







Validation of the Croatian version of the short form of the Burnout Assessment Tool: Findings from a nationally representative sample

Jasmina Tomas¹  | Darja Maslić Seršić¹  | Una Mikac¹  | Blaž Rebernjak¹  |
Vesna Buško¹  | Hans De Witte^{2,3} 

¹Department of Psychology, University of Zagreb, Zagreb, Croatia

²Research Group Work, Organizational and Personnel Psychology WOPP-O2L, KU Leuven, Leuven, Belgium

³Optentia Research Unit, North-West University, Vanderbijlpark, South Africa

Correspondence

Jasmina Tomas, Department of Psychology, Faculty of Humanities and Social Sciences, Ivana Lučića 3, 10 000 Zagreb, Croatia.
Email: jasmina.tomas@ffzg.hr

Funding information

University of Zagreb and Research Group Work & Organisational Psychology—KU Leuven

Abstract

Burnout poses severe health-related and financial risks. However, valid and reliable measurement of this occupational phenomenon has been impeded by the conceptual, psychometric and pragmatical shortcomings of the extant burnout instruments. The Burnout Assessment Tool (BAT) is a new measure of burnout that was developed to overcome these deficiencies. The purpose of this study was to validate the Croatian version of its short form, BAT-12, using a representative sample of the Croatian working population. To do so, we examined the factor structure and measurement invariance of BAT-12 across gender, age, and occupational type, as well as convergent and discriminant data on BAT-12 vis-à-vis an alternative burnout measure, the Oldenburg Burnout Inventory, job resources, job demands and job outcomes. The results obtained on online survey data from 966 employees supported the hierarchical structure of BAT-12, although with slight modifications. We also found comparable loading structure across age, gender, and occupational type, whereas the intercept invariance depended on the moderator. Convergent and discriminant validity was supported in relation to all examined variables. Therefore, the psychometric soundness of the Croatian version of BAT-12 adds new evidence to the current validation process of the BAT-12 and supports the application of this measure on the Croatian working population.

KEYWORDS

burnout, Burnout Assessment Tool, convergent and discriminant validity, factorial structure, measurement invariance, representative sample

Practitioner points

- Extant validation studies support the psychometric soundness of the new measure of burnout (Burnout Assessment Tool; BAT) which was developed to overcome crucial shortcomings of the previous burnout measures.
- This study adds to the current validation process of BAT by examining the psychometric properties of its short version (BAT-12) using a nationally representative sample of the Croatian working population.

- The results support the usage of the BAT-12 not only for academic research purposes, but also for practical purposes of the assessment of risks for employee safety and health.

1 | INTRODUCTION

Decades of research of the burnout construct has reached an unequivocal conclusion—burnout should be a cause of concern for employees, organizations, and societies. Reflecting a metaphor for a work-related state of chronic exhaustion (Schaufeli et al., 2020a), burnout has been linked to various indicators of impaired employee health, organizational malfunctioning, and societal costs. For example, research indicates that burnout predicts diverse physical (e.g., coronary heart disease and type 2 diabetes) and psychological problems and diseases (e.g., depressive symptoms and insomnia; Salvagioni et al., 2017), as well as lower organizational commitment, job satisfaction, and productivity (e.g., Leiter & Maslach, 1988; Salvagioni et al., 2017; Taris, 2006). When translated into larger scale societal costs, these individual and organizational consequences impose a burden for entire healthcare systems and countries' economies (de Beer et al., 2022b).

While nobody disputes the harmful nature of burnout, interestingly, there is still an open debate about the most accurate conceptualization and operationalization of this occupational phenomenon (Demerouti et al., 2021; Schaufeli, 2021). To illustrate, the most frequently used measure of burnout—the Maslach Burnout Inventory (MBI)—defines burnout along three correlated dimensions: exhaustion (i.e., feelings of being depleted of one's resources for work), cynicism (i.e., feelings of being indifferent toward the job), and reduced professional efficacy (i.e., decreased sense of work-related achievement; Maslach et al., 2001). However, some researchers argue that reduced professional efficacy does not represent a constituting component of this work-related state (e.g., de Beer & Bianchi, 2019; Schaufeli & Taris, 2005). Rather, it seems more plausible that employees' sense of inefficacy and lack of achievement appear as a result of burnout (Taris et al., 2005). On the other hand, although extant studies indicate that burned-out employees suffer from deteriorated cognitive functioning, including impaired concentration, attention, and working memory (Deligkaris et al., 2014; Van Dam et al., 2012), a dimension reflecting cognitive impairment is not included neither in MBI, nor in the large proportion of other well-known burnout instruments (cf. Schaufeli et al., 2020b). This lack of agreement about what burnout precisely entails hampers high-quality research that could provide evidence-based guidelines for prevention, diagnosis, and treatment of this harmful phenomenon.

To combat these inconsistencies and provide a more comprehensive and accurate conceptualization of burnout, a group of researchers has recently developed a new definition and measure of burnout—the Burnout Assessment Tool (BAT; Schaufeli et al., 2020a, 2020b). The construction of the BAT combined inductive (i.e., in-depth interviews and review of existing burnout

questionnaires) and deductive approaches (i.e., utilizing extant theoretical framework of burnout [Schaufeli & Taris, 2005], to cluster, reduce, and interpret the compiled list of symptoms).¹ This procedure resulted with a final selection of four core burnout dimensions referring to exhaustion (i.e., a severe loss of physical and mental energy), mental distance (psychological withdrawal and detachment from one's work), emotional impairment (reduced ability to regulate one's emotional reactions, such as sadness and anger), and cognitive impairment (reduced ability to regulate one's cognitive processes, such as memory and attention; Schaufeli et al., 2020b). Accordingly, the authors proposed a new up-to-date definition of burnout describing it as “a work-related state of exhaustion that occurs among employees, which is characterized by extreme tiredness, reduced ability to regulate cognitive and emotional processes, and mental distancing.” (p. 4, Schaufeli et al., 2020b).

Given the need for an updated conceptualization of burnout, the accompanying measure BAT-23 has sparked a surging research interest resulting in validation studies conducted in various languages and cultural settings (e.g., the Dutch speaking Flemish region in Belgium, Netherlands, South Africa, Japan, Italy, Portugal, and Brazil, to name a few). The extant empirical evidence shows that the BAT-23 enables a valid and reliable assessment of employees' burnout complaints (e.g., de Beer et al., 2022b; Hadžibajramović et al., 2020; Sakakibara et al., 2020; Schaufeli et al., 2020b). More recently, researchers have proposed a shortened version of BAT—the BAT-12 (Hadžibajramović et al., 2022), which consists of 12 items measuring the four burnout dimensions. Conditional on demonstrating sound psychometric properties, shorter psychological instruments have several advantages over longer ones—they reduce respondents' attrition, fatigue, and boredom, which can yield responses of higher quality (Edwards, 2019).

In comparison to its longer counterpart, research validating the BAT-12 is still relatively scarce, especially validations conducted on nationally representative samples. Therefore, accounting for the advantages of shorter instruments, the present study aims to provide a comprehensive validation of the BAT-12 using a representative sample of the Croatian working population. More specifically, we examine the factor structure, measurement invariance, convergent and discriminant validity, and reliability of the Croatian version of BAT-12. Below we outline the aims of this study in greater detail and describe how addressing each contributes to the current validation process of the BAT-12.

1.1 | Validation of the Croatian version of BAT-12

The first aim of this study was to test the factor structure of the BAT-12. In doing so, we comport with Schaufeli et al. (2020b) according to

which burnout is a syndrome characterized by the simultaneous experience of several symptoms. Such conceptualization dictates the specification of a hierarchical model in which four dimensions of burnout represent indicators of one second-order factor (i.e., burnout syndrome). More specifically, the second-order factor is predicted to account for the covariation between four first-order factors (Brown, 2015), and as such, represents the reason why exhaustion, mental distance, cognitive and emotional impairment are related and refer to one underlying psychological condition (Schaufeli et al., 2020b). As a result, the BAT enables researchers and practitioners to calculate a single burnout score, as well as a score for each of its dimensions (Schaufeli et al., 2020a). Although a single burnout score facilitates the development of clinically validated cut-off values, previous multidimensional burnout questionnaires provide scores only for separate dimensions (e.g., MBI; Maslach et al., 2017).

