Scandinavian Journal of Psychology, 2023 DOI: 10.1111/sjop.12996

Empirical Article

The psychometric properties of the Burnout Assessment Tool in Norway: A thorough investigation into construct-relevant multidimensionality

LEON T. DE BEER,[1,2] MARIT CHRISTENSEN,[1] TORHILD A. SØRENGAARD,[3] SIW T. INNSTRAND[1] and WILMAR B. SCHAUFELI[4,5]

[1]*Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway*
[2]*WorkWell Research Unit, North-West University, Potchefstroom, South Africa*
[3]*Department of Leadership and Organizational Behaviour, BI Norwegian Business School, Trondheim, Norway*
[4]*Department of Psychology, Utrecht University, Utrecht, the Netherlands*
[5]*Research Unit Occupational & Organizational Psychology and Professional Learning, KU Leuven, Leuven, Belgium*

De Beer, L. T., Christensen, M., Sørengaard, T. A., Innstrand, S. T. & Schaufeli, W. B. (2023). The psychometric properties of the Burnout Assessment Tool in Norway: A thorough investigation into construct-relevant multidimensionality. *Scandinavian Journal of Psychology.*

Objective

The World Health Organization recognizes burnout as an occupational issue. Nevertheless, accurately identifying employee burnout remains a challenging task. To complicate matters, current measures of burnout have demonstrated limitations, prompting the development of the Burnout Assessment Tool (BAT). Given these circumstances, conducting an in-depth examination of the BAT's construct-relevant multidimensionality is crucial.

Method

This study focuses on both the original 23-item BAT and the short 12-item version, using modern factor analytic methods to investigate reliability, validity, and measurement invariance in a representative sample from Norway (n = 493; 49.54% women).

Results

Our findings revealed that the bifactor exploratory structural equation modeling solution (burnout global factor and four specific burnout component factors) best explained the data for both BAT versions. All factors demonstrated adequate omega coefficients, with the global factor showing exceptional strength. Both BAT versions correlated highly with each other and with another burnout measure, suggesting convergent validity. Furthermore, both BAT versions achieved full (strict) measurement invariance based on gender. Finally, our results showed that burnout acts as a mediator in our proposed job demands–resources model as preliminary evidence of predictive validity.

Conclusions

The study validates the Burnout Assessment Tool in the Norwegian context. The study supports the reliability, validity, and unbiased nature of the tool across genders. The findings also reinforce the importance of job demands and resources, along with burnout as a key mediator, in understanding workplace dynamics in accordance with job demands–resources theory.

Key words: Burnout, burnout assessment tool, exploratory structural equation modeling, measurement invariance, job demands, job resources.

Leon T. De Beer, Department of Psychology, Dragvoll Campus, Edvard Bulls veg 1, Trondheim 7491, Norway. E-mail: leondb@gmail.com

INTRODUCTION

The aftermath of the COVID-19 pandemic has sparked renewed emphasis on employee mental health and well-being (Peters, Dennerlein, Wagner & Sorensen, 2022). In the occupational health psychology domain, no topic is perhaps as popular as burnout, which has been the subject of research for almost half a century. The World Health Organization (WHO) acknowledged burnout as a syndrome stemming from chronic workplace stress that is difficult to manage (WHO, 2019). However, burnout has not been classified as a medical condition but rather as an occupational phenomenon (WHO, 2019). One possible explanation for the difficulty in classifying a case of burnout includes the inherent challenges in its diagnosis. While some countries, like the Netherlands, have established national guidelines for assessing burnout, there are no standardized, globally accepted diagnostic criteria for making a formal diagnosis (Nadon, De Beer & Morin, 2022; Parker & Tavella, 2022). This absence of uniform criteria means that occupational health practitioners often depend on self-report surveys to identify employees *at risk* of burnout. These surveys provide only a risk level, not a categorical diagnosis, but remain a crucial initial step in identifying employees in distress. If the survey results indicate potential burnout risk, practitioners can then refer employees for further evaluation and support, which may include employee assistance programs or other relevant support services. This approach highlights the critical role that self-report measures can play in identifying organizational burnout risk. Therefore, it is essential that these measures show robust psychometric properties in the context in which they are used by the appropriate professionals.

However, over the last few decades burnout measures have been shown to have some limitations, including differences in the conceptualization of burnout being measured (exhaustion-only or multidimensional [e.g., including cynicism, professional efficacy]; see Guseva-Canu *et al.*, 2021 and Schaufeli, 2021), the prescribed factor structure, which remains unclear (three-factor vs. two-factor; e.g., Worley, Vassar, Wheeler, & Barnes, 2008), the apparent divergent role of professional efficacy (e.g., De Beer &

Bianchi, 2019), and the inability of these measures to provide a global burnout score as is ideally required when measuring a syndrome (Schaufeli, Desart & De Witte, 2020).

THE BURNOUT ASSESSMENT TOOL: DEVELOPMENT AND CONVERGENT VALIDITY

Due to these afore-mentioned limitations, Schaufeli, Desart, and De Witte (2020) endeavored to create a new instrument, the Burnout Assessment Tool (BAT). The development of the BAT is unique in that, unlike the development of other burnout measures, both inductive and deductive approaches were used to decide on the final list of items included in the survey. The specific inductive approach employed by Schaufeli and colleagues is uncommon, because the Dutch authorities acknowledge burnout as an occupational disease, and therefore experts who have experience working with (and categorizing) these patients could be interviewed. As for the deductive phase, a list of 357 items and 66 dimensions for burnout were considered. The initial list was constructed by examining 12 existing burnout questionnaires by means of a literature review (see Schaufeli, Desart & De Witte, 2020, for the exact procedure). After analyzing the list, the authors concluded that (1) exhaustion is the core of burnout, (2) all multidimensional scales included both exhaustion and mental distance, (3) positively phrased items are the exception, and (4) Likert scales are used with between four and seven anchors. Considering these conclusions and expert discussions between the authors, an initial pool of 33 items before formal factor analyses were chosen.

Subsequently, after the factor analyses of the development study, the BAT-assessed burnout syndrome showed 23 items measuring four components at its core; three of these components relate to the inability to invest energy: exhaustion, cognitive impairment, and emotional impairment, which represent extreme tiredness, loss of cognitive control, and loss of emotional control, respectively. The fourth component, mental distance, represents an unwillingness to invest energy (Schaufeli, Desart & De Witte, 2020; Schaufeli & Taris, 2005). This culminated in an updated definition of burnout as "a work-related state of exhaustion that occurs among employees, characterised by extreme tiredness, reduced ability to regulate cognitive and emotional processes, and mental distancing" (Schaufeli, Desart & De Witte, 2020, p. 4). Notably, the BAT-assessed burnout definition does not include a professional (in)efficacy component, in line with past research questioning its role (e.g., De Beer & Bianchi, 2019). Consequently, the four components of the BAT (three energy components and a withdrawal component) align with the theoretical conceptualization of burnout containing components of both exhaustion and mental distance (Schaufeli, 2021; Schaufeli & Taris, 2005).

However, the BAT must demonstrate that it is on a par with other burnout measures by means of convergent validity. Convergent validity considers whether different methods of measuring the same concept produce similar results (Abma, Rovers & van der Wees, 2016), and past research on the BAT in other contexts has shown its convergent validity with other well-known burnout measures: the Oldenburg Burnout Inventory (OLBI) and the Maslach Burnout Inventory (e.g., De Beer, Schaufeli & Bakker, 2022; De Beer, Schaufeli & De Witte, 2022; Schaufeli, Desart & De Witte, 2020).

