ASSESSING INTEGRITY AT WORK: DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF THE OCCUPATIONAL INTEGRITY SCALE

MARTIN SEITL, KONSTANTINOS KAFETSIOS, KRISTÝNA SCHNEIDEROVÁ, DANIEL DOSTÁL, TOMÁŠ KRAM, TOMÁŠ DOMINIK

Department of Psychology, Faculty of Arts, Palacký University Olomouc, Krčíkova 10, 771 80 Olomouc, Czech Republic. Email: martin.seitol@upol.cz, konstantinos.kafetsios@upol.cz, kristyna.schneiderova01@upol.cz, daniel.dostal@upol.cz, tomas.kram@gmail.com, tomasdominik01@gmail.com

This work was supported by Czech Ministry of Education, Youth, and Sports for specific research, granted in 2021 to Palacký University Olomouc (RGA_FF_2021_008 SPP 31-032102021)

Abstract: Interest in the diagnostics of integrity by work and organizational psychology researchers and practitioners alike has been steadily increasing. In two studies we report results on the development of the Occupational Integrity Scale (OIS), a novel approach to integrity diagnostics, and we test the psychometric properties of the OIS in a sample of 870 adult public education employees. The second study replicated the factor structure of the OIS in a sample of 147 company employees and tested the validity against self-reported Big Five and supervisor-assessor-rated employee CWB two months later. A three-factor solution was identified. The OIS demonstrated consistent relationships with personality criteria and limited associations to supervisor assessment of CWB.

Keywords: counterproductive work behaviour, personality, integrity at work, honesty, public organizations, private organizations.

1 Introduction

Integrity at work has become one of the most distinctive evidence-based research areas for understanding the relationships between personality and work outcomes (Cortina & Luchman, 2013). Integrity at work, i.e., the occurrence of honesty, trustworthiness, truthfulness, fairness, and commitment to morality (Schlenker, 2008), is considered the third most important personality predictor of overall job performance after conscientiousness and general mental ability (Schmidt et al., 2016). Moreover, integrity at work is one of the most robust predictors of counterproductive work behavior (CWB; Ones et al., 1993). Specifically, lack of morale, honesty, conscientiousness, and other positive traits associated with integrity at work are more likely to result in harmful behavior toward others or the organization itself (Cohen et al., 2014; van Iddekinge, 2012). An experience of harmful behavior was experienced by 98% of employees, and even half of them were targets of bad behavior once a week (Porath & Pearson, 2013). Therefore, access to appropriate integrity diagnostics remains a key issue.

The choice of appropriate diagnostics approach is exacerbated by the relative lack of non-commercially validated methods for integrity diagnostics, which limits opportunities for further research. This is particularly true for the Central European region, where there is a lack of integrity testing and little information on testing from public or private organizations. Therefore, the aim of this article is to a) present the development and testing of the Occupational Integrity Scale, a novel, open-access integrity test, b) determine the construct and criterion predictive validity of the new scale using prospectively determined CWB and personality criteria. In doing so, we discuss integrity tests for different types of counterproductive work behaviour in public and private organizations and provide a link to related research and practice.

1.1 Defining of Counterproductive Work Behavior and Integrity

Counterproductive work behavior (CWB) is characterized by some degree of dishonesty and deliberateness. It is behavior that can harm the organization and/or its members (Camara & Schneider, 1994), including the circle of customers, suppliers, and others associated with the organization. Given the nature of CWB, there is no doubt that it is in the interest of everyone associated with the organization to prevent CWB. However, although a general CWB factor is described (Berry et al., 2012), it is more of a hierarchical structure in which a common factor encompasses substantially different behavioral dimensions (Schmitt & Highhouse, 2013). The behavioral dimensions of CWB differ in their severity (e.g., theft vs. misuse of information) and in terms of the target harmed by the behavior (e.g., verbal aggression harms the individual vs. poor work quality harms the organization). In predicting dishonesty, not only should a common CWB factor be considered, but also separate factors representing different counterproductive behaviors in the workplace.

