Routledge
Taylor & Francis Group

Work & Stress

An International Journal of Work, Health & Organisations

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/twst20

Measurement invariance of the Occupational Depression Inventory: a study of 12,589 participants across 14 countries

Renzo Bianchi, Irvin Sam Schonfeld, James F. Sowden, Danísio C. Cavalcante, Cristina Queirós, Vera M. Hebel, Judith Volmer, Caterina Fiorilli, Giacomo Angelini, Krystyna Golonka, Guadalupe Manzano-García, Pilar Montañés-Muro, Markus Jansson-Fröjmark & Leon T. De Beer

To cite this article: Renzo Bianchi, Irvin Sam Schonfeld, James F. Sowden, Danísio C. Cavalcante, Cristina Queirós, Vera M. Hebel, Judith Volmer, Caterina Fiorilli, Giacomo Angelini, Krystyna Golonka, Guadalupe Manzano-García, Pilar Montañés-Muro, Markus Jansson-Fröjmark & Leon T. De Beer (07 Jun 2024): Measurement invariance of the Occupational Depression Inventory: a study of 12,589 participants across 14 countries, Work & Stress, DOI: 10.1080/02678373.2024.2364594

To link to this article: https://doi.org/10.1080/02678373.2024.2364594

WORK & STRESS
https://doi.org/10.1080/02678373.2024.2364594

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Measurement invariance of the Occupational Depression Inventory: a study of 12,589 participants across 14 countries

Renzo Bianchi [a], Irvin Sam Schonfeld[b], James F. Sowden[c], Danísio C. Cavalcante[d,e], Cristina Queirós[d], Vera M. Hebel[f], Judith Volmer[f], Caterina Fiorilli[g], Giacomo Angelini[g], Krystyna Golonka[h], Guadalupe Manzano-García[i], Pilar Montañés-Muro[i], Markus Jansson-Fröjmark[j] and Leon T. De Beer[a]

[a]Department of Psychology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway; [b]City University of New York, New York, NY, USA; [c]Flinders University, Adelaide, Australia; [d]University of Porto, Porto, Portugal; [e]National Institute of Social Security, Fortaleza-Ceará, Brazil; [f]University of Bamberg, Bamberg, Germany; [g]LUMSA University, Rome, Italy; [h]Jagiellonian University, Kraków, Poland; [i]University of La Rioja, Logroño, Spain; [j]Karolinska Institute, Stockholm, Sweden

ABSTRACT

The Occupational Depression Inventory (ODI) reflects a novel approach to job-related distress anchored in depression research. To date, the extent to which the ODI exhibits measurement invariance across countries, languages, and demographics is unclear. Measurement invariance refers to whether a measure has the same structure, or meaning, across groups of interest. Measurement invariance is thus crucial for between-group comparisons and study replicability. This study estimated the measurement invariance of the ODI across 14 countries – Australia, Brazil, France, Germany, Italy, New Zealand, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, and the USA – and 10 languages as well as across sexes and age groups (pooled $N = 12{,}589$). We found evidence for complete measurement invariance (configural, weak, strong, and strict) across countries, languages, sexes, and age groups. Looking into the invariance of structural parameters, we found latent variance-covariance invariance to hold across countries, languages, and sexes and to be equivocal across age groups. Expectedly, the levels of occupational depression, as indexed by latent means, varied within the four categories. Our results indicate that the ODI behaves similarly across countries, languages, sexes, and age groups. Our findings support the use of the ODI with respondents having different cultural backgrounds and individual characteristics.

ARTICLE HISTORY
Received 31 August 2023
Accepted 15 April 2024

KEYWORDS
Bifactor analysis; cultural differences; job stress; measurement; occupational health

The Occupational Depression Inventory (ODI) was recently developed to more efficiently assess individual differences in job-related distress (Bianchi & Schonfeld, 2020;

CONTACT Renzo Bianchi renzo.bianchi@ntnu.no
Supplemental data for this article can be accessed http://doi.org/10.1080/02678373.2024.2364594.

Bianchi, Verkuilen, et al., 2023; Schonfeld & Bianchi, 2022). The instrument is designed to (a) quantify work-attributed depressive symptoms (dimensional approach) and (b) screen for occupational depression using a clinically informed algorithm (categorical approach). The ODI was devised with reference to the nine core diagnostic symptoms of major depression found in the fifth edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-5*; American Psychiatric Association, 2013).[1] The measure thus assesses anhedonia, depressed mood, sleep alterations, fatigue/loss of energy, appetite alterations, feelings of worthlessness, cognitive impairment, psychomotor alterations, and suicidal ideation. The ODI incorporates causal attributions to the respondent's job, which differentiates the instrument from classical, "cause-neutral" depression scales – scales such as the Beck Depression Inventory, the Centre for Epidemiologic Studies Depression scale, and the PHQ-9, to name just a few. Causal attributions have been commonly employed in stress research, for instance, to identify sources of stress among the general public (American Psychological Association, 2015). Causal attributions are also central to diagnostic categories such as acute stress disorder and posttraumatic stress disorder (American Psychiatric Association, 2013).

The ODI was created in response to limitations in how job-related distress has been conceptualised and measured (Nadon et al., 2022; Rotenstein et al., 2018; Schwenk & Gold, 2018). Indicators such as burnout, for instance, have raised considerable controversy over the years (e.g. Bianchi & Schonfeld, 2023; Sen, 2022; Taris, 2006). Anchored in depression research, the ODI benefits from solid clinical and theoretical foundations. The instrument assesses crucially informative symptoms, such as work-related suicidal thoughts, that are generally ignored by measures of job-related distress (Bianchi & Sowden, 2022). Its dual-lens approach to job-related distress (dimensional/categorical) makes it a flexible tool and offers the possibility of producing meaningful prevalence estimates. In addition, the ODI is brief, straightforward in its use, and available to researchers and practitioners at no cost.