CROSS-CULTURAL VALIDITY AND NORWEGIAN WORKING LIFE

The 23-item version of the BAT (BAT-23) has shown strong evidence for cross-cultural validity with studies of measurement invariance from South Africa (De Beer, Schaufeli & Bakker, 2022; De Beer, Schaufeli & De Witte, 2022), Austria, Belgium, Germany, Ireland, Finland, the Netherlands, and Japan (De Beer *et al.*, 2020). Moreover, the 12-item short version (BAT-12) has also shown the requisite validity and reliability (Hadžibajramović, Schaufeli & De Witte, 2022; Mazzetti *et al.*, 2022; Oprea, Iliescu & De Witte, 2021) and evidence of equivalence in other contexts (e.g., De Beer, Schaufeli & Bakker, 2022; De Beer, Schaufeli & De Witte, 2022; Sinval, Vazquez, Hutz, Schaufeli & Silva, 2022). No studies have yet investigated the psychometric properties and the measurement invariance of the BAT in Norway, and it is important to do so. Norway is an egalitarian society, with great equality between the different social groups. Income inequality is smaller than in most other countries, the education system is generally open to all, and everyone has equal rights to medical care and hospitals. Norwegian working life is characterized by the Nordic model (Gustavsen, 2011), including a high degree of participation and a tripartite collaboration between the parties (workers, unions, and employers) in working life. Norway is one of the countries with the highest employment rates among both women and men. Few countries have higher female labor force participation than Norway. Consequently, demonstrating measurement invariance of the BAT across genders is crucial to ensure its equitable application and to underpin valid comparisons of burnout levels between genders (see Meuleman *et al.*, 2023). In fact, adequate measurement invariance based on gender has previously been found in countries such as Brazil and Portugal (Sinval, Vazquez, Hutz, Schaufeli & Silva, 2022), Croatia (Tomas *et al.*, 2023), and South Africa (De Beer, Schaufeli & Bakker, 2022; De Beer, Schaufeli & De Witte, 2022).

Regarding mental health, Norway does well compared with the EU average. Norway scores among the countries with the lowest prevalence of self-reported anxiety, and there is a lower proportion of workers in Norway than in the EU who are mentally exhausted after work. In terms of mental well-being, Norway is on a par with the EU average. The proportion reporting that their health is at risk because of their work is somewhat lower in Norway than in the EU (STAMI, 2023). While Norwegian employees generally report a better work-life balance than the EU average (STAMI, 2023), many employees are still in distress due to work, and thus it remains an important issue to research.

THE JOB DEMANDS–RESOURCES APPROACH TO BURNOUT

Arguably one of the more popular models in the work and organizational psychology domain to explain how the work environment may affect employee health is the job demands–

resources (JD-R) model (Bakker & Demerouti, 2007; Bakker, Demerouti & Sanz-Vergel, 2023). This model has been applied to the situation in Norway (Kaiser, Patras, Adolfsen, Richardsen & Martinussen, 2020). In the JD-R model, burnout is part of what is called a "health impairment process" that explains the development of burnout mainly as an imbalance between job demands and job resources, creating stress that becomes unmanageable, leading to burnout and eventual undesired outcomes such as sleep disturbances (Sørengaard & Saksvik-Lehouillier, 2022), turnover (Søbstad, Pallesen, Bjorvatn, Costa & Hystad, 2021), and psychological health problems (Burke & Mikkelsen, 2006). Moreover, burnout has also been shown to negatively impact both job satisfaction (e.g., Kaiser, Richardsen & Martinussen, 2021) and life satisfaction (e.g., Hombrados-Mendieta & Cosano-Rivas, 2013). Therefore, burnout has been found to act as a mediator between job demands, job resources, and expected outcomes (e.g., Hakanen, Bakker & Schaufeli, 2006; Hakanen, Schaufeli & Ahola, 2008; Schaufeli & Bakker, 2004) – and this has also been shown when measured with the BAT (see De Beer, Schaufeli & Bakker, 2022; De Beer, Schaufeli & De Witte, 2022). We investigate these variables in a mediation model as part of an initial investigation into predictive validity – see Fig. 1 for a conceptual model.

Specifically, we use three job demands that have been shown to be important in the Norwegian context: work-home conflict (Innstrand, 2022; Innstrand, Langballe, Espnes, Falkum & Aasland, 2008; Jensen, 2016), workload (Svedahl *et al*., 2019), and emotional load (Langballe, Innstrand, Hagtvet, Falkum & Gjerløw Aasland, 2009). Contrastingly, we focus on role clarity (Kaiser, Richardsen & Martinussen, 2021) and supervisor support (Martinussen, Richardsen & Burke, 2007) as job resources. To our knowledge, this is the first study to investigate these factors' relationship with BAT-assessed burnout.

RESEARCH OBJECTIVE AND HYPOTHESES

Recognizing the critical importance of culturally and contextually relevant diagnostic tools, it becomes imperative to assess the applicability of BAT beyond the contexts in which it has initially been validated. The distinct work-life balance and social systems in Norway, particularly its advanced egalitarian values, may influence the manifestation and measurement of burnout in ways that may not have been captured in previous studies. With the high participation of both genders in the workforce and the specific job demands and resources characteristic of the Norwegian workplace, this research seeks to fill that gap.

Therefore, this study aims to investigate the validity, reliability, and measurement invariance of the original BAT-23 and shortened BAT-12 in Norway with a representative sample. The mediating role of burnout in the health impairment process is also considered in the study.

Hypothesis 1: The BAT can be modeled as a total score with four specific component factors.

Hypothesis 2: The BAT shows convergent validity with the Maslach Burnout Inventory – General Survey (MBI-GS).

Hypothesis 3: The BAT shows measurement invariance across gender groups.

Hypothesis 4: Burnout acts, in accordance with the JD-R model, as a mediator between job demands (work-home conflict, workload, emotional demands), job resources (supervisor support, role clarity), and outcome variables (job satisfaction, life satisfaction, and turnover intention).

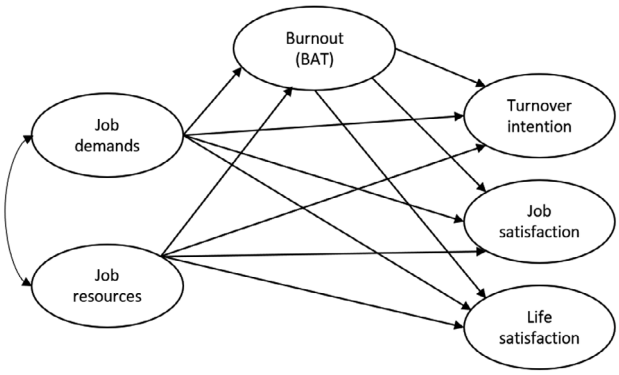

Fig. 1. Conceptual JD-R health model used in this study.

MATERIALS AND METHODS

Participants and procedure

The data research company Bilendi was used to recruit a representative sample of approximately 500 Norwegian workers according to age and gender to complete a survey on the QuestionPro platform. Therefore, the researchers had no direct contact with the participants. Given that we have access to the anonymous data but not the participants, it is not possible to identify any specific individuals, as both Bilendi and QuestionPro adhere to GDRP regulations. Once participants entered the QuestionPro survey, they had to read and accept the informed consent letter explaining the purpose of the study and confirm their voluntary participation before they could continue. The survey contained two attention check questions, which instructed participants to select a specific answer on the scale to demonstrate that they were still paying attention. In the end, there were 530 complete responses, but 37 participants failed the attention checks. As these data could not be fully trusted, they were excluded, resulting in a final sample of 493 employees (n = 493). However, due to the removal of these participants, the data had to be slightly weighted by age and gender to maintain the representativeness of these categories. The average participant was 45.55 years old (SD = 11.54 years) and had, on average, 2.42 children. We provide a further weighted breakdown of the available participants' characteristics in Table 1.

