Validation of the Utrecht Work Engagement Scale (UWES) in the Czech Republic

Heveri Martin1, Lukas Novak1, Iva Polackova Solcova1,2, & Peter Tavel1

1 Palacky University Olomouc - Social Health Institute

2 The Czech Academy of Sciences, Institute of Psychology, Prague, Czech Republic

Author note

The authors made the following contributions. Heveri Martin: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Lukas Novak: Writing - Review & Editing,, Writing - Original Draft Preparation, Writing - Review & Editing, statistical analysis; Iva Polackova Solcova: Writing - Review & Editing; Peter Tavel: Writing - Review & Editing.

Correspondence concerning this article should be addressed to Lukas Novak, Univerzitni 244/22, 771 11, Olomouc, Czech Republic. E-mail: [lukas.novak@oushi.upol.cz](mailto:lukas.novak@oushi.upol.cz)

Abstract

Introduction

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline. Sample of 707 employees (Age: M = 43.65, SD = 10.08, Females: 38.47%) recruited from different companies in the Czech Republic was used for purpouses of this study. Neuroticism, extraversion, self-efficacy, spirituality, chronic health diseases and frequency of health risk behavior were measured. Higher UWES total score was reported in professional workers, chief workers and in people with higher vocational school or university. The confirmatory factor analysis (CFA) supported the original three-factor solution: χ2 (24) = 75.373; p < 0.001; CFI = 0.999; TLI = 0.999; RMSEA = 0.058; SRMR = 0.021. Measrement equivalence suggested that on configural, metric, scalar and strict level, the UWES assess work engagement invariantly between males and females. The UWES had an excelent internal consistency (α = 0.96, McDonald’s ω = 0.96) and its convergent validity was supported by positive association with extraversion, self-efficacy and by negative association with neuroticism. Logistic regression revealed that higher score in the UWES was associated with lower chance of developing skin diseases and pain of unclear origin. There was no association of the UWES and health risk behaviours such as smoking, alcohol drinking or illegal drug use. Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

*Keywords:* keywords

*Word count:* X

Validation of the Utrecht Work Engagement Scale (UWES) in the Czech Republic

# Introduction

Based on theoretical assumptions and previous empirical evidence ([Chan, Ho, Ip, & Wong, 2020](#ref-Chan_Ho_Ip_Wong_2020)) we expected significant positive association between self-efficacy and UWES total score (Hypotheses x).

# Methods

## Participants

From the survey (*n* = 1662), we excluded participants being either without work (*n* = 187), pensioners (*n* = 468) or those who did not answer a question regarding economical status (*n* = 223). This resulted in 784 participants. To increase data quality, we removed subjects finishing the survey in a short period of time i.e. < 15 minutes (*n* = 6). The survey typically lasted > 30 minutes. We also excluded respondents answering discrepantly to quality check items (*n* = 71). These items included information about weight, height and age. Tolerance in these control questions was set on 2 kilograms, 2 centimeters, and 2 years respectively. After removal of these subjects, the final number of participants was 707 (Age: *M* = 43.65, *SD* = 10.08, Females: 38.47%).

## Measures

### Utrecht Work Engagement Scale (UWES).

### Daily Spiritual Experience Scale (DSES).

Internal consistency of the DSES was excellent: Cronbach’s = 0.96 95% CI[0.95 - 0.97] and McDonald’s = 0.96 95% CI[0.95 - 0.97].

### General Self Efficacy Scale (GSES).

Internal consistency of the GSES was excellent: Cronbach’s = 0.95 95% CI[0.94 - 0.95] and McDonald’s = 0.95 95% CI[0.94 - 0.95].

### Big Five Inventory - Neuroticism subscale (BFI\_N).

Internal consistency of the BFI\_N was good: Cronbach’s = 0.87 95% CI[0.86 - 0.89] and McDonald’s = 0.87 95% CI[0.86 - 0.89].

### Big Five Inventory - Extraversion subscale (BFI\_E).

Internal consistency of the BFI\_E was good: Cronbach’s = 0.85 95% CI[0.84 - 0.87] and McDonald’s = 0.85 95% CI[0.84 - 0.87].

## Data analysis

Inspection of histograms and results of the Mardia test of multivariate skewness and kurtosis indicated that the normality assumption is violated in the UWES items. Moreover, examination of residual plots and the result of the Breusch-Pagan test ( = 7.21, *df* = 1, p = 0.007) suggested heteroscedasticity. Examination of the polychoric correlation matrix revealed no extreme multicollinearity among items, with all inter-item correlations below 0.90. Thus, methods not requiring parametric assumptions were used. The Little MCAR test provided evidence that missing values are missing on random. Thus, as there was not a large number of missing values (*n* = 60) incomplete cases were deleted listwise.