Available evidence suggests that the ODI has robust psychometric and structural properties. Despite covering nine different symptoms, the instrument has consistently exhibited essential unidimensionality[2] as well as strong total-score reliability (e.g. Bianchi, Verkuilen, et al., 2023; Hill et al., 2021). Moreover, the measure has been found to meet Mokken scaling requirements (Sijtsma & van der Ark, 2017), including scalability, monotonicity, local independence, and invariant item ordering (e.g. Bianchi, Manzano-García, et al., 2022). The ODI has shown a balance of convergent and discriminant validity vis-à-vis a variety of cause-neutral depression scales (e.g. Bianchi & Schonfeld, 2020; Hill et al., 2021; Jansson-Fröjmark et al., 2023). Such a balance is expected given that, in contrast to classical depression scales that have no etiological foci, the ODI is meant to assess *work-attributed* depressive symptoms (Bianchi, Verkuilen, et al., 2023). Regarding its criterion validity, the ODI has been associated with various work and nonwork variables such as workplace violence, work engagement, sick leave, antidepressant intake, financial strain, and objective (task-based) cognitive performance (e.g. Bianchi & Schonfeld, 2021, 2022; Bianchi, Fiorilli, et al., 2022; Hill et al., 2021). The ODI was recently used in computational science to develop a deep-learning framework able to rate online job-related content in terms of an occupational depression score. The study, which involved over 350,000 employee reviews and 100 US companies, found occupational depression to be (a) negatively linked to companies' stock growth and (b) positively linked to states' economic deprivation (Sen et al., 2022). While the ODI has been validated in countries on different continents, the issue of whether the ODI has the same structure, or meaning, across languages and countries remains unclear. In a similar vein, the consistency of the instrument's characteristics across demographic categories requires further attention.

Establishing whether a measure has the same structure, or meaning, across groups of interest is crucial to estimating between-group comparability (Maassen et al., 2023; Meuleman et al., 2023). Without measurement invariance, observed differences between groups can be due to the measure functioning differently across groups, rather than reflecting *genuine* group-related differences in the construct being measured. Let us take the example of an anxiety scale administered to male and female respondents. Let us assume that the female group scores higher on anxiety than the male group. If the scale does not exhibit measurement invariance across sexes, then the differences in scores could result from differences in how males and females interpret or respond to the scale rather than actual differences in the levels of anxiety. Such muddled findings could then lead, for instance, to incorrect diagnoses and misguided interventions. Far from being a concern reserved for psychometrics purists, measurement invariance is of critical importance for theory-building and knowledge accumulation (Hofmans et al., 2009; Morin, 2023). Unnoticed violations of measurement invariance likely play a role in the widespread failure to replicate findings across samples and settings in psychological research (Flake et al., 2022; Maassen et al., 2023).

The present study estimated the measurement invariance of the ODI across 14 countries – Australia, Brazil, France, Germany, Italy, New Zealand, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, and the USA – and 10 languages – English, Brazilian–Portuguese, French, German, Italian, Norwegian, Polish, Portuguese, Spanish, and Swedish – as well as across sexes and age groups. Measurement invariance analysis allows investigators to ascertain whether a measure behaves similarly across categories of interest (Millsap, 2011; Morin, 2023). We focused on the measurement invariance of the bifactor structure found to best characterise the ODI in past research (e.g. Bianchi, Verkuilen, et al., 2023). A bifactor model partitions the total covariance among a scale's items into a general factor underlying all items and specific factors explaining additional covariance not captured by the general factor.

We conducted a stringent examination by sequentially investigating configural, weak, strong, and strict forms of measurement invariance. Additionally, we inquired into structural parameters by focusing on latent variance-covariance invariance and latent mean invariance. We examined omega reliabilities from the most constrained models in which invariance held. Cortina et al. (2020) noted that "[t]he distance between actual and recommended scale development and evaluation practices may have reached a magnitude that should lead us to question our conclusions regarding organizational phenomena … " (p. 1352). The importance of correcting this trend is difficult to overstate. The robustness of the knowledge that we generate is contingent upon the quality of our measures. Some scales are extensively used for many years before one realises that their psychometric properties are, in fact, inadequate (Hussey & Hughes, 2020). Such unnoticed shortcomings can threaten the validity of vast segments of research, not to mention the waste of resources involved. It is thus pivotal that recently developed instruments – such as the ODI – undergo a thorough examination.

Methods

Study samples

We relied on 14 different samples of employed individuals (pooled $N = 12{,}589$; 69.7% female) enrolled in studies managed by our consortium of researchers. Characteristics of the samples are displayed in Table 1. Convenience sampling was employed in all countries except Germany and the USA, in which quota sampling was used. Respondi/Bilendi recruited the German sample. Prolific Academic recruited the US sample. Respondi/Bilendi and Prolific Academic are trusted international panel providers (Munzert et al., 2021; Palan & Schitter, 2018; Peer et al., 2022). As previously noted, the samples were recruited in 14 different countries and involved 10 different languages. The countries are located on four different continents. The inclusion of participants living on four different continents was likely an asset in terms of sample diversity. There is evidence that the lifetime prevalence of depressive disorders differs greatly across the countries under scrutiny. As an illustration, Kessler and Bromet (2013)

Table 1. Samples Under Examination.

Sample	Country	*N*	Language	% female	Age (*M*, *SD*)	Occupation	ODI score (*M*, *SD*)	Study
1	Australia	1,485	English	90.7	40, 10	Educators	1.544, 0.708	Bianchi, Verkuilen, et al. (2023)
2	Brazil	1,612	Brazilian-Portuguese	59.5	44, 9	Civil servants	1.060, 0.768	Bianchi, Cavalcante, et al. (2023a)
3	France	1,450	French	84.3	44, 10	Educators	0.976, 0.730	Bianchi and Schonfeld (2020)
4	Germany	1,000	German	50.0	46, 11	Mixed	0.488, 0.581	Unpublished
5	Italy	963	Italian	69.9	40, 11	Mixed	0.649, 0.538	Bianchi, Fiorilli, et al. (2022)
6	New Zealand	492	English	79.9	47, 12	Educators	1.076, 0.746	Bianchi and Schonfeld (2020)
7	Norway	838	Norwegian	61.1	42, 13	Mixed	0.609, 0.591	Unpublished
8	Poland	526	Polish	46.8	40, 10	Mixed	0.768, 0.734	Golonka et al. (2024)
9	Portugal	708	Portuguese	66.9	39, 10	Mixed	0.782, 0.635	Unpublished
10	South Africa	327	English	59.9	N/A*	Mixed	0.886, 0.775	Hill et al. (2021)
11	Spain	386	Spanish	70.7	46, 9	Mixed	0.719, 0.683	Bianchi, Manzano-García, et al. (2022)
12	Sweden	365	Swedish	88.2	43, 11	Mixed	1.197, 0.818	Jansson-Fröjmark et al. (2023)
13	Switzerland	1,971	French	71.6	36, 12	Mixed	0.703, 0.599	Bianchi, Verkuilen, et al. (2023)
14	USA	466	English	51.5	45, 16	Mixed	0.458, 0.581	Unpublished

Notes. *SD*: standard deviation; ODI: Occupational Depression Inventory. *: in the South African sample, age was assessed using five categories.

found major depression to affect about one in five individuals in countries such as the USA and France, and only one in ten individuals in countries such as Germany and Italy. Readers willing to learn more about the cultural differences across the countries under consideration may want to use the Country Comparison Tool.[3] The tool focuses on between-country cultural differences through the prism of six dimensions: power distance; individualism; motivation towards achievement and success; uncertainty avoidance; long-term orientation; and indulgence. All samples used in the present study were recruited in the context of online surveys conducted in accordance with the ethical standards of the main investigators' home institution.