Robinson and Bennett (1995) were among the first to propose, using multidimensional scaling techniques, a commonly cited subdivision of CWB: the severity and goal of harm to organizations or individuals. Later, Grusy (1999) introduced another taxonomy consisting of 11 different CWB types. Sackett (2002) combined and tested (Grusy & Sackett, 2003) the results of the precursors into a new taxonomy that is probably one of the most influential at present (Schmitt & Highhouse, 2013). This latter taxonomy was applied in the present study.

Typically, integrity at work is assessed using personality-based and overt integrity tests (Ones & Viswesvaran, 2001), which have emerged particularly from the empirical tradition. Personality-based tests (e.g., the Personnel Reaction Blank, Gough, 1971) measure personality traits that correlate with a wide range of counterproductive behaviors (Sackett et al., 1989).

Conversely, overt integrity tests aim to capture applicants’ or employees’ attitudes toward various forms of dishonest behavior by directly targeting items related to experiences or attitudes about theft, absenteeism, and the like; examples of such measures include the London House Personnel Selection Inventory (Boye & Wasserman, 1996). A somewhat different perspective, encompassing broader theoretical sources, is offered by tests drawn from the philosophical tradition (e.g., Moral Character Questionnaire, Furr et al., 2022 or Integrity Scale, Schlenker, 2008). These tests have also attracted research attention (Ampuni et al., 2019; Krettenauer et al., 2021; Okan & Eksi, 2020) and demonstrated their utility in predicting CWB.

Although an association between integrity and CWB has been demonstrated (van Iddekinge et al., 2012; Ones & Viswesvaran, 2001; Stehlik, 2015), the explained variance occasionally appears to be small (Coyne & Bartram, 2002; Van Hein et al., 2007). Therefore, access to appropriate integrity diagnostics remains a key issue.
Although CWB is relatively well defined, integrity is a complex concept and lacks a consistent definition (Wanek et al., 2003). As argued here, this inconsistency is partly the result of different traditions. From a philosophical perspective, integrity is a complex organization of personality (Dunn, 2009), whereas from the empirical perspective of overt and personality-based tests, the characteristics that most closely resemble the label of integrity are those of honesty (Ash, 1991; Frost & Rafilson, 1989). However, the next distinction between overt and personality-based integrity tests may also arise from their validity against behavioral criteria. Marcus et al. (2016) established the predictive validity of integrity tests using indicators of behavioral integrity beyond standard personality measures, with the definition of behavioral integrity being "proactive adherence to general rules of ethical and interpersonally fair behavior" (p. 64). Behavioral integrity, as the criterion for tests evolved from empirical tradition, is then close to the philosophical perspective. In summary, despite the differences noted between the various research traditions and operationalizations of integrity, there is also considerable overlap that informed the delineation of integrity in this study.

In the present study, integrity was considered as a commitment to moral principles that leads individual behavior to be aligned with a set of moral values derived from cultural and social norms that become evident over time and in different social contexts (Dunn, 2009; Schlenker, 2008). These include trustworthiness, honesty, and respect for promises and one's principles, as well as the fulfillment of tasks and roles for which individuals have accepted responsibility. A key component of this approach is the moral component of integrity advocated by Dunn (2009).

1.2 Empirical evidences for types of integrity tests

Tests with a clear objective and with direct questions about attitudes toward and experience with honest behavior, labelled as overt integrity tests, were developed after World War II (Ash, 1991). The original primary purpose of developing an integrity test was to help predict workplace theft, so items were often worded along these lines (Sackett et al., 1989). Accordingly, Frost and Rafilson (1989) found that theft had a stronger relationship with overt integrity scores than did the general CWB factor. However, the predictive power of overt integrity tests is not limited to theft, and it would be a mistake to focus only on the predictive power of this one CWB factor (Frost & Rafilson, 1989; Ones & Viswesvaran, 2001). The current goal of open-ended tests, based on criterion validity summarized in the meta-analyses, is to capture attitudes toward various forms of dishonest behavior through proper item design (Ones et al., 1993; Ones & Viswesvaran, 2001).