In line with the presented conceptual underpinning, previous studies validating BAT-12 showed that the second-order model is supported on samples from Belgium and the Netherlands (Schaufeli et al., 2020a), Brazil and Portugal (Sinval et al., 2022), Ecuador (Vinueza-Solórzano et al., 2021), Italy (Mazzetti et al., 2022), Romania (Oprea et al., 2021), and South Africa (de Beer et al., 2022a). This study contributes to these research efforts by testing the proposed model for the first time in Croatia. In doing so, we provide a rigorous psychometric test by comparing the proposed model with several theoretically plausible alternative factorial structures (one-factor, four-factor, and bifactor model). To our knowledge, such a research approach in validating the BAT-12 has so far been undertaken only in one study (Mazzetti et al., 2022). Comporting with presented conceptual and empirical arguments, we hypothesize:

H1. The factorial structure of the Croatian version of BAT-12 is best represented by a hierarchical model consisting of four first-order factors (Exhaustion, Mental Distance, Cognitive Impairment, and Emotional Impairment) loading on one second-order factor (burnout syndrome).

The next aim was to examine the measurement invariance of the BAT-12 across gender, occupational type (i.e., manual and nonmanual workers), and age. We chose these three employee characteristics because comparing associations between variables (which is justified by at least partial invariance of factor loadings) or mean levels of burnout (which is justified on a latent level by at least partial invariance of intercepts) across gender, age, and occupational type is often of substantive research interest. Previous studies demonstrated that BAT-12 functions invariantly across countries (Hadžibajramović et al., 2022; Sinval et al., 2022), ethnicities (de Beer et al., 2022a), gender (de Beer et al., 2022a; Hadžibajramović et al., 2022; Sinval et al., 2022), time (before and during the COVID-12 pandemic; Mazzetti et al., 2022), and age (Hadžibajramović et al., 2022). However, to our knowledge, this study is the first to examine the measurement invariance of the BAT-12 (or BAT-23) across occupational type. In all, we hypothesize that:

H2. The Croatian version of the BAT-12 shows measurement invariance across gender, occupational type, and age.

The third aim was to demonstrate that the BAT-12 converges with other burnout measures, and at the same time has a unique contribution in measuring this construct as compared with extant measures. Therefore, we examined the convergent and discriminant validity of the BAT-12 vis-à-vis the Oldenburg Burnout Inventory (OLBI; Demerouti et al., 2003) by using a confirmatory factor analysis (CFA) multitrait-multimethod (MTMM) approach. According to OLBI, burnout consists of two dimensions—exhaustion (loss of affective, physical, and cognitive energy) and disengagement at work (distancing from one's work and negative attitudes towards the work; Demerouti et al., 2003). Extant studies demonstrated convergence of BAT-12 with MBI for employees from South Africa (de Beer et al., 2022a) and Romania (Oprea et al., 2021). We hypothesize similar with regards to OLBI:

H3. The Croatian version of the BAT-12 will partly, but not completely overlap with OLBI.

Finally, we also examined associations of the general score and four subdimensions of burnout with another measure of burnout (OLBI), job demands, job resources, and job outcomes, aiming to provide additional convergent and discriminant data at the manifest level. In case of external variables, we draw upon the job demands-resources (JD-R) model which outlines the expected nomological network of these constructs (Demerouti et al., 2001). According to the JD-R model, job demands can induce stress and instigate a health impairment process leading to strain (e.g., burnout; Bakker & Demerouti, 2007). In contrast, job resources “are functional in achieving work goals, reduce job demands and the associated physiological and psychological costs, or stimulate personal growth, learning, and development” (Bakker & Demerouti, 2007, p. 312). Accordingly, burnout, particularly its motivational component, may also stem from low levels of job resources (Schaufeli & Taris, 2014). Whereas job demands and resources are one of the most powerful antecedents of burnout, employee work-related behaviors (e.g., job performance) and attitudes (e.g., turnover intention) represent the proximal outcomes of burnout. Burnout relates positively with job demands and turnover intention (e.g., Sakakibara et al., 2020) and negatively with job resources and job performance (de Beer et al., 2020; Taris, 2006). In this study, we examine associations between burnout and diverse job demands (i.e., role ambiguity, work overload, interpersonal conflicts, and emotional demands), job resources (i.e., job autonomy, supervisor support, opportunities for learning), and job outcomes (turnover intention and in-role performance) by hypothesizing that:

H4. Burnout, measured with the Croatian version of the BAT-12, is positively associated with OLBI, job demands and employee turnover intention, and negatively associated with job resources and in-role performance.

2 | METHOD

2.1 | Procedure

To collect a representative sample of the Croatian working population, we employed a professional data collection agency. Cross-sectional dataset was collected via an online survey. The introduction to the survey included information about the study purpose and the voluntary nature of participation. Informed consent was obtained from all respondents. Each respondent received reward points replaceable for vouchers for completing the survey. The study was approved by the Ethics committee of the first author's institution.

2.2 | Participants

The sample was representative of the Croatian working population with regards to gender, age, region, and industrial sector (with a deviation of up to 5% from the population allowed in most cases²). The target population comprised adult persons who received a payment for their work regardless of the number of working hours and the type of employment arrangement (including employees affiliated with an organization, self-employed persons, part-time employees, entrepreneurs etc.). The agency used the most recent data available from the Croatian Bureau of Statistics (<https://dzs.gov.hr/en>) to establish the population quota. In total, 1011 participants were randomly drawn from the web panel³ in a way to adhere to the quota as precisely as possible, and comprised the 20.25% of all contacted. To detect careless respondents, we used several criteria in conjunction (Meade & Craig, 2012): careless number of identical responses in a row, psychometrical synonyms, Mahalanobis distance, and responses to bogus items. The exclusions were validated against the total time needed to finish the survey. As a result, we excluded 45 careless respondents yielding a final dataset of 966 employees. Slightly more than half of the participants were female (57.2%). Participants were 41.53 years old on average ($SD = 9.66$). Approximately half of them were highly educated (53.9%), 45.1% completed secondary, and 0.9% completed primary education. The vast majority of participants were permanently employed (86.9%) and most of them (58.3%) worked for private sector organizations. With regards to occupational type, the sample was composed of 81.7% nonmanual and 18.3% manual employees. Mean length of employees' current employment arrangement was 10.70 years ($SD = 9.50$). Participants were employed in 18 diverse industrial sectors with the five most represented being manufacturing (12.4%), education (12.1%), public administration and defense (11.1%), wholesale and retail (10.5%), and healthcare and social care system (8.1%).

2.3 | Measures

Burnout was measured with two measures. The first measure was a short form of the BAT, the BAT-12 (Hadžibajramović et al., 2022). To ensure a valid translation of BAT-12, we conducted a translation and back-translation. First, the English version was translated into

Croatian language by the four authors of the current study who are native in Croatian and proficient in the English language. Then a bilingual psychologist translated our Croatian translation back into English. The final step consisted of an iterative procedure through which the study authors and the bilingual psychologist compared and harmonized the original and back-translated versions. BAT-12 consists of three items per each subscale (Exhaustion, Mental Distance, Cognitive Impairment, Emotional Impairment). Respondents rated each item on a 5-point scale (1 = *never*; 2 = *rarely*; 3 = *sometimes*; 4 = *often*; 5 = *always*). The Croatian version of BAT-12 is presented in Appendix B. The second measure of burnout used in the current study was the OLBI (Demerouti et al., 2003). OLBI includes two dimensions—Exhaustion (8 items, e.g., “After my work, I regularly feel worn out and weary”; $\alpha = .81$, $\omega = 0.84^4$) and Disengagement (8 items; e.g., “I frequently talk about my work in a negative way”; $\alpha = .83$, $\omega = 0.87$). In contrast to BAT-12, OLBI consist of both negatively and positively worded items (e.g., “I always find new and interesting aspects in my work”; reversely coded item measuring disengagement). Positively worded items are recoded so that a higher score reflects higher levels of exhaustion and disengagement. Responses were indicated on a scale ranging from 1 (*strongly agree*) to 4 (*strongly disagree*).