Measure of interest

Our measure of interest was the ODI (Bianchi & Schonfeld, 2020). The ODI comprises nine core items rated on a 4-point frequency scale (from 0 for "never or almost never" to 3 for "nearly every day"). Consistent with *DSM-5*'s diagnostic criteria for major depression (American Psychiatric Association, 2013), respondents are asked to report on symptoms experienced over the past two weeks. A sample ODI item is: "My experience at work made me feel like a failure." The ODI is accompanied by instructions to respondents that play an important role in the scale's administration. Respondents are invited to consider various sources for their symptoms. If a respondent attributes a symptom to a problem unrelated to work (personal problems, marital problems, family problems, health problems, etc.) or to a source he or she cannot identify, the respondent is asked to select "0" when answering. This precaution is intended to discourage hasty attributions of symptoms to work. Responses to the ODI involved no missing values. The ODI can be used free of charge and can be found in most ODI-related articles, including Bianchi and Schonfeld's (2020). The content of the instrument is displayed in Supplemental Material 1. Characteristics of the ODI in the pooled sample are available in Table 2. All translated versions of the ODI were generated based on the original English version using back-translation procedures (Streiner et al., 2015). The translation process of the instrument was *not* accompanied by an "adaptation" process involving more profound modifications of the scale's content (Geisinger, 1994). There are at least three justifications for this choice. First, the original (English) version of the ODI was devised with attention to content simplicity and phrasing clarity (e.g. by avoiding jargon and using words that virtually all respondents can be expected to understand; Bradburn et al., 2004). Second, the symptoms targeted by the ODI (e.g. depressed mood, fatigue/loss of energy, cognitive impairment) are assumed to be species-level, not culture-specific, phenomena. Third, the reference frame of the ODI is expected to be shared across all the groups represented in the study samples. The references to "work" and people's "job," for instance, are shared across the countries sampled in the study.

Data analyses

We analyzed the data in Mplus 8.9 (Muthén & Muthén, 1998-2023). We examined measurement invariance within an exploratory structural equation modelling (ESEM) bifactor analytic framework (Marsh et al., 2014; Rodriguez et al., 2016). We used an

Table 2. Characteristics of the Occupational Depression Inventory in the Pooled Sample.

	ODI1	ODI2	ODI3	ODI4	ODI5	ODI6	ODI7	ODI8	ODI9	OD
Mean	1.062	0.887	1.091	1.453	0.877	0.725	0.822	0.838	0.204	0.884
Median	1.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	0.000	0.667
Mode	1	0	0	1	0	0	0	0	0	0.000
Interquartile range	2	1	2	1	2	1	1	1	0	1.111
Standard deviation	0.970	0.941	1.024	1.019	1.032	0.930	0.929	0.952	0.580	0.741
Skewness ($SE = 0.022$)	0.567	0.794	0.534	0.140	0.821	1.100	0.876	0.860	3.187	0.791
Kurtosis ($SE = 0.044$)	−0.677	−0.353	−0.878	−1.096	−0.630	0.163	−0.237	−0.335	10.157	−0.215
Loevinger's H*	0.689	0.668	0.663	0.694	0.639	0.632	0.695	0.672	0.610	0.666
Minimum	0	0	0	0	0	0	0	0	0	0.000
Maximum	3	3	3	3	3	3	3	3	3	3.000
N	12,589	12,589	12,588	12,589	12,589	12,589	12,588	12,589	12,588	12,589

Notes. * Derived from Mokken scale analysis (all SEs ≤ 0.008). SE = standard error; ODI1: anhedonia; ODI2: depressed mood; ODI3: sleep alterations; ODI4: fatigue/loss of energy; ODI5: appetite alterations; ODI6: feelings of worthlessness; ODI7: cognitive impairment; ODI8: psychomotor alterations; ODI9: suicidal ideation; OD: occupational depression.

invariance syntax generator tool for ESEM bifactor analysis in Mplus that limits tedious coding and reduces the risk of human error (De Beer & Morin, 2022). Our ESEM bifactor analysis involved two specific factors in addition to the general Occupational Depression factor. Two specific factors (or group factors) were extracted because the ODI assesses "anhedonic-somatic" and "dysphoric" symptoms. This bifactor structure has been found to best characterise the ODI in past research (e.g. Bianchi & Schonfeld, 2020; Bianchi, Fiorilli, et al., 2022; Hill et al., 2021). We treated the ODI items as ordinal, used the weighted least squares – mean and variance adjusted – estimator, and relied on a target rotation. The Anhedonic-Somatic specific factor involved items 1 (anhedonia), 3 (sleep alterations), 4 (fatigue/loss of energy), 5 (appetite alterations), 7 (cognitive impairment), and 8 (psychomotor alterations); the Dysphoric specific factor involved items 2 (depressed mood), 6 (feelings of worthlessness), and 9 (suicidal ideation). The model is represented in Figure 1. Using a target rotation, ESEM bifactor analysis allows the investigator to specify a model a priori (i.e. to adopt a confirmatory approach) without endorsing the somewhat unrealistic zero cross-loading assumption linked to common-practice confirmatory factor analysis (Marsh et al., 2014). As recommended, we focused primarily on (changes in) the Comparative Fit Index (CFI), and secondarily on (changes in) the Root Mean Square Error of Approximation (RMSEA) and the Tucker-Lewis Index (TLI), to estimate measurement invariance (Houle et al., 2022; Lohbeck et al., 2022; Maïano et al., 2022).