Furthermore, using the pooled cutoff criteria for the BAT-23, provided by Schaufeli, De Witte, Hakanen, Kaltiainen, and Kok (2023), the total burnout score of this sample indicates that 12.9% (n = 63) of the sample was at high risk of becoming burned out. However, it must be emphasized that in the continued absence of clinical diagnostic criteria to identify burnout cases, this estimate should be taken not at face value but as a guideline to identify employees who may need further assistance.

Measures

Burnout was assessed with the full 23-item version of the BAT (BAT-23; Schaufeli, Desart & De Witte, 2020) and with the shortened 12-item version (BAT-12; Hadžibajramović, Schaufeli & De Witte, 2022). Precisely, the following four underlying aspects were measured: exhaustion, by eight items (e.g., "At work, I feel mentally exhausted"), mental distance, by five items (e.g., "I feel indifferent about my job"), cognitive impairment, by five items (e.g., "When I'm working, I have trouble concentrating"), and emotional impairment, by five items (e.g., "At work, I feel unable to control my emotions"). The BAT-12 comprises three items for each of the four components, and the example items from the

Table 1. *Characteristics of the participants*

Category	Sub-category	Count	%
Gender	Male	247	50.06
	Female	244	49.54
	Non-binary	1	0.20
	Other	1	0.20
Education	Primary school	29	5.93
	Secondary school	154	31.17
	3-year higher education	210	42.52
	4-year higher education or more	100	20.38
Sector	Health care	124	25.23
	Trade, business, commerce	52	10.64
	Industry	31	6.30
	Construction	38	7.78
	Education	59	11.91
	Office administration	55	11.14
	Professional services	38	7.72
	Information technology and media	42	8.46
	Oil and gas	24	4.77
	Finance and insurance	22	4.44
	Missing values	8	1.60
Employment	Permanent employee	451	91.26
	Temporary employee	32	6.58
	Self-employed	10	2.16
Full-time equivalent %	100% (Full-time)	387	78.28
	51–100%	78	15.88
	1–50%	28	5.70
Remote working	Never, or almost never	292	59.22
	Sometimes	123	24.91
	Often	43	8.76
	Always, or almost always	35	7.11

Note: Weighted counts and percentages. Count rounded to an appropriate integer. Percentages are unchanged.

BAT-23 above are all included in the BAT-12 (see Hadžibajramović, Schaufeli & De Witte, 2022). All items were measured on a five-point Likert scale ranging from *Never* (1) to *Always* (5). See Table S3 for both BAT versions' English and Norwegian translations.

Moreover, to test convergent validity, we also measured burnout with the MBI-GS (Schaufeli, Leiter, Maslach & Jackson, 1996), which is considered the "gold standard" (Schaufeli, Desart & De Witte, 2020). The MBI-GS is a 16-item instrument that measures three components of burnout: emotional exhaustion (5 items), cynicism (5 items), and professional efficacy (6 items).[1]

The first *job demand* we measured was *work-home conflict* with four items on a five-point scale ranging from *Strongly disagree* to *Strongly agree* (Innstrand, Langballe, Falkum, Espnes & Aasland, 2009; Wayne, Musisca & Fleeson, 2004). An example item from this scale was "Worries or problems at work distract me at home." For the remaining variables, we used one-item indicators to promote the parsimony of our model. Indeed, research (e.g., Fisher, Matthews & Gibbons, 2016; Gilbert & Kelloway, 2014; Williams & Smith, 2016) has shown the validity of one-item indicators. Furthermore, Matthews, Pineault, and Hong (2022) not only show the validity of one-item indicators but also encourage researchers to proactively consider how leveraging one-item measures in their own research might be applicable – as we have done here.

We also measured two additional demands along with work-home conflict: *workload* – "I have too much work to do"; and *emotional demands* – "My work puts me in emotionally upsetting situations").

Conversely, the following *job resources* were measured: *role clarity* – "I have a clear understanding of what is expected of me in my job" and *supervisor support* – "I can count on my supervisor for support when I need it"; Job Demands-Resources Scale (JDRS) (see Rothmann, Mostert & Strydom, 2006).

Finally, we also measured the three outcome variables: *job satisfaction* ("Overall, I am satisfied with my job"; a seven-point scale ranging from *Extremely dissatisfied* to *Extremely satisfied*), *life satisfaction* ("As a whole, I am satisfied with my life"; a seven-point scale ranging from *Extremely dissatisfied* to *Extremely satisfied*; Fisher, Matthews & Gibbons, 2016), and *turnover intention* ("I am actively looking for other jobs"; a five-point scale ranging from *Strongly disagree* to *Strongly agree*; Sjöberg & Sverke, 2000).

Analyses

Mplus 8.9 (Muthén & Muthén, 2023) was used to model the data. Specifically, we used latent variables within a confirmatory factor analysis (CFA) framework for the measurement model with the mean- and variance-adjusted weighted least squares (WLSMV) estimation procedure. First, we systematically estimated six sequential model specifications for each version of the BAT: (1) one-factor CFA, (2) four-factor CFA, (3) second-order factor CFA, (4) bifactor CFA (BCFA), (5) four-factor exploratory structural equation modeling (ESEM), and (6) bifactor exploratory structural equation modeling (BESEM) to ascertain the most appropriate model (Hoyle, 2023). The BESEM model differed from the four-factor ESEM model by adding a global factor to the model and specifying all factors as orthogonal. To assess the adequacy of these models, we employed the comparative fit index (CFI) and the Tucker-Lewis index (TLI), where values of at least 0.90 are indicative of acceptable fit and values exceeding 0.95 are considered to reflect excellent fit. Moreover, for the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR), values up to 0.08 are recognized as denoting acceptable fit (see Kline, 2011; Van de Schoot, Lugtig & Hox, 2012; Wang & Wang, 2020). Second, based on the most appropriate model for the data, we conducted tests of measurement invariance between genders on both versions of the BAT, including configural (factor structure), metric (factor loadings), strong (thresholds), and strict invariance (uniquenesses) (Morin, 2023; Putnick & Bornstein, 2016). Strong invariance indicates that the latent mean scores can be fairly compared between groups, whereas strict invariance (full measurement invariance) indicates that observed scores can also be compared between groups. Regarding the evaluation of the tests for measurement invariance, a decline in CFI and TLI values of −0.01 or greater, or an increase in RMSEA of +0.015 or more when juxtaposing a model with the previous model in the sequence, is indicative of a failure to uphold that level of invariance (Chen, 2007; Cheung & Rensvold, 2002). Therefore, if the cutoff criterion is violated, the level of invariance is considered at the previous acceptable level.

To consider the reliability of the scales, we calculated McDonald's omega coefficients (McDonald, 1970). This approach is especially apt when estimating (B)ESEM models (Morin, 2023). We also tested correlations for convergent validity between the BAT-23, BAT-12, and MBI-GS scores. A correlation approaching one suggests that the constructs are converging (Brown, 2015).

Furthermore, for parsimony, we also tested a structural model (mediation model) based on the observed scores of the variables. From the analyses, we specifically considered the statistical significance ($p < 0.05$) and direction of the standardized beta coefficients. To test for burnout's mediating role in the model, we also bootstrapped the model parameters 10,000 times to generate a range of lower and upper 95% confidence intervals for the indirect effects. If a confidence interval did not contain zero (change sign), it indicates that the indirect effect can be seen as meaningful (Hayes, 2022). This implies that the variable is functioning as a mediator in the model.

