
ORIGINAL RESEARCH REPORT

Validating the Workplace Dignity Scale

Casey Scott-Campbell* and Matt Williams†

Workplace Dignity has long been the subject of scholarly enquiry, although until recently the body of research has been dominated by ethnographic work. Recently, Thomas and Lucas (2019) developed the first quantitative, direct measure of perceptions of workplace dignity: the Workplace Dignity Scale (WDS). Given the importance of understanding dignity in the workplace, this study sought to replicate the initial scale validation study conducted by Thomas and Lucas, so as to further test the validity of the WDS and the reliability of the scores it produces. Moreover, the current study contributes to the ongoing methodological reform of psychology towards a transparent and rigorous science by preregistering the method and analysis script prior to collecting data. A large sample of workers ($N = 812$) from the United States were recruited through Prolific Academic and completed an online questionnaire that included the WDS, as well as theoretically related scales (e.g., workplace incivility). Confirmatory factor analyses indicated that the model specified by Thomas and Lucas had reasonable global fit, although it did not meet all of our criteria for good fit, and estimates of reliability (ω_c) indicated that responses to items making up the two subscales of the WDS, Dignity and Indignity, had high internal consistency. Nomological analyses revealed that the Dignity subscale of the WDS was significantly correlated in the expected directions with theoretically related variables. Furthermore, the Dignity and Indignity factors of the WDS were found to highly correlate with one another, and an exploratory analysis suggested that the Indignity factor might be a methodological artefact, posing questions as to whether the two factors are qualitatively different phenomena as was argued by Thomas and Lucas. It is concluded that the WDS is a promising tool for measuring workplace dignity although refinement of the proposed measurement model may be necessary.

Keywords: Workplace Dignity; reproducibility; organizational psychology; psychometrics; validity; reliability

Purpose of the Study

Until recently, a scale that measured workplace dignity had not been developed, despite many years of research that employed this concept. This has slowed extending knowledge on workplace dignity because valuable quantitative enquiry has been stilted. Nonetheless, Thomas and Lucas (2019) have recently developed the Workplace Dignity Scale (WDS) which seeks to address this gap. The WDS is a promising tool for organisational psychology, and it has shown the potential to predict important organisational outcomes (e.g., engagement, burnout, turnover intentions) above and beyond other organisational variables (e.g., meaningful work, organisational respect) (see Study 3 of Thomas & Lucas, 2019). The purpose of the current study was to provide a close replication of the validation of the WDS with preregistered inferential criteria for validity and reliability. Not only did this allow for an assessment of whether

Thomas and Lucas' findings could be reproduced, but it also meant that there was less ambiguity about whether the data collected supported the validity of the scale and the reliability of the scores it produces.

Introduction

It has long been thought that dignity is a core human characteristic that distinguishes human beings from other animals (Bolton, 2007). Broadly defined as the sense of worth and respect deserved by all people (Hodson, 2001), the concept of dignity is referenced across a wide range of disciplines – from medical ethics to the law – and is considered the ultimate human value, to the extent that it forms a foundation of the United Nations universal declaration of human rights (United Nations, 1948). As such, it is important to explore how dignity manifests and is influenced in the many facets of human life. In particular, the workplace is suggested as an environment in which dignity can be both realised and violated.

Research undertaken on workplace dignity thus far has primarily relied on ethnographic methods, exploring, for example: the lived experience of nurses facing dignity violations (Khademi et al., 2012), discursive analyses of

* Massey University, Wellington, NZ

† Massey University, Auckland, NZ

Corresponding author: Casey Scott-Campbell
(casey.scott-campbell.1@uni.massey.ac.nz)

neoliberal discourse of dirty work (Purser, 2009), and the protection strategies that minority groups use when facing workplace dignity violations (Baker & Lucas, 2017). These idiographic approaches demonstrate the breadth of manifestations of workplace dignity. However, a body of quantitative work focussing explicitly on workplace dignity has yet to be developed to complement these findings. The small number of studies throughout the organisational literature (e.g., Lucas et al., 2017) that have sought to quantitatively investigate workplace dignity have relied on variables from a single dataset which are theorised to relate to dignity (e.g., meaningful work, absence of supervisor conflict) but which do not directly assess workplace dignity itself. Thus, a measure of workplace dignity has been lacking until recently.

A nomothetic approach (i.e., discovering generalisable knowledge) to investigation alongside idiographic approaches (e.g., ethnographic approaches) would lead to a more comprehensive understanding because it provides a different *type* of knowledge about workplace dignity, and opens up the possibility for different types of questions to be answered about workplace dignity. For example, a quantitative approach, could assess the extent to which other factors influence and are influenced by workplace dignity across a population of people (e.g., job enlargement, turnover intentions). In turn, new insights may be used in applied settings. For instance, a quantitative measure of workplace dignity could form part of a baseline assessment of an organisation's climate, from which the effects of an intervention (such as a diversity training programme) could be measured.

What is Workplace Dignity?

Lucas (2017) has defined workplace dignity as: "the self-recognised and other-recognised worth acquired from (or injured by) engaging in work activity" (p. 2549). The study of workplace dignity, as a phenomenon of interest itself, is only in its infancy. It is, nonetheless, difficult to conceptualise because it is drawn upon under many different circumstances, its meaning is often assumed, and explicit descriptions of it tend to invoke not entirely overlapping concepts, such as autonomy, status, humanity, or work contribution (Lucas, 2017). One related concept to workplace dignity is decent work. Having a right to decent work, itself, appeals to notions of human dignity, and moreover *workplace* dignity may be upheld or violated depending on the decency of one's work. Modelling their work off guidelines set out by the International Labour Organization, Duffy et al. (2017) identified five factors underlying their Decent Work Scale, and it is not unreasonable to suggest that these would provide the grounds for the promotion or degradation of dignity. For example, when a worker is provided physically and interpersonally safe working conditions (i.e., factor 1 of the Decent Work Scale), their dignity may be more likely to be upheld, whereas when one's work does not permit them sufficient hours for free time and rest (i.e., factor 4), this might lead to a violation of dignity. Decent work may act as a means through which workers' inherent human worth is recognised (Lucas, 2017). Despite being related,

workplace dignity is not the same concept as decent work, rather it is a possible consequence. Other related but somewhat distinct concepts to workplace dignity are workplace wellbeing (Weziak-Bialowolska, Bialowolski, Sacco, VanderWeele, & McNeely, 2020) and organisational justice (Silva & Caetano, 2016).

Since the purpose of this paper is to replicate the original validation work conducted by Thomas and Lucas (2019), here we relay the four key principles that they suggested might underlie workplace dignity which can, to some extent, be traced back to religious teachings on human dignity (Sison et al., 2016). Nonetheless, Lucas (2015) emphasises that further theory development is necessary, and that the principles described below do not constitute a comprehensive theory of workplace dignity (for a more extensive discussion of these principles, see Lucas, 2015).

Firstly, Thomas and Lucas (2019) suggest that workplace dignity is self-construed: while dignity judgements are influenced by discourses surrounding work, "it is the individual who is the ultimate arbiter of her or his experience of workplace dignity" (p. 76). Secondly, workplace dignity is thought to be 'communicatively bound' whereby it depends on one's own assessment of their worth and the extent to which others display recognition and acknowledgement of this worth through social interaction (Lucas, 2015). Workplace dignity may, therefore, be considered similar to the construct of organisational-based self-esteem which involves the self-evaluation of one's worth and competence in one's organisation (Pierce & Gardner, 2004). Where the two constructs diverge is that workplace dignity also depends on others to communicate one's worth. Indeed, evidence from focus groups indicates that the violation or promotion of dignity occurs largely through communication, to the extent that demonstrating competence or making a contribution was not itself enough to affect the experience of dignity (Lucas, 2015).

Thirdly, Thomas and Lucas (2019) suggested that there are two components to workplace dignity. The first component is *inherent* dignity. This is often described as human dignity, and is the belief that without exception, all people have a baseline level of unconditional and absolute worth simply because they are human (Brennan & Lo, 2007). In contrast, and somewhat paradoxically, dignity may be *earned* through instrumental contributions in the workplace (Hodson, 2001). Earned dignity, then, is conditional and unequal because it is afforded on the basis of an individual's efforts and abilities in the workplace. The earned dignity component is argued as being central to distinguishing the construct of workplace dignity, from (inherent) dignity more broadly. Because one can earn dignity through their behaviours in the workplace, workplace dignity might be thought of as distinct from simply inherent dignity operating in a workplace context (Lucas, 2015). Finally, Thomas and Lucas (2019) suggested that workplace dignity might be bivalent, in that it has positive elements, from which dignity can be promoted, and negative elements, from which dignity must be protected.

