

Original Research Article

Cross-cultural adaptation and psychometric evaluation of the Malayalam version of the Oldenburg burnout inventory among secondary school teachers

Ritwik Sivadas, Jissa Vinoda Thulaseedharan*

Achutha Menon Centre for Health Science Studies, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India

Received: 13 May 2026

Revised: 21 July 2026

Accepted: 24 July 2026

*Correspondence:

Dr. Jissa Vinoda Thulaseedharan,

E-mail: jissa@sctimst.ac.in

Copyright: © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: The Oldenburg burnout inventory (OLBI) is a widely used instrument for measuring burnout. It consists of 16 items grouped into two dimensions: exhaustion and disengagement. This study aimed to translate and validate the English version of the OLBI into Malayalam among secondary school teachers in Kerala.

Methods: After obtaining permission from the author, we translated the OLBI into Malayalam and validated it following standard guidelines, including forward translation, back translation, and content validation with cultural adaptation. Test-retest reliability was assessed among 40 teachers using SPSS version 28. The internal consistency, exploratory and confirmatory factor analyses were performed in STATA/SE 17.0 software, using data from 255 teachers.

Results: The OLBI-Malayalam version retained 16 items, all of which demonstrated excellent item-content validity indices (0.83-1.00) and a scale-content validity index of 0.97. The scale demonstrated good internal consistency, with Cronbach's α values of 0.84 for the total scale, 0.80 for exhaustion, and 0.61 for disengagement. Exploratory factor analysis (EFA) suggested a one-factor model that explained 31.2% of the variance. Confirmatory factor analysis (CFA) indicated no significant difference between the one-factor and two-factor models, with moderately acceptable fit indices.

Conclusions: The translated version of the OLBI offers a validated and culturally adapted tool for assessing burnout among Malayalam-speaking populations. The tool demonstrated a near-stable one-factor structure over the originally conceptualized two-factor model. It holds potential for use in other occupational groups, though further validation is recommended to account for differing job demands and work environments.

Keywords: Burnout, Oldenburg burnout inventory, Reliability, Teachers, Translation, Validity

INTRODUCTION

In recent years, stress, anxiety, and depression have become increasingly common, often attributed to the fast-paced, high-pressure environments of modern life.^{1,2} Burnout, a related but distinct occupational phenomenon, has gained increasing attention due to its adverse impacts on personal and professional well-being. Defined in the ICD-11 as a syndrome resulting from chronic workplace stress that has not been successfully managed, burnout is

specifically related to work environments and typically improves with time off or workplace changes.³

First described by Freudenberg in the 1970s as a social problem, burnout was later conceptualized by Christina Maslach as a prolonged response to chronic, emotional and interpersonal stressors at work. Exhaustion, cynicism, and a sense of inefficacy are the main characteristics of burnout, which develops progressively through the interplay of various workplace factors and manifests in

physical, emotional, and behavioural symptoms.^{4,5} The Maslach burnout inventory (MBI), developed by Christina Maslach and Susan Jackson, remains the most widely used tool to assess burnout. It has a strong theoretical foundation, with three dimensions such as: emotional exhaustion, depersonalization, and reduced personal accomplishment.⁶

However, to overcome certain limitations of the MBI, the Oldenburg burnout inventory (OLBI) was developed and validated by Demerouti and Bakker (1999) in German, and later it was translated into English and many other languages. Compared to MBI, OLBI evaluates two core dimensions, exhaustion and disengagement, based on the Job demands-resources (JD-R) model, using both positively and negatively worded items. Exhaustion refers to the physical, cognitive, and affective fatigue resulting from prolonged exposure to job demands, whereas disengagement refers to distancing from one's work, characterized by negative attitudes and a reduced commitment.⁷ As a freely available tool, OLBI was validated across various professions and cultures, including healthcare professionals, teachers, lawyers, Information Technology professionals, and other diverse occupations. Its use in these diverse occupational groups demonstrated how different occupational demands contribute to burnout in distinct ways.

As elsewhere in India, teachers in Kerala manage a broad spectrum of academic, administrative, and extracurricular duties, making them vulnerable to burnout due to heavy workloads, limited resources, and emotional strain. However, research on burnout among teachers is limited, and a validated tool in the regional language is lacking to assess it. Given this, the researchers aimed to translate the English version of OLBI into Malayalam and evaluated its validity, reliability, and cultural relevance for assessing burnout among secondary school teachers in Kerala. The paper presents findings from the translation and validation process of the OLBI tool in the Malayalam language.