Very similar to the format of the open integrity tests, but coming from a different tradition of integrity research, are relatively new tests (Furr et al., 2022; Schlenker, 2008) based on the philosophical tradition. Schlenker (2008, p. 1086) found that high scores on the Integrity scale also reflect a stronger assertion of being committed to ethical principles. The items target more general moral principles from which respondents can choose what is right or wrong for them. However, a portion of the Integrity Scale items also focus on attitudes toward behaviors of the disruptive kind, such as lying or cheating, as these issues are inherent to integrity (Schlenker, 2008). In this sense, the Integrity Scale is close to the approach of open-ended integrity tests (empirical tradition with items directly targeting dishonesty) and also to the concept of morality in character and moral maturity (Cohen et al., 2014). The approach advocated by the Integrity Scale seems to offer a broader picture than the overt tests in the empirical tradition, as it also focuses on the philosophical foundations of integrity and on applied attitudes.

The origins of personality-based integrity testing can be traced to the early 20th century in the form of covert purpose tests that could detect dishonesty (Ash, 1991). After World War II, dishonesty tests evolved into measures of personality traits that were largely related to dishonest work behavior. However, there are mixed results supporting the applicability of personality-based integrity tests. While Cullen and Sackett (2003) suggest that personality-based integrity tests can be expected to predict specific subordinate factors of counterproductive work behavior such as turnover or absenteeism, a meta-analysis by Ones et al. (1993) provides evidence that the predictive validity of personality-based integrity diagnostics for the general CWB factor. Another meta-analysis (van Iddekinge et al., 2012) presented validity coefficients corrected for non-reliability ρ = .27 for all types of designs with personality-based integrity tests and concluded that differences in validity results for personality-based tests depend on several moderating factors, including the type of participants, the source of the criterion, and the type of research design. Therefore, in the two studies reported here, we chose to control for the aforementioned variables and set the same conditions for both tests used.

The most commonly used personality traits in the context of CWB, which provide the theoretical background for the development of personality-based integrity tests, include: conformity, sociability, impulse control, reliability, and credibility (Ones et al., 1993; Wanek et al., 2003). Yet, despite the identification of several such integrity indicators, the five-factor model (FFM) dominates the operationalization of personality-based integrity traits (Salgado & de Frutos, 2005; Schmidt et al., 2016), with a meta-analytic review of studies linking CWB to FFM dimensions (Hoffman & Dilchert, 2012) suggests that CWB is mainly associated with Conscientiousness (validity coefficients corrected for non-reliability [p] ranging from .26 to .35), Agreeableness (p ranging from .28 to .40), and Emotional Stability (p ranging from .20 to .35).

2 Present study

Several overt and personality-based integrity tests are available, both for commercial and research use (van Iddekinge et al., 2012; Ones & Viswesvaran, 2001). Widely known overt integrity tests include the London House Personnel Selection Inventory (Boye & Wasserman, 1996), the Reid Report (Weiss, 2010), and to some extent the Integrity Scale (Schlenker, 2008), as discussed above. Personality-based tests typically include tests such as the Personal Outlook Inventory (see Sackett et al., 1989) or the Employment Inventory (Pajanan, 1986, cited in Frost & Rafilson, 1989).

Most studies testing the validity of integrity tests to predict CWB have used commercial instruments (van Iddekinge et al., 2012), and fewer studies have used noncommercial integrity tests, especially open ones, making independent research more difficult. Non-commercial open-ended integrity tests include the Employee Attitude Index (EAI, Ahart & Sackett, 2004) or the Workplace Productivity Questionnaire (WPQ, Nicol & Paunonen, 2002). However, the latter list is short and provides insight into the scarcity of current open integrity tests. The availability of open integrity tests in different national languages is also limited. In particular, contemporary open and personality-based integrity tests for non-commercial use are still not available in the Czech and Slovak Republics (Seifl, 2015), while commercial instruments are generally used instead. It is noteworthy that much of this work has been conducted in private organizations.

