th>
<th>% Var. S.E.</th>
<th>% Var acc. for</th>
<th>90% CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All integrity tests</td>
<td>13 972</td>
<td>28</td>
<td>0.14</td>
<td>0.1364</td>
<td>0.1208</td>
<td>0.20</td>
<td>0.19</td>
<td>15.9</td>
<td>21.5</td>
<td>-0.02</td>
</tr>
<tr>
<td>2. Personality-based tests</td>
<td>5435</td>
<td>18</td>
<td>0.25</td>
<td>0.0851</td>
<td>0.0000</td>
<td>0.36</td>
<td>0.00</td>
<td>56.3</td>
<td>100.0</td>
<td>0.36</td>
</tr>
<tr>
<td>3. Overt integrity tests</td>
<td>8537</td>
<td>10</td>
<td>0.06</td>
<td>0.1120</td>
<td>0.1022</td>
<td>0.09</td>
<td>0.16</td>
<td>15.2</td>
<td>16.8</td>
<td>-0.10</td>
</tr>
</tbody>
</table>

$K =$ number of validities, $r_{\text{mean}} =$ mean observed validity, $SD_r =$ observed standard deviation, $SD_{\text{res}} =$ residual standard deviation, $\rho =$ mean operational validity, $SD_\rho =$ true score standard deviation, %Var. S.E. = % variance due to sampling error, % Var. acc. for = % variance due to all corrected statistical artifacts (sampling error, predictor unreliability, criterion reliability, and range restriction), 90% CV = lower 90% credibility value.
tests is constant, across studies and situations for the criterion of absenteeism. (Although second order sampling error may have affected our variability estimates, the mean operational validity is relatively less affected by second order sampling error (Hunter & Schmidt, 1990a, p. 411)).

For overt integrity tests, the results across ten validities and 8537 employees showed that the best estimate of overt integrity tests’ validity in absenteeism was 0.09. The 90% credibility value of −0.10 indicated that the validity was not positive across overt tests, studies and situations. The percentage variance accounted for by corrected statistical artifacts was low at 17%, and the standard deviation of the true validity (SDρ) was 0.16, a fairly large value relative to the mean validity. These results confirmed the second hypothesis. That is, personality-based integrity tests have a higher mean operational validity in predicting absenteeism than overt integrity tests.

The next set of meta-analyses was conducted to study the effect of validation strategy (predictive versus concurrent) on both the overt and personality-based integrity test validities. Results for fully hierarchical moderator analyses are reported in Table 4. Predictive validities of personality-based integrity tests had a mean true validity of 0.33, with an SDρ of zero; while concurrent studies had a mean of 0.50, with an SDρ of 0.16. For personality-based tests hypothesis 3 was confirmed: concurrent validities appear to overestimate predictive validities in predicting absenteeism, using personality-based integrity tests. For overt integrity tests, there were nine predictive studies but only one concurrent study. Therefore a completely parallel analysis cannot be reported for overt tests. However, the single estimate of concurrent validity was considerably higher than mean corrected, predictive validity (0.44 versus 0.09), as hypothesized. The mean

Table 4. Hierarchical moderator analyses of integrity test validities for predicting lack of absenteeism