RESULTS

Model fit, correlations, and omega coefficients (composite reliability)

As shown in Table 2, our systematic investigation of the factor structure of both versions of the BAT revealed that a

Table 2. *Fit statistics for the model specification investigation of the BAT-23 and BAT-12*

Model (BAT-23) BAT-23	χ^2	*df*	CFI	TLI	RMSEA [90% CI]	SRMR
One-factor	2432.89	230	0.836	0.819	0.139 [0.134–0.144]	0.089
Four-factor	1120.21	224	0.933	0.924	0.090 [0.085–0.095]	0.055
Second-order factor	1082.85	226	0.936	0.928	0.088 [0.082–0.093]	0.055
Bifactor CFA	666.27	211	0.966	0.959	0.066 [0.060–0.072]	0.041
ESEM	519.15	167	0.974	0.960	0.065 [0.059–0.072]	0.026
Bifactor ESEM	452.86	148	0.977	0.961	0.065 [0.058–0.071]	0.023
Model (BAT-12) BAT-12	χ^2	*df*	CFI	TLI	RMSEA [90% CI]	SRMR
One-factor	800.53	54	0.865	0.835	0.167 [0.157–0.177]	0.074
Four-factor	192.15	48	0.974	0.964	0.078 [0.067–0.090]	0.034
Second-order factor	178.65	50	0.977	0.969	0.072 [0.061–0.084]	0.035
Bifactor CFA	195.85	47	0.973	0.962	0.080 [0.069–0.092]	0.039
ESEM	43.83	24	0.996	0.990	0.041 [0.021–0.060]	0.012
Bifactor ESEM	11.57	16	0.999	0.999	0.000 [0.000–0.029]	0.005

Note: χ^2 = 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. Graphical representation of the models is provided in Figs. S1–S6.

unidimensional factor structure was not a good fit. Furthermore, even though some other models showed acceptable fit statistics, the bifactor ESEM models best fitted the data, supporting Hypothesis 1. The online supplementary material provides all models' accompanying factor loadings and uniquenesses (Tables S1 and S2).

However, our selection of the bifactor ESEM model for both versions of the BAT was based not solely on the fit statistics but also on other considerations such as reductions in correlations and the omega coefficients. Specifically, Table 3 shows the size of the correlations between the factors reduced from the CFA model to the ESEM model in both BAT versions ($\Delta Mr_{23} = 0.155$; $\Delta Mr_{12} = 0.119$),[2] providing additional evidence that an underlying global factor is present (Morin, 2023). Moreover, all correlations were statistically significant in the expected directions and showed large effect sizes, except for the correlation between exhaustion (EX) and emotional impairment (EI) in the BAT-12 model, which was a borderline case ($r = 0.494$; medium effect). All omega coefficients were also above 0.700.

The results of the BESEM models revealed a strongly defined and reliable global factor for both versions of the BAT ($\omega_{23} = 0.966$; $\omega_{12} = 0.945$), consistent with the presence of a strong common core to all BAT items. Beyond this global factor, the results also revealed that meaningful specificity remained[3] at the level of the EX ($\omega_{23} = 0.836$; $\omega_{12} = 0.744$), CI ($\omega_{23} = 0.822$; $\omega_{12} = 0.776$), and EI ($\omega_{23} = 0.743$; $\omega_{12} = 0.712$) factors. However, the MD specific factor seemed to retain less specificity in the BAT-12 ($\omega_{12} = 0.455$), where these items primarily seem to reflect participants' global levels of burnout. Nonetheless, it is crucial to emphasize that this specific factor in the BAT-12 still maintains meaningful specificity by accounting for variance not accounted for in its global burnout factor. Indeed, the specific factor retained more meaningful levels of specificity in the BAT-23 ($\omega_{23} = 0.646$).

Table 3. *Descriptive statistics, omega coefficients, and CFA/ESEM correlation matrix*

Factors	*M*	*SD*	1	2	3	4
BAT-23						
1. EX	2.69	0.73	(0.916/0.913)	0.517	0.519	0.554
2. MD	2.24	0.75	0.744	(0.844/0.819)	0.594	0.522
3. CI	2.13	0.72	0.676	0.705	(0.931/0.926)	0.607
4. EI	1.88	0.66	0.700	0.689	0.726	(0.901/0.872)
BAT-12						
1. EX	2.62	0.82	(0.858/0.846)	0.656	0.517	0.494
2. MD	2.18	0.78	0.760	(0.744/0.716)	0.631	0.600
3. CI	2.12	0.73	0.631	0.731	(0.884/0.881)	0.584
4. EI	1.78	0.69	0.643	0.730	0.696	(0.865/0.833)

Note: Four-factor CFA correlations below the diagonal; ESEM correlations above the diagonal; omega coefficients on the diagonal in brackets (CFA/ESEM); M = observed weighted mean; *SD* = observed weighted standard deviation of the mean; all correlations $p < 0.001$; EX = exhaustion; MD = mental distance; CI = cognitive impairment; EI = emotional impairment.

Table 4. *Results of the BESEM measurement invariance testing for gender*

	χ^2	*df*	CFI	TLI	RMSEA	CM	ΔCFI	ΔTLI	ΔRMSEA
BAT-23: Gender									
M1: configural	560.48	296	0.982	0.969	0.060 [0.053–0.068]	–			
M2: metric (λ)	610.01	386	0.985	0.980	0.049 [0.041–0.056]	M1	+0.003	+0.011	−0.011
M3: strong (λ, ν)	668.24	450	0.985	0.983	0.044 [0.037–0.051]	M2	0.000	+0.003	−0.005
M4: strict (λ, ν, δ)	740.00	473	0.982	0.980	0.048 [0.041–0.054]	M3	−0.003	−0.003	+0.004
BAT-12: gender									
N1: configural	44.77	32	0.998	0.991	0.040 [0.000–0.066]	–			
N2: metric (λ)	105.72	67	0.994	0.988	0.048 [0.030–0.065]	N1	−0.004	−0.003	+0.008
N3: strong (λ, τ)	138.14	98	0.993	0.991	0.041 [0.023–0.056]	N2	−0.001	+0.003	−0.007
N4: strict (λ, τ, δ)	167.97	110	0.991	0.989	0.046 [0.032–0.060]	N3	−0.002	−0.002	+0.005

Note: χ^2 = robust chi-square; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation with 90% confidence interval of the RMSEA; λ = factor loadings; τ = thresholds; δ = uniquenesses; CM = comparison model; ΔCFI = change in CFI; ΔTLI = change in TLI; ΔRMSEA = change in RMSEA.

In terms of the BAT versions, convergent validity with the different version of itself ($r = 0.979$) and the MBI-GS ($r_{23} = 0.870$; $r_{12} = 0.856$) were achieved. This supported Hypothesis 2.

Measurement invariance

Table 4 provides the results of the measurement invariance tests for both versions of the BAT. As can be seen, full measurement (strict) invariance was obtained as the CFI and TLI never maximally worsened by −0.010 or the RMSEA by +0.015. Specifically, strict invariance means that the latent means of the BAT-23 and BAT-12 can be compared between sexes *and* the corresponding observed scores. These findings supported Hypothesis 3. Furthermore, the mean results showed neither significant difference between the sexes on the global burnout score nor the specific MD, CI, and EI factors. However, in both the BAT-23 and BAT-12, men scored lower on the specific EX component ($M_{23} = -0.576$, $p < 0.001$; $M_{12} = -0.540$, $p = 0.003$) than women, but also on the CI component in the BAT-12 ($M_{12} = -0.302$, $p = 0.026$).

Direct and indirect path results

Table 5 provides the standardized beta coefficients and related statistics for the direct and indirect paths.