METHODS

Type of the study

This was a cross-sectional methodological study involving cross-cultural adaptation and psychometric evaluation of the Malayalam version of OLBI among secondary school teachers.

Study duration

The study was conducted from August 2024 to June 2025. The initial translation and cross-cultural adaptation of the instrument were completed between August and December 2024. Data collection among secondary school teachers was carried out from January to March 2025. Psychometric validation, data analysis, and report writing were performed between March and April 2025. The

methods used in this study are described in detail in the following sections.

Part-1 scale adaptation, translation, and initial validation of OLBI

The English version of the 16-item OLBI is evenly divided into two dimensions, such as Exhaustion and Disengagement, each of which contains positively and negatively worded items on a 4-point Likert scale, with negative items reverse-scored (1=4, 2=3, 3=2, 4=1). Higher scores reflect higher burnout.

The process began after obtaining permission from the original author of OLBI. Initially, the OLBI was translated into Malayalam by three independent translators, and a reconciled version (version 1) was prepared, followed by a back-translation of the same by two independent translators unfamiliar with the original. The authors reviewed the version 1 by comparing the back-translation with the original OLBI and made necessary modifications to improve clarity, resulting in version 2. The translators and back-translators were selected purposively based on their language proficiency and translation experience.

For content validity assessment, nine experts were invited via personal contacts based on their expertise and experience in the relevant subject area, instrument development, or psychometric evaluation. Of these, six with a public health, psychology, and social work background, including two with prior scale validation experience, participated. Google forms were used to capture the evaluation of the experts. English and Malayalam versions of each item were presented alongside five structured questions to assess conceptual and cultural equivalence. The questions asked were: "Does the translated statement indicate the same as that of the original English statement?", "If 'no,' please explain why it is different and suggest how it should be modified?", "Is there a chance of cultural variation in meaning?", "If yes, can you explain the possible cultural variation and how this can be addressed or resolved?", and "overall, how would you rate the relevance of the translation?" respectively. Experts rated the relevance of translation on a 4-point scale (1 = not relevant, 2=somewhat relevant, 3= quite relevant, and 4 = highly relevant).

The item-content validity index (I-CVI), probability of chance agreement (Pc), modified Kappa, and scale-level content validity index (S-CVI) were calculated. The S-CVI is the Sum of I-CVI scores divided by the Number of items(S-CVI/Ave), and a value of 0.90 or higher was considered acceptable.⁸⁻¹⁰ Based on these indicators and expert feedback, the authors revised the second version, resulting in version 3.

Cognitive interviewing was employed to ensure the clarity and relevance of the translated items from the

perspective of the target population. In-person interviews were conducted with four eligible teachers (both male and female) selected purposively from the target population, all of whom had more than 10 years of teaching experience. The researcher read each item aloud and asked probing questions such as, “what did you understand when I asked this question?” and “can you help me understand the reason for your response?” Participants restated items in their own words and provided feedback on clarity and meaning. Based on their input, additional revisions were made, resulting in the final Malayalam-translated field version (version 4).

Part-2 reliability assessment and validation of the translated scale

Further, the reliability and validity of the translated scale were assessed using data from a survey on burnout among secondary school teachers in Kottayam district, Kerala, conducted after the initial validation of the translated OLBI. The primary objective of this survey was to estimate the prevalence of burnout among secondary school teachers using a cross-sectional study design. The translated and validated OLBI was used as the study instrument to assess burnout, and the survey data were subsequently used to evaluate the psychometric properties of the translated OLBI scale.

The target population for the survey comprised all secondary school teachers working in Government, aided, and unaided high schools in Kottayam district. A total of 255 teachers from 36 schools (18 from each of the two educational districts of Kottayam) participated in the study through a multi-stage cluster sampling method. Teachers who were employed full-time, had at least one year of teaching experience, were available at the time of data collection, and provided informed consent were included in the study. The detailed methodology and findings of the prevalence study will be presented in a separate paper.