<table>
<thead>
<tr>
<th>Validation strategy</th>
<th>Overt integrity tests</th>
<th>Personality-based integrity tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N</td>
<td>8508</td>
<td>4922</td>
</tr>
<tr>
<td>K</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>r&lt;sub&gt;mean&lt;/sub&gt;</td>
<td>0.06</td>
<td>0.23</td>
</tr>
<tr>
<td>SD&lt;sub&gt;r&lt;/sub&gt;</td>
<td>0.1105</td>
<td>0.0538</td>
</tr>
<tr>
<td>SD&lt;sub&gt;res&lt;/sub&gt;</td>
<td>0.1014</td>
<td>0.0000</td>
</tr>
<tr>
<td>ρ</td>
<td>Predictive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>0.33</td>
</tr>
<tr>
<td>SDρ</td>
<td>0.16</td>
<td>0.00</td>
</tr>
<tr>
<td>% var. acc. for</td>
<td>15.7</td>
<td>100.0</td>
</tr>
<tr>
<td>90% CV</td>
<td>−0.10</td>
<td>0.33</td>
</tr>
<tr>
<td>Total N</td>
<td>29</td>
<td>513</td>
</tr>
<tr>
<td>K</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>r&lt;sub&gt;mean&lt;/sub&gt;</td>
<td>0.33</td>
<td>0.35</td>
</tr>
<tr>
<td>SD&lt;sub&gt;r&lt;/sub&gt;</td>
<td>—</td>
<td>0.1522</td>
</tr>
<tr>
<td>SD&lt;sub&gt;res&lt;/sub&gt;</td>
<td>—</td>
<td>0.0988</td>
</tr>
<tr>
<td>ρ</td>
<td>Concurrent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.44</td>
<td>0.50</td>
</tr>
<tr>
<td>SDρ</td>
<td>—</td>
<td>0.16</td>
</tr>
<tr>
<td>% var. acc. for</td>
<td>—</td>
<td>57.9</td>
</tr>
<tr>
<td>90% CV</td>
<td>—</td>
<td>0.32</td>
</tr>
</tbody>
</table>

K = number of validities, r<sub>mean</sub> = mean observed validity, SD<sub>r</sub> = observed standard deviation, SD<sub>res</sub> = residual standard deviation, ρ = mean operational validity, SDρ = true score standard deviation, % var. acc. for = % variance due to all corrected statistical artifacts (sampling error, predictor unreliability, criterion reliability, and range restriction), 90% CV = lower 90% credibility value.
predictive true validity of overt tests for predicting absenteeism was 0.09, a value much smaller than the personality-based test predictive operational validity of 0.33.

**DISCUSSION**

Our findings indicate that the validity of personality-based integrity tests for predicting absenteeism is sizable and generalizable. Personality-based integrity tests are specifically designed to predict counterproductive behaviours other than theft, while overt tests are designed specifically to predict theft. The results of this study indicate that, in the prediction of absenteeism using integrity tests, concurrent designs may lead to overestimates of predictive validity. Based on predictive validation studies, personality-based tests are estimated to have a mean true validity of 0.33, a value likely to translate into sizable cost savings from reduced absenteeism if such tests are used in selection. Similar statements cannot be made for overt integrity tests for the criterion of voluntary absenteeism. As previous research shows, they are valid predictors of job performance and a broad spectrum of counterproductive behaviours, but their validity for predicting absenteeism alone appears to be low (0.09), and may not be positive across all overt scales, settings and organizations. In interpreting this finding, it should be borne in mind that overt integrity tests were not intended by their developers to predict absenteeism.

**Comparisons with integrity test validities for other criteria**

It is interesting to compare the results of our study with those found by Ones et al. (1993) in their meta-analyses. Ones et al. (1993) found that integrity tests predict overall job performance better than theft, the criterion they were designed to predict. In our study, we found that overt tests predict absenteeism at a validity level of 0.09, much lower than their validity for overall job performance (0.41). The implication seems to be that overt tests focus on attitudes and traits which are good determinants of a broad criterion such as overall job performance. However, as the criterion width becomes narrower, such as theft or absenteeism, their predictive validity declines. The important principle here is that the broader the criterion, the better the validity of overt tests in prediction (see Ones & Viswesvaran, 1996, for a full discussion of bandwidth/fidelity dilemma in prediction).

Why was the same phenomenon not observed with personality-based tests? There are two potential reasons for this. First, Ones (1993) meta-analytically showed that overt and personality-based integrity tests share a common personality core (conscientiousness, agreeableness, and emotional stability), yet the unreliability corrected correlation between the two types of test is around 0.60, clearly less than unity. The differential results for overt versus personality-based tests in predicting absenteeism might suggest better capturing of volitional aspects of absence behaviour by personality-based tests. Second, differences in findings could be due to differences in item selection strategies and other test development features. Personality-based tests often included absenteeism as a criterion of interest in test development, while overt tests did not. Future research attention needs to be directed to this issue.