The instrument’s factor structure was investigated via Confirmatory Factor Analysis (CFA). A series of competing models was specified, reflecting those most prominent in the validation literature. For the full overview and model definitions, see Figure 1 below. The analysis included the original correlated three-factor model (Vigor, Dedication, Absorption) and a unidimensional one-factor model (Schaufeli, Bakker, & Salanova, 2006). To account for the consistently high inter-factor correlations, a hierarchical model, with the three factors loading onto a single second-order Work Engagement factor, was also tested (Domínguez-Salas et al., 2022). Additionally, several alternative configurations were examined, including various two-factor solutions (e.g., Chaudhary, Rangnekar, & Barua, 2012), a partial and full bi-factor model (de Bruin & Henn, 2013), and modified three-factor models with correlated error terms (e.g., Simbula et al., 2013).

*Figure* 1. Confirmatory Factor Analysis Models for UWES-9. All 14 unique tested models are displayed with their respective factor structures and correlated errors where applicable.

Kaiser-Meyer-Olkin (KMO) measure, together with the Bartlett test of sphericity, was applied to assess the factorability of the UWES data. Five indices were used to inspect model fit: 1) Mean Square Error of Approximation (RMSEA); 2) Standardised Root Mean Square Residual (SRMR); 3) chi-square test; 4) Comparative Fit index (CFI); and 5) Tucker-Lewis index (TLI). While these traditional fit indices are widely reported, their fixed cutoff values (e.g., RMSEA < .06, CFI > .95) have been criticized for poor generalizability, particularly because their performance is highly dependent on the specific characteristics of the model being tested (McNeish & Wolf, 2022). Therefore, to provide a more rigorous and accurate assessment of model fit, we adopted the Dynamic Fit Index (DFI) approach (McNeish & Wolf, 2021). This method generates customized fit index cutoffs tailored to the unique characteristics of each model, such as the number of factors, items, sample size, and the magnitude of factor loadings. Using Monte Carlo simulation, the DFI method generates a distribution of fit indices for the researchers’ model under the assumption that it is correctly specified (Level-0) and compares it to distributions where the model contains a degree of hypothetical misspecification (e.g., Level-1, containing minor omitted cross-loadings). This process yields a direct, empirical benchmark (the Level-0 cutoff) representing the expected fit for a well-fitting model with the current data’s characteristics. The Weighted Least Squares Mean and Variance adjusted (WLSMV) on the polychoric correlation matrix was used to fit CFA models.

The invariance of a measurement was explored between males and females. Configural, metric, scalar, and strict invariance were supported if CFI was < 0.01 between invariance models. The scale reliability was measured by McDonald’s and also by Cronbach’s . In addition to these, model-based Composite Reliability (CR) was calculated to assess the internal consistency of each factor, with values > 0.70 indicating good reliability (Fornell & Larcker, 1981).

Construct validity was evaluated in several ways. First, we explored internal convergent validity during CFA, which assessed the degree to which items of a specific factor are related. More specifically, this was tested using the Average Variance Extracted (AVE), with values > 0.50 considered acceptable (Fornell & Larcker, 1981). Next, we assessed internal discriminant validity, which evaluated whether the factors are statistically distinct. For this purpose, the Heterotrait-Monotrait Ratio of Correlations (HTMT) test was used. If values from this test do not reach 0.90, discriminant validity is supported (Henseler et al., 2015). In the further step, we tested convergent and divergent validity externally, i.e., via zero-order Spearman rank correlations with self-efficacy, neuroticism, and extraversion (convergent), and with spirituality (divergent). Relatedly, to provide a formal test of external discriminant validity, the magnitudes of the dependent correlations between the UWES subscales and these external criteria were compared using the z-test from Hittner, May, and Silver (2003). To quantify the magnitude of these differences, we calculated Cohen’s q as an effect size. Cohen’s q values are interpreted as follows: q 0.10 indicates a small effect, q 0.30 a medium effect, and q 0.50 a large effect.

Due to the non-normal distribution of the data, an association between the chronic health illnesses, health risk behaviour, and UWES was calculated using logistic regression. In the logistic models, the outcome variable was the presence of an individual chronic illness or the practice of health risk behaviour. The UWES score was set as a predictor. Education and work position were covariates. Both crude and adjusted effects were estimated. The p-values were corrected by the Bonferroni correction.

Comparisons between socio-demographic groups in the UWES total and subscale scores were performed using the Kruskal-Wallis test. Significant omnibus tests were followed by post-hoc analysis using the Games-Howell test for groups with unequal variances or Dunn’s test where variances were equal. To quantify the magnitude of effects, epsilon-squared () was reported for the overall Kruskal-Wallis test, while Cohen’s *d* and the Rank-Biserial Correlation (*rbc*) were reported for the respective post-hoc comparisons. The interpretation of effect sizes followed established conventions: for epsilon-squared, values of 0.01, 0.06, and 0.14 are considered small, medium, and large effects, respectively. For Cohen’s *d*, benchmarks of 0.2, 0.5, and 0.8 were used for small, medium, and large effects, and for the rank-biserial correlation, values of 0.1, 0.3, and 0.5 were used to interpret small, medium, and large effects. All statistical calculations were conducted in R [Version 4.4.1; R Core Team ([2021](#ref-R-base))]. Primary packages used for analysis included: *lavaan* [@R-lavaan], *papaja* [@R-papaja], *psych* [@R-psych], and *usf* [@R-ufs].