Test-retest reliability

To assess the temporal stability, 40 participants from the original survey who consented to repeat participation completed the OLBI twice with a minimum two-week interval based on the same eligibility criteria used previously. Data were collected using Open Data Kit (ODK) and analysed with SPSS version 28.0. The significant difference between the responses at two time points was assessed using the Wilcoxon signed-rank test, a non-parametric method for paired samples, where a significance level of 0.05 was considered the cutoff for deciding the significance. The intra-class correlation coefficient (ICC) was computed for each item, the overall scale, and its subscales to quantify the consistency of responses over time. While the Wilcoxon test identifies statistical differences between two time points, the ICC provides a direct measure of reliability, with clearly defined cutoffs and interpretation. Following Koo and Li,

ICC values between 0.50 and 0.90 were considered moderate to good, and values above 0.90 were considered excellent.¹¹

The internal consistency, exploratory, and confirmatory factor analysis were performed in STATA/SE 17.0, Texas, USA, referring to the textbook “discovering structural equation modelling using STATA” by Acock.¹²

Internal consistency

Internal consistency was evaluated among the 255 responses obtained using Cronbach’s alpha for each item, the overall scale, and its two dimensions. It reflects how well the items on a scale measure the same underlying construct, and a Cronbach’s alpha of ≥ 0.7 indicates acceptable reliability, while values between 0.6 and 0.7 are considered questionable or only borderline acceptable.¹³

Exploratory factor analysis (EFA)

We conducted EFA using the principal component factor analysis (PCFA) approach in STATA, applying the unrotated method to examine the dimensionality of the translated OLBI. Eigenvalues, the primary output of EFA, represent how much of the total variance across all items is explained by each extracted factor. We followed the commonly used Kaiser criterion (eigenvalue ≥ 1) to determine the optimal number of factors to retain.

Confirmatory factor analysis (CFA)

While exploratory factor analysis (EFA) identifies the underlying structure of a dataset without predefined hypotheses, confirmatory factor analysis (CFA) tests a specified theoretical model to assess how well it fits the observed data, thereby confirming the instrument’s reliability and construct validity.¹⁴

We conducted CFA to compare the fit of two models: a one-factor model representing burnout as a single construct, and a two-factor model representing the subscales exhaustion and disengagement. Before testing the two-factor structure, separate CFAs were performed on each subscale to verify that each factor was independently meaningful. In CFA, latent variables (e.g., exhaustion, disengagement) are treated as independent (exogenous), while the observed items are considered dependent (endogenous) variables. The standardized coefficients indicate the strength and direction of the relationship between observed items and their corresponding latent constructs on a standardized scale.

In CFA, there are different measures to assess the model fit. The Chi-squared test compares the model to a saturated model that has no degrees of freedom. A significant p value (typically $p < 0.05$) suggests a poor model fit, and the observed differences are unlikely to be due to random sampling error alone. The comparative fit

index (CFI) compares the observed model with a baseline model, assuming no relationships among the observed variables. A CFI value of 0.90 or above generally indicates an acceptable model fit, with higher values suggesting a better fit. The Tucker-Lewis Index (TLI) is interpreted similarly to the CFI. Other commonly used fit indices in CFA, such as the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR), were also reported. Both provide measures of model fit based on residuals. Values less than 0.08 generally indicate a reasonable fit, while values below 0.05 suggest a good fit. The coefficient of determination (CD) summarizes how well the model explains the observed variance, with values ranging from 0 to 1. Although higher values indicate a better fit, the CD has no universally accepted cutoff.¹²

Ethical consideration and approval

The institutional ethics committee approved the study protocol before its implementation. Permission for the translation and validation of the tool was obtained from the original author.

Informed consent was obtained from participants at each relevant stage of the validation process and for the survey among teachers.

RESULTS

Translation and initial validation of OLBI

Version 1 was a reconciled form of the translations by three translators. Comparison of the two independent back-translations of version 1 with the original OLBI showed reasonable similarity in wording and meaning, leading to minor rewording and rearrangements in four items (3, 5, 11, and 13), with all 16 items retained in version 2. While assessing the content validity of version 2, most experts agreed that the translated items in OLBI-Malayalam reflected the original meaning with minor cultural differences. I-CVI for each item ranged from 0.83 to 1, the modified kappa from 0.81 to 1, and the S-CVI/Ave was 0.97, all exceeding the recommended threshold (Table 1).

Table 1: The relevance rating and content validity index of OLBI-Malayalam.