Therefore, Thomas and Lucas (2019) suggest that the psychometric measure of workplace dignity should attend to these principles by using self-report, assessing social interactions, including positive and negative items, and enquiring about both inherent and earned dignity. It is through these principles that Thomas and Lucas developed the first psychometric measure of workplace dignity: The Workplace Dignity Scale (WDS). As for any psychological construct, it is important that the supposed measurement of workplace dignity is valid and reliable.

The Workplace Dignity Scale

In the following paragraphs, we present the results of the initial validation of the WDS by Thomas and Lucas (2019). The WDS is the first psychometric measure of employee perceptions of dignity in the workplace. The 18-item scale consists of self-report items such as “I feel respected when I interact with people at work”, and “I am treated as less valuable than objects or pieces of equipment” which are responded to on a 7-point rating scale (1 = *Strongly Disagree* to 7 = *Strongly Agree*).

Initial scale development

Initially, four general themes of workplace dignity (respectful interaction, recognition of competence and contribution, equality, and inherent value) emerged following focus group interviews with 62 working adults (Lucas, 2015).

These participants were asked about their personal definitions of ‘dignity at work’ and any examples of where their dignity was either affirmed or denied. Conceptual definitions of the four themes and an additional general workplace dignity definition were generated prior to the development of a bank of positively- and negatively-valenced items (in accordance with the theorised bivalent nature of workplace dignity). The content validity of 97 items was then evaluated by 11 experts who rated the extent to which each item was essential for measuring its intended theme and gave open-ended feedback. Items yielding a consistent response from experts were retained (Ayre & Scally, 2014; Lawshe, 1975).

Construct and nomological validity

Thomas and Lucas (2019) analysed the subsequent 61-item scale based on the responses of 401 participants who passed attention checks and met the inclusion criteria of working at least 30 hours per week, being over 21 years-old, and having at least 2 years paid work experience.

Thomas and Lucas settled on a six-factor measurement model indicated by 18 items using exploratory structural equation modelling (E-SEM) whereby models with an increasing number of factors were tested successively. The six-factor model was yielded because, relative to other models, this model was more parsimonious, had stronger factor loadings, fewer cross-loaded items, better global fit (based on the comparative fit index [CFI], Tucker–Lewis index [TLI], and root mean square error approximation [RMSEA]), and was supported by modification indices and background theory. Four factors reflected the previously identified themes, one factor reflected general workplace

dignity, and the remaining factor consisted of negatively-valenced items, subsequently labelled workplace indignity.

Based on previous literature indicating that workplace dignity is multidimensional, and the high correlations between factors involving positively valenced items, Thomas and Lucas tested a second-order model using CFA whereby Dignity was reflected by five first-order factors and could covary with the separate Indignity factor. This model was argued to provide the best fit to the data, $\chi^2(128) = 392.97$, CFI = .955, TLI = .946, RMSEA = .072 [.064, .080], compared with first-order, one-, and two-factor solutions, based on CFI, TLI, and RMSEA values, and modification indices.

Internal consistency values (α) for the Dignity and Indignity factors, were .96 and .88, respectively. We note here, however, that a second-order model whereby Dignity is reflected by six factors (i.e., Indignity is a sub-factor loading on the Dignity higher-order factor) would have had identical fit to Thomas and Lucas’ second-order model. In Study 3, Thomas and Lucas (2019) used the E-SEM procedure with a new sample ($N = 532$) to assess whether they could replicate the second-order measurement model established in Study 2. The identical model was, indeed, supported, and a separate CFA on the new sample also supported the model as having good fit: $\chi^2(128) = 493.01$, $p < .001$, CFI = .958, TLI = .951, RMSEA = .073 [.066, .080]. Internal consistency estimates here for Dignity and Indignity were, again, .96 and .88, respectively.

Thomas and Lucas (2019) then evaluated the nomological validity of the WDS by exploring relationships between the WDS and theoretically related variables using SEM. Dignity was significantly and positively related to Workplace Status, Need for Competence, and Interpersonal Justice, and negatively related to Workplace Alienation and Workplace Incivility (each of which yielded single-factor solutions). Dignity was also negatively related to the two factors yielded from Workplace Objectification: Objectification and Humanization (reverse coded to align with Objectification). Indignity was negatively related to Competence and Interpersonal Justice, and positively related to Incivility, Alienation, and Objectification, but not related to Status or Humanization. These relationships were largely the same when they calculated correlations using observed scale scores, with the exception of Dignity and Humanization which were not significantly correlated with one another. The authors took the different pattern of relationships that Dignity and Indignity had with other variables as further evidence for treating them as qualitatively different factors.

Finally, discriminant validity was assessed (and considered established) with Fornell and Larcker’s (1981) criterion whereby the average variance extracted in WDS items (.72) was higher than the highest squared correlation with the other latent variables ($r^2_{\text{dignity-justice}} = .59$).

Conceptualising validity

Thomas and Lucas (2019) refer to validity as a property of the WDS scale itself (e.g., p. 87, “The results of Study 2 support the construct validity of the WDS”). In contrast,

the Standards for Education and Psychological Testing (American Educational Research Association et al., 2014) suggest that validity is a property not of scales but of *interpretations* of scores for specific uses. While Thomas and Lucas's framing differs from that advocated in the Standards, it is consistent with the theory of validity of Borsboom et al. (2004), which posits that a test is a valid measure of an attribute if the attribute exists and the attribute causes variation in the test scores (see also Borsboom et al., 2009). Treating validity as a property of scales rather than interpretations of scores is also consistent with Cronbach and Meehl (1955), who described how the validity of tests could be evaluated by reference to a nomological network (a strategy utilised by Thomas and Lucas, 2019). In this article, we thus follow Thomas and Lucas (2019) in referring to validity as a (possible) property of the WDS itself, rather than of interpretations of WDS scores.

Psychometric Research and Reproducibility

When testing reliability and validity, there are many possible analyses that can be conducted, and ambiguity about what constitutes adequate evidence for reliability and validity. This means that there are 'researcher degrees of freedom' when evaluating the reliability and validity of scales, which may lead to evidence that contradicts reliability and validity being hidden. Subjecting 15 social and personality psychology scales to a more comprehensive assessment of reliability and validity than is typically reported in the field (Flake, Pek, & Hehman, 2017), and with a large sample ($N = 81,986$), Hussey and Hughes (2019) found that only 4% of the measures demonstrated 'good' structural validity: that is, they passed recommended thresholds for internal consistency, test-retest reliability, confirmatory factor structure and measurement invariance tests.

A partial solution to the problem of hidden invalidity is to restrict researcher degrees of freedom by preregistering tests of reliability and validity. Although we know of no equivalent review of validation practices in the work psychology literature, the possibility of researcher degrees

of freedom when testing reliability and validity is a potential problem in all areas of psychological research (for a discussion of questionable research practices, more generally, in work psychology, see Banks, Rogelberg, Woznyj, Landis, & Rupp, 2016; Bosco, Aguinis, Field, Pierce, & Dalton, 2016; Kepes & McDaniel, 2013). Furthermore, while we have no reason to suspect that Thomas and Lucas' (2019) results were biased by researcher degrees of freedom, their study was not preregistered, and therefore it is important to conduct a preregistered replication. Indeed, replication studies are intrinsically valuable in light of the ongoing problems with replicability in psychology (Chambers, 2017; Schmidt, 2009).