2.3.1 | Job demands

To measure role ambiguity, we used three items from Bowling et al.'s (2017) Role Ambiguity Scale (e.g., “I am not sure what is expected of me at work.”; $\alpha = .77$, $\omega = 0.79$). Respondents indicated their responses on a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Work overload was measured with a four-item scale (e.g., “Do you have too much work to do?”; $\alpha = .81$, $\omega = 0.83$) from the Energy Compass, a questionnaire based on a JD-R model (Schaufeli, 2015). Items were scored on a scale ranging from 1 (*never*) to 5 (*always*). Interpersonal conflicts were measured using three items from the Interpersonal Conflicts at Work Scale (Spector & Jex, 1998). These items assessed how often employees experienced arguments with and rudeness from their coworkers (e.g., “How often do you get into arguments with others at work?”; $\alpha = .83$, $\omega = 0.85$). Responses were indicated on a 5-point scale ranging from 1 (*less than once per month or never*) to 5 (*several times per day*). To measure emotional demands, we used the respective three items ($\alpha = .81$, $\omega = 0.82$) from the third version of the Copenhagen Psychosocial Questionnaire (COPSOQ III; Burr et al., 2019). Items were scored on a 5-point scale ranging from 1 (*never*) to 5 (*always*). The sample item is “Does your work put you in emotionally disturbing situations?”. For each scale measuring job demands, translation to Croatian was conducted by the four study authors who are native in Croatian and proficient in the English language.

2.3.2 | Job resources

Job autonomy was measured using the existing translation of items from the Psychological Climate Questionnaire originally developed by

James and James (1989) and adapted by Tomas et al. (2019; $\alpha = .79$, $\omega = 0.80$). The sample item is "I have autonomy to make decisions within my own work area." Responses were indicated on a 5-point scale (1 = *completely disagree*; 5 = *completely agree*). Supervisor support and opportunities for learning were each measured with three items from the Energy Compass (Schaufeli, 2015; $\alpha = .89$, $\omega = 0.89$ and $\alpha = .84$, $\omega = 0.85$, respectively). Sample items were: "Can you count on your supervisor for help and support, when needed?" (supervisor support) and "In my work I always learn new things." (opportunities for learning). Items were scored on a 5-point scale ranging from 1 (*never*) to 5 (*always*) for supervisor support and from 1 (*completely disagree*) to 5 (*completely agree*) for opportunities for learning. With the exception of the scale measuring job autonomy, all remaining scales measuring job resources were translated by the same four study authors as noted above.

2.3.3 | Outcomes of burnout

Turnover intention was measured with a three-item scale developed by Irving et al. (1997; e.g., "I will probably be looking for a new job within a year"; $\alpha = .89$, $\omega = 0.9$). The items were assessed on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). To measure in-role performance, we used three items from the Williams and Anderson's (1991) scale (e.g., "I fulfil the responsibilities specified in my job description."; $\alpha = .73$, $\omega = 0.75$). Respondents provided responses on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). For both scales, Croatian translations were already available.

2.4 | Data analysis

To assess the factor structure of the instrument, we performed CFAs for the Croatian version of the BAT-12 using the robust MLR estimator. To evaluate model fit, we used the following cut-off criteria: comparative fit index (CFI) values over 0.95 and root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) values below 0.05 indicated good fit; while CFI values over 0.90 and RMSEA and SRMR below 0.08 were considered to indicate adequate model fit (Little, 2013).

Since the theoretical model describes a higher-order burnout syndrome with core symptoms as first-order factors (Schaufeli et al., 2020b), we followed the guidelines developed by Credé and Harms (2015) when specifying the models to be tested. We first compared the theoretical second-order model (SO) with a more parsimonious one (one-factor model, 1F), to see if it fits the data better, as well as two more complex alternative models [four correlated factors model (4F) and bifactor model (BF)] to see if the SO fits the data at least as well. We also took into account the fit of these models. Due to the results of these comparisons, we specified a modified version of the second-order model, including the theoretically substantiated covariance between the two first-order factors, and ran the comparisons again. In particular, the following models

were specified and compared using the scaled chi-square difference test for nested models (Satorra & Bentler, 2001):

1. A second-order model (SO), describing the theoretical structure with a second-order factor (burnout syndrome) explaining the covariation between the first-order factors (Exhaustion, Mental Distance, Emotional Impairment, and Cognitive Impairment).
2. A modified second-order model (modSO), identical to SO, but additionally including the covariance between the disturbances of the first-order factors Exhaustion and Mental Distance. These two subscales are considered to be core and inherently linked dimensions of burnout because both are necessary to describe the psychological process underlying burnout: psychological distance represents the (inefficient) coping strategy burned-out employees use to cope with feelings of exhaustion (Schaufeli & Taris, 2005; Schaufeli et al., 2020b). Such theoretical reasoning corresponds to the conceptual framework by Schaufeli and Taris (2005) which argues that employees' inability (i.e., exhaustion) and unwillingness (i.e., mental distance) to invest effort into work lie at the heart of the burnout syndrome. Therefore, we hypothesized that they could exhibit covariation even after controlling for overall burnout, i.e., that the part of variance they share (above what is shared between the four dimensions) is the inner core of the burnout symptom.
3. A one-factor model (1F), with all 12 items as indicators of a single latent dimension.
4. A correlated four-factor model (4F), with the hypothesized four core symptoms as latent dimensions. All items loaded only on their respective factor, and the correlations between the latent dimensions were freely estimated.
5. A bifactor model (BF), with four uncorrelated factors representing the core symptoms, as well as a general factor uncorrelated with the specific ones. This model allows for explicit decomposition of manifest variances into general and specific parts.

For the model chosen we tested measurement invariance. Gender and occupational type invariance was tested with multigroup CFA by comparing a sequence of nested models based on the differences in CFI, RMSEA, and McDonald's noncentrality index (Putnick & Bornstein, 2016). If full invariance was not achieved, partial invariance was tested by determining noninvariant parameters with a forward method using confidence intervals (Jung & Yoon, 2016). We chose the referent items based on previous research as to ease comparison with it (Vinueza-Solórzano et al., 2021). Age invariance was tested using local structural equation modeling (LSEM, Hildebrandt et al., 2016), an approach better suited than multigroup CFA or moderated factor analysis for continuous moderators for which an exact hypothesis on the nature of the moderation effect does not exist (Greiff & Scherer, 2018). LSEM includes testing a series of models, one for each focal point (in our case a different age). Each model is fitted to all the data but the observations in each model are weighted proportional to their distance from the focal point. We thus obtain information on the

variation in the model parameters across different ages. LSEM includes two significance tests based on permutations ($k = 1000$): a global test for each parameter that informs us whether there are significant deviations across age from the average value of the parameter; which can, when significant, be followed by pointwise tests for significance of the deviations of each focal point's value from the average (due to the large number of comparisons we used $p < .01$; Hildebrandt et al., 2016). A bandwidth factor $h = 2$ was used for LSEM in accordance with Hildebrandt et al.'s (2016) simulation study.

We then assessed the internal reliability of the global score and each subdimension of BAT-12 by calculating Cronbach's α and McDonald's ω total where appropriate.

To compare our measure with another burnout measure, we used the CFA MTMM approach with burnout dimensions as traits and BAT-12 and OLBI as methods, which allows us to conclude whether the compared measures measure the same, similar, or different traits (convergent and discriminant validity). CFA MTMM includes the comparison of a set of nested models (Kyriazos, 2018; Widaman, 1985). First, we specified the least restrictive of MTMM models—the correlated traits-correlated methods (CTCM). In this model each of 28 items loaded on two factors—the corresponding trait (Exhaustion, Mental Distance/Disengagement, Cognitive or Emotional Impairment) and the corresponding method factor (BAT or OLBI). The trait factors were allowed to covary between themselves, as well as the method ones, but method factors were unrelated to the trait factors. Other models were then specified and compared with CTCM. Second, we specified the no traits-correlated methods (NTCM) model, which differs from the CTCM in the fact that no trait factors exist. If NTCM fits the data better than CTCM, this would imply that the items reflect only the method used, and not the trait they were supposed to measure. Better fit of CTCM compared to NTCM implies the items measure common traits, i.e., that different measures of the same construct have a common factor, thereby informing us on the convergent validity of BAT-12 and OLBI. Third, we specified the correlated traits-no methods (CTNM) model, which differs from CTCM in the fact that no method factors exist. If CTNM fits the data better, this would indicate that there is no method variance, i.e., that it does not matter which instrument is used when measuring traits. Better fit of CTCM compared to CTNM means that there are some differences between what is measured with these two instruments, indicating discriminant validity across methods. Fourth, we compared the correlated traits-uncorrelated methods (CTUM), which differs from CTCM in the fact that methods factors do not covary. A better fit of CTUM to the data would mean that the methods used do not have anything in common. Better fit of CTCM compared to CTUM would indicate common method variance, i.e., convergent validity of BAT-12 and OLBI. In the fifth model we specified one common method factor across all items while keeping the trait factors, i.e., the correlated traits-perfectly correlated method (CTPCM) model. Better fit of CTPCM compared to CTCM indicates that methods cannot be differentiated, while better fit of CTCM indicates discriminant validity of the methods.