Items	Experts*						I-CVI†	Pc‡	K§	Evaluation
	1	2	3	4	5	6				
item1	1	1	1	1	1	1	1	0.016	1	Excellent
item2	1	1	1	1	1	1	1	0.016	1	Excellent
item3	1	1	1	0	1	1	0.83	0.094	0.81	Excellent
item4	1	1	1	1	1	1	1	0.016	1	Excellent
item5	1	1	1	1	1	1	1	0.016	1	Excellent
item6	1	1	1	1	1	1	1	0.016	1	Excellent
item7	1	1	1	1	1	0	0.83	0.094	0.81	Excellent
item8	1	1	1	1	1	1	1	0.016	1	Excellent
item9	1	1	1	1	1	0	0.83	0.094	0.81	Excellent
item10	1	1	1	1	1	1	1	0.016	1	Excellent
item11	1	1	1	1	1	1	1	0.016	1	Excellent
item12	1	1	1	1	1	1	1	0.016	1	Excellent
item13	1	1	1	1	1	1	1	0.016	1	Excellent
item14	1	1	1	1	1	1	1	0.016	1	Excellent
item15	1	1	1	1	1	1	1	0.016	1	Excellent
item16	1	1	1	1	1	1	1	0.016	1	Excellent

* Experts who rated the item 1 or 2 are recoded as 0, and those who rated 3 or 4 are recoded to 1. †I-CVI, item-level content validity index = ((Number of experts rating 3 or 4) ÷ Total no. of experts), acceptable cutoff ≥ 0.83 .^{9,21,22} ‡ Pc, probability of chance agreement, $Pc = [N! / A!(N-A)!] \times 0.5^N$, where N = number of experts, A = number of experts agreeing on relevance §K = modified kappa value on the relevance, which is $K = (ICVI - Pc) / (1 - Pc)$, where Pc is the probability of chance agreement, || Evaluation criteria for kappa: Fair: 0.40-0.59 Good: 0.60-0.74 Excellent: ≥ 0.75 .^{23,24}

Based on expert feedback, minor revisions were made to seven items (2, 5, 6, 8, 9, 10, and 11), and all 16 items were retained in version 3. Cognitive interviews were further conducted with four experienced teachers (three females, one male), each with over 10 years of teaching experience. The interviews averaged 16 minutes and focused on assessing the clarity and appropriateness of each item. While most items were well understood, minor

revisions were made based on feedback to improve clarity, specificity, and contextual relevance (items 2, 10, and 12). Changes included refining temporal terms, enhancing vocabulary precision, and aligning meanings more closely with the original English version (version 4). These adjustments improved the comprehensibility in terms of conveying the intended concepts and improving participants' understanding of the Malayalam OLBI.

Table 2: Test-retest results for individual items, dimensions, and Total OLBI score using the Malayalam version.

Items	Test	Retest	Wilcoxon signed-rank test	Intra-class correlation
	Median (range*)	Median (range)	P values	ICC (95% CI)†
Item1	2 (1-3)	2 (1-3)	0.025	0.677 (0.455-0.817)
Item2	3 (1-4)	3 (1-4)	0.527	0.787 (0.632-0.881)
Item3	2 (1-3)	2 (1-3)	0.414	0.703 (0.505-0.831)
Item4	3 (1-4)	3 (1-4)	0.999	0.769 (0.603-0.871)
Item5	2 (1-4)	2 (1-3)	0.248	0.540 (0.282-0.726)
Item6	2 (1-4)	2 (1-4)	0.166	0.701 (0.504-0.829)
Item7	2 (1-3)	2 (1-3)	0.655	0.781 (0.623-0.878)
Item8	2.5 (1-4)	2 (1-4)	0.317	0.790 (0.639-0.883)
Item9	3 (1-4)	3 (1-4)	0.705	0.790 (0.639-0.883)
Item10	3 (1-4)	3 (1-4)	0.317	0.758 (0.589-0.864)
Item11	3 (1-4)	3 (1-4)	0.999	0.562 (0.304-0.742)
Item12	3 (1-4)	3 (1-4)	0.593	0.604 (0.363-0.769)
Item13	3 (1-4)	3 (1-4)	0.134	0.598 (0.360-0.764)
Item14	2 (1-3)	2 (1-3)	0.705	0.579 (0.328-0.754)
Item15	2 (1-3)	2 (1-3)	0.48	0.693 (0.490-0.825)
Item16	2 (1-3)	2 (1-3)	0.999	0.748 (0.570-0.858)
Exhaustion	20 (10-27)	19 (10-27)	0.174	0.851 (0.737-0.918)
Disengagement	18.5 (10-24)	18 (11-24)	0.683	0.785 (0.629-0.880)
Total OLBI score	38 (23-50)	37.5 (21-46)	0.212	0.857 (0.747-0.922)

*The range of individual items were from 1 to 4; †P-values of all ICC values were <0.001.