Specifically, the results showed that work-home conflict ($\beta_{23} = 0.523$, $SE = 0.038$, $p < 0.001$; $\beta_{12} = 0.486$, $SE = 0.039$, $p < 0.001$) and emotional demands ($\beta_{23} = 0.175$, $SE = 0.043$, $p < 0.001$; $\beta_{12} = 0.190$, $SE = 0.046$, $p < 0.001$), but not workload ($p_{23} = 0.894$; $p_{12} = 0.708$), had positive paths to burnout for both versions of the BAT. In terms of job resources, supervisor support had a negative path to burnout ($\beta_{23} = -0.110$, $SE = 0.046$, $p < 0.001$; $\beta_{12} = -0.123$, $SE = 0.047$, $p < 0.001$); role clarity had a negative path to burnout only in the BAT-12 ($\beta_{12} = -0.092$, $SE = 0.041$, $p = 0.023$) but not in the BAT-23 ($p_{23} = 0.094$).

Interestingly, none of the job demands or job resources had significant direct paths to the outcome variable turnover intention ($p > 0.05$) except supervisor support, which had a negative relationship ($\beta_{23} = -0.209$, $SE = 0.048$, $p < 0.001$; $\beta_{12} = -0.201$, $SE = 0.047$, $p < 0.001$). The situation was similar for the job satisfaction outcome, for which only supervisor support to turnover intention showed a positive path ($\beta_{23} = 0.257$, $SE = 0.049$, $p < 0.001$; $\beta_{12} = 0.252$, $SE = 0.050$, $p < 0.001$). Moreover, for life satisfaction, only work-home conflict showed a significant, negative, relationship ($\beta_{23} = -0.143$, $SE = 0.071$, $p = 0.043$; $\beta_{12} = -0.188$, $SE = 0.069$, $p = 0.007$).

Burnout revealed a significant positive path to turnover intention ($\beta_{23} = 0.335$, $SE = 0.056$, $p < 0.001$; $\beta_{12} = 0.336$, $SE = 0.052$, $p < 0.001$) and significant negative paths to job satisfaction ($\beta_{23} = -0.407$, $SE = 0.065$, $p < 0.001$; $\beta_{12} = -0.396$, $SE = 0.066$, $p < 0.001$) and life satisfaction ($\beta_{23} = -0.277$, $SE = 0.073$, $p < 0.001$; $\beta_{12} = -0.205$, $SE = 0.074$, $p < 0.001$) in both versions of the BAT. All in all, among the direct paths in the model, the results were relatively consistent between the BAT-23 and BAT-12 except for the path of role clarity to burnout in the BAT-23 relative to the BAT-12, described earlier.

The indirect effects showed a similar pattern in which the direct paths were significant. Therefore, Hypothesis 4 was only partially supported, even though most indirect effects were meaningful. For example, work-home conflict and emotional demands indirectly affected turnover through burnout in both versions of the BAT.

An online application was created for Norwegian employees to consider their own burnout risk results, for entertainment purposes, based on the 23-item BAT data from this study to benchmark against. The application can be accessed at the following location: https://theburnout.app/?mod=no.

DISCUSSION

This study investigated the construct-relevant multidimensionality and measurement invariance of the BAT-23 and BAT-12, an updated tool to measure burnout. Results broadly supported that the scales have robust properties: validity and reliability.

First, Hypothesis 1 was supported as it was established that the best representation of the BAT-related data was the BESEM latent variable representation. That is a model that included a strong global burnout factor, with four specific components of the BAT. This is in line with other recent studies on the BAT, which have

Table 5. *Direct and indirect path results for the structural models*

Direct structural path (BAT-23/BAT-12)	β_{23}/β_{12}	SE_{23}/SE_{12}	p_{23}/p_{12}
Work-home conflict → Burnout	0.523*/0.486*	0.038/0.039	<0.001/<0.001
Workload → Burnout	−0.005/-0.014	0.037/0.038	0.894/0.708
Emotional demands → Burnout	0.175*/0.190*	0.043/0.046	<0.001/<0.001
Role clarity → Burnout	−0.066/-0.092*	0.039/0.041	0.094/0.023
Supervisor support → Burnout	−0.110*/-0.123*	0.046/0.047	0.017/0.009
Work-home conflict → Turnover intention	0.029/0.025	0.063/0.062	0.652/0.680
Workload → Turnover intention	0.015/0.018	0.050/0.050	0.769/0.714
Emotional demands → Turnover intention	0.074/0.063	0.057/0.056	0.189/0.262
Role clarity → Turnover intention	−0.057/-0.045	0.039/0.038	0.144/0.242
Supervisor support → Turnover intention	−0.209*/-0.201*	0.048/0.047	<0.001/<0.001
Work-home conflict → Job satisfaction	0.003/-0.017	0.063/0.060	0.959/0.772
Workload → Job satisfaction	−0.020/-0.023	0.047/0.047	0.675/0.628
Emotional demands → Job satisfaction	−0.080/-0.076	0.052/0.052	0.128/0.147
Role clarity → Job satisfaction	0.024/0.014	0.042/0.043	0.570/0.736
Supervisor support → Job satisfaction	0.257*/0.252*	0.049/0.050	<0.001/<0.001
Work-home conflict → Life satisfaction	−0.143*/-0.188*	0.071/0.069	0.043/0.007
Workload → Life satisfaction	0.095/0.094	0.057/0.057	0.097/0.102
Emotional demands → Life satisfaction	0.067/0.057	0.059/0.060	0.255/0.340
Role clarity → Life satisfaction	0.025/0.024	0.050/0.051	0.621/0.635
Supervisor support → Life satisfaction	0.054/0.059	0.052/0.053	0.294/0.262
Burnout → Turnover intention	0.335*/0.366*	0.056/0.052	<0.001/<0.001
Burnout → Job satisfaction	−0.407*/-0.396*	0.065/0.066	<0.001/<0.001
Burnout → Life satisfaction	−0.277*/-0.205*	0.073/0.074	<0.001/<0.001
Indirect effect (BAT-23/BAT-12)	**β_{23}/β_{12}**	**L 95% CI**	**U 95% CI**
Work-home conflict → Burnout → Turnover	0.175*/0.178*	0.112/0.121	0.239/0.237
Workload → Burnout → Turnover	−0.002/-0.005	−0.026/-0.033	0.024/0.023
Emotional demands → Burnout → Turnover	0.058*/0.070*	0.026/0.033	0.096/0.112
Role clarity → Burnout → Turnover	−0.022/-0.034*	−0.053/-0.069	0.003/-0.005
Supervisor support → Burnout → Turnover	−0.037*/-0.045*	−0.067/-0.078	−0.007/-0.012
Work-home conflict → Burnout → Job satisfaction	−0.231*/-0.192*	−0.287/-0.263	−0.139/-0.124
Workload → Burnout → Job satisfaction	0.002/0.006	−0.028/-0.024	0.033/0.037
Emotional demands → Burnout → Job satisfaction	−0.071*/-0.075*	−0.116/-0.123	−0.033/-0.036
Role clarity → Burnout → Job satisfaction	0.027/0.036*	−0.004/0.005	0.064/0.076
Supervisor support → Burnout → Job satisfaction	0.045*/0.049*	0.008/0.011	0.089/0.094
Work-home conflict → Burnout → Life satisfaction	−0.145*/-0.100*	−0.226/-0.176	−0.066/-0.028
Workload → Burnout → Life satisfaction	0.001/0.003	−0.019/-0.012	0.025/0.022
Emotional demands → Burnout → Life satisfaction	−0.048*/-0.039*	−0.089/-0.078	−0.018/-0.010
Role clarity → Burnout → Life satisfaction	0.018*/0.019*	−0.003/0.001	0.047/0.047
Supervisor support → Burnout → Life satisfaction	0.030*/0.025*	0.005/0.004	0.064/0.055

Note: β = standardized beta coefficient; SE = standard error; p = two-tailed statistical significance.
*Significant.

found a bifactor representation to be most suitable in other contexts (e.g., Basińska, Gruszczyńska & Schaufeli, 2023; De Beer, Schaufeli & Bakker, 2022; De Beer, Schaufeli & De Witte, 2022). Furthermore, the evidence shows that the BAT can be modeled as a total score, with its subcomponent scores, solving one of the main concerns against other burnout measures purportedly measuring a syndrome.