Finally, to further evaluate the discriminant and convergent validity of the instrument, correlations of the four core symptoms and the total BAT-12 score with OLBI scores, job demands, job resources, and job outcomes were computed.

Analyses were conducted with R language (R Core Team, 2022), packages *lavaan* (Rosseel, 2012), *psych* (Revelle, 2022) and *sirt* (Robitzsch, 2022).

3 | RESULTS

3.1 | Factor structure

Table 1 shows the fit parameters of the five specified models, as well as model comparisons between the chosen nested models following the logic outlined above. In absolute terms, we can see that 1F shows poor fit to the data, while all other models show good (according to CFI and SRMR values) or adequate fit (according to RMSEA values).

The theoretically expected second-order model SO showed significantly better fit compared to the more parsimonious 1F. This suggests that focusing on just a single dimension is not a valid strategy for dealing with BAT-12 scores. Compared to the more complex models, however, SO, although showing an adequate fit itself, demonstrated significantly worse model fit than both the 4F and the BF (Table 1).

Importantly, the difference between SO and 4F was very small in terms of compared fit indices. Still, our results suggest that for the Croatian version of BAT-12, four correlated factors representing the core burnout symptoms provide a better model. The only possible source of misfit between SO and 4F is a covariation between first-order factors not accounted for by the general burnout dimension. In other words, our results seem to suggest that there is some unexplained covariation between core symptoms of BAT-12, after controlling for the general burnout factor.

In line with this observation and the theoretical argumentation presented above, we specified a modified hierarchical model with correlated disturbances between Exhaustion and Mental Distance (modSO). As presented in Table 1, this model showed an improvement over the more parsimonious 1F, but also performed just as well as the more complex 4F. In addition, modSO provided significantly better fit compared to the original hierarchical model (SO). Note that modSO was not compared directly with the BF since they are not nested. However, information criteria presented in Table 1 hint at the fact that modSO fits the data better than BF. These results suggest that the modified higher-order model shows adequate fit and the best level of fit with minimum complexity among all tested models. Furthermore, modSO shows good model fit and high levels of factor saturation, both for first-order factors (factor loadings ranging from 0.54 to 0.84) and the second-order one (factor loadings ranging from 0.77 to 0.86). Accounting for these findings and comporting H_1 , modSO was chosen as a model on which further analyses were based. This justifies using the core symptoms subscales in addition to the general burnout scale score.

TABLE 1 Comparison of structure models.

Model	χ^2	df	Model fit			Comparison with SO			Comparison with modSO		
			CFI	RMSEA [90% CI]	SRMR	AIC	BIC	$\Delta\chi^2$	Δdf	$\Delta\chi^2$	Δdf
Second-order model (SO)	265.640***	50	0.958	0.067 [0.059–0.075]	0.042	24,393	24,530				
Modified second-order model (modSO)	238.994***	49	0.963	0.063 [0.055–0.072]	0.035	24,369	24,510	19,808***	1		
One-factor model (1F)	1121.222***	54	0.792	0.143 [0.136–0.150]	0.078	25,241	25,358	552.279***	4	584,882***	5
Four correlated factors model (4F)	238.793***	48	0.963	0.064 [0.056–0.072]	0.035	24,371	24,517	21,701***	2	0.179	1
Bifactor model (BF)	234.498***	42	0.962	0.069 [0.060–0.078]	0.041	24,378	24,554	25,758***	8		

Note: χ^2 = Chi-square statistic.

Abbreviations: CFI, comparative fit index; CI, confidence interval; df, degrees of freedom; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual. *** $p < .001$.

3.2 | Measurement invariance

To test the invariance regarding gender and occupational type, we compared the models suggested for testing the invariance of higher order models (Chen et al., 2005; Rudnev et al., 2018). Due to the established covariance between the first-order factors, we decided to additionally test the structural invariance of this relationship by testing first the invariance of the variances of all of the first-order factors, and in the final step, the model with the covariance between disturbances of Exhaustion and Mental Distance set to be equal across groups (Somaraju et al., 2022). There were no significant differences between different invariance models regarding gender. Accordingly, BAT-12 showed full loading and intercept invariance across gender for both first- and second-order factors, as well as invariance of the correlation between disturbances (Table 2). As for the occupational type, setting all item intercepts to be the same led to a significantly worse fit as indicated by the change in CFI (1st order intercepts model in Table 2). The forward method using confidence intervals indicated two intercepts as non-invariant between groups. However, the confidence interval of one of them was close to 0 [0.078, 0.411]. Therefore, we tested a model in which the intercept of only one item (EX3, *At work, I feel physically exhausted*; which was 0.496 in the manual group and -0.010 in the nonmanual) was allowed to vary between groups. This partial invariance model did not show worse fit according to change in CFI and RMSEA, nor did the following models (Table 2). Therefore, BAT-12 showed full loading invariance across occupational type for both first- and second-order factors, and almost full intercept invariance for both first- and second-order factors, except for the Exhaustion factor which achieved partial invariance (two out of three intercepts invariant).

Age invariance was tested via LSEM. Since large differences in the number of observations at each focal point (age) can lead to bias (Olaru et al., 2019), LSEM was performed on focal points (ages) for which the effective sample size (which takes into account the age-specific weights and number of participants) was larger than 100 (age: 23–59, $n = 933$). This resulted in 37 focal points (37 models), in which the sample size at each focal point varied from 5 to 45 and the effective sample size used in the models ranged from 120.44 to 382.74. The model fit did not differ between models with different ages as focal points: CFI ranged from 0.899 to 0.962, but although some of these do not indicate an acceptable fit, the global permutation test indicated they did not differ significantly from the average CFI of 0.945 ($p = .376$). The same was true for RMSEA (range: 0.064–0.110, $M = 0.078$, $p = .305$) and SRMR (range: 0.033–0.048, $M = 0.038$, $p = .558$). The global permutation tests indicated that the loadings of both first- and second-order factors were the same for all ages (see Appendices C and Da), as well as the variances of all first-order factors and the correlation between Exhaustion and Mental Distance factor (range: 0.164–0.556, $M = 0.426$, $p = .706$). However, both the global and pointwise test indicated that the intercepts of all factor indicators (except for reference variables) varied significantly across age (see Appendices C and Db). For the second-order factor intercepts were mostly higher than average in older age groups, and about average in respondents in their late thirties. Similar

TABLE 2 Comparison of invariance models regarding gender ($n_m = 413$; $n_f = 553$) and occupational type ($n_{\text{manual}} = 177$; $n_{\text{nonmanual}} = 789$).

Invariance model	Model fit					Model difference				
	χ^2	df	CFI	RMSEA [90% CI]	SRMR	$\Delta\chi^2$	Δdf	ΔCFI	$\Delta RMSEA$	ΔMc
Gender										
Configural	257.757***	98	0.963	0.063 [0.054, 0.073]	0.037					
1st order loadings	268.961***	106	0.962	0.061 [0.052, 0.071]	0.041	11.453	8	0.001	0.002	0.003
2nd order loadings	269.897***	109	0.963	0.060 [0.051, 0.069]	0.042	0.718	3	-0.001	0.001	-0.001
1st order intercepts	283.144***	117	0.962	0.059 [0.050, 0.067]	0.043	12.239	8	0.001	0.001	0.002
2nd order intercepts	290.313***	120	0.961	0.059 [0.050, 0.067]	0.044	7.159	3	0.001	0	0.002
1st order variances	287.421***	124	0.962	0.057 [0.048, 0.065]	0.044	0.292	4	-0.001	0.002	0.002
1st order covariance	287.648***	125	0.962	0.056 [0.048, 0.065]	0.044	0.441	1	0	0	0
Occupational type										
Configural	236.738***	98	0.968	0.059 [0.049, 0.068]	0.034					
1st order loadings	249.235***	106	0.967	0.057 [0.048, 0.067]	0.038	12.246	8	0.001	0.001	0.003
2nd order loadings	250.073***	109	0.968	0.056 [0.047, 0.065]	0.038	0.933	3	0	0.001	-0.001
1st order intercepts—partial ^a	276.902***	116	0.963	0.058 [0.049, 0.067]	0.039	28.421***	7	0.004	-0.002	0.010
1st order intercepts ^b	331.581***	117	0.951	0.066 [0.058, 0.075]	0.042	93.61***	8	0.016	-0.01	0.038
2nd order intercepts ^c	297.309***	119	0.960	0.060 [0.051, 0.069]	0.041	23.249***	3	0.004	-0.002	0.009
1st order variances	301.137***	123	0.959	0.059 [0.051, 0.068]	0.042	5.063	4	0	0.001	0.002
1st order covariance	302.790***	124	0.959	0.060 [0.051, 0.068]	0.042	1.647	1	0	0	0

Note: All models are compared to the model above them, except ^bwhich is compared to the 2nd order loadings and ^cwhich is compared to 1st order intercepts—partial. Model differences are calculated as the fit index of the parent model minus the fit index of the nested model.