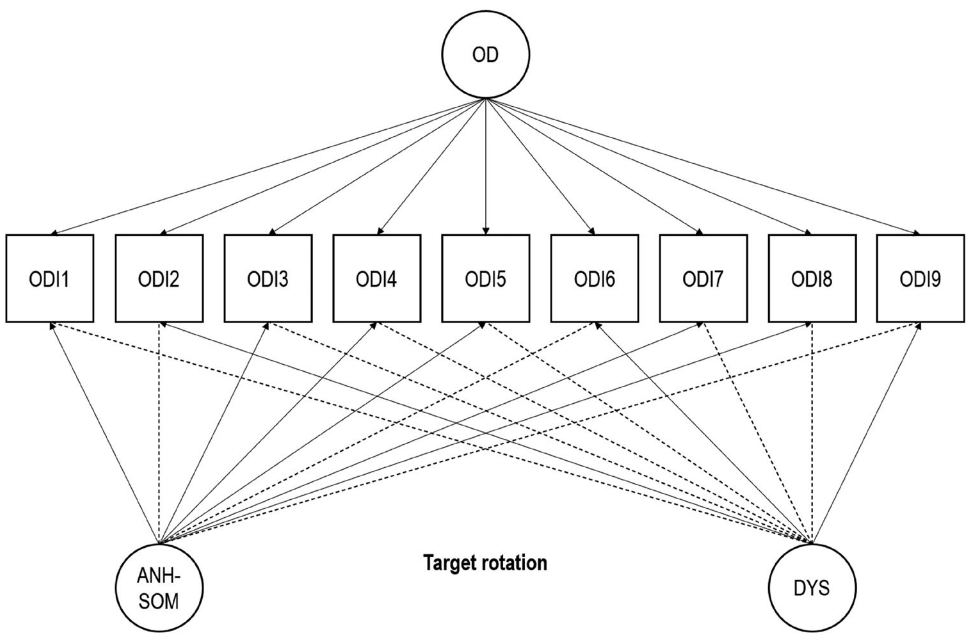

Figure 1. Exploratory structural equation modelling bifactor structure under examination. The solid lines indicate target loadings. OD: general Occupational Depression factor; ANH-SOM: Anhedonic-Somatic specific factor; DYS: Dysphoric specific factor; ODI1: anhedonia; ODI2: depressed mood; ODI3: sleep alterations; ODI4: fatigue/loss of energy; ODI5: appetite alterations; ODI6: feelings of worthlessness; ODI7: cognitive impairment; ODI8: psychomotor alterations; ODI9: suicidal ideation.

We investigated measurement invariance across languages and countries (cultural backgrounds) as well as sexes and age groups (individual characteristics). Sexes involved two categories (male/female) and age groups, three categories (18-34 [early career]; 35–49 [mid-career]; and 50+ [late career]). Our primary focus was on configural, weak, strong, and strict forms of invariance (Morin et al., 2020). The invariance constraints are cumulative. Configural invariance refers to equivalence in factor structures. Weak invariance adds an equivalence in factor loadings. Strong invariance adds an equivalence in item thresholds (i.e. in how the response scale is used). Strict invariance adds an equivalence in item residuals (i.e. in uniquenesses). The implications of strong invariance and strict invariance are worth spelling out further. The absence of strong invariance indicates that the groups (e.g. males and females) use the response scale differently. In other words, item scores can differ across the groups irrespective of differences in the latent variable. Strong invariance is a prerequisite to between-group comparisons involving *latent means*. The absence of strict invariance indicates that the reliability (i.e. the measurement error) with which the construct is assessed differs across the groups. Strict invariance is a prerequisite to between-group comparisons involving *observed scores*. Where strict invariance was reached, we inquired into the invariance of structural parameters, with a focus on latent variance-covariance invariance and latent mean invariance (Morin, 2023). Latent variance-covariance invariance considers equivalence at the level of factor variance-covariance matrices (i.e. between-factor correlations), and latent mean invariance considers equivalence at the level of factor means. We applied commonplace standards for identifying deviations from measurement invariance. Regarding RMSEA, invariance violations were signalled by increases exceeding .015; regarding CFI and TLI, invariance violations were signalled by decreases exceeding .010 (Chen, 2007; Houle et al., 2022; Maïano et al., 2022).

We computed omega coefficients (McDonald, 1999) related to the general (omega$_{\mathrm{OD}}$) and specific (omega$_{\mathrm{ANH\text{-}SOM}}$ and omega$_{\mathrm{DYS}}$) factors based on our ESEM bifactor analytic outputs (Morin, 2023; Morin et al., 2020). Omega coefficients were calculated from the most constrained models in which invariance held. For instance, if invariance were to be observed from configural to strict, the strict model would serve the calculation of omegas. Omegas are computed from factor analytical outputs and constitute recommended reliability indicators (Cho & Kim, 2015; Cortina et al., 2020).

Results

The latent means, together with their 95% confidence intervals, are available in Supplemental Material 1. A visual illustration is displayed in Figure 2 for latent means across countries. Visual illustrations for latent means across languages, sexes, and age groups are provided in Supplemental Material 1. Supplemental Material 1 also contains the standardised factor loadings and uniquenesses related to the most constrained models in which invariance was observed.

Measurement invariance across countries

The results of our measurement invariance analysis across countries are summarised in Table 3. Regarding configural, weak, strong, and strict forms of invariance, CFI never

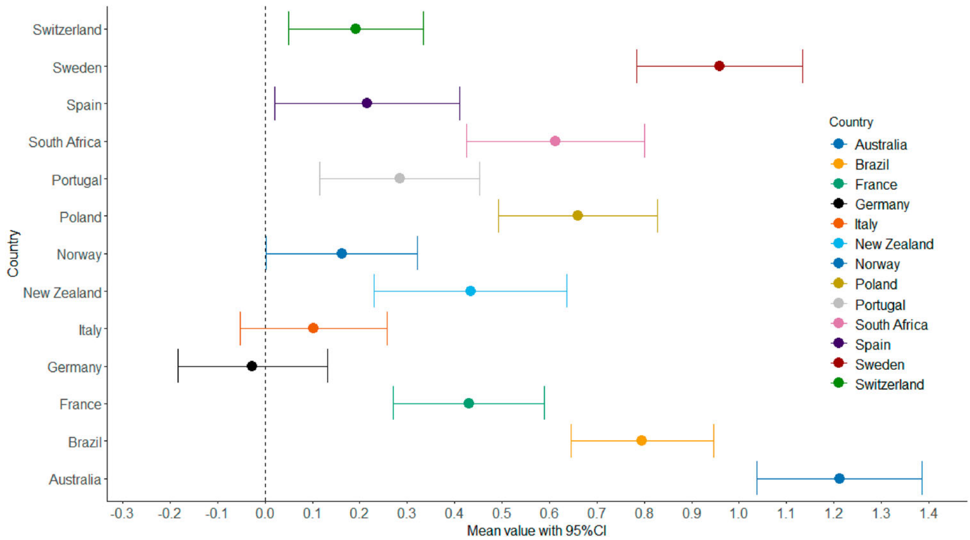

Figure 2. Latent means related to countries – the dashed vertical line represents the reference point (the USA).

decreased by more than .004, RMSEA never increased by more than .002, and TLI did not decrease at all. Strict invariance was thus reached. Latent variance-covariance invariance was observed as well, with CFI decreasing by only .002, RMSEA increasing by only .001, and TLI remaining virtually identical. Latent means were not invariant (ΔCFI = -.047; ΔRMSEA = .065; ΔTLI = -.028).