# Results

## Socio-demographic results

Results of the Kruskal-Wallis test, followed by the Games-Howell and the Dunn test, revealed that there are significant differences in socio-demographic groups in the UWES total and subscale scores: professional workers had significantly higher scores in the UWES total, Dedication, and Absorption subscales as compared with workers. Similarly, managers reported higher UWES total scores and also higher Dedication, Absorption, and Vigor subscale scores compared with workers (see Table 1).In terms of education, individuals with a higher vocational school or university degree had a significantly higher total UWES score compared to those with a non-graduation high school or lower education. For the Absorption subscale, the university-educated group scored significantly higher than both the non-graduation high school and the high school-educated groups (Table 1). Furthermore, a significant difference was found based on family status, where individuals in a relationship reported a higher score on the Absorption subscale than those not in a relationship. There were no other significant differences between the tested socio-demographic groups.

Table 1  
*Descriptive Statistics and Group Comparisons for Utrecht Work Engagement Scale (UWES) Scores Across Socio-Demographic Variables*

| variable | n(%) | UWES Group difference | UWES\_D Group difference | UWES\_A Group difference | UWES\_V Group difference | UWES M(SD) | UWES\_D M(SD) | UWES\_A M(SD) | UWES\_V M(SD) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Gender |  |  |  |  |  |  |  |  |  |
| 1. Male | 406(62.65) |  |  |  |  | 39.02 (11.83) | 13.21 (4.38) | 13.11 (4.12) | 12.7 (4.07) |
| 2. Female | 242(37.35) |  |  |  |  | 40.8 (12.94) | 13.66 (4.87) | 13.7 (4.51) | 13.45 (4.25) |
| Education |  |  |  |  |  |  |  |  |  |
| 1. Basic school | 19(2.93) | H(3) = 10.47, p = 0.015 , epsilon2 = 0.01 (2 vs 4, t = 3.28, p = 0.021, d = -0.27) | H(3) = 9.97, p = 0.019, epsilon2 = 0.01 | H(3) = 13.29, p = 0.004 , epsilon2 = 0.02 (2 vs 4, t = 1.46, p = 0.001, d = -0.35; 3 vs 4, t = 1.03, p = 0.047, d = -0.27) |  | 33.68 (14.67) | 11.05 (5.03) | 11.68 (5.23) | 10.95 (4.85) |
| 2. Non graduation high school or lower | 235(36.27) |  |  |  |  | 38.45 (13.16) | 12.96 (4.81) | 12.77 (4.68) | 12.72 (4.42) |
| 3. High school | 188(29.01) |  |  |  |  | 39.59 (12.32) | 13.39 (4.5) | 13.2 (4.3) | 13 (4.21) |
| 4. Higher vocational school or University | 206(31.79) |  |  |  |  | 41.73 (10.56) | 14.06 (4.21) | 14.23 (3.47) | 13.43 (3.62) |
| Family\_status |  |  |  |  |  |  |  |  |  |
| 1. Not in relationship | 108(16.67) | H(4) = 10.12, p = 0.038, epsilon2 = 0.01 |  | H(4) = 10.51, p = 0.033 , epsilon2 = 0.01 (1 vs 2, z = 2.88, p = 0.040, rbc = 0.19) | H(4) = 11.62, p = 0.020, epsilon2 = 0.01 | 36.63 (12.07) | 12.38 (4.46) | 12.23 (4.34) | 12.02 (4.05) |
| 2. In relationship | 132(20.37) |  |  |  |  | 40.25 (11.79) | 13.62 (4.37) | 13.8 (4.11) | 12.83 (4.1) |
| 3. Married | 297(45.83) |  |  |  |  | 40.15 (11.98) | 13.47 (4.53) | 13.54 (4.12) | 13.14 (4.03) |
| 4. Divorced | 101(15.59) |  |  |  |  | 40.29 (13.52) | 13.69 (4.95) | 13.13 (4.64) | 13.47 (4.52) |
| 5. Widow/Widower | 10(1.54) |  |  |  |  | 45.5 (12.64) | 15.2 (4.66) | 14.8 (4.87) | 15.5 (3.69) |
| Religiosity |  |  |  |  |  |  |  |  |  |
| 1. Yes, I am a member of church | 50(7.72) |  |  |  |  | 40.68 (11.1) | 13.8 (4.44) | 13.56 (3.86) | 13.32 (3.64) |
| 2. Yes, but I am not a member of a church | 131(20.22) |  |  |  |  | 38.59 (12.71) | 12.98 (4.65) | 13.09 (4.53) | 12.52 (4.23) |
| 3. No | 344(53.09) |  |  |  |  | 40.11 (12.2) | 13.54 (4.5) | 13.43 (4.23) | 13.14 (4.14) |
| 4. No, I am convinced atheist | 123(18.98) |  |  |  |  | 39.27 (12.51) | 13.18 (4.74) | 13.21 (4.32) | 12.88 (4.29) |
| Work\_position |  |  |  |  |  |  |  |  |  |
| 1. Worker | 337(52.01) | H(2) = 17.08, p < .001 , epsilon2 = 0.02 (1 vs 2, t = 3.52, p = 0.002, d = -0.29; 1 vs 3, t = 5.32, p < .001, d = -0.42) | H(2) = 16.14, p < .001 , epsilon2 = 0.02 (1 vs 2, t = 1.33, p = 0.001, d = -0.29; 1 vs 3, t = 1.88, p = 0.001, d = -0.4) | H(2) = 20.99, p < .001 , epsilon2 = 0.03 (1 vs 2, t = 1.41, p < .001, d = -0.33; 1 vs 3, t = 1.8, p = 0.001, d = -0.41) | H(2) = 10.88, p = 0.004 , epsilon2 = 0.01 (1 vs 3, t = 1.64, p = 0.001, d = -0.38) | 37.76 (13.12) | 12.67 (4.84) | 12.6 (4.54) | 12.49 (4.44) |
| 2. Professional worker | 227(35.03) |  |  |  |  | 41.28 (10.95) | 14 (4.11) | 14.01 (3.76) | 13.26 (3.81) |
| 3. Chief worker | 84(12.96) |  |  |  |  | 43.08 (10.85) | 14.55 (4.17) | 14.4 (4.04) | 14.13 (3.53) |