Thus, the aim of the present study was to develop a new self-report scale for assessing integrity at work, the Occupational Integrity Scale, and to investigate its basic psychometric properties (factor structure, consistency, and validity) in samples of employees from public and private organizations. In addition, we aimed to assess the criterion predictive validity of the new scale against supervisor-rated CWB and construct validity against personality traits. The Occupational Integrity Scale was compiled in accordance with the above outlined constructs. It is an originally developed inventory with 35 items. In defining the items, descriptions and examples of items from the philosophical approach (e.g., Dunn, 2009; Schlenker, 2008) were used as a theoretical source. Subsequently, the designed items were discussed with three experts—a director of HR and two academic scale construction specialists. The first set of items was adapted.
and new attitude items were added to the scale in terms of open-ended integrity tests. The Integrity Scale also inspired us to take this step, however, given the need to use more job-specific than general attitudinal items, we focused more on applied attitudes toward dishonesty in the workplace. The final goal was to obtain two specific and one general latent factor for workplace integrity. The first specific factor, moral integrity, which refers to general moral issues and commitment to them, consisted of 21 items. The items focus on the respondent's principles and moral values, adherence to their word, credibility, reliability, reluctance to violate their own established principles, and belief in general rules. The remaining 14 items, representing the second factor, were selected to reflect the demonstration of honest behaviors at work (Schneiderová, 2017).

2.1 Study 1

The aim of the first study was to observe the basic psychometric characteristics, select from the proposed items the subset that maximizes factor reliability, and test the functionality of the new scale with a large sample of public sector employees in the Czech Republic.

Method

Sample and procedure. To test the psychometric properties of the OIS, teaching staff from elementary schools in the seven Czech regions were used as test participants. Educators were chosen mainly because of relatively good accessibility and high degree of homogeneity. A total of 992 participants completed the test administration. Following information about the aims, right to withdraw at any point, and anonymity, the first page of the survey stated that completion of the questionnaire constituted proof of participants’ informed consent to take part in the study. When completing the study, participants also had the right to withdraw, in which case, their data were immediately deleted. A total of 43 respondents selected to be deleted from the database, and a further nine sets of answers were discarded as meaningless. The final sample thus comprised 870 respondents, of whom 105 (12%) were men. The age span ranges from 19 to 73 years (median 44 years). More than half (52%) of the respondents reported working for more than 16 years.

Measure. The scale comprised 35 items scored on a five-point scale ("I totally agree" -5- to "I totally disagree" 1); Therefore, lower scores on the Occupational Integrity Scale indicate lower scale ("I totally agree" -5- to "I totally disagree" 1). The average seniority was  M = 16.4 years, SD = 9.7, ages ranged from 22 to 61 years (M = 42.8, SD = 9.1, median = 43). The sample comprised a total of 147 male respondents, all employees of a manufacturing company in the Czech Republic. All respondents were from one shift and contained enough items to avoid compromising the test’s reliability. The final solution used 23 items (12 fall under the first factor, six under the second, and the remaining five under the third) specific factors which are saturated with selected items. Thus, two factors affect the score for one item – one general and one specific. All factors are modelled as orthogonal.

The result of a model that included all 35 items (to be attributed to the highest factor saturation obtained by EFA) had an acceptable data match, χ2 (493) = 1075.92, CFI = .935, RMSEA = .037 (.034; .040). However, after reviewing the amount of error variance for each item, we found that a number of items did not perform any function, since their communalities reached just one or a few per cent, generating what was, basically, error variance. As a result, we removed individual items according to their communality until we determined a solution that demonstrated good fit with the data and, at the same time, contained enough items to avoid compromising the test’s reliability. The final solution used 23 items (12 fall under the first factor, six under the second, and the remaining five under the third) and demonstrated an excellent match with the data, χ2 (207) = 387.96, CFI = .971, RMSEA = .032 (.027; .037). The factor loadings are shown in Table 1. For more information, see Table S3 (see the list of the tables in the repository before the list of references).