Furthermore, the correlations from the convergent validity assessment showed that both versions of the BAT are highly correlated with one another – showing convergent validity (R^2 = 95.84%) and supporting Hypothesis 2. This is important, as it shows that the shorter version of the BAT reflects an excellent representation of the scores one would achieve on the full 23-item version. Similarly, but less strongly, both versions of the BAT correlated highly with the MBI-GS, said to be the "gold standard" of burnout measurement (Schaufeli, Desart & De Witte, 2020).

Regarding equivalence, both versions of the BAT achieved full (strict) measurement invariance as pertaining to gender – supporting Hypothesis 3. This means that not only the latent scores but also the observed scores of BAT-assessed burnout can be compared between men and women. This aligns with other research on the BAT, which has shown it to be invariant across countries (De Beer *et al*., 2020) and in specific contexts (De Beer, Schaufeli & De Witte, 2022). Specifically, the results showed no statistically significant difference in the mean level of global burnout between men and women in this representative sample. This contrasts with other studies, which have found levels of global burnout to be higher among women (e.g., De Beer, Schaufeli & De Witte, 2022). One potential explanation for this result within the Norwegian context could be the more egalitarian nature of working life, so similar scores are apparent. Nevertheless, the higher level of specific exhaustion found among

women using both the BAT-23 and the BAT-12 align with previous studies conducted in Norway using the Oldenburg Burnout Inventory (OLBI) as a burnout measure (Innstrand, Langballe, Falkum & Aasland, 2011). Innstrand and colleagues suggested that the gender differences might vary across occupational groups. Unfortunately, we could not test gender differences across occupational groups in the present study, due to under-representation.

Finally, in the structural model, both versions of the BAT performed almost identically in predicting outcomes, in line with the JD-R model. Of the direct paths, it is important to underline the strong effect of work-home conflict in contributing to burnout. Although this aligns with previous findings from Norway (Innstrand, Langballe, Espnes, Falkum & Aasland, 2008), it underlines the need to respect work-home balance even in the Nordic work model (Gustavsen, 2011). Specifically, burnout mediated between work-home conflict and emotional load and the outcome variables, and between job resources and job satisfaction and life satisfaction, supporting Hypothesis 4. However, an interesting finding was that workload did not have any significant direct effects on burnout in the presence of work-home conflict and emotional load. Our cross-sectional result contradicts stronger findings from a meta-analytic study that found a small effect of workload over time (Guthier, Dormann & Voelkle, 2020). This could be explained by fair workloads in Norway, as the meta-analytic effect was already small – or it could be an artifact in this sample. Be that as it may, work-home conflict and emotional load did show effects, and these should be managed in the workplace. Furthermore, the importance of supervisor support as a job resource is also underlined in this study, as it has been in a meta-analysis (Aronsson *et al.*, 2017).

In general, all the results supported the role of BAT-assessed burnout in the health impairment process of the JD-R model (Bakker, Demerouti & Sanz-Vergel, 2023). Specifically, job demands in the form of work-home conflict and emotional load contributed to burnout and the outcomes of this study, whereas supervisor support as a job resource minimized the mediating impact of burnout on the outcomes. Therefore, leaders of organizations in Norway should be cognizant of these demands and resource to promote optimal outcomes.

LIMITATIONS

This study is subject to limitations that merit attention. While the demographic composition of the sample approximated the general Norwegian population in terms of age group and gender, it remained cross-sectional in nature. Longitudinal research designs are recommended for future studies to establish causality among the variables of interest. Moreover, due to the plurality of health-care workers in this sample, it would be beneficial for future research to expand sample sizes and aim to include occupational group as an additional dimension of representativeness of the Norwegian workforce. This approach would also facilitate in-depth explorations of intra- and inter-occupational group dynamics by using multigroup analyses and similar techniques, which may yield insightful dynamics about burnout. Second, this study did not include any self-reported depression-related variables that could have been used to contribute to the burnout-depression debate (e.g., Nadon, De Beer & Morin, 2022). Lastly, our data did not contain objective data such as actual performance or productivity, which would have been an interesting feature. Therefore, future studies should endeavor to obtain and include performance, sickness absence, and health-related data to investigate the finer nuances of BAT-assessed burnout in individual and organizational outcomes. These efforts could include investigating at what levels of burnout severity or score the BAT may be useful to predict depression, anxiety, or other health-related outcomes that have binary, diagnostic classifications.

CONCLUSIONS

This study provided robust evidence for the validity, reliability, and gender equivalence of both the BAT-23 and BAT-12 within Norway. Therefore, the BAT-23 and BAT-12 can be used with confidence to ascertain organizational employee burnout risk. The results showed that the BAT-23 captures nuances beyond the global burnout score more thoroughly than the shorter BAT-12. Specifically, with the BAT-12, mental distance (withdrawal) is reflected more by the global burnout factor even though the specific mental distance does maintain some meaningful specificity beyond the global factor. However, the primary advantage of the BAT is its deference to a total score. Our results also supported BAT-assessed burnout's role as a mediator in the health impairment process of a JD-R model.

ACKNOWLEDGEMENTS

None.

CONFLICTS OF INTEREST

None declared.

ETHICS STATEMENT

Respondents gave informed consent by accepting the voluntary nature of participation before the online survey commenced and were free to cease participation at any time. Due to online anonymity and a limited set of background information, it was not possible for the researchers to identify who was responding to the survey. The project therefore followed the requirements of the Norwegian Centre for Research Data (NSD).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ENDNOTES

[1] The MBI-GS is copyrighted, and therefore example items are not shared here.

[2] Delta mean (change) in the correlations from the CFA model to the ESEM model.

[3] Due to the division of item-level true score variance between two sets of factors (G and S) in the bifactor ESEM solutions, it is common for the S factors to have weaker definition compared with the CFA or ESEM (see

Morin, Myers & Lee, 2020). This has led to arguments that composite reliability coefficients of 0.500 can still be deemed acceptable for S factors (Perreira *et al*., 2018).

REFERENCES

Abma, I., Rovers, M. & van der Wees, P.J. (2016). Appraising convergent validity of patient-reported outcome measures in systematic reviews: Constructing hypotheses and interpreting outcomes. *BMC Research Notes*, *9*, 226. https://doi.org/10.1186/s13104-016-2034-2.

Aronsson, G., Theorell, T., Grape, T., Hammarström, A., Hogstedt, C., Marteinsdottir, I. *et al*. (2017). A systematic review including meta-analysis of work environment and burnout symptoms. *BMC Public Health*, *17*, 264. https://doi.org/10.1186/s12889-017-4153-7.

Bakker, A.B. & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309–328. https://doi.org/10.1108/02683940710733115.

Bakker, A.B., Demerouti, E. & Sanz-Vergel, A. (2023). Job demands–resources theory: Ten years later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 25–53. https://doi.org/10.1146/annurev-orgpsych-120920-053933.