Abbreviations: CFI, comparative fit index; CI, confidence interval; df, degrees of freedom; Mc, McDonald's noncentrality index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

^aThe intercept of EX3 *At work, I feel physically exhausted* varied across groups.

*** $p < .001$.

TABLE 3 Descriptives, reliability, and Pearson correlations for the study variables.

	Descriptives		Reliability		OLBI		Job demands				Job resources			Job outcomes	
	M	SD	α	ω (total)	Exh.	Dis.	RA	WO	IC	ED	JA	SS	OL	TI	IP
Burnout	2.06	0.58	-	0.92	0.741	0.627	0.537	0.412	0.360	0.327	-0.307	-0.408	-0.360	0.363	-0.241
Exhaustion	2.44	0.85	.83	0.84	0.721	0.506	0.406	0.466	0.322	0.356	-0.304	-0.370	-0.291	0.303	-0.114
Mental distance	2.10	0.78	.75	0.77	0.632	0.732	0.536	0.293	0.295	0.198	-0.316	-0.463	-0.504	0.450	-0.187
Cognitive impairment	2.01	0.63	.79	0.79	0.505	0.386	0.413	0.262	0.213	0.208	-0.169	-0.198	-0.159	0.187	-0.292
Emotional impairment	1.69	0.63	.76	0.78	0.477	0.341	0.360	0.270	0.320	0.277	-0.163	-0.237	-0.154	0.187	-0.215

Note: α was not calculated for the entire BAT-12, as the scale is not unidimensional. All of the observed correlations are significant at the $p < .01$ level.

Abbreviations: Dis, Disengagement (OLBI subscale); ED, Emotional Demands; Exh, Exhaustion (OLBI subscale); IC, Interpersonal Conflicts; IP, In-role Performance; JA, Job Autonomy; OL, Opportunities for Learning; RA, Role Ambiguity; SS, Supervisor Support; TI, Turnover Intention; WO, Work Overload.

Correlations between corresponding burnout measures between BAT and OLBI are presented in **bold**.

findings were observed for indicators of Mental Distance. For indicators of other first-order factors the trend was not clear (Appendix Db). Therefore, LSEM indicates the existence of loading invariance of BAT-12 across age, but lack of intercept invariance. In all, H_2 was partially supported.

3.3 | Reliability

Table 3 shows the reliability coefficients for the BAT-12 total score and the four subscales, which was similar to previous studies (e.g., de Beer et al., 2022a; Oprea et al., 2021).

3.4 | Convergent and discriminant data

The comparison of BAT-12 to OLBI was performed via CFA MTMM. The CTCM model showed adequate fit to the data and statistically better fit than all the other models (Table 4), thus implying (a) that there are traits common to the items regardless of the instrument they are measured with (CTCM vs. NTCM), but also that (b) these instruments differ in what they measure (CTCM vs. CTNM and vs. CTPCM), and although they differ, (c) there is a significant part of the method variance that is common (CTCM vs. NTCM). This is further corroborated by a correlation between the BAT and OLBI method factors of $r = .861$ ($p < .001$) in the best fitting CTCM model. All the loadings of BAT-12 were significant and between 0.2 and 0.7 for the method factor, 0.6 and 0.7 for the trait factors of Exhaustion, Cognitive and Emotional Impairment, and 0.3–0.4 for Mental Distance/Disengagement (Appendix E), probably due to the differences in the definition of this dimension between measures. The loadings of OLBI were between 0.3 and 0.8 for the method factor, and a bit lower for Exhaustion (0.1–0.7), while four of the items did not load significantly on Mental Distance/Disengagement (others being between 0.2 and 0.5, Appendix E). Thus, in line with H_3 , we can conclude that BAT-12 and OLBI measure the same constructs, as implied by variance common to items regardless of the instrument (a), common method variance (c), the size of the method factors correlation, and the Exhaustion loadings, indicating convergent validity; but also that they do not measure these constructs in the same way, as implied by existence of instrument specific variance (b), lack of perfect correlation between method factors, and significant loadings on the method factor, indicating a certain level of discriminant validity.

Finally, the correlations of the four subscales and the total score with the OLBI subscales, as well as job demands, job resources, and work-related outcomes are presented in Table 3. As suggested by H_4 , correlations between corresponding burnout measures were positive, and highest of all correlations. Also, higher values of burnout are reported by individuals perceiving more demanding jobs, those perceiving less job resources, and those exhibiting poorer job

outcomes (as indicated by their higher turnover intention and lower in-role performance). In general, a similar pattern can be observed for the core symptoms as well. Of the four symptoms, job demands show the strongest relations with Exhaustion, job resources and turnover intention with Mental Distance, and in-role performance with Cognitive Impairment.

4 | DISCUSSION

In the present study we have investigated the psychometric properties of the BAT-12, a short version of the new burnout measurement tool BAT-23. Using a representative sample of the Croatian working population, we found several indicators that support the validity of this instrument: (a) CFA supported the hypothesized hierarchical structure of the syndrome consisting of four primary symptoms; (b) the construct validity was additionally supported by measurement invariance tests which showed comparable loading structure across age, gender, and occupational type; (c) the convergent and discriminant validity of BAT-12 was demonstrated vis-à-vis an alternative burnout measure (OLBI), job demands, job resources, and job outcomes. Below, we discuss these general findings in greater detail highlighting how they contribute to the extant literature.

In addition to providing the first empirical test of the validity and reliability of the Croatian version of the questionnaire, the present study complements the findings of earlier validations of BAT-12 conducted in other countries. To our knowledge, only five empirical validations of the BAT-12 have been conducted to date (i.e., De Beer et al., 2022a; Mazzetti et al., 2022; Oprea et al., 2021; Sinal et al., 2022; Vinuesa-Solórzano et al., 2021), and this is the first study that meets two important methodological standards: (a) use of a representative sample; (b) basing the CFA results on testing five alternative models.

When testing the factorial structure of the Croatian version of BAT-12 (H_4) we have introduced a modified second-order model including a theoretically elaborated correlation between the disturbances of two

TABLE 4 Comparison of multitrait-multimethod (MTMM) models of burnout dimensions measured by BAT-12 and OLBI.

MTMM model	Model fit					Model difference			
	χ^2	df	CFI	RMSEA [90% CI]	SRMR	$\Delta\chi^2$	Δdf	ΔCFI	$\Delta RMSEA$
CTCM	1032.194***	315	0.937	0.051 [0.048, 0.055]	0.045				
NTCM	3009.332***	349	0.759	0.095 [0.092, 0.098]	0.078	1636.6***	34	-0.178	0.044
CTNM	1978.057***	344	0.852	0.075 [0.072, 0.078]	0.067	789.1***	29	-0.084	0.024
CTUM	1153.824***	316	0.925	0.056 [0.052, 0.059]	0.047	45.195***	1	-0.011	0.004
CTPCM	1186.539***	317	0.923	0.057 [0.053, 0.06]	0.042	78.118***	2	-0.014	0.005

Note: All the models are compared to CTCM.