Reliability. To estimate the reliability of total test and individual factors, we used the omega coefficients based on the factor weights of the bifactorial model. The reliability of the twenty-three-item OIS version was equal to .84. The reliability factor has a reliability of .78, the adherence to principles factor one of .72, and the moral sense factor one of .65.

2.2 Study 2

The aim of the second study was to replicate results from the first study regarding the factorial structure of the scale, and to test the criterion predictive validity of the OIS via supervisor-reported CWB and construct validity via self-reported personality traits.

Method

Sample and procedure. The sample comprised a total of 147 male respondents, all employees of a manufacturing company in the Czech Republic. All respondents were from one shift and represented 100% of this shift. Their voluntary involvement resulted from the study’s ethical clearance. The respondents’ ages ranged from 22 to 61 years (M = 42.8, SD = 9.1, median = 43). The average seniority was M = 16.4 years, SD = 9.7, median =16 years.
Respondents were recruited with the use of the intentional selection method, since voluntary choice in testing for integrity can lead to the greater involvement of people with higher integrity and therefore reduce the validity of results. Each respondent filled in an informed consent form, followed by the NEO-PI-R questionnaire, the Occupational Integrity Scale and standard demographic data (age and the length of their employment in the company). Codes on each respondent’s questionnaire were placed by the line managers and passed to the research team so that only the line manager knew the link between the employee’s name and the code. The line managers knew the respondents’ identity, which the research team did not know, and, on the other hand, the line managers did not have the results of the questionnaires after their evaluation. In the second phase, that took place two months after the respondents were tested, their superiors were invited to assess the incidence of counterproductive behaviour among the respondents. In the second phase of the study. The NEO-PI-R (Costa & McCrae, 2008; Hřeblíčková, 2004) was used because of its verified relation to CWB (Hoffman & Dilchert, 2012; Ones et al., 1993). Specifically, in line with Hoffman & Dilchert’s (2012) meta-analysis, we utilized three FFM dimensions from the NEO-PI-R (conscientiousness, agreeableness, neuroticism). The NEO-PI-R is made up of 240 items, on which the respondent describes on a five-point scale to what extent it corresponds with them.

The evaluation of employees’ CWB by their line management contained 30 statements, each describing one particular manifestation of CWB. To create the statements, we applied the list of counterproductive work behaviour (Sackett, 2002). These were: theft, destruction of property, misuse of information, misuse of time and resources, unsafe behaviour, poor attendance, poor quality (poor quality of work), alcohol use, drugs use, inappropriate verbal action, and inappropriate physical action. Managers assessed each statement, depending on how often the behaviour of the particular employee occurred in the time interval that was observed, on a 0 – 4 scale (with a verbal anchoring, never, rarely, occasionally, often, very often).

Results

Firstly, we tested whether the psychometric properties of the OIS observed in the group of teachers in study 1 could also be observed in a more diverse sample. We therefore aimed to verify factor invariance between the first and the second respondent group. The model proposed in the first study was repeatedly fitted to both groups, and restrictions on the match of parameter values between the two groups were gradually added. The model configuration invariance, which assumes the same structure but does not place any conditions on the coefficients of factor weights, showed a very good match with the data, $\chi^2 (415) = 714.49$, CFI = .970, RMSEA = .038 (.033; .042). As a result of attaching the restriction to the factor saturation concordance, a small but statistically significant decrease in accuracy occurred. Adding other restrictions did not result into a deterioration in the model quality. The residual variance model still demonstrated a high fit, $\chi^2 (487) = 1025.382$, CFI = .963, RMSEA = .047 (.043; .051). The results, therefore, suggest that the OIS retains the same factor structure for respondents from a diverse work environment.