The Current Study

In this study, we aimed to examine the internal factor structure, reliability and nomological validity of the WDS, thus conducting a replication of Thomas and Lucas' (2019) Study 2. The main differences between the current study and that of Thomas and Lucas are that: (a) we conducted a CFA to assess the fit of their model on new data, given that we are seeking to confirm their proposed factor structure, whereas they used E-SEM prior to CFA (our single CFA was the only SEM-based analysis in our study); (b) in addition to reporting the global fit statistics they reported, we report the standardised root mean square residual (SRMR); (c) we did not include the Workplace Objectification Scale, given that it has not been subject to thorough validation work prior to in Thomas and Lucas' study; (d) we only replicated their observed variable/correlational approach to assessing nomological validity, whereas they additionally calculated SEM relationships with latent factors (for correlations calculated using sum scores, see above-diagonal area on Thomas and Lucas' Table 3, for nomological relationships estimated using SEM see below-diagonal area on Thomas and Lucas' Table 3 and Figure 1 of Thomas and Lucas) – this was in order to simplify our analyses and address the limitation of amplified SEM-based correlations; (e) in addition to reporting α as in Thomas and Lucas, we report McDonald's Omega hierarchical (ω_h) and Omega total (ω_t)

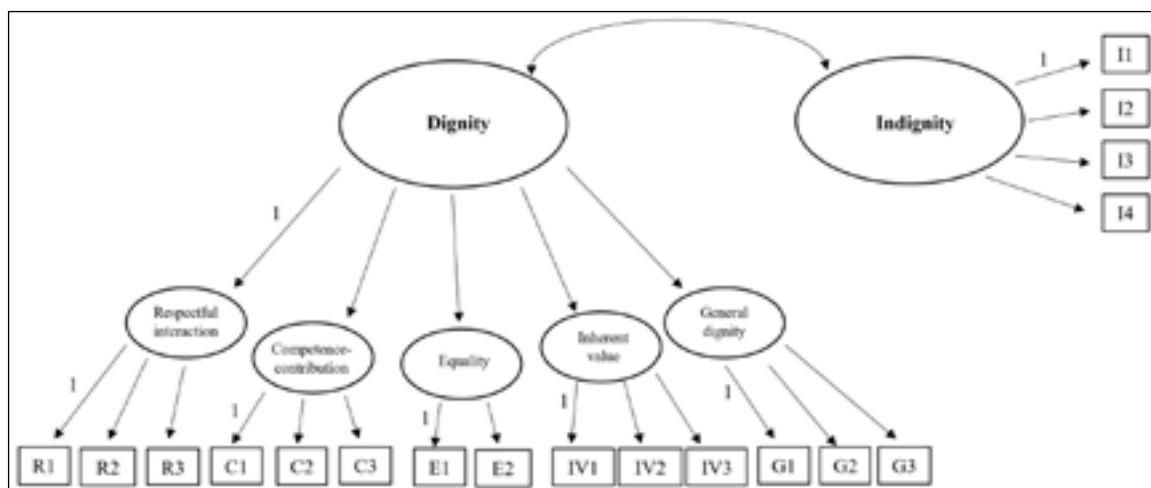


Figure 1: Thomas and Lucas' (2019) best fitting model that was preregistered to be identified in the current study using CFA. Note that each item will have its own measurement error term, with no covariances between error terms. Error terms are not displayed for the sake of brevity.

reliability estimates, and; (f) we preregistered our data collection and analysis plans prior to collecting data, whereas Thomas and Lucas did not.

Hypotheses

We assessed the fit of a second-order CFA model that includes a Dignity factor reflected by five sub-factors and an Indignity factor that covaries with Dignity, which was the best fitting model in Thomas and Lucas (2019). We hypothesised that this model would provide good fit (see inference criteria in Method section) to the data (*hypothesis 1a*).

Thomas and Lucas suggested that because Dignity correlated with Indignity at $-.59$ using observed scale scores and $-.64$ using SEM estimation, and because the relationships between Dignity and other variables were not simply mirrored in the relationships between Indignity and those variables, Dignity and Indignity were related but operated on separate continuums. Therefore, we hypothesised that Dignity would negatively correlate (using observed scale scores) with Indignity at $-.7 < \rho < -.3$ (*hypothesis 1b*). That is, Dignity and Indignity would correlate negatively, but the magnitude would not be so large as to indicate that they measure the same thing.

To assess the nomological validity of the WDS we computed correlations (treating each of the scales as observed variables) between Dignity and Indignity and theoretically related variables. Specifically, based on Thomas and Lucas (2019), we hypothesised that Dignity would be positively correlated at $\rho > .3$ with Need for Competence (*hypothesis 2a*), Interpersonal Justice (*2b*), and Workplace Status (*2c*). Need for Competence refers to individuals' desires to interact effectively with their workplace environment (Van den Broeck et al., 2010), and should, thus, be theoretically related to the *earned* component of workplace dignity (for more detail on the theoretical rationale for each nomological hypothesis, see Thomas and Lucas, 2019). Interpersonal Justice relates to the quality of interpersonal interactions an individual has, with respect being a key factor (Bies & Moag, 1986), and should be related to workplace dignity given the importance of respectful interaction for upholding inherent dignity (Sayer, 2007). It has been argued that an individual's status within their group is closely linked with the level of respect that they perceive to be treated with (Blader & Yu, 2017); one's workplace status, then, should be similarly related to workplace dignity given the importance of respect to dignity.

Similarly, as per Thomas and Lucas, we hypothesised that Dignity would be negatively correlated at $\rho < -.3$ with Alienation (*2d*) and Incivility (*2e*). Alienation and Dignity are closely related with one another: to be psychologically damaged by or disconnected from one's work is often thought of as an automatic violation to one's dignity (Hodson, 2001). Incivility should be inversely related to Dignity because uncivil behaviours typically involve displaying a lack of respect (Andersson & Pearson, 1999).

The null hypotheses of $\rho = .3$ and $\rho = -.3$ were chosen because a correlation of $.3$ would indicate a medium effect size (Cohen, 1988), and because such correlations had a magnitude of approximately $.5$ in Thomas and

Lucas (2019). This allowed us to directly test whether the true correlations have magnitudes greater than $.3$ in each case above.

If workplace dignity is to be considered bivalent with Indignity and Dignity being related but "operating somewhat independently" (Thomas & Lucas, 2019, p. 100), there needs to be *a priori* theory about Indignity's relationship with nomological variables. However, Thomas and Lucas did not hypothesise about how Indignity would relate to nomological variables, nor outline a rationale *a priori* as to why it would or would not. Thus, in line with Thomas and Lucas, we made no formal predictions about Indignity's relationship with nomological variables. Finally, we hypothesised that Dignity would demonstrate acceptable reliability (*hypothesis 3a*), as would Indignity separately (*3b*), with the lower bound of the 95% confidence intervals for $\omega_i > 0.7$.

Method

Sample Characteristics

Participants were sought from the Prolific Academic database that were aged 21 years or over, worked at least 31 hours per week, worked in the USA, and had a Prolific approval rating of 95% or higher. Thus, for the purposes of replication, participants were recruited from the same population as in Thomas and Lucas (2019). We employed a stopping rule that data collection would cease when we reached 850 participant responses. This is approximately twice the sample size of Thomas and Lucas' (2019) Study 2. Furthermore, based on a power analysis for computing a correlation between Dignity/Indignity and a theoretically related variable, for 90% power, an N of 809 would be required if $\rho = .4$ when the null hypothesis is that $\rho = .3$. The smallest significant correlation for Thomas and Lucas (2019) (calculated using observed scale scores) between either Dignity or Indignity and the nomological variables included in the current study was $r = .48$ and was between Dignity and Status (also see inference criteria in the Nomological Validity subsection). Moreover, when using the maximum likelihood (ML) estimation method with SEM for the proposed model, which includes 60 statistical estimates, a sample size of 850 gives a ratio of approximately 14:1 participants per parameter. This ratio exceeds the minimum recommendation of 10:1 provided by Jackson (2003).

In total, there were 853 participants who completed the study (three more than our preregistered stopping rule), owing to an unanticipated feature of the Prolific data collection process. Three extra participants were included to fill in for those who Prolific determined to have taken too long to complete the survey but whose data were nonetheless suitable. Thus, we obtained slightly more data than our set target.

Measures

The demographic variables measured included age, gender, ethnicity, country of work, highest educational level completed, occupation type, hours worked per week, years in current job, and years of prior work experience. These demographic variables, however, were not used in the main set of analyses.

Scales

The scales used in this study included the 18-item WDS (Thomas & Lucas 2019), the 4-item Interpersonal Justice Scale (Colquitt, 2001), the 6-item Need for Competence Scale (Van den Broeck et al., 2010), the 5-item Workplace Status Scale (Djurdjevic et al., 2017), the 7-item Workplace Incivility Scale (Cortina et al., 2001), and 5 items (as per Thomas & Lucas, 2019) from the 8-item Work Alienation Scale (Nair & Vohra, 2009). These scales were presented to participants using the same stems as in Thomas and Lucas, some of which had been modified from their original papers; for instance, participants were asked “During the past FIVE years of employment, how often have you been in a situation where any of your supervisors or coworkers...” for the Workplace Incivility Scale. These were the only measured variables used in the main analyses.

Scale coding

Each of the scales had the same number of response options as presented in Thomas and Lucas (2019). Scales were coded also according to Thomas and Lucas (e.g., Strongly Disagree and Strongly Agree on an item of the WDS were coded as 1 and 7, respectively). Items 1 and 4 of the Need for Competence Subscale were reverse coded. Greater scores on each scale indicate higher levels of the construct they intend to measure.