*Note.* M = Mean; SD = Standard Deviation; n = count. Group differences were assessed using the Kruskal-Wallis test, reported as the *H* statistic with its degrees of freedom (df), p-value, and epsilon-squared effect size. Significant omnibus tests were followed by pairwise post-hoc comparisons. For groups with unequal variances, the Games-Howell test is reported (t-statistic and Cohen’s *d* effect size). For groups with equal variances, Dunn’s test is reported (z-statistic and Rank-Biserial Correlation, *rbc*, effect size). UWES = Utrecht Work Engagement Scale total score; UWES\_D = Dedication subscale; UWES\_A = Absorption subscale; UWES\_V = Vigor subscale.

### Confirmatory Factor Analysis.

Bartlett test ( (36) = 6,322.72, p < .001) as well as KMO (0.96) revealed that UWES data are sufficiently correlated to perform CFA. Next, a series of CFAs were conducted to determine the optimal factorial structure of the UWES-9 for the Czech sample. Results of these analyses are presented in Table 2. Initial models based on standard one-factor, two-factor, hierarchical, and the original correlated three-factor structures demonstrated poor fit to the data. For instance, the standard three-factor model yielded an RMSEA of 0.09, which was substantially higher than its dynamically-generated Level-0 cutoff of 0.03, indicating a significant degree of misfit. This led to the testing of more complex models that would be theoretically meaningful and, at the same time, account for sources of local misfit. Thus, in the next step, we tested bi-factor models.

Fitting bi-factor models, however, resulted in severe problems with estimation, making the results of these models not possible to interpret. Therefore, in the further step, we tested factor three models that reached the highest model fit in past studies. The first was the three-factor model, firstly proposed by Domínguez-Salas et al. (2022), which included correlated errors between items 1 and 2 (Vigor) and items 8 and 9 (Absorption). The second, a more complex model following Zecca et al. (2015), included these same two correlations between residuals plus a third between items 3 and 4 (Dedication). The third and the most complex was the model of Balducci et al. (2010), who proposed the same correlations between error terms as Domínguez-Salas et al. (2022), with additional correlations between items 2 and 5 (Vigor) and 6 and 8 (Absorption).

It was revealed that all three of these models exhibit the best fit with the data as compared with other tested models. Thus, in the further step, we examined which of these three factor models provided the most optimal fit. The Chi-square difference test comparing these models revealed that the simpler Domínguez-Salas et al. model (diff(2) = 13.82, *p* = p < .001) and model of Zecca et al. (2015) (diff(1) = 0.04, *p* = 0.84) represented a statistically significant loss of fit compared to the more complex model of Balducci et al. (2010). Therefore, the Balducci et al. (2010) model can be considered as the best-fitting in our data.

Further evaluation of this final model (see Figure 2) revealed an excellent values of most of fit indices: (20)=158.853; p < .001; CFI=0.987; TLI=0.976; RMSEA=0.077; SRMR= 0.017. However, its RMSEA was higher than its dynamically generated cutoff of 0.038, it still contains a non-trivial degree of misfit. This suggests that while this three-factor structure with modifications is the best representation of the data, the UWES-9 scale may not perfectly capture the construct of work engagement within the present study sample. Factor loadings () in the Balducci et al. (2010) model were high (ranging from: 0.80 to 0.93) as were correlations between the three factors (see Figure 1). Modification indices did not suggested high change in in case of releasing constrains between UWES items. Correlation between residuals in manifest variables was low: *r* = range(-0.04 - 0.04). Correlation matrix depicting relationships between item residuals can be found in the Supplementary Table 1.

*Figure* 2. Path diagram of the UWES structural equation model with standardized coefficients and residual variances.