Basińska, B., Gruszczyńska, E. & Schaufeli, W.B. (2023). The polish adaptation of the burnout assessment tool (BAT-PL). *Psychiatria Polska*, *57*(1), 223–235. https://doi.org/10.12740/pp/onlinefirst/141563.

Brown, T.A. (2015). *Confirmatory factor analysis for applied research*, 2nd edition. New York, NY: Guilford Press.

Burke, R.J. & Mikkelsen, A. (2006). Burnout among Norwegian police officers: Potential antecedents and consequences. *International Journal of Stress Management*, *13*(1), 64–83. https://doi.org/10.1037/1072-5245.13.1.64.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834.

Cheung, G.W. & Rensvold, R.B. (2002). Evaluating goodness-of fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

De Beer, L.T. & Bianchi, R. (2019). Confirmatory factor analysis of the Maslach burnout inventory: A Bayesian structural equation modelling approach. *European Journal of Psychological Assessment*, *35*(2), 217–224. https://doi.org/10.1027/1015-5759/a000392.

De Beer, L.T., Schaufeli, W.B. & Bakker, A.B. (2022). Investigating the validity of the short form burnout assessment tool: A job demands-resources approach. *African Journal of Psychological Assessment*, *4*, a95. https://doi.org/10.4102/ajopa.v4i0.95.

De Beer, L.T., Schaufeli, W.B. & De Witte, H. (2022). The psychometric properties and measurement invariance of the Burnout Assessment Tool (BAT-23) in South Africa. *BMC Public Health*, *22*(1), 1555. https://doi.org/10.1186/s12889-022-13978-0.

De Beer, L.T., Schaufeli, W.B., De Witte, H., Hakanen, J.J., Shimazu, A., Glaser, J. *et al*. (2020). Measurement invariance of the Burnout Assessment Tool (BAT) across seven cross-national representative samples. *International Journal of Environmental Research and Public Health*, *17*(15), 5604. https://doi.org/10.3390/ijerph17155604.

Fisher, G.G., Matthews, R.A. & Gibbons, A.M. (2016). Developing and investigating the use of single-item measures in organizational research. *Journal of Occupational Health Psychology*, *21*(1), 3–23. https://doi.org/10.1037/a0039139.

Gilbert, S. & Kelloway, E.K. (2014). Using single items to measure job stressors. *International Journal of Workplace Health Management*, *7*(3), 186–199. https://doi.org/10.1108/IJWHM-03-2013-0011.

Guseva-Canu, I., Marca, S.C., Dell'Oro, F., Balázs, Á., Bergamaschi, E., Besse, C. *et al*. (2021). Harmonised definition of occupational burnout: A systematic review, semantic analysis, and Delphi consensus in 29 countries. *Scandinavian Journal of Work, Environment & Health*, *47*(2), 95–107. https://doi.org/10.5271/sjweh.3935.

Gustavsen, B. (2011). The Nordic model of work organization. *Journal of the Knowledge Economy*, *2*, 463–480. https://doi.org/10.1007/s13132-011-0064-5.

Guthier, C., Dormann, C. & Voelkle, M.C. (2020). Reciprocal effects between job stressors and burnout: A continuous time meta-analysis of longitudinal studies. *Psychological Bulletin*, *146*(12), 1146–1173. https://doi.org/10.1037/bul0000304.

Hadžibajramović, E., Schaufeli, W. & De Witte, H. (2022). Shortening of the burnout assessment tool (BAT): From 23 to 12 items using content and Rasch analysis. *BMC Public Health*, *22*, 560. https://doi.org/10.1186/s12889-022-12946-y.

Hakanen, J., Bakker, A.B. & Schaufeli, W.B. (2006). Burnout and work engagement among teachers. *The Journal of School Psychology*, *43*, 495–513. https://doi.org/10.1016/j.jsp.2005.11.001.

Hakanen, J.J., Schaufeli, W.B. & Ahola, K. (2008). The job demands-resources model: A three-year cross-lagged study of burnout, depression, commitment, and work engagement. *Work & Stress*, *22*, 224–241. https://doi.org/10.1080/02678370802379432.

Hayes, A.F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, 3rd edition. New York, NY: The Guilford Press. https://apastyle.apa.org/blog/publisher-locations-in-book-references.

Hombrados-Mendieta, I. & Cosano-Rivas, F. (2013). Burnout, workplace support, job satisfaction and life satisfaction among social workers in Spain: A structural equation model. *International Social Work*, *56*(2), 228–246. https://doi.org/10.1177/0020872811421620.

Hoyle, R.H. (2023). *Handbook of Structural Equation Modeling*, 2nd edition. New York, NY: Guilford Press.

Innstrand, S.T. (2022). Burnout among health care professionals during COVID-19. *International Journal of Environmental Research and Public Health*, *19*(18), 11807. https://doi.org/10.3390/ijerph191811807.

Innstrand, S.T., Langballe, E.M., Espnes, G.A., Falkum, E. & Aasland, O.G. (2008). Positive and negative work–family interaction and burnout: A longitudinal study of reciprocal relations. *Work & Stress*, *22*(1), 1–15. https://doi.org/10.1080/02678370801975842.

Innstrand, S.T., Langballe, E.M., Falkum, E. & Aasland, O.G. (2011). Exploring within- and between-gender differences in burnout: 8 different occupational groups. *International Archives of Occupational and Environmental Health*, *84*, 813–824. https://doi.org/10.1007/s00420-011-0667-y.

Innstrand, S.T., Langballe, E.M., Falkum, E., Espnes, G.A. & Aasland, O.G. (2009). Gender-specific perceptions of four dimensions of the work/family interaction. *Journal of Career Assessment*, *17*(4), 402–416. https://doi.org/10.1177/1069072709334238.

Jensen, M.T. (2016). A two-wave cross-lagged study of work-role conflict, work-family conflict and emotional exhaustion. *Scandinavian Journal of Psychology*, *57*(6), 591–600. https://doi.org/10.1111/sjop.12328.

Kaiser, S., Patras, J., Adolfsen, F., Richardsen, A.M. & Martinussen, M. (2020). Using the job demands–resources model to evaluate work-related outcomes among Norwegian health care workers. *SAGE Open*, *10*(3), 2158244020947436. https://doi.org/10.1177/2158244020947436.

Kaiser, S., Richardsen, A.M. & Martinussen, M. (2021). Burnout and engagement at the Northernmost University in the World. *SAGE Open*, *11*(3), 21582440211031552. https://doi.org/10.1177/21582440211031552.

Kline, R.B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

Langballe, E.M., Innstrand, S.T., Hagtvet, K.A., Falkum, E. & Gjerløw Aasland, O. (2009). The relationship between burnout and musculoskeletal pain in seven Norwegian occupational groups. *Work*, *32*(2), 179–188. https://doi.org/10.3233/WOR-2009-0804.

Martinussen, M., Richardsen, A.M. & Burke, R.J. (2007). Job demands, job resources, and burnout among police officers. *Journal of Criminal Justice*, *35*(3), 239–249. https://doi.org/10.1016/j.jcrimjus.2007.03.001.

Matthews, R.A., Pineault, L. & Hong, Y.H. (2022). Normalizing the use of single-item measures: Validation of the single-item compendium for organizational psychology. *Journal of Business and Psychology*, *37*(4), 639–673. https://doi.org/10.1007/s10869-022-09813-3.