Scale summation

Mirroring Thomas and Lucas (2019), responses to each of the scales were summed into single scale scores, aside from the WDS, which was split into two subscales: Dignity (which included all items aside from I1, I2, I3, and I4) and Indignity (I1, I2, I3, I4) (see **Figure 1**). Scale scores for Dignity, Indignity, Interpersonal Justice, Need for Competence, Status, Incivility, and Alienation were created by summing the responses to the items within each scale for each participant.

Data Exclusions

Only participants who answered “Yes” to the study consent question were permitted to participate. Respondents who answered “No” were directed out of the survey using survey flow settings, and their responses discarded. Prolific Academic only advertised this study to those who had specified that they were 21 years or over, work 31 or more hours per week, work in the USA, and had a Prolific approval rating of 95% or higher (as per our prescreening request). However, if a participant indicated that they did not meet these demographic criteria in response to the demographic questions early in the survey, they were directed out of the survey and their response discarded.

As a quality check, an attention check was included at the end of the Workplace Incivility Scale reading: “Please demonstrate that you are paying attention by ticking Often”. Participants who gave any other response or did not respond to this item were excluded during data processing. Our preregistered exclusion criteria specified that in the event that there were any duplicate responses from the same Prolific worker (as indicated by a matching

Prolific ID), and these duplicate responses were still present after applying the exclusion criteria specified above, then only the most recent response from each Prolific worker would be retained and the remainder excluded. Furthermore, participants were to be excluded if the computer returned a response outside the available range of responses on any item.

We did not specify any exclusion criteria at the level of data within participants because our use of an online survey with items in a rating scale format limited the possibility of extreme outliers. We prespecified that participants who had 11 or more of the 45 items in the main scales missing were to be excluded. After participants were excluded based on these and the earlier criteria, we conducted a single imputation using the expectation maximisation method to impute missing values. Only the main study variables in the dataset were included in the imputation model. Single imputation was chosen (rather than a more complex imputation method) because we anticipated that there would be very few missing data points (which is typical when recruiting through Prolific Academic; e.g., Margolis et al., 2019). Although the full information maximum likelihood procedure can handle missing data for the CFA, the remaining analyses required imputation. Thus, for the sake of simplicity, we used the imputed dataset for all the analyses, including the CFA. The above rule for missing data was specified to cover all instances where participants failed to complete the survey including those in which technical issues arose.

Of the 853 original participants, two participants did not work in the US, 32 did not work at least 31 hours per week, and 10 failed the attention check. There were no duplicates nor responses outside of the range of possible responses. The resulting number of participants included for analysis was 812 (as opposed to 809), as three of the excluded participants failed more than one of the exclusion criteria. One item was missing for one participant, and, thus, no one was excluded based on the exclusion criteria for missing data.

Procedure

We conducted an observational (non-experimental) study in an online survey format to assess the validity of the WDS. As such, there were no randomisation or blinding procedures. Participants completed the full survey using the Qualtrics survey platform. Demographic information was collected first followed by the six main scales, the order of which was randomised.

Data Analysis

Confirmatory factor analysis

We conducted a CFA with the maximum likelihood (ML) estimation method using the lavaan package (version 0.6-3; Rosseel, 2012) within the R software environment (version 3.6.0; R Core Team, 2019) to assess the factor structure of the 18-item WDS. The first item that each latent variable loaded on to was set to 1 (the ‘marker item’ approach), and standardised parameter estimates are reported.

Assumptions

Items were treated as continuous in line with the ML estimation method. This has been suggested as appropriate when items have at least five response options (Rhemtulla et al., 2012). ML has the following assumptions:

1. Independence of errors; this was not directly tested as it is captured in the model output.
2. Linear relationships between constructs; this is not feasibly testable.
3. Multivariate normality; we tested for multivariate skew and kurtosis using Mardia's tests in order to demonstrate how well the data approximated a multivariate normal distribution. However, our preregistration specified that the analysis method was not to be changed based on the results of these diagnostics.

Model specification and identification

The best fitting model identified by Thomas and Lucas (2019) was fitted to our data whereby five dignity factors: Respectful Interaction, Competence-Contribution, Equality, Inherent Value, and General Dignity (and their respective indicators) reflect a higher-order Dignity factor, and with Indignity (and its respective indicators) as a second factor allowed to covary with Dignity.

The WDS includes 18 items (observed variables) and, hence, has 171 distinct sample moments. This figure was calculated using the formula: $[v(v+1)]/2$, where v is the number of observed variables (Brown, 2015). There were 42 parameters to be estimated, thus, giving 129 degrees of freedom. One factor loading for each of the seven factors (i.e., both the first- and second-order factors) was fixed (known), with the remaining (unknown) parameters estimated including: 18 error terms (one per item), 7 factor variances, 16 factor loadings, and one factor covariance (see **Figure 1**). We only identified one CFA model as per our preregistered plan.

Inference criteria

To assess the global fit of the CFA model we planned to report the values of each of the following fit indices (values indicating 'good' fit are given in parentheses, based on recommendations by Hu & Bentler, 1999): Root Mean Square Error of Approximation (RMSEA < 0.06), Standardised Root Mean Square of the Residual (SRMR < 0.09), Tucker Lewis Index (TLI > 0.95), and Comparative Fit Index (CFI > 0.95). A confidence interval for the RMSEA value was also reported. The Chi-squared test statistic was also reported accompanied by a p value with a significance threshold of $p < 0.05$. However, *hypothesis 1a* was to be considered supported if the SRMR, CFI, TLI, and RMSEA statistics *all* demonstrate good fit, in line with the criteria used by Hussey and Hughes (2019). In the case that the single CFA model being tested returned a negative result (i.e., hypothesis 1a was not supported), this result is still meaningful in the context of replication and, thus, we did not have contingency plans for probing unexpected CFA results (any analyses we did conduct in this regard, such

as investigating modification indices, were to be explicitly labelled as exploratory).

Hypothesis 1b was to be considered supported if the lower bound of the 95% confidence interval for the Spearman's correlation between Dignity and Indignity was greater than $-.7$ and the upper bound less than $-.3$. Regardless of whether hypotheses 1a and 1b were supported, in the spirit of replication we planned to and did proceed with subsequent nomological (and then reliability) analyses.

Nomological validity

Each of the scales were to be treated as observed variables and nomological validity was assessed by computing Spearman's correlations with 95% confidence intervals between Dignity (and Indignity) with each of: Need for Competence, Interpersonal Justice, Status, Alienation, and Incivility.

Inference criteria

Each of the sub-hypotheses of *hypothesis 2* that involved testing $\rho > .3$ (i.e., Dignity positively correlating with Need for Competence, Interpersonal Justice, and Status) were to be considered supported if the lower bound of the confidence interval was greater than $.3$, for the null hypothesis that $\rho = .3$. Similarly, each of the sub-hypotheses of *hypothesis 2* that involved testing $\rho < -.3$ (i.e., Dignity negatively correlating with Alienation and Incivility) were to be considered supported if the upper bound of the confidence interval was less than $-.3$, for the null hypothesis that $\rho = -.3$. The null hypotheses of $\rho = .3$ and $\rho = -.3$ were chosen because a correlation of $.3$ would indicate a medium effect size, and because such correlations had a magnitude of approximately $.5$ in Thomas and Lucas (2019). This allowed us to directly test whether the true correlations had magnitudes greater than $.3$ in each case above.

Familywise corrections for Type 1 error were not calculated because we followed a specific preregistered set of analyses and applying such a correction would have increased the risk of incorrectly finding that particular claims in the original study had not been replicated.

Reliability

We calculated Cronbach's α and McDonald's (1999) ω_h and ω_t as estimates of reliability for the Dignity, Indignity, Interpersonal Justice, Workplace Status, Need for Competence, Incivility, and Alienation scales. These estimates were calculated using a hierarchical factor analysis, with one general factor, three "group" factors (factors specific to subsets of items), factors extracted using the minimum residual method, a maximum of 100 iterations, and direct oblimin rotation (Revelle, 2019). For each reliability statistic, the lower bound of the 95% confidence interval should be greater than 0.7 to demonstrate acceptable reliability (Hussey & Hughes, 2019). The factor model assumed in the omega calculation differs from the CFA model specified in hypothesis 1a in the sense that it involves direct effects of the general "dignity" factor (as well as effects of the group/lower-order

factors) on all items, and in that the “group” factors in the omega calculation are extracted based on a data-driven process rather than prespecified. This reflects the differing purposes of the models: The CFA model in hypothesis 1a is intended to represent a prespecified, parsimonious and testable model of variation in item responses, whereas the model assumed in the omega calculation is intended purely to decompose variation such that it is possible to determine the proportion of variability due to a single general factor underlying all items (ω_h) and the proportion due to general and group factors (ω_t).