Investigating the presence of CWB and monitoring the factors that may indicate its presence brings a number of likely pitfalls.

The first is the great diversity of its manifestations and its relative scarcity within some of the areas being monitored. If one follows the division of CWB into the 12 areas proposed by Grays (1999) and Sackett (2002), she will observe almost exclusively zero values for several of the areas. The second issue concerns possible differences in the style and quality of the evaluation of the employees by the line managers. After records obtained were reviewed, it became clear that the data from one of the evaluators was strikingly different from that from the rest – the number of registered CWB occurrences in this case was within the 26-40 range (median 28), while in the other evaluators it was within the 0-25 range (median 4). Most of the records of this assessor were almost or totally identical, so we did not consider them valid and they were disregarded in analyses. This led to a reduction in the sample size to 113 respondents (22 to 61 years, mean age of respondents 43.1 years, median 44 years, standard deviation 9.8 years).

Table S4 depicts the means, standard deviations, and intercorrelations between variables (see the list of the tables in the repository before the list of references). The first part of the table summarizes also descriptive statistics for study 1. Attention should be paid especially to intercorrelations among OIS subscales as well as to relations between OIS and personality test in the meaning of N, A, C factors of NEO-PI-R. On the level of OIS, the relation of Moral sense factor seems to be interesting, because this factor, unlike the others, has just weak or no relation to the rest of OIS factors. With the exception of relations between OIS and Neuroticism, the results represent expected relations of OIS to Agreeableness and Conscientiousness. The results support the construct validity of the OIS. We should point out however, that the presentation of relations of CWB’s to OIS and personality measures is done as a general summary of the data and does not account for organization level differences, something that is controlled for in the regression analyses.
Because of the characteristic arrangement of CWB scores with their peak around zero, dependent variables in the regression models were logged transformed. For each of the CWB dimensions where we observed a sufficient non-zero value occurrence, we created two statistical models: one predicting the number of transferred CWB points on the basis of the three subscales of the twenty-three-item version of the OIS and the other predicting the same using the three NEO-PI-R factors. Apart from the stated regressors, a nominal regressor, “evaluator”, was also included in both models to control for the non-independence of measurement in the assessment of CWB. In both models we observed their predictive validity and especially how their predictive validity depends on the scales of the diagnostic method beyond the information of who rated the employee. Details of how much information on the CWB the individual tests bring and whether the results are statistically significant are summarized in Table 1. Subsequently, we examined the effect of the various factors of the OIS, or the FFM dimensions, on various manifestations of CWB. Lastly, in a test of construct validity, relationship between the OIS factors and the FFM dimensions was examined. These results are shown in Tab. 2.

The obtained results provided limited evidence of the possibility of predicting CWBs on the basis of the OIS factors. A significant association with the OIS factors was only detected in the case of passive aggressiveness, where a negative relation with the scale of moral sense was also observed. The contribution above the framework of variance, which can be explained by the regressor of the evaluator, was, however, only 4%. A higher CWB prediction was observed by means of the three FFM dimensions. This includes, especially, the dimension of Agreeableness, which, with its negative weight, links both with the total CWB score and with specific demonstrations, such as Misuse of time and resources, Poor Attendance, and Inappropriate verbal actions. The explained variance does not exceed 6%. With the exception of Agreeableness, a significant relation was also observed in Conscientiousness, which has a positive weight in Passive aggressiveness.

Finally, the analysis brought findings of the construct validity of OIS. The Reliability factor had a negative effect on Neuroticism, all three OIS factors had a positive effect on Agreeableness, which is true also for Conscientiousness, with exception of Moral sense effect. Preliminary conclusions to the construct validity were confirmed and supplemented in detail.