Measurement invariance across languages

Measurement invariance analysis across languages is summarised in Table 3. Regarding configural, weak, strong, and strict forms of invariance, CFI never decreased by more than .003, RMSEA never increased by more than .004, and TLI never decreased by more than .001. Again, strict invariance was attained. Latent variance-covariance invariance was supported as well with (a) CFI decreasing by only .001, (b) no increase in RMSEA, and (c) no decrease in TLI. By contrast, requirements for latent mean invariance were not met (ΔCFI = -.024; ΔRMSEA = .047; ΔTLI = -.015).

Measurement invariance across sexes

The results of our measurement invariance analysis across sexes are summarised in Table 4. Regarding configural, weak, strong, and strict forms of invariance, CFI never decreased by more than .001, RMSEA never increased, and TLI did not decrease at all. Strict invariance was thus achieved. Latent variance-covariance invariance was supported as well, with a nondecreasing CFI, a nonincreasing RMSEA, and a nondecreasing TLI. Despite a ΔCFI of – .009 and a ΔTLI of – .008, latent means could not be regarded as invariant with a ΔRMSEA as large as .048.

Table 3. Measurement Invariance Across Countries and Languages.

Invariance model	χ^2 (*df*)	CFI	TLI	RMSEA [90% CI]	CM	$\Delta\chi^2$	Δ*df*	ΔCFI	ΔTLI	ΔRMSEA
Countries										
1. Configural	858.935 (168)	.996	.989	.068 [.063, .072]	–	–	–	–	–	–
2. Weak (λ)	1564.664 (402)	.994	.992	.057 [.054, .060]	1	830.741*	234	-.002	.003	-.011
3. Strong (λ, τ)	2438.586 (597)	.990	.992	.059 [.056, .061]	2	943.232*	195	-.004	.000	.002
4. Strict (λ, τ, δ)	2856.050 (714)	.989	.992	.058 [.056, .060]	3	572.231*	117	-.001	.000	-.001
5. Latent variance-covariance (λ, τ, δ, ξ/φ)	3252.656 (792)	.987	.992	.059 [.057, .061]	4	646.275*	78	-.002	.000	.001
6. Latent mean (λ, τ, δ, ξ/φ, η)	12353.365 (831)	.940	.964	.124 [.122, .126]	5	3837.449*	39	-.047	-.028	.065
Languages										
1. Configural	1271.208 (120)	.995	.984	.087 [.083, .092]	–	–	–	–	–	–
2. Weak (λ)	1196.994 (282)	.996	.995	.051 [.048, .054]	1	224.996*	162	.001	.011	-.036
3. Strong (λ, τ)	2007.941 (417)	.993	.994	.055 [.053, .057]	2	844.335*	135	-.003	-.001	.004
4. Strict (λ, τ, δ)	2277.151 (498)	.992	.994	.053 [.051, .055]	3	418.269*	81	-.001	.000	-.002
5. Latent variance-covariance (λ, τ, δ, ξ/φ)	2395.172 (552)	.991	.994	.052 [.049, .054]	4	458.656*	54	-.001	.000	-.001
6. Latent mean (λ, τ, δ, ξ/φ, η)	7682.826 (579)	.967	.979	.099 [.097, .101]	5	2253.704*	27	-.024	-.015	.047

Notes. * $p < .01$; $\Delta\chi^2$ chi-square difference test calculated using the Mplus DIFFTEST option (reported for descriptive purposes); CM = comparison model; CI = confidence interval; λ = factor loadings; τ = thresholds; δ = uniquenesses; ξ = factor variances; φ = factor covariances; η = factor means.

Table 4. Measurement Invariance Across Sexes and Age Groups.

Invariance model	χ^2 (*df*)	CFI	TLI	RMSEA [90% CI]	CM	$\Delta\chi^2$	Δdf	ΔCFI	ΔTLI	ΔRMSEA
Sexes										
1. Configural	259.708 (24)	.999	.997	.040 [.035, .044]	–	–	–	–	–	–
2. Weak (λ)	379.140 (42)	.998	.997	.036 [.033, .039]	1	141.520*	18	-.001	.000	-.004
3. Strong (λ, τ)	419.298 (57)	.998	.998	.032 [.029, .035]	2	80.238*	15	.000	.001	-.004
4. Strict (λ, τ, δ)	406.761 (66)	.998	.998	.029 [.026, .031]	3	42.637*	9	.000	.000	-.003
5. Latent variance-covariance (λ, τ, δ, ξ/φ)	241.422 (72)	.999	.999	.019 [.017, .022]	4	21.096*	6	.001	.001	-.010
6. Latent mean (λ, τ, δ, ξ/φ, η)	2166.556 (75)	.990	.991	.067 [.064, .069]	5	707.212*	3	-.009	-.008	.048
Age groups										
1. Configural	314.670 (36)	.999	.996	.044 [.040, .049]	–	–	–	–	–	–
2. Weak (λ)	314.998 (72)	.999	.998	.029 [.026, .032]	1	57.395	36	.000	.002	-.015
3. Strong (λ, τ)	387.033 (102)	.999	.999	.026 [.024, .029]	2	100.436*	30	.000	.001	-.003
4. Strict (λ, τ, δ)	486.056 (120)	.998	.999	.028 [.025, .030]	3	110.884*	18	-.001	.000	.002
5. Latent variance-covariance (λ, τ, δ, ξ/φ)	1303.088 (132)	.995	.996	.047 [.045, .050]	4	352.108*	12	-.003	-.003	.019
6. Latent mean (λ, τ, δ, ξ/φ, η)	1484.575 (138)	.994	.995	.049 [.047, .052]	5	150.115*	6	-.001	-.001	.002

Notes. * $p < .01$; $\Delta\chi^2$ chi-square difference test calculated using the Mplus DIFFTEST option (reported for descriptive purposes); CM = comparison model; CI = confidence interval; λ = factor loadings; τ = thresholds; δ = uniquenesses; ξ = factor variances; φ = factor covariances; η = factor means. Age-related data were collected based on five categories in the South African sample, leading us to exclude that sample from the analyses of age groups.

Measurement invariance across age groups

Our measurement invariance analysis across age groups is summarised in Table 4. Regarding configural, weak, strong, and strict forms of invariance, CFI never decreased by more than .001, RMSEA never increased by more than .002, and TLI never decreased. Strict invariance was once again established. The results pertaining to latent variance-covariance invariance were somewhat equivocal. While CFI and TLI each decreased by only .003, RMSEA increased by .019, in which case latent mean invariance could not be considered to hold.