Inference criteria

Hypothesis 3a was to be considered supported if the lower bound of the 95% confidence interval for ω_t was greater than 0.7 for the *Dignity* scale. *Hypothesis 3b* was to be considered supported if the lower bound of the 95% confidence interval for ω_t was greater than 0.7 for the *Indignity* scale. We chose ω_t as our main estimate of reliability, while also reporting α and ω_h in line with modern recommendations (Dunn et al., 2014; Hussey & Hughes, 2019; Revelle & Condon, 2018).

Results

Descriptive Statistics and Demographic Information

Table 1 displays the means and standard deviations of the scale sum scores on each of the measured variables. **Table 2** displays the demographic characteristics of the participants included for analysis. Moreover, respondents

reported that they had worked in their current jobs for an average of 6.5 years ($SD = 5.2$) and had an average total work experience of 15.9 years ($SD = 10.0$).

Hypothesis 1a

To assess the global fit of the CFA model, four fit indices were calculated: RMSEA = .073 [.068, .078], SRMR = .030, TLI = .958, CFI = .965. Each of these fit indices apart from the RMSEA met the prespecified inference criteria for the model to demonstrate good fit.

Thus, the hypothesis that the model will provide good fit to the data was not supported as the inference criteria required every statistic to meet their specified thresholds.

Table 1: Descriptive Statistics of Sum Scores for Dignity, Indignity, and Nomological Variables.

Variables	M (Range of possible scores)	SD
Dignity	78.87 (14–98)	14.73
Indignity	8.77 (4–28)	5.28
Need for Competence	27.84 (6–42)	2.55
Interpersonal Justice	17.63 (4–20)	3.07
Workplace Incivility	14.00 (7–35)	5.50
Workplace Status	15.28 (5–25)	5.21
Workplace Alienation	16.61 (5–35)	8.44

Table 2: Demographic Information.

Demographic variable	Category	% respondents
Gender	Male	57.8
	Female	41.3
	Non-binary	0.7
	Prefer not to say	0.2
Age	21–30	34.0
	31–40	40.1
	41–50	17.1
	Over 50	8.7
Ethnicity	White	74.0
	Black or African American	6.6
	Asian	6.0
	Other ethnicity ^a	13.4
Educational attainment	Some or all high school	5.9
	Some college no degree	15.3
	Bachelor’s, Associate or Technical degree	57.0
	Master’s, Professional, or Doctorate degree	21.8
Occupation	Computer/Mathematical Management	16.6
		12.4
	Business and Financial Operations	9.5
	Other occupation ^b	61.5

Note: ^aOther ethnicity includes those of mixed ethnicity, American Indian or Alaska Native, Latino, and Hispanic. ^bOther occupation includes those from a wide range of Department of Labor classifications; the most common three are presented.

In addition, we also calculated χ^2 , although this did not form part of our inference criteria: $\chi^2(129) = 687.98$, $p < .001$. Mardia's test statistics for multivariate skew and kurtosis were 10848.03 ($p < .001$) and 162.68 ($p < .001$), respectively, as such there was evidence of multivariate nonnormality. Path estimates for the measurement model are presented in **Figure 2**.

Hypothesis 1b

The Spearman's correlation calculated between Dignity and Indignity (using observed scale scores) was $-.75 [-.78, -.70]$. While the upper bound of the confidence interval was less than $-.3$, the lower bound was less than $-.7$. Thus, the hypothesis that Dignity would negatively correlate with Indignity at $-.7 < \rho < -.3$ was not supported (see **Table 3**).

Hypothesis 2

Spearman's correlations between Dignity and nomological variables, as well as between Indignity and nomological variables are presented in **Table 3**. Correlations calculated between Dignity and the nomological variables indicated positive relationships between Dignity and Need for Competence ($r_s = .43 [.36, .49]$), Interpersonal Justice ($r_s = .63 [.58, .67]$), and Workplace Status ($r_s = .51 [.45, .56]$).

The lower bounds of the confidence intervals for each of these correlations were greater than $.3$, which supported the hypotheses 2a, 2b, and 2c. Furthermore, Spearman's correlations indicated negative relationships between Dignity and Workplace Alienation ($r_s = -.64 [-.68, -.59]$), and Workplace Incivility ($r_s = -.54 [-.59, -.48]$). The upper bounds of the confidence intervals for these correlations were less than $-.3$, which supported hypotheses 2d and 2e.

Hypothesis 3

The reliability of the scores produced by the WDS was assessed by estimating ω_t and 95% confidence intervals for the Dignity and Indignity factors, separately. These estimates and those of α and ω_h are presented for each factor, as well as for the nomological variables in **Table 4**. The ω_t reliability estimates and confidence intervals for Dignity and Indignity, respectively, were $.98 [.97, .99]$ and $.95 [.92, .98]$. As the lower bound of the 95% confidence intervals were greater than $.7$, hypotheses 3a and 3b were supported.

Exploratory Analyses

To examine the possibility that the Indignity factor might be a methodological artefact given that it included only negatively worded items (Gnambs & Schroeders, 2020;

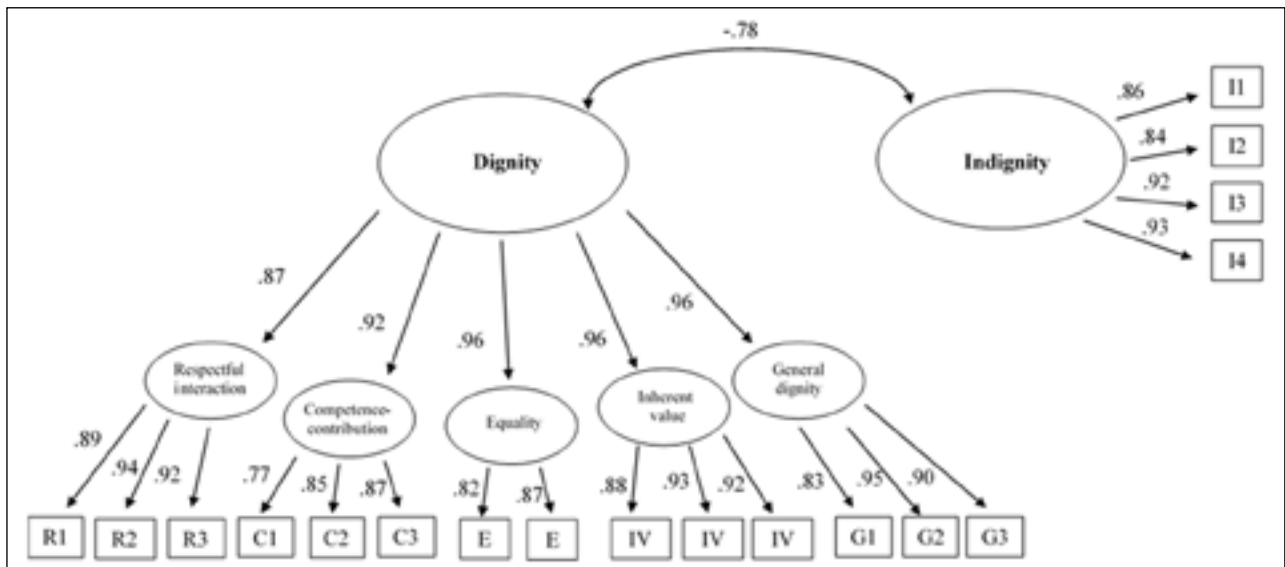


Figure 2: Measurement model including standardised path estimates. Error variances are omitted for brevity.

Table 3: Correlations between Dignity, Indignity, and Nomological Variables.