3 Discussion

The aim of the studies presented here was to test the basic psychometric properties of the Occupational Integrity Scale (OIS), a new, freely available integrity scale. In addition, we aimed to determine the predictive validity of the OIS with respect to supervisors' prospective assessment of various types of counterproductive work behaviors (CWB) and construct validity with respect to employees' personality traits.

The OIS was constructed as an instrument to assess the level of integrity in employment in a public and a private organization. After a conceptual analysis of the various approaches to integrity in the workplace, we expected to find two factors in the structure of the new scale. One factor representing moral integrity and commitment to principles based on the philosophical tradition of integrity, and a second factor representing overt integrity through items focusing on attitudes toward a variety of counterproductive behaviors. The expected two-factor solution was found, as well as the presence of a general factor, which is consistent with the results of other integrity tests, such as the Squares (Cut-e Czech s.r.o., 2015). However, a closer look at the results suggests that a three-factor structure would be more appropriate, corresponding to the factors ‘Reliability’, ‘Adherence to principles’ and ‘Moral sense’.

The final three-factor structure provided a better solution and an opportunity to better align with the theoretical background of the concept of integrity, including the distinction between the importance of the second and third factors. For example, the second factor - adherence to principles - monitors a person's identification with general or higher principles that go beyond the work environment. Adherence to principles relates to the person him/herself; the related OIS elements also relate to the person's self-concept. As described, adherence to principles is closer to Schlenker's (2008) commitment to principles. Principles are without content, items focus only on adherence to them, and respondents can think of different principles. In contrast, the third factor, moral sense, reflects applied moral principles, and this usually involves anonymous application in social space, again beyond the specific work setting. Moral sense represents attitudes or experiences on specific issues related to integrity. Finally, the first factor, called 'Reliability', is similar in both factor solutions. Reliability is best described as the image of a 'good employee' in relation to work and other people. The items are attitudinal and are applied to the work context, and in this context provide an obvious test of integrity.

Overall, the three-factor solution was supported by a CFA in a large sample of public educators in the Czech Republic and additionally verified in a second group of employees of private companies. We consider the good interpretability of the three-factor solution, the fit of the model to the data of both groups, and the satisfactory reliability as supporting elements that allow the OIS to be used for further research. Nevertheless, the OIS factor structure could be considered partially heterogeneous given its conceptual origins: It combines the philosophical and empirical traditions. From this perspective, we also suggest the possibility of selecting individual OIS factors depending on the specific goals of the research.

The results of the second study confirmed the construct validity of the new scale. The expected relationships between the OIS and conscientiousness and agreeableness were found (Ones & Viswesvaran, 2001; Schlenker, 2008). Also expected (Hoffman & Dilchert, 2012) were correlations with neuroticism, although to a lesser extent, only with the reliability factor. However, we

<table>
<thead>
<tr>
<th>CWB</th>
<th>Reliability</th>
<th>Adherence to principles</th>
<th>Moral sense</th>
<th>NEO-PI-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misuse of time and resources</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>Unsafe behaviour</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.09</td>
<td>0.12*</td>
</tr>
<tr>
<td>Poor attendance</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.08</td>
</tr>
<tr>
<td>Poor quality of work</td>
<td>-0.09</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.03*</td>
</tr>
<tr>
<td>Inappropriate verb.</td>
<td>0.01</td>
<td>0.05</td>
<td>-0.07</td>
<td>-0.08</td>
</tr>
<tr>
<td>Passive aggressiveness</td>
<td>0.06</td>
<td>0.01</td>
<td>-0.18*</td>
<td>0.02</td>
</tr>
<tr>
<td>Total CWB</td>
<td>0.03</td>
<td>0.01</td>
<td>-0.11</td>
<td>-0.11*</td>
</tr>
<tr>
<td>Neuroticism (N)</td>
<td>-0.20*</td>
<td>-0.04</td>
<td>-0.18</td>
<td>-0.11*</td>
</tr>
<tr>
<td>Agreeableness (A)</td>
<td>0.31**</td>
<td>0.19*</td>
<td>0.27**</td>
<td>0.11</td>
</tr>
<tr>
<td>Conscientiousness (C)</td>
<td>0.43***</td>
<td>0.29**</td>
<td>0.12</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Note. *,** and *** indicate statistical significance at levels of significance of .05, .01, and .001.
acknowledge some limitations. In the present study, three specific Big Five personality traits were used to test the construct validity of the OIS. Future research could extend these tests to other personality factors that have shown solid associations with tests of overt integrity in previous research (e.g., modesty, Marcus et al., 2007). In terms of criterion validity, the current results do not allow us to assess the overlap between the OIS and other standardized integrity tests because the design of the current study lacks a second integrity test. This limitation may affect the utility of the new scale, and future research should address this issue.