**Comparisons with other personality scales**

As already noted, Salgado (2002) conducted a meta-analysis of the Big Five personality scales and alternative criteria, including absenteeism. The number of validity coefficients for absenteeism ranged from eight to 12 ($n = 1339–2491$), depending on the Big Five
dimension examined. In his meta-analysis, Salgado (2002) found that none of the Big Five dimensions predicted absenteeism substantially (operational $r$ values ranged from $-0.06$ to $0.08$). The results of our research are much more encouraging for personality-based integrity tests (operational $r = 0.33$). Yet, the results of the two meta-analyses may not be fully comparable. In this paper, we focused on the volitional aspects of absenteeism. That is, the criterion of interest was voluntary absences. Personality traits may better predict motivationally determined criteria. Salgado’s (2002) database did not allow him to differentiate between voluntary and involuntary absences in his meta-analysis.

**Methodological, substantive, and practical contributions**

Methodologically, our research has three strengths. First, we incorporated the recent advancements in psychometric meta-analyses in cumulating across studies (e.g. use of mean observed correlation to correctly estimate sampling error variance, nonlinear range restriction). Second, our approach to estimating the reliability of absenteeism is a novel approach, that cumulates across studies after controlling for temporal variations. Our reliability estimates are robust and take into account the many problems discussed in this literature. Finally, our study was able to avoid self-reported criteria in the assessment of absenteeism.

Substantively our research contributes to the theoretical knowledge of absenteeism in three main ways. First, we attempted to answer the question of why absenteeism occurs. We investigated one potential personality variable that explains absence behaviour. Second, by relating absence-taking behaviour to integrity we attempted to build a bridge between the extant absenteeism literature and the vast body of knowledge on personality traits. Finally, our research laid the foundation for future research to examine the phenomenological nature of absence taking behaviour, and focuses attention on the various meanings of absence (volitional versus non-volitional).

From a practical point of view, the literature on absenteeism has tended to focus on variables that can be manipulated by organizational interventions (see e.g. Dalton & Mesch, 1991). Such interventions have focused, among other things, on alternative working schedules (Baltes, Briggs, Huff, Wright, & Neuman, 1999), financial incentives (see e.g. Schlotzhauer & Rosse, 1985), and self-management (Frayne & Latham, 1987). Our understanding of absenteeism may be enhanced by comparing the effectiveness of other organizational interventions to reduce absenteeism with the effectiveness of personnel selection using personality-based integrity tests. Most approaches to understanding and preventing absenteeism conceptualize it as a withdrawal behaviour that occurs as a result of unfavourable work attitudes such as job dissatisfaction (Johns, 2001). For example, Hackett (1989) reported a meta-analysis of the satisfaction and absenteeism relationship. Corrected correlations with absence frequency were $-0.09$ for both satisfaction with pay and promotion, and $-0.07$ for satisfaction with work itself. In two meta-analyses (Farrell & Stamm, 1988; Spector, 1986), job control and absenteeism correlated around $-0.20$. Given these modest relationships, organizational interventions such as improving pay, work redesign, and increasing job control may not have the strongest influences in reducing absenteeism. The use of personality-based integrity tests in personnel staffing decisions appears to be somewhat more promising.