Table 2  
*Empirical Fit Indices for Competing CFA Models of the UWES-9, supplemented with a Dynamic Fit Indices (DFI) Analysis of Robustness to Misspecification.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Model 1 One Factor  Schaufeli** |  |  |  |  | **Model 2 Hierarchical  Dom nguez Salas** |  |  |  |  | **Model 3 Three Factor Standard  Schaufeli** |  |  |  |
| Misspecification | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |
| Level-0 | 0.012 | 0.026 | 0.999 |  |  | 0.012 | 0.039 | 0.996 | NONE |  | 0.011 | 0.032 | 0.997 | NONE |
| Level-1 | 0.013 | 0.036 | 0.999 |  |  | NONE | NONE | NONE | 0.177 |  | NONE | NONE | NONE | 0.177 |
| Level-2 | 0.017 | 0.059 | 0.997 |  |  | NONE | NONE | NONE | 0.143 |  | NONE | NONE | NONE | 0.143 |
| Level-3 | 0.022 | 0.09 | 0.994 |  |  |  |  |  |  |  |  |  |  |  |
| Fitted Model | 0.026 | 0.139 | 0.986 |  |  | 0.021 | 0.092 | 0.978 |  |  | 0.021 | 0.092 | 0.978 |  |
|  | **Model 4a Two Factor Vigor Dedication  Willmer** |  |  |  |  | **Model 4b Two Factor Vigor Absorption  Chaudhary** |  |  |  |  | **Model 4c Two Factor Dedication Absorption  Panthee** |  |  |  |
| Misspecification | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |
| Level-0 | 0.012 | 0.033 | 0.997 | NONE |  | 0.012 | 0.034 | 0.997 | NONE |  | 0.012 | 0.035 | 0.996 | NONE |
| Level-1 | NONE | NONE | NONE | 0.174 |  | NONE | NONE | NONE | 0.187 |  | NONE | NONE | NONE | 0.19 |
| Fitted Model | 0.024 | 0.106 | 0.968 |  |  | 0.022 | 0.094 | 0.974 |  |  | 0.024 | 0.107 | 0.967 |  |
|  | **Model 5a Modified 3F  Balducci** |  |  |  |  | **Model 5b Modified 3F  Chaudhary** |  |  |  |  | **Model 5c Modified 3F  Dominguez** |  |  |  |
| Misspecification | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |
| Level-0 | 0.011 | 0.038 | 0.997 | NONE |  | 0.011 | 0.033 | 0.997 | NONE |  | 0.011 | 0.036 | 0.997 | NONE |
| Level-1 | NONE | NONE | NONE | 0.122 |  | NONE | NONE | NONE | 0.178 |  | NONE | NONE | NONE | 0.15 |
| Level-2 |  |  |  |  |  | NONE | NONE | NONE | 0.144 |  | NONE | NONE | NONE | 0.121 |
| Fitted Model | 0.017 | 0.077 | 0.987 |  |  | 0.021 | 0.093 | 0.978 |  |  | 0.017 | 0.076 | 0.986 |  |
|  | **Model 5d Modified 3F  Littman** |  |  |  |  | **Model 5e Modified 3F  Seppala** |  |  |  |  | **Model 5g Modified 3F  Zecca** |  |  |  |
| Misspecification | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |  | SRMR | RMSEA | CFI | Magnitude |
| Level-0 | 0.011 | 0.035 | 0.997 | NONE |  | 0.011 | 0.034 | 0.997 | NONE |  | 0.011 | 0.037 | 0.997 | NONE |
| Level-1 | NONE | NONE | NONE | 0.178 |  | NONE | NONE | NONE | 0.15 |  | NONE | NONE | NONE | 0.15 |
| Level-2 | NONE | NONE | NONE | 0.121 |  | NONE | NONE | NONE | 0.143 |  | NONE | NONE | NONE | 0.121 |
| Fitted Model | 0.018 | 0.077 | 0.985 |  |  | 0.02 | 0.091 | 0.979 |  |  | 0.017 | 0.078 | 0.986 |  |

*Note.* SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index. Bold values indicate the fitted model results. NONE indicates parameter was not estimated for that specification level. Magnitude = reflects the extent of model misspecification of the simulated error at each DFI threshold.

## Item statistic and reliability

Internal consistency of the UWES total score was excellent: Cronbach’s = 0.96 95% CI[0.96 - 0.96] and McDonald’s = 0.96 95% CI[0.96 - 0.96]. When assessing the internal consistency of the UWES subcales, the highest values yielded Dedication subscale: Cronbach’s = 0.93 95% CI[0.92 - 0.94] and McDonald’s = 0.93 95% CI[0.92 - 0.94] followed by the Vigor subscale: Cronbach’s = 0.90 95% CI[0.89 - 0.91] and McDonald’s = 0.90 95% CI[0.89 - 0.91]. The lowest internal consistency was observed in the Absorption factor: Cronbach’s = 0.88 95% CI[0.86 - 0.89] and McDonald’s = 0.88 95% CI[0.86 - 0.89]. Reliability of the UWES was supported by the CR: its values for Vigor (0.86), Dedication (0.91), and Absorption (0.82), were excellent as they all range above the 0.70 threshold for good reliability. However, the omega hierarchical analysis ( = 0.92) revealed that most of the subscale reliability (73-89%) was attributable to the general factor rather than specific dimensions, with group-specific factors contributing only 5-11% of reliable variance. In contrast, the general engagement factor accounted for 90% of reliable variance, providing strong empirical support for using the UWES-9 as a total score. The Table 3 illustrates statistics of UWES items. In general, correlations between these items and item-total correlations were high. The lowest item-total correlation was for item 9.