Omega reliability

Omega coefficients are reported in Supplemental Material 1. For countries, languages, and sexes, omegas were computed based on the latent variance-covariance invariance models. For age groups, omegas were computed based on the strict invariance model. Omega$_{\mathrm{OD}}$ coefficients were all ≥ .948, indicating very high reliability. Expectedly, omega$_{\mathrm{ANH\text{-}SOM}}$ and omega$_{\mathrm{DYS}}$ had much lower values and did not exceed .723. Regarding countries and languages, the “anhedonic-somatic” specific factor was relatively well-delineated. By contrast, the “dysphoric” specific factor retained only a limited amount of specificity beyond the variance explained by the general factor. The observation of a much weaker "dysphoric" specific factor suggested that the "dysphoric" items were more clearly indicative of respondents' general level of occupational depression than of respondents' specific levels of dysphoric symptoms. An opposite pattern of results was observed for sexes and age groups. The "dysphoric" specific factor was better delineated than the "anhedonic-somatic" specific factor. The general factor accounted for 85% to 93% of the common variance extracted across the groups of interest, supporting the ODI's essential unidimensionality.

Discussion

The present study examined the measurement invariance of the ODI across national and linguistic groups as well as key demographic characteristics. Measurement invariance allows investigators to estimate whether the relationships between indicators and latent variables are consistent across groups. Measurement invariance is a pivotal property for between-group comparisons and study replicability. By investigating configural, weak, strong, and strict forms of invariance as well as the parameters of structural invariance (latent variance-covariance invariance and latent mean invariance), we submitted the instrument to close scrutiny.

Main findings

We found evidence for complete measurement invariance – configural, weak, strong, and strict, meaning that the ODI behaved equivalently across countries, languages, sexes, and age groups. These results are auspicious given the prospect of comparisons on a global scale. While strict invariance is a prerequisite to comparisons involving observed scores (Morin, 2023), it implies a very high level of constraint and is extremely difficult to achieve in practice (Luong & Flake, 2022). Omega coefficients indicated that the ODI's reliability was mostly attributable to the general Occupational Depression factor. These findings are consistent with the well-established finding that the ODI meets the requirements for essential unidimensionality (e.g. Bianchi, Verkuilen, et al., 2023). All in all, our results comport with the notion that depression (whether attributed to work or not) can be understood as a universal condition in *homo sapiens* (Sapolsky, 2021; Willner et al., 2013).

Looking into the parameters of structural invariance, we found latent variance-covariance invariance to hold across countries, languages, and sexes. Latent variance-covariance invariance was unclear for age groups, possibly suggesting a degree of variation in how the factors were related to each other across adulthood. This finding can be interpreted in light of the divergent expressions of depressive symptoms in older individuals (e.g. Hybels et al., 2012). It is of note that latent variance-covariance invariance, which bears on between-factor correlations, is not expected to constitute a key feature for scales exhibiting essential unidimensionality, such as the ODI.

Unsurprisingly, latent means varied across countries, languages, sexes, and age groups. Regarding countries, there would be little reason to expect the levels of occupational depression to be similar cross-nationally, especially when a heterogeneous set of nations is under scrutiny – as was the case in our study. Many factors may account for this variance, including cross-national differences in labour law, the emphasis placed on workplace safety and employee well-being, specific working conditions, and general socioeconomic conditions. The language variable partly overlapped with the country variable. The absence of latent mean invariance across languages likely reflects differences in occupational depression across geographic areas.

Our latent mean invariance analysis indicated that ODI scores were higher among women than among men. This finding is consistent with the well-documented tendency of women to (a) report more depressive symptoms than men and (b) more frequently receive diagnoses of depressive disorders in comparison to men (American Psychiatric Association, 2013). Regarding age groups, the absence of latent mean invariance is consistent with the finding that depressive disorders are more prevalent during early adulthood and among younger generations (American Psychiatric Association, 2013; Kessler et al., 2003; Villarroel & Terlizzi, 2020).

Based on a systematic review of 426 psychology papers, Maassen et al. (2023) underscored the "dire disregard of measurement invariance testing in psychological science." The authors further observed that, when tested, measurement invariance is generally violated in major ways. This state of affairs likely calls into question large amounts of statistical inferences and study conclusions. The situation does not appear less disquieting when one specifically focuses on measures employed in research on job-related distress. As an illustration, the Maslach Burnout Inventory – the most widely used measure of burnout – has seldom been examined from the standpoint of measurement invariance (Maslach et al., 2001). On the rare occasions when measurement invariance was addressed, non-invariance was found (e.g. across countries and occupational groups; Aboagye et al., 2018; De Beer et al., 2024; Vanheule et al., 2007). These findings lend credence to the concern that the burnout construct may, in its dominant conceptualisation, be quite ethnocentric (Schaufeli, 2017). Given this backdrop, the full measurement invariance across countries, languages, sexes, and age groups displayed by the ODI stands out as highly encouraging.

Limitations

Our study has at least five limitations. First, although we were able to examine the ODI's measurement invariance across 10 languages and 14 countries found on four different continents, it would have been an added advantage if more languages and countries had been included, notably Asian languages and countries. The ODI's measurement invariance will have to be examined further when the instrument is employed in other geographic areas. Second, most of our samples were highly diverse regarding the occupations represented, preventing an analysis of measurement invariance across occupational groups. Third, our samples showed nontrivial variability in size. For example, country-level sample sizes varied from 327 (for South Africa) to 1,971 (for Switzerland). Such variability merits attention. Unequal sample sizes can affect the sensitivity of fit indices (Chen, 2007). In addition, variations in sample sizes might partly account for the differences in latent means that we observed. Fourth, most of the samples under consideration were convenience samples having unclear representativeness. In this respect, we note that the implementation of probability sampling techniques, such as random sampling, is logistically challenging and frequently unfeasible in practice (e.g. because the population of interest cannot be accurately circumscribed or exhaustively contacted). Moreover, nonresponses can alter representativeness ex post, an issue that can be difficult to handle. Unsurprisingly, the use of probability sampling techniques has been the exception rather than the rule in occupational health psychology (Sinclair et al., 2013). Fifth, our reliance on cross-sectional data prevented us from estimating temporal measurement invariance.

Conclusions

Measurement invariance is crucial for theory-building and knowledge accumulation. Our findings suggest that the ODI's measurement invariance holds across languages, countries, sexes, and age groups, providing empirical evidence for the comparability of ODI raw scores across these categories. This study further supports the use of the ODI by researchers and practitioners interested in job-related distress and mental health at work.