Variables	Dignity	95% CI		Indignity	95% CI	
		Lower	Upper		Lower	Upper
Dignity	–	–	–	–	–	–
Indignity	–.75	–.78	–.70	–	–	–
Need for Competence	.43	.36	.49	–.42	–.48	–.36
Interpersonal Justice	.63	.58	.67	–.61	–.65	–.56
Workplace Status	.51	.45	.56	–.33	–.39	–.26
Workplace Alienation	–.64	–.68	–.59	.54	.48	.59
Workplace Incivility	–.54	–.59	–.48	.52	.46	.58

Marsh, 1986, 1996), we tested a model with a nested method factor (see **Figure 3**). In this model there was no Indignity factor; instead, the Indignity items loaded directly on to the second-order Dignity factor and there was an additional method factor that only loaded on to the Indignity items and that was orthogonal to the other latent variables. The global fit statistics were as follows: RMSEA = .071 [.066, .077], SRMR = .026, TLI = .960, CFI = .967. Further, to examine whether this new model fit the data better than the original model, the Bayesian Information Criterion (BIC) was calculated for both models, and their difference converted to a Bayes Factor (see formula 10 in Wagenmakers, 2007). There was a BIC difference of 22.7, which translated to a Bayes Factor of 82,867, meaning that the data we observed would be 82,867 times as probable if the nested model was true than if the original model was true. A comparison of the

Akaike Information Criterion (AIC) values also favoured the nested model over the original model.

Discussion

The purpose of this study was to conduct a close replication of Study 2 of Thomas and Lucas (2019). Thus, we investigated the validity and the reliability of the scores produced by the WDS by conducting a CFA to assess whether the best-fitting model suggested by Thomas and Lucas could be replicated in a larger sample, by calculating correlations between the Dignity and Indignity components of the WDS, as well as between Dignity and theoretically related variables (using observed scale scores), and by estimating reliability using a psychometrically sound measure of internal consistency, ω_t . Furthermore, by preregistering inferential criteria for validity and reliability, we sought to provide more clarity with respect to whether our collected data could be said to support the validity of the scale and reliability of the scores produced by the WDS.

Table 4: Reliability Estimates for the WDS and Nomological Variables.

Variables	ω_t	ω_h	α
Dignity	.98	.91	.97
Indignity	.95	.91	.94
Need for Competence	.94	.86	.91
Interpersonal Justice	.93	.92	.92
Workplace Status	.96	.95	.95
Workplace Alienation	.96	.90	.95
Workplace Incivility	.93	.85	.92

Confirmatory Findings

With respect to the fit of the CFA model, three out of four of the model fit statistics (SRMR, CFI, and TLI) met our prespecified criteria for indicating good fit as recommended by Hu and Bentler (1999) with only the RMSEA (=0.073) failing to meet the recommended threshold (of less than .06; although we acknowledge that the RMSEA value was close to this threshold). Thus, *hypothesis 1a*, which required all four of the fit statistics to meet the prespecified criteria (Hussey & Hughes, 2019), was not supported. Nonetheless, the RMSEA, CFI, and TLI values described in this study were comparable to those

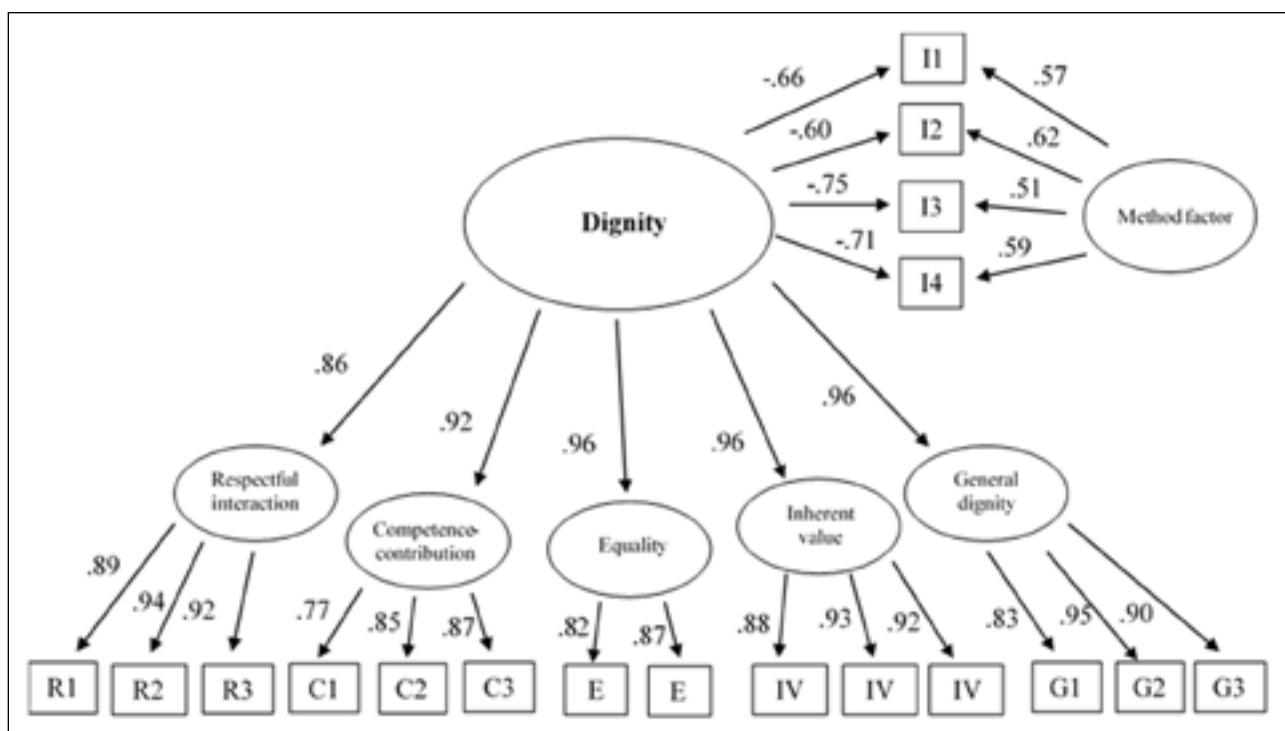


Figure 3: Measurement model with method factor for Indignity items including standardised path estimates. Error variances are omitted for brevity.

identified by Thomas and Lucas (who did not calculate SRMR). However, in neither study did the RMSEA value (nor the lower bound of its 90% confidence interval) meet the recommended threshold suggesting acceptable fit (Hu & Bentler, 1999).

The next hypothesis proposed that the Dignity and Indignity components of the WDS would be negatively correlated with one another, but that the magnitude of the correlation would not be so large as to indicate that they, in fact, measure the same thing. This hypothesis (1b) was not supported: we found that the estimated correlation had a magnitude (-.75) that was greater than predicted. This result is perhaps inconsistent with Thomas and Lucas' (2019) theoretical notion that Dignity and Indignity are qualitatively distinct constructs. Interestingly, this estimate was somewhat larger than that found in the original validation study, albeit in the same direction. Nonetheless, we did not plan *a priori* to test whether our correlations significantly differed from those of Thomas and Lucas (2019). However, given that our de-identified data is publicly available on an online depository (OSF.io/svpgf), future researchers may wish to explore this.

Each of the sub-hypotheses comprising *hypothesis 2* regarding correlations between the Dignity component of the WDS and theoretically related variables were supported. Dignity was positively correlated with Need for Competence, Interpersonal Justice, and Workplace Status, and negatively correlated with Workplace Alienation and Incivility. Furthermore, the magnitudes of the positive correlations were all significantly larger than .3, and the negative correlations larger than -.3. Here, our findings were very similar to the findings of Thomas and Lucas (2019) in terms of the size of these correlations, with the greatest difference in size between studies being .10 (for the correlation between Dignity and Interpersonal Justice). Again, a preregistered objective assessment of whether the correlations differ between studies may help to establish the theoretical position of workplace dignity within its proposed nomological network. Nonetheless, our results seem to support the nomological validity of the WDS.

Finally, we found that subscale scores on the WDS demonstrated strong internal consistency, as evidenced by particularly high ω_i estimates for both Dignity and Indignity (of .98, and .95, respectively). Thus, *hypothesis 3* was supported. Unlike Thomas and Lucas (2019) who relied on α as their estimate of reliability, we used ω_i – which measures the total reliability of a scale by calculating the proportion of variance that is not attributed to measurement error – as our main estimate in line with recent recommendations in the literature (e.g., Dunn et al., 2014). Furthermore, we also reported α values that are consistent with those estimated by Thomas and Lucas (2019).

Are Dignity and Indignity Distinct Phenomena?

Overall, in this study, our results were similar to some of the findings of Thomas and Lucas' (2019) Study 2. One question, however, that should be further probed is whether Dignity and Indignity should be considered as separate factors. Our findings provide some indication

that Indignity is not a qualitatively different phenomenon to Dignity.