Table 3  
*Item statistic and Polychoric correlations between the UWES items*

| Items | UWES\_1 | UWES\_2 | UWES\_3 | UWES\_4 | UWES\_5 | UWES\_6 | UWES\_7 | UWES\_8 | UWES\_9 | ITC | Skewness | kurtosis | M(SD) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| UWES\_1 | 1 |  |  |  |  |  |  |  |  | 0.84 | -0.35 | -0.28 | 4.31 (1.48) |
| UWES\_2 | 0.79\*\*\* | 1 |  |  |  |  |  |  |  | 0.81 | -0.43 | -0.34 | 4.56 (1.49) |
| UWES\_3 | 0.75\*\*\* | 0.73\*\*\* | 1 |  |  |  |  |  |  | 0.89 | -0.3 | -0.56 | 4.43 (1.59) |
| UWES\_4 | 0.73\*\*\* | 0.73\*\*\* | 0.84\*\*\* | 1 |  |  |  |  |  | 0.85 | -0.28 | -0.87 | 4.23 (1.71) |
| UWES\_5 | 0.75\*\*\* | 0.71\*\*\* | 0.82\*\*\* | 0.76\*\*\* | 1 |  |  |  |  | 0.84 | -0.22 | -0.75 | 4.11 (1.65) |
| UWES\_6 | 0.75\*\*\* | 0.72\*\*\* | 0.74\*\*\* | 0.7\*\*\* | 0.71\*\*\* | 1 |  |  |  | 0.81 | -0.6 | -0.22 | 4.76 (1.54) |
| UWES\_7 | 0.7\*\*\* | 0.69\*\*\* | 0.82\*\*\* | 0.78\*\*\* | 0.77\*\*\* | 0.7\*\*\* | 1 |  |  | 0.83 | -0.43 | -0.6 | 4.72 (1.66) |
| UWES\_8 | 0.71\*\*\* | 0.71\*\*\* | 0.73\*\*\* | 0.72\*\*\* | 0.68\*\*\* | 0.73\*\*\* | 0.69\*\*\* | 1 |  | 0.82 | -0.55 | -0.44 | 4.6 (1.63) |
| UWES\_9 | 0.66\*\*\* | 0.65\*\*\* | 0.72\*\*\* | 0.73\*\*\* | 0.68\*\*\* | 0.67\*\*\* | 0.7\*\*\* | 0.72\*\*\* | 1 | 0.78 | -0.14 | -0.82 | 3.96 (1.66) |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001, M = Mean, SD = Standard Deviation, ITC = Item-total correlation corrected for scale reliability and item overlap

### Construct Validity Evaluation.

With the best-fitting model established, we conducted an evaluation of its construct validity. We first assessed convergent validity by calculating the AVE for each factor. The results for Vigor (0.73), Dedication (0.82), and Absorption (0.67) all exceeded the 0.50 treshold, indicating that for each factor, a majority of the variance in its items was captured by the construct itself rather than by measurement error.

However, given the high inter-factor correlations (see Figure 2), a specific test of discriminant validity was necessary to determine if the factors were empirically distinct. Thus, we employed the HTMT for this purpose. The analysis yielded values of 0.95 (Vigor-Dedication), 0.96 (Vigor-Absorption), and 0.94 (Dedication-Absorption). As all values substantially exceeded the 0.90 threshold, this provided definitive evidence that the factors, while structurally sound, are not internally distinct from one another.

This finding led to our final validity test, where we examined the scale’s relationships with external criteria to see if the subscales behaved differently despite their internal overlap. As shown in Table 4, the UWES scores behaved as expected: the total score and all subscales were positively associated with extraversion and self-efficacy, and negatively associated with neuroticism. For divergent validity, no significant correlation was found with spirituality, with the exception of the Dedication subscale.

To formally test whether the magnitude of these correlations differed across subscales, we used the z-test and calculated Cohen’s q to quantify the effect size of the differences. The analysis revealed specific differences in the subscales’ external relationships. For example, the correlation between Vigor and extraversion was significantly stronger than that of Absorption and extraversion, representing a small effect (q = 0.13, p < .001). In contrast, the difference in their correlations with neuroticism was not always statistically significant; the difference between Vigor and Dedication was negligible (q = 0.03, p = 0.132), while the difference between Vigor and Absorption was significant with a small effect (q = 0.11, p < .001). Taken together, despite a high overlap between UWES-9 dimensions, subscales of the UWES exhibit some statistically distinct relationships with external criteria.