The ODI has been scrutinised multiple times based on some of the most sophisticated analytical techniques available (e.g. ESEM factor analysis, Mokken scale analysis). Few measures of job-related distress have undergone and passed such stringent examinations. This study enhances our understanding of the properties of the ODI by suggesting that the instrument can serve in the context of global surveys targeting international and demographic comparisons.

Notes

1. These diagnostic symptoms are consistent with those found in the latest edition of the *International Classification of Diseases* (https://icd.who.int/browse11/l-m/en#/http%3a%2f%2fid.who.int%2ficd%2fentity%2f1563440232).
2. Essential unidimensionality warrants the use of a scale's total score despite the presence of a small degree of multidimensionality; in other words, a single factor accounts for a critical part of the common variance extracted among the scale's items. Essential unidimensionality is typically examined in the context of bifactor models.
3. https://www.hofstede-insights.com/country-comparison-tool

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Renzo Bianchi http://orcid.org/0000-0003-2336-0407

References

Aboagye, M. O., Qin, J., Qayyum, A., Antwi, C. O., Jababu, Y., & Affum-Osei, E. (2018). Teacher burnout in pre-schools: A cross-cultural factorial validity, measurement invariance and latent mean comparison of the Maslach Burnout Inventory, Educators Survey (MBI-ES). *Children and Youth Services Review*, *94*, 186–197. https://doi.org/10.1016/j.childyouth.2018.09.041

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

American Psychological Association. (2015). *Stress in America™: Paying with our health*. Author.

Bianchi, R., Cavalcante, D. C., Queirós, C., Santos, B. D. M., Verkuilen, J., & Schonfeld, I. S. (2023a). Validation of the occupational depression inventory in Brazil: A study of 1,612 civil

servants. *Journal of Psychosomatic Research*, *167*, 111194. https://doi.org/10.1016/j.jpsychores.2023.111194
Bianchi, R., Fiorilli, C., Angelini, G., Dozio, N., Palazzi, C., Palazzi, G., Vitiello, B., & Schonfeld, I. S. (2022). Italian version of the occupational depression inventory: Validity, reliability, and associations with health, economic, and work-life characteristics. *Frontiers in Psychiatry*, *13*, 1061293. https://doi.org/10.3389/fpsyt.2022.1061293
Bianchi, R., Manzano-García, G., Montañés-Muro, P., Schonfeld, E. A., & Schonfeld, I. S. (2022). Occupational depression in a Spanish-speaking sample: Associations with cognitive performance and work-life characteristics. *Journal of Work and Organizational Psychology*, *38*(1), 59–74. https://doi.org/10.5093/jwop2022a5
Bianchi, R., & Schonfeld, I. S. (2020). The occupational depression inventory: A new tool for clinicians and epidemiologists. *Journal of Psychosomatic Research*, *138*, 110249. https://doi.org/10.1016/j.jpsychores.2020.110249
Bianchi, R., & Schonfeld, I. S. (2021). Occupational depression, cognitive performance, and task appreciation: A study based on Raven's Advanced Progressive Matrices. *Frontiers in Psychology*, *12*(4276), 695539. https://doi.org/10.3389/fpsyg.2021.695539
Bianchi, R., & Schonfeld, I. S. (2022). Is the Occupational Depression Inventory predictive of cognitive performance? A focus on inhibitory control and effortful reasoning. *Personality and Individual Differences*, *184*, 111213. https://doi.org/10.1016/j.paid.2021.111213
Bianchi, R., & Schonfeld, I. S. (2023). Examining the evidence base for burnout. *Bulletin of the World Health Organization*, *101*(11), 743–745. https://doi.org/10.2471/BLT.23.289996
Bianchi, R., & Sowden, J. F. (2022). A paradigm shift from burnout to occupational depression. *Journal of Affective Disorders*, *303*, 230–232. https://doi.org/10.1016/j.jad.2022.02.049
Bianchi, R., Verkuilen, J., Sowden, J. F., & Schonfeld, I. S. (2023). Towards a new approach to job-related distress: A three-sample study of the Occupational Depression Inventory. *Stress and Health*, *39*(1), 137–153. https://doi.org/10.1002/smi.3177
Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design—for market research, political polls, and social and health questionnaires* (2nd ed.). Jossey-Bass.
Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834
Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well known but poorly understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994
Cortina, J. M., Sheng, Z., Keener, S. K., Keeler, K. R., Grubb, L. K., Schmitt, N., Tonidandel, S., Summerville, K. M., Heggestad, E. D., & Banks, G. C. (2020). From alpha to omega and beyond! A look at the past, present, and (possible) future of psychometric soundness in the Journal of Applied Psychology. *Journal of Applied Psychology*, *105*, 1351–1381. https://doi.org/10.1037/apl0000815
De Beer, L. T., & Morin, A. J. S. (2022). (B) ESEM invariance syntax generator for Mplus. https://figshare.com/articles/software/_B_ESEM_invariance_syntax_generator_for_Mplus/19360808.
De Beer, L. T., van der Vaart, L., Escaffi-Schwarz, M., De Witte, H., & Schaufeli, W. B. (2024). Maslach Burnout Inventory – General Survey: A systematic review and meta-analysis of measurement properties. *European Journal of Psychological Assessment. Advance online publication*, https://doi.org/10.1027/1015-5759/a000797
Flake, J. K., Davidson, I. J., Wong, O., & Pek, J. (2022). Construct validity and the validity of replication studies: A systematic review. *American Psychologist*, *77*(4), 576–588. https://doi.org/10.1037/amp0001006
Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, *6*(4), 304–312. https://doi.org/10.1037/1040-3590.6.4.304

Golonka, K., Malysheva, K. O., Fortuna, D., Gulla, B., Lytvyn, S., De Beer, L. T., Schonfeld, I. S., & Bianchi, R. (2024). A validation study of the occupational depression inventory in Poland and Ukraine. *Scientific Reports*, *14*(1), 4403. https://doi.org/10.1038/s41598-024-54995-w

Hill, C., De Beer, L. T., & Bianchi, R. (2021). Validation and measurement invariance of the occupational depression inventory in South Africa. *PLoS One*, *16*(12), e0261271. https://doi.org/10.1371/journal.pone.0261271

Hofmans, J., Pepermans, R., & Loix, E. (2009). Measurement invariance matters: A case made for the ORTOFIN. *Journal of Economic Psychology*, *30*(4), 667–674. https://doi.org/10.1016/j.joep.2009.06.002

Houle, S. A., Rich, B. L., Comeau, C. A., Blais, A. R., & Morin, A. J. (2022). The Job engagement scale: Development and validation of a short form in English and French. *Journal of Business and Psychology*, https://doi.org/10.1007/s10869-021-09782-z