Abbreviations: CI, confidence interval; CFI, comparative fit index; CTCM, Correlated traits-correlated methods; CTNM, Correlated traits-no methods; CTPCM, Correlated traits-perfectly correlated methods; CTUM, Correlated traits-uncorrelated methods; NTCM, No traits-correlated methods; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

*** $p < .001$.

first-order factors—Exhaustion and Mental Distance. As a result, we base the results on the comparison of five theoretical models of BAT-12. A similar approach was used only by Mazzetti et al. (2022), Oprea et al. (2021), and Schaufeli et al. (2020a). In particular, Schaufeli et al. (2020a) found that SO fitted the data as good as the 4F. In contrast, using nonprobabilistic convenience samples Mazzetti et al. (2022) and Oprea et al. (2021) found that the 4F fitted the data better than SO.⁵ The improvement of BF over SO is of similar magnitude. However, we cannot statistically discern whether 4F or BF fit the data better, since they are not nested and cannot be compared directly (Credé & Harms, 2015). Validation studies on BAT-23 did not face the choice between these two non-nested models because, although in some studies BF had better fit than SO (e.g., de Beer et al., 2022b), SO was in most studies preferred above 4F (e.g., Schaufeli et al., 2020b). So the main dissimilarity between the current BAT-12 validation and previous ones is this difference between SO and 4F, i.e., explicitly allowing for all the correlations between the first-order factors (4F) results in a better fit for the BAT-12.

By adopting a presented approach, the present study introduces additional theoretical elaboration of the latent structure of BAT-12. The model that we propose relies on earlier definitions of burnout, understanding it as a syndrome that is primarily described by the appearance of mental exhaustion followed by the symptom of psychological distancing from work (Schaufeli et al., 2020a, 2020b). Early on, Maslach and Jackson (1981) defined burnout as emotional exhaustion accompanied by depersonalization and reduced personal accomplishment. High emotional exhaustion and depersonalization and a low personal accomplishment was supposed to be indicative for burnout in the MBI, the first well-known burnout questionnaire. Since the JD-R model (Bakker & Demerouti, 2007; Bakker et al. 2023; Demerouti et al., 2001) gave burnout a central mediating role through which high demands and lack of resources at work lead to negative work outcomes, reduced personal accomplishment is omitted from the more recent definition of burnout and is treated as an outcome of the syndrome instead (Taris et al., 2005). However, this is by no means the case with the depersonalization symptom. Analyzing multidimensional measures of burnout that were developed later, Schaufeli et al. (2020b) concluded that they all include some measure of exhaustion and mental distancing. Theoretically, these two symptoms distinguish burnout from fatigue and make burnout a unique syndrome of impaired work-related well-being.

Accounting for these theoretical arguments, we added and found a positive correlation between the disturbances of Exhaustion and Mental Distance (i.e., parts of their variances that were unexplained by the second-order factor), after establishing that the initial SO fitted the data worse than the 4F and the BF. The modSO did not differ significantly from the 4F but was more parsimonious than this model. Importantly, it also more closely aligned with the theoretical underpinning of the BAT-12 in comparison to the competing models. Accordingly, the modSO shows the best level of fit with minimum complexity between all tested models.

Invariance testing (H_2) indicated that the modSO had a comparable loading structure across age, gender, and occupational

type, thus justifying the comparison of relations across these variables. Full intercept invariance across gender also justified the comparison of latent and manifest means across this sample characteristic. However, the item “At work, I feel physically exhausted” contributed differently to the mean of the Exhaustion factor of manual and nonmanual workers, and via that probably to the second-order burnout factor. This is probably due to the nature of job tasks of manual workers which are usually more physically demanding, therefore contributing to exhaustion, which might contribute less to the level of exhaustion in nonmanual work. According to the classification of the sources of invariance by Somaraju et al. (2022), this could be classified as different group experiences as sources of invariance. Since only one intercept was noninvariant, authors should be careful when comparing the manifest levels of Exhaustion and Burnout of manual and nonmanual workers, but the analyses of the latent means regarding occupational type are justified for all factors since partial invariance was achieved (Steinmetz, 2013). For item intercepts across different ages, invariance was not achieved. Therefore, caution is required when comparing burnout dimensions across age. Differences (or lack thereof) could be the result of this item contributing differently to the mean level of the scale, and not of an actual difference in the burnout. Since there is no clear pattern regarding age, it is hard to judge what the sources of this invariance could be. Possible candidates could be different norms or frame-of-reference across different ages, as well as different experiences and differences in developmental stage (Somaraju et al., 2022). These results add to the existing literature by providing one of the first empirical evidence on the invariance regarding occupational type and continuous age.

The reported findings corresponding to the H_1 and H_2 complement previous studies testing the latent structure of the BAT-12, which are still relatively scarce. With the exception of previously described studies testing the alternative latent structures of BAT-12 (Mazzetti et al., 2022; Oprea et al., 2021; Schaufeli et al., 2020a), other studies either did not test alternative measurement models (de Beer et al., 2022a; Sinval et al., 2022), or they compared a second-order model to a one-factor model only (Vinueza-Solorzano et al., 2021). In summary, our results support the usage of the total score of the Croatian version of the BAT-12 questionnaire, as well as the score for each of the four symptoms. Consequently, the results of the study have practical relevance, giving the green light to the application of the BAT-12 on the Croatian population. Future research should further test the latent structure of the BAT-12 by comparing the hypothesized higher-order model with alternative models and test its invariance across different demographic groups (e.g., age, employment contract). Age invariance should be rechecked on a larger sample, more representative of younger and older individuals, because the size of certain age groups, specifically the extreme ones, were small in the current study. Analyzing age as continuous moderator is less often used in research than using age groups. However, treating age as a continuous variable has the advantage of avoiding arbitrary categorization and gives more detailed information on the nature of relationships (Greiff &

Scherer, 2018). For future research that decides using different age groups, our results can be informative for establishing age cut-off points.

To test the convergent and discriminant validity of the BAT-12 relative to other commonly used burnout questionnaires (H_3), the present study used both the CFA MTMM approach and correlations to examine the relationship between the BAT-12 and OLBI (Demerouti et al., 2003). We are aware of only two validation studies comparing BAT-12 and other burnout questionnaires (de Beer et al., 2022a; Oprea et al., 2021). However, none of these studies used the MTMM approach. In this study, we have chosen the OLBI that includes two symptoms of burnout: exhaustion and disengagement as a negative and detached response to one's job (Demerouti & Bakker, 2008; Demerouti et al., 2003). OLBI was chosen because it is multidimensional (hence, burnout is not exclusively limited to exhaustion), has sufficient data on its psychometric quality (cf. Schaufeli et al., 2020a), is most commonly used (together with MBI; Maslach, & Jackson, 1981), and is publicly available for free (like BAT-23 and BAT-12). Using CFA MTMM, Schaufeli et al. (2020a) demonstrated the convergence of the BAT-23 with the OLBI and MBI regarding exhaustion and mental distance, but also found sufficient divergence to conclude that each instrument independently contributes to the measurement of burnout.

In our case, the BAT-12 and OLBI had traits in common, but did not represent one common trait. This supports the notion that both instruments measure burnout, but BAT-12 adds to the field by representing a more complex measure of burnout including emotional and cognitive impairment. However, the convergent and discriminant validity of the BAT-12 questionnaire in relation to other questionnaires should be further explored. Recently, de Beer et al. (2022a) and Oprea et al. (2021) reported that the BAT-12 showed appropriate convergent and discriminative validity in relation to the MBI, based on bivariate correlations between the dimensions of the BAT-12 and its total score and the MBI score. When explored this way, our data again show evidence of validity, with two dimensions of BAT-12 which are parallel to the dimensions of OLBI (i.e., Exhaustion and Mental Distance) as well as the total score of BAT-12 being related with OLBI more strongly than with other variables. The two remaining BAT-12 dimensions (i.e., Cognitive and Emotional Impairments) show lower correlations with the OLBI dimensions, and they are in the ranges of the observed bivariate correlations with job demands and resources.

Finally, based on JD-R theory (Bakker & Demerouti, 2017; Schaufeli & Bakker, 2004), correlations between the BAT-12 and job demands, job resources, and job outcomes were analyzed as additional convergent and discriminant data (H_4). The correlations between the measures of the four core symptoms and the total score of the BAT-12 and selected job demands, job resources, and job-related outcomes supported the validity of the BAT-12. Relevant to the multidimensional definition of burnout, however, some indicative differences were found between the correlations. Exhaustion was most strongly related to job demands, mental distance to job resources as well as turnover intention, and cognitive impairment

to in-role performance. These results indirectly support the interpretation of burnout according to social exchange theory (Robinson & Rousseau, 1994; Taris et al., 2005), emphasizing the significant and divergent role of job demands and job resources in the appearance of burnout. While job demands primarily influence the experience of exhaustion, the employee's perception of a lack of job resources may play a dominant role in the experience of related mental distancing that could lead to employee turnover. Vinuela-Solórzano et al. (2021) found a similar pattern of correlations: job demands were more strongly related to exhaustion than mental distance, but social and content job resources were approximately equally strongly related to both core symptoms. While previous research has focused on demonstrating burnout as a higher-order factor defined by four core symptoms, future research could contribute to knowledge about the incremental validity of the four core dimensions of BAT.