Correlation analysis indicated that there is a significant positive association between all UWES subscales and total score and extraversion. The highest correlation was found in the Vigor subscale. In addition, there was a significant negative correlation between all UWES subscales and total score with neuroticism. The highest association was also found in the Vigor subscale. Moreover, the UWES total and all its subscales were associated with self-efficacy. The strongest association was observed in the Vigor subscale. Finally, there was no correlation between the UWES composite and subcale score with spirituality, with the exception of the Dedication subscale (see Table 4).

Table 4  
*Correaltion matrix of the UWES, personality characteristics and socio-demographic indicators*

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | M(SD) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. UWES | - |  |  |  |  |  |  |  |  | 39.69 (12.27) |
| 2. UWES\_V | .94\*\*\* | - |  |  |  |  |  |  |  | 12.98 (4.15) |
| 3. UWES\_D | .95\*\*\* | .85\*\*\* | - |  |  |  |  |  |  | 13.38 (4.57) |
| 4. UWES\_A | .93\*\*\* | .82\*\*\* | .83\*\*\* | - |  |  |  |  |  | 13.33 (4.28) |
| 5. BFI\_E | .19\*\*\* | .23\*\*\* | .18\*\*\* | .13\*\*\* | - |  |  |  |  | 24.20 (5.21) |
| 6. BFI\_N | -.19\*\*\* | -.23\*\*\* | -.18\*\*\* | -.12\*\* | -.27\*\*\* | - |  |  |  | 23.02 (5.70) |
| 7. Age | .03 | .06 | .01 | .02 | -.01 | -.10\*\* | - |  |  | 43.65 (10.08) |
| 8. Gender | .07 | .08 | .05 | .07 | .06 | .20\*\*\* | .08\* | - |  | 1.38 (0.49) |
| 9. DSES | .13 | .09 | .17\* | .11 | .09 | -.06 | -.02 | .10 | - | 2.39 (1.10) |
| 10. GSES | .28\*\*\* | .30\*\*\* | .26\*\*\* | .25\*\*\* | .31\*\*\* | -.44\*\*\* | .08\* | -.09\* | .13 | 28.43 (4.95) |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001; SD = standard deviation, M = mean, UWES = Utrecht Work Engagement Scale, BFI\_N = Big Five Inventory - Neuroticism subscale, BFI\_E = Big Five Inventory - Extraversion subscale, UWES\_A = Utrecht Work Engagement Scale - Absorption subscale, UWES\_D = Utrecht Work Engagement Scale - Dedication subscale, UWES\_V = Utrecht Work Engagement Scale - Vigor subscale. DSES = Daily Spiritual Experience Scale, GSES = General Self Efficacy Scale

## Invariance testing and factor loadings

The results of the measurement invariance testing are summarized in Table 5). In this analysis, we report the scaled fit indices, as the robust versions could not be computed for the scalar and strict invariance models. It was revealed that in female group, factor intercorrelations exceeded 1.0 (Vigor-Absorption r=1.02). To ensure model stability, correlated residuals from the original Italian validation were removed, as they caused identification issues in the smaller female subsample. After this removal, the model included correlated errors only between items 1 and 2 (Vigor) and items 8 and 9 (Absorption). After this modification the model was the same as the Domínguez-Salas et al. (2022) model. In this simplified three-factor model, the change in the scaled Comparative Fit Index (CFI) was less than 0.01 across all model comparisons. This finding strongly supports full measurement invariance and indicates that the UWES assesses work engagement equivalently in males and females.

Table 5  
*Measurement equivalence of the UWES between genders*

| Model | x2 | df | p-value | CFI | TLI | RMSEA | SRMR |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Baseline | 168.916 | 22 | p < .001 | 0.994 | 0.99 | 0.102 90% CI (0.088-0.116) | 0.017 |
| Configural | 216.885 | 44 | p < .001 | 0.993 | 0.988 | 0.11 90% CI (0.096-0.125) | 0.02 |
| Metric | 174.424 | 50 | p < .001 | 0.995 | 0.993 | 0.088 90% CI (0.074-0.102) | 0.022 |
| Scalar | 187.198 | 92 | p < .001 | 0.996 | 0.997 | 0.057 90% CI (0.045-0.068) | 0.021 |
| Strict | 187.198 | 92 | p < .001 | 0.996 | 0.997 | 0.057 90% CI (0.045-0.068) | 0.021 |

*Note.* x2 = chi-square, df = degrees of freedom, CFI = Comparative Fit Index, TLI = Tucker-Lewis index, RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square Residual, CI = confidence interval

## Association of the UWES with chronic health ilnesses

Results of the regression analysis revealed that work engagement is significantly related with chronic diseases. Specifically, higher work engagement was significantly related to a lower probability of developing skin diseases or eczema (in crude effect), pain of unclear origin (both crude and adjusted effect (see Table 6).