Hussey, I., & Hughes, S. (2020). Hidden invalidity among 15 commonly used measures in social and personality psychology. *Advances in Methods and Practices in Psychological Science*, *3*(2), 166–184. https://doi.org/10.1177/2515245919882903

Hybels, C. F., Landerman, L. R., & Blazer, D. G. (2012). Age differences in symptom expression in patients with major depression. *International Journal of Geriatric Psychiatry*, *27*(6), 601–611. https://doi.org/10.1002/gps.2759

Jansson-Fröjmark, M., Badinlou, F., Lundgren, T., Schonfeld, I. S., & Bianchi, R. (2023). Validation of the occupational depression inventory in Sweden. *BMC Public Health*, *23*(1), 1505. https://doi.org/10.1186/s12889-023-16417-w

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E., & Wang, P. S. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *JAMA*, *289(23)*, 3095–3105. https://doi.org/10.1001/jama.289.23.3095

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual Review of Public Health*, *34*(1), 119–138. https://doi.org/10.1146/annurev-publhealth-031912-114409

Lohbeck, A., Tóth-Király, I., & Morin, A. J. S. (2022). Disentangling the associations of academic motivation with self-concept and academic achievement using the bifactor exploratory structural equation modeling framework. *Contemporary Educational Psychology*, *69*, 102069. https://doi.org/10.1016/j.cedpsych.2022.102069

Luong, R., & Flake, J. K. (2022). Measurement invariance testing using confirmatory factor analysis and alignment optimization: A tutorial for transparent analysis planning and reporting. *Psychological Methods*, Advance online publication, https://doi.org/10.1037/met0000441

Maassen, E., D'Urso, E. D., van Assen, M. A. L. M., Nuijten, M. B., De Roover, K., & Wicherts, J. M. (2023). The dire disregard of measurement invariance testing in psychological science. *Psychological Methods*, Advance online publication, https://doi.org/10.1037/met0000624

Maïano, C., Morin, A. J. S., April, J., Webster, E. K., Hue, O., Dugas, C., & Ulrich, D. (2022). Psychometric properties of a French-Canadian version of the Test of Gross Motor Development – Third Edition (TGMD-3): A bifactor structural equation modeling approach. *Measurement in Physical Education and Exercise Science*, *26*(1), 51–62. https://doi.org/10.1080/1091367X.2021.1946541

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*(1), 397–422. https://doi.org/10.1146/annurev.psych.52.1.397

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.

Meuleman, B., Żółtak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D. L., Billiet, J., & Schmidt, P. (2023). Why measurement invariance is important in comparative research. A response to Welzel et al. (2021). *Sociological Methods & Research*, *52*(3), 1401–1419. https://doi.org/10.1177/00491241221091755

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Taylor & Francis.

Morin, A. J. S. (2023). Exploratory structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed, pp. 503–524). Guilford Press.

Morin, A. J. S., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM), and bifactor-ESEM. In G. Tenenbaum, & R. C. Eklund (Eds.), *Handbook of sport psychology* (Vol. 2 (4th ed, pp. 1044–1073). Wiley.

Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F., & Lowe, W. (2021). Tracking and promoting the usage of a COVID-19 contact tracing app. *Nature Human Behaviour*, *5*(2), 247–255. https://doi.org/10.1038/s41562-020-01044-x

Muthén, L. K., & Muthén, B. O. (1998-2023). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nadon, L., De Beer, L. T., & Morin, A. J. S. (2022). Should burnout be conceptualized as a mental disorder? *Behavioral Sciences*, *12*(3), 82. https://doi.org/10.3390/bs12030082

Palan, S., & Schitter, C. (2018). Prolific.ac-A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rotenstein, L. S., Torre, M., Ramos, M. A., Rosales, R. C., Guille, C., Sen, S., & Mata, D. A. (2018). Prevalence of burnout among physicians: A systematic review. *JAMA*, *320*(*11*), 1131–1150. http://doi.org/10.1001/jama.2018.12777

Sapolsky, R. M. (2021). Glucocorticoids, the evolution of the stress-response, and the primate predicament. *Neurobiology of Stress*, *14*, 100320. https://doi.org/10.1016/j.ynstr.2021.100320

Schaufeli, W. B. (2017). Burnout: A short socio-cultural history. In S. Neckel, A. K. Schaffner, & G. Wagner (Eds.), *Burnout, fatigue, exhaustion: An interdisciplinary perspective on a modern affliction* (pp. 105–127). Palgrave Macmillan.

Schonfeld, I. S., & Bianchi, R. (2022). Distress in the workplace: Characterizing the relationship of burnout measures to the Occupational Depression Inventory. *International Journal of Stress Management*, *29*(3), 253–259. https://doi.org/10.1037/str0000261

Schwenk, T. L., & Gold, K. J. (2018). Physician burnout—a serious symptom, but of what?. *JAMA*, *320*(*11*), 1109–1110. http://doi.org/10.1001/jama.2018.11703

Sen, S. (2022). Is it burnout or depression? Expanding efforts to improve physician well-being. *New England Journal of Medicine*, *387*(18), 1629–1630. https://doi.org/10.1056/NEJMp2209540

Sen, I., Quercia, D., Constantinides, M., Montecchi, M., Capra, L., Scepanovic, S., & Bianchi, R. (2022). Depression at work: Exploring depression in major US companies from online reviews. *Proceedings of the ACM on Human-Computer Interaction*, *6*, 438. https://doi.org/10.1145/3555539

Sijtsma, K., & van der Ark, L. A. (2017). A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *British Journal of Mathematical and Statistical Psychology*, *70*(1), 137–158. http://doi.org/10.1111/bmsp.2017.70.issue-1

Sinclair, R. R., Wang, M., & Tetrick, L. E. (2013). *Research methods in occupational health psychology: Measurement, design and data analysis*. Routledge. https://doi.org/10.4324/9780203095249.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press.

Taris, T. W. (2006). Bricks without clay: On urban myths in occupational health psychology. *Work & Stress*, *20*(2), 99–104. https://doi.org/10.1080/02678370600893410

Vanheule, S., Rosseel, Y., & Vlerick, P. (2007). The factorial validity and measurement invariance of the Maslach Burnout Inventory for human services. *Stress and Health*, *23*(2), 87–91. https://doi.org/10.1002/smi.1124

Villarroel, M. A., & Terlizzi, E. P. (2020). *Symptoms of depression among adults: United States, 2019.* NCHS Data Brief, no 379. National Center for Health Statistics.

Willner, P., Scheel-Krüger, J., & Belzung, C. (2013). The neurobiology of depression and antidepressant action. *Neuroscience & Biobehavioral Reviews*, *37*(10), 2331–2371. https://doi.org/10.1016/j.neubiorev.2012.12.007