4.1 | Limitations and suggested research lines

Several notes with regard to our results and interpretations should be given. First, concerning the sample composition presented in the Methods section, the structure of the study sample by age, gender, region, and industry, appears to approximate reasonably well Croatian working population. However, the percentage of highly educated participants by far exceeds the figures expected in the population. Albeit this observation is generally not surprising, this might have influenced our results and conclusions to the extent that the level and the effects of burnout vary depending on the education level. Second, the response rate in this study is relatively low, which may raise concerns about the potential bias in our sample with regards to burnout levels (i.e., burned-out employees may be less prone to fill out the surveys). However, the mean levels of burnout obtained in our sample are comparable to the mean burnout levels of healthy employees reported in other studies (Schaufeli et al., 2023). Therefore, we deem this explanation less plausible than alternative explanations. One potential reason for the obtained response rate is that participants had to fit into a relatively complex combination of population quota. As a result, in some cases the agency was compelled to invite participants who are generally less inclined to participate in various surveys.

Third, we are aware that the proposed hierarchical structure of the BAT-12 scores suggested by our results may not be the only possible solution of the data composition examined and that at least two other competing models offer similarly acceptable yet conceptually different description of the data. Moreover, given the empirical character of the finally accepted second-order latent structure, including a residual covariance parameter added at the first-level, the described solution should certainly be evaluated on new samples.

Fourth, in terms of relevant validity evidence, the key information provided by this study seems to be contained in the structure of correlations between the BAT-12 and its subdimensions with theoretically related sets of variables reflecting job demands, job resources, and job outcomes. The similarity of the pattern, signs and

magnitude of these relationships in our and previous validation studies (e.g., de Beer et al., 2022a), suggests that the Croatian adaptation is equivalent to the original version of the BAT. Nevertheless, bearing in mind the recent conceptualization of the burnout concept, we suggest that future validation studies additionally include measures of workplace stress and specific coping strategies, as these are more proximal correlates of burnout than the examined JD-R concepts.

5 | CONCLUSIONS

The validation data presented in this study on the Croatian version of the BAT-12 appear to offer robust support for the overall psychometric quality of this instrument in line with the underlying conceptualization of the burnout construct. The expected hierarchical structure of the instrument was generally confirmed by the exploratory and confirmatory tests performed on the Croatian employee sample data. Moreover, the obtained latent structure was shown to be invariant across age, gender, and occupation-type related groups, suggesting a reasonably high level of score comparability over the groups defined by the basic sociodemographic features examined. Furthermore, the comparisons with OLBI scores offered convincing indices of convergent and discriminant validity of the BAT-12 subscale scores. The highest correlations were found for BAT-12 Exhaustion and Mental Distance scores with corresponding OLBI subscale scores and significant but lower values for all other intercorrelations in the matrix. Finally, additional validity evidence was provided by the pattern of relationships obtained between the BAT-12 syndrome and its 4 core symptoms with a set of variables selected to represent the nomological network of the burnout construct according to the JD-R model. Significant correlations of expected size and sign were observed between the BAT-12 subscales and selected job demands, job resources, and job-related outcome measures. Overall, the findings presented on the Croatian version of BAT-12 add new evidence to the current validation process of the BAT-12 supporting its usage for scientific and applied purposes.

ACKNOWLEDGMENTS

The project was funded by the University of Zagreb and Research Group Work & Organisational Psychology—KU Leuven.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

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

ORCID

Jasmina Tomas  <http://orcid.org/0000-0002-7945-5328>

Darja Maslić Seršić  <http://orcid.org/0000-0003-4968-9243>

Una Mikac  <http://orcid.org/0000-0001-9369-6462>

Blaž Rebernjak  <http://orcid.org/0000-0002-4674-3083>

Vesna Buško  <http://orcid.org/0000-0001-8644-0688>

Hans De Witte  <http://orcid.org/0000-0002-6691-517X>

ENDNOTES

- ¹ For a more comprehensive overview of the development of the BAT, see Schaufeli et al. (2020b).
- ² See Appendix A for a more detailed account of the deviations between the study sample and the population. In a few categories referring to gender (male and female), age (35–39 age category), and region (two Croatian regions) we accepted deviations ranging between 5% and 10%. Accordingly, the observed deviations in our study correspond to the “soft” selection criterion (cf. Schaufeli et al., 2020a).
- ³ A 75% of the panel was randomly drawn from the population. Additional recruitment was conducted for less accessible demographic groups (i.e., less educated and older respondents) through public calls in the media and targeted recruitment. The panel includes 31,185 respondents (of 4,105,267 Croatian residents and 1,635,034 employed persons in Croatia at the time of data collection; <https://podaci.dzs.hr/2022/hr/29224>).
- ⁴ All reliability coefficients reported for each scale refer to coefficients obtained in this study.
- ⁵ Even though Oprea et al. (2021) do not report explicit model comparisons.

REFERENCES

- 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. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology*, 22(3), 273–285. <https://doi.org/10.1037/ocp0000056>
- 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, 25–53. <https://doi.org/10.1146/annurev-orgpsych-120920-053933>
- de Beer, L. T., & Bianchi, R. (2019). Confirmatory factor analysis of the Maslach Burnout Inventory: A Bayesian structural equation modeling approach. *European Journal of Psychological Assessment*, 35(2), 217–224. <https://doi.org/10.1027/1015-5759/a000392>
- de Beer, L. T., Schaufeli, W., & Bakker, A. (2022a). Investigating the validity of the short form Burnout Assessment Tool: A job demands-resources approach. *African Journal of Psychological Assessment*, 4(0), a95. <https://doi.org/10.4102/ajopa.v4i0.95>
- de Beer, L. T., Schaufeli, W. B., & De Witte, H. (2022b). The psychometric properties and measurement invariance of the Burnout Assessment Tool (BAT-23) in South Africa. *BMC Public Health*, 22 Article 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., Seubert, C., Bosak, J., Sinval, J., & Rudnev, M. (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>
- Bowling, N. A., Khazon, S., Alarcon, G. M., Blackmore, C. E., Bragg, C. B., Hoepf, M. R., Barelka, A., Kennedy, K., Wang, Q., & Li, H. (2017). Building better measures of role ambiguity and role conflict: The validation of new role stressor scales. *Work & Stress*, 31(1), 1–23. <https://doi.org/10.1080/02678373.2017.1292563>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