Firstly, Thomas and Lucas (2019) claimed that the relationships between Dignity and nomological variables were not simply mirrored in the relationships between Indignity and nomological variables (but with opposite signs), thus suggesting that Dignity and Indignity were qualitatively different. However, we found that the pattern of nomological relationships with Dignity compared to that of Indignity were practically mirrored, showing mostly similar magnitudes though in opposite directions. Additionally, the correlation between Dignity and Indignity was particularly large (-.75), indicating that perhaps Indignity, for the most part, is simply the reverse of Dignity, rather than a separate phenomenon.

Secondly, the items that make up the Indignity factor are very similar to those that make up the Dignity factor, in that they seem simply to be negatively worded Dignity items. For example, the Indignity item: "I am treated in undignifying ways at work" seems to be a reversal of the Dignity item: "I am treated with dignity at work". It is possible that in the original EFA by Thomas and Lucas (2019), the Indignity items emerged as a separate factor simply because they were all negatively worded. Factor analysis can sometimes cause spurious factors to emerge that reflect negatively worded items (Schmitt & Stuits, 1985; Spector et al., 1997; Woods, 2006). This explanation is consistent with the result of our exploratory analysis, which suggested that the Indignity factor may be a methodological artefact. Researchers need to further investigate this idea by conducting preregistered confirmatory analyses.

Limitations and Future Directions

A limitation of the current study is that we did not attempt to replicate Study 3 of Thomas and Lucas (2019) which explored potential antecedents and consequences of workplace dignity, thus, further testing the criterion-related validity of the WDS. As such, we did not assess the validity of the WDS *comprehensively*. We also acknowledge that by replicating Study 3 we would have been able to further test the claim that workplace dignity is bivalent, given that Thomas and Lucas (2019) used the pattern of external relationships with Dignity and with Indignity as evidence.

A second limitation is that the age of our sample was skewed towards younger workers (i.e., approximately 87% of the sample was aged between 21 and 50), whereas a substantial proportion (approximately 30%) of the US workforce is composed of older workers (U.S. Bureau of Labor Statistics, 2019). Moreover, with an aging workforce this proportion continues to grow, meaning that as it currently stands, the WDS may lack generalisability to older workers. Future research should investigate whether the current structural validity of the WDS extends to older workers by testing for measurement invariance between age groups; it may be the case that there are generational differences in how people interpret the items of the WDS.

Third, we did not assess the factor structure of each nomological variable given that we wanted to keep a

narrow focus for this replication attempt and the primary focus was to assess the factor structure of the WDS. However, we acknowledge the need to continually assess the factor structure of scales each time they are used with a new sample and encourage future researchers not only to reassess the validity of WDS scores but also to reassess other measures used in conjunction with the WDS (Hussey & Hughes, 2019). Moreover, by using SEM to test nomological validity we could have accounted for the effects of measurement error on estimates of the relationships between WDS factors and nomological variables (Westfall & Yarkoni, 2016).

Finally, we did not establish norms or any alternative scheme for the interpretation of individual scores. The mixed evidence for validity suggests that there is not yet a sufficient empirical basis to use the WDS to make real-world decisions about individual people or organisations. On the other hand, use of the WDS for substantive *research* purposes may be justifiable (depending on the needs of a given study).

Future work should continue to probe the validity of the WDS, not only by conducting studies similar to this one, but also by validating the scale with populations from different countries to the US to assess measurement invariance. Indeed, it may be case that different cultures have different conceptualisations of dignity in the workplace (or simply dignity) (e.g., in non-Western countries) and that the WDS is not interpreted by different populations in a conceptually similar manner. One other way workplace dignity could be explored is for future researchers to interview a new, larger group of working adults and follow the same item generation process as in Thomas and Lucas (2019) to see whether the same types of items emerge. In turn, this would indicate whether the conceptual themes that underpin workplace dignity (as measured by the WDS) are reflective of an understanding of workplace dignity that is common across working adults. Finally, because the WDS is the first quantitative measure of workplace dignity, there are many other measures in the work psychology literature worth contrasting the WDS against. One relationship to explore is with the Decent Work Scale (Duffy et al., 2017) described at the outset of this paper. We predict that workplace dignity would correlate positively but not perfectly with decent work, given that the constructs do overlap but do not appear to be the same. This investigation would constitute a further test of the nomological validity of workplace dignity.

Conclusion

Overall, we found evidence in support of the nomological validity and reliability of the scores produced by the WDS, although the proposed measurement model was not supported, suggesting that more work is needed to determine whether it requires refinement. The results of the present study in conjunction with those of Thomas and Lucas (2019) merit continued investigation of the factor structure of the WDS, and replication and extension of the criterion-related validity work conducted in Study 3 of Thomas and Lucas (2019). It is our hope

that such investigations are preregistered to contribute to the ongoing methodological reform of scientific psychology.

Data Accessibility Statement

The preregistered method protocols, participant data (before exclusions), and analysis scripts can be found on this paper's project page on the [OSF.io/svpgf].

The preregistration for this study was in the form of a Word document which included a list of hypotheses and a "method protocols" section. The method section included in this manuscript is very similar to the preregistered "method protocols" section, but with a change from future to past tense, the addition of some information that could only be included after the study had been completed (e.g., final sample size), and some extra explanatory text (not altering the substantive plan for the study).

Funding Information

CS-C – Massey University School of Psychology Postgraduate Research Fund.

MNW – Massey University Strategic Research Excellence Fund.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

- Contributed to conception and design: CS-C, MNW
- Contributed to acquisition of data: CS-C
- Contributed to analysis and interpretation of data: CS-C, MNW
- Drafted and/or revised the article: CS-C, MNW
- Approved the submitted version for publication: CS-C, MNW

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education.** (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Andersson, L. M., & Pearson, C. M.** (1999). Tit for tat? The spiraling effect of incivility in the workplace. *Academy of Management Review*, *24*(3), 452–471. DOI: <https://doi.org/10.5465/amr.1999.2202131>
- Ayre, C., & Scally, A. J.** (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79–86. DOI: <https://doi.org/10.1177/0748175613513808>
- Baker, S. J., & Lucas, K.** (2017). Is it safe to bring myself to work? Understanding LGBTQ experiences of workplace dignity: LGBTQ experiences of workplace dignity. *Canadian Journal of Administrative Sciences/Revue Canadienne Des Sciences de l'Administration*, *34*(2), 133–148. DOI: <https://doi.org/10.1002/cjas.1439>
- Banks, G. C., Rogelberg, S. G., Woznyj, H. M., Landis, R. S., & Rupp, D. E.** (2016). Editorial: Evidence on