Table 6  
*Logistic regression table depicting associations (in odds ratios) between the UWES and chronic diseases*

|  | Skin diseases eczema | Pain of unclear origin | Hypertension | Diabetes | Arthritis |
| --- | --- | --- | --- | --- | --- |
| Crude effect | 0.98\* (0.95, 1.00) | 0.93\*\* (0.89, 0.97) | 1.01 (0.99, 1.03) | 1.00 (0.98, 1.03) | 0.97 (0.95, 1.01) |
| Adjusted effect | 0.98 (0.96, 1.00) | 0.94\*\* (0.90, 0.98) | 1.01 (0.99, 1.03) | 1.01 (0.98, 1.03) | 0.98 (0.95, 1.01) |
|  | Depression/Anxiety | Migraine | Cancer | Thyroid disease | Astma |
| Crude effect | 0.99 (0.96, 1.02) | 1.00 (0.97, 1.03) | 1.00 (0.95, 1.07) | 1.01 (0.99, 1.04) | 0.98 (0.96, 1.00) |
| Adjusted effect | 1.00 (0.97, 1.02) | 1.00 (0.97, 1.04) | 1.00 (0.94, 1.07) | 1.02 (1.00, 1.05) | 0.98 (0.96, 1.01) |
|  | Gastric or duodenal ulcers | Chronic lung disease | Skin diseases eczema | Allergy | Pain in the small pelvis |
| Crude effect | 1.01 (0.95, 1.10) | 0.97 (0.92, 1.02) | 0.98\* (0.95, 1.00) | 0.99 (0.97, 1.01) | 1.00 (0.96, 1.05) |
| Adjusted effect | 1.01 (0.94, 1.10) | 0.97 (0.93, 1.02) | 0.98 (0.96, 1.00) | 0.99 (0.97, 1.01) | 1.01 (0.97, 1.05) |
|  | Ischemic heart disease | Obesity | Stroke | Back pain |  |
| Crude effect | 1.00 (0.93, 1.08) | 0.99 (0.97, 1.01) | 0.95 (0.87, 1.04) | 0.99 (0.97, 1.00) |  |
| Adjusted effect | 0.99 (0.92, 1.07) | 0.99 (0.97, 1.01) | 0.95 (0.86, 1.04) | 0.99 (0.98, 1.01) |  |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001, results are reported in odds ratios; Education and Work position were covariates in adjusted effect; values in brackets refers to 95% confidence interval for odds ratios

# Association of the UWES with health risk behaviour

Results of logistic regression suggested that there is no relationship between work engagement and smoking, alcohol drinking, drug abuse, coffee drinking, or using a computer or television for recreation in both crude and adjusted effects (Supplementary Table 2). Variable smoking was the closest to the significance threshold.

# Discussion

# References

Chan, E. S. S., Ho, S. K., Ip, F. F. L., & Wong, M. W. Y. (2020). Self-efficacy, work engagement, and job satisfaction among teaching assistants in hong kong’s inclusive education. *SAGE Open*, *10*(3), 2158244020941008. <https://doi.org/10.1177/2158244020941008>

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

# (Supplementary Table 1)

Table 7  
*Lower-triangular correlations between residuals of the final factor solution*

|  | UWES\_1 | UWES\_2 | UWES\_5 | UWES\_3 | UWES\_4 | UWES\_7 | UWES\_6 | UWES\_8 | UWES\_9 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| UWES\_2 | 0 |  |  |  |  |  |  |  |  |
| UWES\_5 | 0.01 | -0.01 |  |  |  |  |  |  |  |
| UWES\_3 | -0.01 | -0.02 | 0.03 |  |  |  |  |  |  |
| UWES\_4 | -0.01 | 0.01 | -0.01 | 0 |  |  |  |  |  |
| UWES\_7 | -0.02 | -0.01 | 0.02 | 0 | -0.01 |  |  |  |  |
| UWES\_6 | 0.04 | 0.03 | -0.03 | -0.02 | -0.02 | -0.01 |  |  |  |
| UWES\_8 | 0.01 | 0.02 | -0.05 | -0.01 | 0.01 | -0.01 | 0.01 |  |  |
| UWES\_9 | -0.01 | -0.02 | -0.02 | 0 | 0.03 | 0.02 | -0.02 | 0 |  |

*Note.* UWES = Utrecht Work Engagement Scale

# (Supplementary Table 2)

Table 8  
*Logistic regression table depicting associations (in odds ratios) between the UWES and health risk behaviours*

|  | Smoked | Drunk alcohol | Used illegal drugs | Drunk coffee | Used television or computer for recreation |  |
| --- | --- | --- | --- | --- | --- | --- |
| Crude effect | 1.00 (0.99, 1.02) | 1.00 (0.98, 1.01) | 0.97 (0.92, 1.02) | 1.01 (1.0, 1.02) | 1.01 (0.99, 1.03) |  |
| Adjusted effect | 1.01 (1.00, 1.03) | 1.00 (0.98, 1.01) | 0.98 (0.93, 1.04) | 1.01 (1.00, 1.03) | 1.01 (0.99, 1.03) |  |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001, results are reported in odds ratios; Education and Work position were covariates in adjusted effect; values in brackets refers to 95% confidence interval for odds ratios