- Burr, H., Berthelsen, H., Moncada, S., Nübling, M., Dupret, E., Demiral, Y., Oudyk, J., Kristensen, T. S., Llorens, C., Navarro, A., Lincke, H.-J., Bocéréan, C., Sahan, C., Smith, P., Pohrt, A., & international COPSOQ Network. (2019). The third version of the Copenhagen Psychosocial Questionnaire. *Safety and Health at Work*, 10(4), 482–503. <https://doi.org/10.1016/j.shaw.2019.10.002>
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7
- Credé, M., & Harms, P. D. (2015). 25 years of higher-order confirmatory factor analysis in the organizational sciences: A critical review and development of reporting recommendations. *Journal of Organizational Behavior*, 36(6), 845–872. <https://doi.org/10.1002/job.2008>
- Van Dam, A., Keijsers, G. P. J., Eling, P. A. T. M., & Becker, E. S. (2012). Impaired cognitive performance and responsiveness to reward in burnout patients: Two years later. *Work & Stress*, 26(4), 333–346. <https://doi.org/10.1080/02678373.2012.737550>
- Deligkaris, P., Panagopoulou, E., Montgomery, A. J., & Masoura, E. (2014). Job burnout and cognitive functioning: A systematic review. *Work & Stress*, 28(2), 107–123. <https://doi.org/10.1080/02678373.2014.909545>
- Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *The Journal of applied psychology*, 86(3), 499–512. <https://doi.org/10.1037/0021-9010.86.3.499>
- Demerouti, E., Bakker, A. B., Peeters, M. C. W., & Breevaart, K. (2021). New directions in burnout research. *European Journal of Work and Organizational Psychology*, 30(5), 686–691. <https://doi.org/10.1080/1359432X.2021.1979962>
- Demerouti, E., Demerouti, E., Bakker, A. B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments. *European Journal of Psychological Assessment*, 19(1), 12–23. <https://doi.org/10.1027/1015-5759.19.1.12>
- Edwards, J. R. (2019). Response invalidity in empirical research: Causes, detection, and remedies. *Journal of Operations Management*, 65(1), 62–76. <https://doi.org/10.1016/j.jom.2018.12.002>
- Greiff, S., & Scherer, R. (2018). Still comparing apples with oranges? Some thoughts on the principles and practices of measurement invariance testing. *European Journal of Psychological Assessment*, 34(3), 141–144. <https://doi.org/10.1027/1015-5759/a000487>
- Hadžibajramović, E., Schaufeli, W., & De Witte, H. (2020). A Rasch analysis of the Burnout Assessment Tool BAT. *PLoS One*, 15(11), Article e0242241. <https://doi.org/10.1371/journal.pone.0242241>
- 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), 560. <https://doi.org/10.1186/s12889-022-12946-y>
- Hildebrandt, A., Lüdtke, O., Robitzsch, A., Sommer, C., & Wilhelm, O. (2016). Exploring factor model parameters across continuous variables with local structural equation models. *Multivariate Behavioral Research*, 51(2–3), 257–258. <https://doi.org/10.1080/00273171.2016.1142856>
- Irving, P. G., Coleman, D. F., & Cooper, C. L. (1997). Further assessments of a three-component model of occupational commitment: Generalizability and differences across occupations. *Journal of Applied Psychology*, 82(3), 444–452. <https://doi.org/10.1037/0021-9010.82.3.444>
- James, L. A., & James, L. R. (1989). Integrating work environment perceptions: Explorations into the measurement of meaning. *Journal of Applied Psychology*, 74(5), 739–751. <https://doi.org/10.1037/0021-9010.74.5.739>
- Jung, E., & Yoon, M. (2016). Comparisons of three empirical methods for partial factorial invariance: Forward, backward, and factor-ratio tests. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(4), 567–584. <https://doi.org/10.1080/10705511.2015.1138092>
- Kyriazos, T. A. (2018). Applied psychometrics: The application of CFA to multitrait-multimethod matrices (CFA-MTMM). *Psychology*, 09(12), 2625–2648. <https://doi.org/10.4236/psych.2018.912150>
- Leiter, M. P., & Maslach, C. (1988). The impact of interpersonal environment on burnout and organizational commitment. *Journal of Organizational Behavior*, 9(4), 297–308. <https://doi.org/10.1002/job.4030090402>
- Little, T. (2013). *Longitudinal structural equation modeling*. Guilford press.
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, 2(2), 99–113. <https://doi.org/10.1002/job.4030020205>
- Maslach, C., Leiter, M. P., & Jackson, S. E. (2017). *Maslach Burnout Inventory* (4th ed.). Mind Garden Inc.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, 52, 397–422. <https://doi.org/10.1146/annurev.psych.52.1.397>
- 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), Article 8562. <https://doi.org/10.3390/ijerph19148562>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Olaru, G., Schroeders, U., Hartung, J., & Wilhelm, O. (2019). Ant colony optimization and local weighted structural equation modeling. A tutorial on novel item and person sampling procedures for personality research. *European Journal of Personality*, 33(3), 400–419. <https://doi.org/10.1002/per.2195>
- 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>
- 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>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL <https://www.R-project.org/>
- Revelle, W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University. R package version 2.2.9. <https://CRAN.R-project.org/package=psych>
- Robinson, S. L., & Rousseau, D. M. (1994). Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior*, 15(3), 245–259. <https://doi.org/10.1002/job.4030150306>
- Robitzsch, A. (2022). *sirt: Supplementary Item Response Theory Models*. R package version 3.12–66. <https://CRAN.R-project.org/package=sirt>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., & Zick, A. (2018). Testing measurement invariance for a second-order factor: A cross-national test of the alienation scale. *Methods, Data, Analyses: A Journal for Quantitative Methods and Survey Methodology (mda)*, 12(1), 47–76. <https://doi.org/10.12758/mda.2017.11>
- Sakakibara, K., Shimazu, A., Toyama, H., & Schaufeli, W. B. (2020). Validation of the Japanese version of the Burnout Assessment Tool. *Frontiers in Psychology*, 11, Article 1819. <https://doi.org/10.3389/fpsyg.2020.01819>
- Salvagioni, D. A. J., Melanda, F. N., Mesas, A. E., González, A. D., Gabani, F. L., & Andrade, S. M. (2017). Physical, psychological and occupational consequences of job burnout: A systematic review of

- prospective studies. *PLoS One*, 12(10), Article e0185781. <https://doi.org/10.1371/journal.pone.0185781>
- Satorra, A., & Bentler, P. M. (2001). A scaled difference Chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514. <https://doi.org/10.1007/BF02296192>
- Schaufeli, W. (2021). The burnout enigma solved? *Scandinavian Journal of Work, Environment & Health*, 47(3), 169–170. <https://doi.org/10.5271/sjweh.395010.5271/sjweh.3950>
- Schaufeli, W. B. (2015). Engaging leadership in the job demands-resources model. *Career Development International*, 20(5), 446–463. <https://doi.org/10.1108/CDI-02-2015-0025>
- 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(3), 293–315. <https://doi.org/10.1002/job.248>
- Schaufeli, W. B., Desart, S., & De Witte, H. (2020b). 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., & Taris, T. W. (2005). The conceptualization and measurement of burnout: Common ground and worlds apart. *Work & Stress*, 19(3), 256–262. <https://doi.org/10.1080/02678370500385913>
- Schaufeli, W. B., & Taris, T. W. (2014). A critical review of the job demands-resources model: Implications for improving work and health. In G. F. Bauer & O. Hämmig, (Eds.), *Bridging occupational, organizational and public health: A transdisciplinary approach* (pp. 43–68). Springer Science + Business Media. <https://doi.org/10.1007/978-94-007-5640-3>
- Schaufeli, W. B., De Witte, H., & Desart, S. (2020a). *Manual Burnout Assessment Tool (BAT) - Version 2.0*. KU Leuven.
- Schaufeli, W. B., De Witte, H., Hakonen, 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>
- 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>
- Somaraju, A. V., Nye, C. D., & Olenick, J. (2022). A review of measurement equivalence in organizational research: What's old, what's new, what's next? *Organizational Research Methods* 25, (4) 741–785. <https://doi.org/10.1177/109442812111056524>
- Spector, P. E., & Jex, S. M. (1998). Development of four self-report measures of job stressors and strain: Interpersonal conflict at work scale, organizational constraints scale, quantitative workload inventory, and physical symptoms inventory. *Journal of Occupational Health Psychology*, 3(4), 356–367. <https://doi.org/10.1037//1076-8998.3.4.356>
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology*, 9(1), 1–12. <https://doi.org/10.1027/1614-2241/a000049>
- Taris, T. W. (2006). Is there a relationship between burnout and objective performance? A critical review of 16 studies. *Work & Stress*, 20(4), 316–334. <https://doi.org/10.1080/02678370601065893>
- Taris, T. W., Le Blanc, P. M., Schaufeli, W. B., & Schreurs, P. J. G. (2005). Are there causal relationships between the dimensions of the Maslach Burnout Inventory? A review and two longitudinal tests. *Work & Stress*, 19(3), 238–255. <https://doi.org/10.1080/02678370500270453>
- Tomas, J., Maslić Seršić, D., & De Witte, H. (2019). Psychological climate predicting job insecurity through occupational self-efficacy. *Personnel Review*, 48(2), 360–380. <https://doi.org/10.1108/PR-05-2017-0163>
- Vinueza-Solórzano, A. M., Portalanza-Chavarría, C. A., de Freitas, C. P. P., Schaufeli, W. B., De Witte, H., Hutz, C. S., & Souza Vazquez, A. C. (2021). The Ecuadorian version of the Burnout Assessment Tool (BAT): Adaptation and validation. *International Journal of Environmental Research and Public Health*, 18(13), Article 7121. <https://doi.org/10.3390/ijerph18137121>
- Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9(1), 1–26. <https://doi.org/10.1177/014662168500900101>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: 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>