To test the predictive criterion validity of the new scale, we tested an approach that predicts counterproductive work behavior (CWB). In doing so, we applied several important principles, some of which have also been used in previous research (e.g., van Iddekinge et al., 2012; Ones et al., 1993). First, we deliberately selected a homogeneous group to reduce the risks of other selection types when the integrity factor might be insufficiently distributed. Another principle was the adoption of a prospective study design and a CWB measure by supervisors. Although we consider these principles essential to the aims of the study, they also have potential drawbacks. First, the above principles might be expected to lead to a lower association between integrity and CWB. Although van Iddekinge et al. (2012) reported the overall estimated observed validity, this was mainly related to established samples and the source of the criterion in self-report. In addition, the Berry et al. (2012) meta-analysis concluded that self-reports of CWB provide surprisingly richer information about employee CWB than other reports, including supervisor ratings. Despite research interest in external data (Barclay & Aquino, 2010), the superiority of third-party reports of CWB compared to self-reports can be questioned because supervisors do not have the opportunity to observe CWB of real employees (Berry et al., 2012).

Based on the results of the second study, the OIS appears to have low predictive validity of CWB. In general, the relationships among the variables of interest are relatively low, so they are not consistent with the more optimistic conclusions of previous meta-analyses (cf. Ones et al., 1993) nor with the overall estimated observed validity in meta-analysis with more stringent requirements for the included studies (van Iddekinge et al., 2012). In this sense, our results are closer to those reported by van Iddekinge et al. (2012) specifically for predictive designs, applicant samples, and non-self-report sources of criteria. As Berry et al. (2012) found, third-party reports may lead to a lower frequency of CWB observations, which is consistent with both van Iddekinge et al.’s (2012) and our results. In addition, we attribute the lower predictive power of the integrity test to the research design limitations mentioned earlier. The frequency and variability of types of CWB could naturally be lower in a stable team of employees verified by long-term working relationships, making prediction more difficult.

The results devoted to the CWB subfactors provided limited statistically significant information about the predictive power of our integrity test. However, given the low predictive power of CWB overall, no significant effects were expected. Taken together, these results suggest that the Occupational Integrity Scale can be used to some extent to predict aspects of CWB such as passive aggressiveness.

Following the above comments on the criterion validity of the OIS in units of percent of true variance explained, we cannot readily recommend the use of the scale in practice. Unlike other predictors used in personnel selection, inferences about integrity are used for adverse selection. The possible exclusion of lower integrity applicants in the OIS poses risks for further selection procedures in the absence of better indicators of criterion validity for the OIS.

In conclusion, the newly developed integrity test has good psychometric properties that support its further use in research in both public and private organizations. However, the applicability of the new scale in professional practice requires further validation work. Similarly, the OIS needs to be validated against other integrity tests. The two studies provide important information not only about the psychometric properties of the new scale, but also about the differential relationships between the newly proposed integrity scale and facets of the CWB.

Electronic Supplementary Material

The tables S1 – S4 are freely accessible at the research repository: https://osf.io/7bx3k/?view_only=ea57a5b11548fcfbafe151e913baa

Literature:


Primary Paper Section: A
Secondary Paper Section: AN