Mazzetti, G., Consiglio, C., Santarpia, F.P., Borgogni, L., Guglielmi, D. & Schaufeli, W.B. (2022). Italian validation of the 12-item version of the Burnout Assessment Tool (BAT-12). *International Journal of*

Environmental Research and Public Health, *19*(14), 8562. https://doi.org/10.3390/ijerph19148562.
McDonald, R. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*, 1–21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x.
Meuleman, B., Żółtak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D.L. *et al*. (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.
Morin, A.J.S. (2023). Exploratory structural equation modeling. In R.H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed., pp. 503–524). New York, NY: 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* (pp. 1044–1073). Hoboken, NJ: Wiley. https://doi.org/10.1002/9781119568124.ch51.
Muthén, L. & Muthén, B. (2023). *Mplus user's guide, version 8.9*. Los Angeles, CA: Muthén & Muthén.
Nadon, L., De Beer, L.T. & Morin, A.J.S. (2022). Should burnout be conceptualised as a mental disorder? *Behavioral Sciences*, *12*(3), 82. https://doi.org/10.3390/bs12030082.
Oprea, B., Iliescu, D. & De Witte, H. (2021). Romanian short version of the Burnout Assessment Tool: Psychometric properties. *Evaluation & the Health Professions*, *44*(4), 406–415. https://doi.org/10.1177/01632787211048924.
Parker, G. & Tavella, G. (2022). The diagnosis of burnout: Some challenges. *The Journal of Nervous and Mental Disease*, *210*(7), 475–478. https://doi.org/10.1097/NMD.0000000000001492.
Perreira, T.A., Morin, A.J.S., Hebert, M., Gillet, N., Houle, S.A. & Berta, W. (2018). The short form of the Workplace Affective Commitment Multidimensional Questionnaire (WACMQ-S): A bifactor-ESEM approach among healthcare professionals. *Journal of Vocational Behavior*, *106*, 62–83. https://doi.org/10.1016/j.jvb.2017.12.004.
Peters, S.E., Dennerlein, J.T., Wagner, G.R. & Sorensen, G. (2022). Work and worker health in the post-pandemic world: A public health perspective. *The Lancet Public Health*, *7*(2), e188–e194. https://doi.org/10.1016/S2468-2667(21)00259-0.
Putnick, D.L. & Bornstein, M.H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004.
Rothmann, S., Mostert, K. & Strydom, M. (2006). A psychometric evaluation of the job demands-resources scale in South Africa. *SA Journal of Industrial Psychology*, *32*(4), 76–86. https://hdl.handle.net/10520/EJC89107.
Schaufeli, W.B. (2021). The burnout enigma solved? *Scandinavian Journal of Work, Environment & Health*, *47*(3), 169–170. https://doi.org/10.5271/sjweh.3950.
Schaufeli, W.B. & Bakker, A.B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior*, *25*, 293–315. https://doi.org/10.1002/job.248.
Schaufeli, W.B., De Witte, H., Hakanen, J.J., Kaltiainen, J. & Kok, R. (2023). How to assess severe burnout?: Cutoff points for the Burnout Assessment Tool (BAT) based on three European samples. *Scandinavian Journal of Work, Environment & Health*, *49*(4), 293–302. https://doi.org/10.5271/sjweh.4093.
Schaufeli, W.B., Desart, S. & De Witte, H. (2020). Burnout assessment tool (BAT): Development, validity, and reliability. *International Journal of Environmental Research and Public Health*, *17*(24), 9495. https://doi.org/10.3390/ijerph17249495.
Schaufeli, W.B., Leiter, M.P., Maslach, C. & Jackson, S.E. (1996). Maslach burnout inventory-general survey. In C. Maslach, S.E. Jackson & M.P. Leiter (Eds.), *The Maslach burnout inventory-test manual* (3rd ed.) Palo Alto, CA: Consulting Psychologists Press.
Schaufeli, W.B. & Taris, T.W. (2005). The conceptualisation and measurement of burnout: Common ground and worlds apart. *Work & Stress*, *19*(3), 256–262. https://doi.org/10.1080/02678370500385913.
Sinval, J., Vazquez, A.C.S., Hutz, C.S., Schaufeli, W.B. & Silva, S. (2022). Burnout Assessment Tool (BAT): Validity evidence from Brazil and Portugal. *International Journal of Environmental Research and Public Health*, *19*(3), 1344. https://doi.org/10.3390/ijerph19031344.
Sjöberg, A. & Sverke, M. (2000). The interactive effect of job involvement and organisational commitment on job turnover revisited: A note on the mediating role of turnover intention. *Scandinavian Journal of Psychology*, *41*(3), 247–252. https://doi.org/10.1111/1467-9450.00194.
Søbstad, J.H., Pallesen, S., Bjorvatn, B., Costa, G. & Hystad, S.W. (2021). Predictors of turnover intention among Norwegian nurses: A cohort study. *Health Care Management Review*, *46*(4), 367–374. https://doi.org/10.1097/HMR.0000000000000277.
Sørengaard, T.A. & Saksvik-Lehouillier, I. (2022). Associations between burnout symptoms and sleep among workers during the COVID-19 pandemic. *Sleep Medicine*, *90*, 199–203. https://doi.org/10.1016/j.sleep.2022.01.022.
STAMI. (2023). The National Institute of Occupational Health in Norway https://stami.no/.
Svedahl, E.R., Pape, K., Toch-Marquardt, M., Skarshaug, L.J., Kaspersen, S.L., Bjørngaard, J.H. *et al*. (2019). Increasing workload in Norwegian general practice–a qualitative study. *BMC Family Practice*, *20*, 68. https://doi.org/10.1186/s12875-019-0952-5.
Tomas, J., Maslić Seršić, D., Mikac, U., Rebernjak, B., Buško, V. & De Witte, H. (2023). Validation of the Croatian version of the short form of the Burnout Assessment Tool: Findings from a nationally representative sample. *International Journal of Selection and Assessment*, 1–14. https://doi.org/10.1111/ijsa.12447.
Van de Schoot, R., Lugtig, P. & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740.
Wang, J. & Wang, X. (2020). *Structural equation modelling*, 2nd edition. Hoboken, NJ: Wiley & Sons.
Wayne, J.H., Musisca, N. & Fleeson, W. (2004). Considering the role of personality in the work-family experience: Relationships of the big five to work-family conflict and facilitation. *Journal of Vocational Behavior*, *64*, 108–130. https://doi.org/10.1016/S0001-8791(03)00035-6.
WHO [World Health Organization]. (2019). Burnout QD85. In *International statistical classification of diseases and related health problems* (11th ed.). Geneva: WHO. https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/129180281.
Williams, G.M. & Smith, A.P. (2016). Using single-item measures to examine the relationships between work, personality, and well-being in the workplace. *Psychology*, *7*, 753–767. https://doi.org/10.4236/psych.2016.76078.
Worley, J.A., Vassar, M., Wheeler, D.L. & Barnes, L.L.B. (2008). Factor structure of scores from the Maslach Burnout Inventory: A review and meta-analysis of 45 exploratory and confirmatory factor-analytic studies. *Educational and Psychological Measurement*, *68*(5), 797–823. https://doi.org/10.1177/0013164408315268.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Table S1. *Standardized factor loadings (λ) and uniquenesses (δ) for models of the BAT-23 (*n = *493)*
Table S2. *Standardized factor loadings (λ) and uniquenesses (δ) for models of the BAT-12 (*n = *493)*
Table S3. *The Norwegian (Norsk) BAT-23 and BAT-12*

Fig. S1. The one-factor CFA model.
Fig. S2. The four-factor CFA model.
Fig. S3. The second-order CFA model.
Fig. S4. The bifactor CFA model.
Fig. S5. The four-factor ESEM model.
Fig. S6. The bifactor ESEM model.

Received 29 August 2023, Revised 15 November 2023, accepted 7 December 2023