- questionable research practices: The good, the bad, and the ugly. *Journal of Business and Psychology*, 1–16. DOI: <https://doi.org/10.1007/s10869-016-9456-7>
- Bies, R. J., & Moag, J. F.** (1986). Interactional justice: Communication criteria of fairness. *Research on Negotiations in Organizations*, 1, 43–55.
- Blader, S. L., & Yu, S.** (2017). Are status and respect different or two sides of the same coin? *Academy of Management Annals*, 11(2), 800–824. DOI: <https://doi.org/10.5465/annals.2015.0150>
- Bolton, S. C.** (2007). Dignity in and at work: why it matters. In *Dimensions of Dignity at Work* (pp. 3–16). DOI: <https://doi.org/10.1016/B978-0-7506-8333-3.50006-8>
- Borsboom, D., Cramer, A. O. J., Kievit, R. A., Scholten, A. Z., & Franíć, S.** (2009). The end of construct validity. In *The concept of validity: Revisions, new directions, and applications* (pp. 135–170). Charlotte, NC, US: IAP Information Age Publishing.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J.** (2004). The concept of validity. *Psychological Review*, 111(4), 1061–1071. DOI: <https://doi.org/10.1037/0033-295X.111.4.1061>
- Bosco, F. A., Aguinis, H., Field, J. G., Pierce, C. A., & Dalton, D. R.** (2016). HARKing's threat to organizational research: Evidence from primary and meta-analytic sources. *Personnel Psychology*, 69, 709–750. DOI: <https://doi.org/10.1111/peps.12111>
- Brennan, A., & Lo, Y. S.** (2007). Two conceptions of dignity: honour and self-determination. In J. Malpas & N. Lickiss (Eds.), *Perspectives on Human Dignity: A Conversation* (pp. 43–58). DOI: https://doi.org/10.1007/978-1-4020-6281-0_5
- Brown, T. A.** (2015). *Confirmatory factor analysis for applied research* (Second edition). New York; London: The Guilford Press.
- Chambers, C.** (2017). *The seven deadly sins of psychology: a manifesto for reforming the culture of scientific practice*. Princeton University Press. DOI: <https://doi.org/10.1515/9781400884940>
- Cohen, J.** (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Colquitt, J. A.** (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, 86(3), 386–400. DOI: <https://doi.org/10.1037/0021-9010.86.3.386>
- Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D.** (2001). Incivility in the workplace: incidence and impact. *Journal of Occupational Health Psychology*, 6(1), 64–80. DOI: <https://doi.org/10.1037/1076-8998.6.1.64>
- Cronbach, L. J., & Meehl, P. E.** (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. DOI: <https://doi.org/10.1037/h0040957>
- Djurdjevic, E., Stoverink, A. C., Klotz, A. C., Koopman, J., da Motta Veiga, S. P., Yam, K. C., & Chiang, J. T.-J.** (2017). Workplace status: The development and validation of a scale. *Journal of Applied Psychology*, 102(7), 1124–1147. DOI: <https://doi.org/10.1037/apl0000202>
- Dunn, T. J., Baguley, T., & Brunnsden, V.** (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412. DOI: <https://doi.org/10.1111/bjop.12046>
- Flake, J. K., Pek, J., & Hehman, E.** (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, 8(4), 370–378. DOI: <https://doi.org/10.1177/1948550617693063>
- Fornell, C., & Larcker, D. F.** (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39. DOI: <https://doi.org/10.2307/3151312>
- Gnamb, T., & Schroeders, U.** (2020). Cognitive abilities explain wording effects in the Rosenberg Self-Esteem Scale. *Assessment*, 27(2), 404–418. DOI: <https://doi.org/10.1177/1073191117746503>
- Hodson, R.** (2001). *Dignity at work*. DOI: <https://doi.org/10.1017/CBO9780511499333>
- Hu, L., & Bentler, P. M.** (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. DOI: <https://doi.org/10.1080/10705519909540118>
- Hussey, I., & Hughes, S.** (2019). *Hidden invalidity among fifteen commonly used measures in social and personality psychology* [Preprint]. DOI: <https://doi.org/10.31234/osf.io/7rbfp>
- Jackson, D. L.** (2003). Revisiting sample size and number of parameter estimates: some support for the N:q hypothesis. *Structural Equation Modeling: A Multidisciplinary Journal*, 10(1), 128–141. DOI: https://doi.org/10.1207/S15328007SEM1001_6
- Kepes, S., & McDaniel, M. A.** (2013). How trustworthy is the scientific literature in industrial and organizational psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 252–268. DOI: <https://doi.org/10.1111/iops.12045>
- Khademi, M., Mohammadi, E., & Vanaki, Z.** (2012). Nurses' experiences of violation of their dignity. *Nursing Ethics*, 19(3), 328–340. DOI: <https://doi.org/10.1177/0969733011433926>
- Lawshe, C. H.** (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. DOI: <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Lucas, K.** (2015). Workplace dignity: communicating inherent, earned, and remediated dignity. *Journal of Management Studies*, 52. DOI: <https://doi.org/10.1111/joms.12133>
- Lucas, K.** (2017). *Workplace dignity*. DOI: <https://doi.org/10.1002/9781118955567.wbieoc222>
- Lucas, K., Manikas, A. S., Mattingly, E. S., & Crider, C. J.** (2017). Engaging and misbehaving: How dignity affects employee work behaviors. *Organization Studies*, 38(11), 1505–1527. DOI: <https://doi.org/10.1177/0170840616677634>
- Margolis, S., Schwitzgebel, E., Ozer, D. J., & Lyubomirsky, S.** (2019). A new measure of life

- satisfaction: the riverside life satisfaction scale. *Journal of Personality Assessment*, 101(6), 621–630. DOI: <https://doi.org/10.1080/00223891.2018.1464457>
- Marsh, H. W.** (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology*, 22(1), 37–49. DOI: <https://doi.org/10.1037/0012-1649.22.1.37>
- Marsh, H. W.** (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70(4), 810–819. DOI: <https://doi.org/10.1037//0022-3514.70.4.810>
- McDonald, R. P.** (1999). Test homogeneity, reliability, and generalizability. In *Test theory: A unified approach* (pp. 76–120). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nair, N., & Vohra, N.** (2009). Developing a new measure of work alienation. *Journal of Workplace Rights*, 14(3), 293–309. DOI: <https://doi.org/10.2190/WR.14.3.c>
- Pierce, J. L., & Gardner, D. G.** (2004). Self-esteem within the work and organizational context: A review of the organization-based self-esteem literature. *Journal of Management*, 30, 591–622. DOI: <https://doi.org/10.1016/j.jm.2003.10.001>
- Purser, G.** (2009). The dignity of job-seeking men: Boundary work among immigrant day laborers. *Journal of Contemporary Ethnography*. DOI: <https://doi.org/10.1177/0891241607311867>
- R Core Team.** (2019). R: a language and environment for statistical computing. (Version 3.6.0). Retrieved from <https://www.R-project.org/>
- Revelle, W.** (2019). *Using R and the psych package to find ω* . <http://personality-project.org/r/psych/HowTo/omega.pdf>
- Revelle, W., & Condon, D. M.** (2018). *Reliability from alpha to omega: a tutorial* [Preprint]. DOI: <https://doi.org/10.31234/osf.io/2y3w9>
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V.** (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354–373. DOI: <https://doi.org/10.1037/a0029315>
- Rosseel, Y.** (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(1), 1–36. DOI: <https://doi.org/10.18637/jss.v048.i02>
- Sayer, A.** (2007). Dignity at work: broadening the agenda. *Organization*, 14(4), 565–581. DOI: <https://doi.org/10.1177/1350508407078053>
- Schmidt, S.** (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13(2), 90–100. DOI: <https://doi.org/10.1037/a0015108>
- Schmitt, N., & Stuits, D. M.** (1985). Factors defined by negatively keyed items: the result of careless respondents? *Applied Psychological Measurement*, 9(4), 367–373. DOI: <https://doi.org/10.1177/014662168500900405>
- Sison, A. J. G., Ferrero, I., & Guitián, G.** (2016). Human dignity and the dignity of work: Insights from Catholic Social Teaching. *Business Ethics Quarterly*, 26(4), 503–528. DOI: <https://doi.org/10.1017/beq.2016.18>
- Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P. Y.** (1997). When two factors don't reflect two constructs: how item characteristics can produce artifactual factors. *Journal of Management*, 23(5), 659–677. DOI: <https://doi.org/10.1177/014920639702300503>
- Thomas, B., & Lucas, K.** (2019). Development and validation of the Workplace Dignity Scale. *Group & Organization Management*, 44(1), 72–111. DOI: <https://doi.org/10.1177/1059601118807784>
- United Nations.** (1948). *Universal declaration of human rights*. Geneva.
- U.S. Bureau of Labor Statistics.** (2019). *Employed persons by detailed occupation and age*. Retrieved from <https://www.bls.gov/cps/cpsaat11b.htm>
- Van den Broeck, A., Vansteenkiste, M., Witte, H., Soenens, B., & Lens, W.** (2010). Capturing autonomy, competence, and relatedness at work: construction and initial validation of the Work-related Basic Need Satisfaction scale. *Journal of Occupational and Organizational Psychology*, 83(4), 981–1002. DOI: <https://doi.org/10.1348/096317909X481382>
- Westfall, J., & Yarkoni, T.** (2016). Statistically controlling for confounding constructs is harder than you think. *PLOS ONE*, 11(3), e0152719. DOI: <https://doi.org/10.1371/journal.pone.0152719>
- Weziak-Bialowolska, D., Bialowolski, P., Sacco, P. L., VanderWeele, T. J., & McNeely, E.** (2020). Well-being in life and well-being at work: Which comes first? Evidence from a longitudinal study. *Frontiers in Public Health*, 8, 1–12. DOI: <https://doi.org/10.3389/fpubh.2020.00103>
- Woods, C. M.** (2006). Careless responding to reverse-worded items: implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, 28(3), 186–191. DOI: <https://doi.org/10.1007/s10862-005-9004-7>

Peer review comments

The author(s) of this paper chose the Open Review option, and the peer review comments can be downloaded at:

<http://doi.org/10.1525/collabra.337.pr>

How to cite this article: Scott-Campbell, C., & Williams, M. (2020). Validating the Workplace Dignity Scale. *Collabra: Psychology*, 6(1): 31. DOI: <https://doi.org/10.1525/collabra.337>

Senior Editor: Don Moore

Editor: Wiebke Bleidorn

Submitted: 16 February 2020 **Accepted:** 18 June 2020 **Published:** 20 July 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.