**Directions for future research**

Although absenteeism has had a long history of research, integrative theoretical perspectives have been lacking. As our research suggests, some individuals are more likely to be more
frequently absent than others. Personality-based integrity tests assess relevant traits for predicting voluntary absenteeism. To date, no extensive research has been reported on the role of personality variables, especially integrity, into path analytic causal models of absenteeism. This is an important venue for future research. How do personality traits come to predict absenteeism? Confirming a mediational hypothesis, George (1989) reported that positive affectivity was related to positive mood at work, which correlated negatively with absenteeism. Similarly, Iverson, Olekalns, and Erwin (1998) found that positive affectivity related to feelings of personal accomplishment, which in turn was correlated with reduced absenteeism. Future research is needed to investigate several potential causal process mechanisms explanatory of predictive validity of integrity tests for absenteeism. In predicting voluntary absences using personality-based integrity tests, the main causal paths might be through volition and motivation. Further, it is possible that individuals high on integrity are more planful and deliberate, thereby taking into account the many contingencies of life, and hence reducing absenteeism.

Beyond simple bivariate relationships between personality traits and forms of absenteeism, future research should also examine moderator hypotheses. In this vein, Johns (2001) suggested ‘Some people may be more inclined than others to withdraw in response to job dissatisfaction’. The personality traits of relevance for this moderator hypothesis may be neuroticism, self-esteem, and hostility. Indeed, a Finnish study ($n = 757$) found that lower levels of job control were more strongly linked to higher absences for hostile employees compared with non-hostile employees (Vahtera, Kivimäki, Utela, & Pentti, 2000).

In our work, we did not have the data to study involuntary absences (e.g. sick leave). Even if we had, we would not expect integrity tests to predict involuntary absences. However, another personality construct, neuroticism, might be a distal determinant of involuntary absenteeism. Neurotic individuals tend to be stress prone, anxious, and depressed. This might lead to stress reactions, including physical and/or psychological illness, and hence to absenteeism. Large scale ($n = 79,070$) empirical findings linking perceived stress and illness-related absenteeism would seem to offer initial support for this hypothesis (Jacobson, Aldana, Goetzel, Vardell et al., 1996). Yet, it is interesting that, even for preventing sickness absence, variables such as role overload, role conflict, and role enhancement offer little explanatory power (Mastekaasa, 2000, in a large scale Norwegian study, $n = 94,869$; 10% of all employed adults in that country). On the other hand, neuroticism and related psychopathology (e.g. major depressive disorder) were found to relate to employee sickness absence (Kivimäki, Vahtera, Thomson, Griffiths, Cox, & Pentti, 1997; Laitinen-Krispijn & Bijl, 2000).

The important point here is that voluntary and involuntary absences probably have different personological etiologies and different causal models might be needed to explain the role of personality traits in each. Similar arguments can be offered for voluntary and involuntary lateness (see e.g. Bardsley & Rhodes, 1996).

Finally, as Johns (2001) notes, ‘...absenteeism is “the failure to report for scheduled work.” Marotcchio and Harrison (1993: p. 263) define it as “an individual’s lack of physical presence at a given location and time when there is a social expectation for him or her to be there.”’ Yet, theoretical insights into personality traits associated with absenteeism can also be gained by studying attendance, and even presenteeism, the tendency to attend work despite the presence of legitimate reasons for not attending (e.g. being sick, but yet working) (Aronson, Gustafsson, & Dallner, 2000). Such research would also be valuable in disentangling differences among absenteeism, attendance, and presenteeism.
Absenteeism is costly and disruptive. Our research suggests that an individual differences framework can be used to understand absenteeism and that a personnel selection strategy can be used to reduce absenteeism rates. Previously, research has focused primarily on increasing employee satisfaction and ability to attend (e.g. by providing transportation) as the main options available to management to combat absenteeism. This study investigated a new alternative strategy for reducing absenteeism, specifically, the strategy of screening job applicants using pre-employment integrity tests. The personnel selection approach we recommend for reduced absenteeism aims to deal with the problem before employees are hired and is therefore proactive. Naturally, good human resource management practices are extremely important in keeping absenteeism rates down after employees who are not prone to be absent have been hired.

ACKNOWLEDGEMENTS

We thank numerous test publishers, authors, and colleagues who generously sent us proprietary data on integrity tests; their trust is much appreciated.

REFERENCES