Table 3: Internal consistency of the items in the overall scale and the exhaustion and disengagement dimensions (OLBI-Malayalam).

Items	Cronbach's alpha if the item is removed		
	All items	Exhaustion	Disengagement
Item1	0.83		0.58
Item2	0.83	0.78	
Item3	0.83		0.55
Item4	0.83	0.78	
Item5	0.82	0.76	
Item6	0.82		0.53
Item7	0.84		0.58
Item8	0.82	0.76	
Item9	0.83		0.57
Item10	0.82	0.76	
Item11	0.82		0.55
Item12	0.82	0.77	
Item13	0.85		0.65
Item14	0.82	0.77	
Item15	0.83		0.59
Item16	0.83	0.80	
Cronbach's alpha of the overall scale and two dimensions	0.84	0.80	0.61

Test-retest reliability and internal consistency of OLBI-Malayalam

Descriptive statistics for total OLBI scores and their dimensions at the test and retest time points are presented in Table 2. The Wilcoxon signed-rank test showed a

statistically significant difference ($p=0.025$) only for Item 1 ("I always find new and interesting aspects in my work"). All other items, including the total and subscale scores, showed no significant differences, indicating response consistency over time. ICC values ranged from 0.54 to 0.79 for individual items and 0.79 to 0.86 for the

two dimensions and total OLBI score, demonstrating moderate to good test-retest reliability (Table 2).

The OLBI total score demonstrated good internal consistency with an alpha of 0.84, which slightly improved to 0.85 upon removal of Item 13. The exhaustion subscale also showed good reliability ($\alpha=0.80$), with no increase in alpha upon deletion of any item. In contrast, the disengagement subscale had a lower alpha of 0.61, which improved to 0.65 when Item 13 was removed, suggesting borderline acceptable reliability (Table 3). Additionally, a bivariate correlation between the exhaustion and disengagement dimensions showed a moderate positive correlation ($r=0.67$, $p<0.001$) between the two subscales.

Exploratory factor analysis

The unrotated principal component factor analysis (PCFA) identified four factors with eigenvalues ≥ 1 . The first factor had an eigenvalue of 4.99, accounting for 31.21% of the total variance across the 16 items

(4.99/16), indicating a dominant underlying construct, and the pattern supports a one-dimensional structure of the scale. In the 16-item model, most items met the acceptable factor loading threshold (>0.4), indicating a moderate to good correlation between items and the latent factor, except item 13, which had a mild negative loading (-0.17), suggesting poor alignment. The revised 15-item model without item 13 retained the unidimensional structure, with the first factor's eigenvalue at 4.96, explaining 33.07% of the variance (4.96/15) and showed improved item coherence.

We further assessed the subscale dimensionality. For the exhaustion dimension, a single dominant factor emerged with an eigenvalue of 3.36, explaining 42% of the variance (3.36/8) and strong item coherence (loadings ranged from 0.47 to 0.72). In contrast, the disengagement dimension showed weaker factor dominance with the first factor's eigenvalue at 2.28; however, most items had acceptable loadings (0.46-0.69), except for item 13, which appeared misaligned with the intended construct (Table 4).

Table 4: Summary output from the exploratory factor analysis (principal component factor analysis -unrotated).

PCFA output	Total burnout		Exhaustion	Disengagement	
	16 items	15 items	8 items	8 items	7 items
Factors retained	4	3	1	3	2
LR test p value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Factor1	4.99	4.97	3.36	2.28	2.27
Factor2	1.59	1.57		1.23	1.23
Factor3	1.11	1.00		1.02	
Factor4	1.00				
Factor loadings					
Items	Factor1	Factor1	Factor1	Factor1	Factor1
Item1	0.41	0.41		0.46	0.46
Item2	0.51	0.50	0.57		
Item3	0.52	0.53		0.65	0.66
Item4	0.52	0.51	0.58		
Item5	0.70	0.70	0.72		
Item6	0.68	0.68		0.69	0.69
Item7	0.37	0.37		0.49	0.49
Item8	0.70	0.70	0.71		
Item9	0.48	0.48		0.55	0.55
Item10	0.69	0.69	0.71		
Item11	0.62	0.62		0.61	0.60
Item12	0.67	0.67	0.69		
Item13	-0.17			-0.15	
Item14	0.67	0.68	0.68		
Item15	0.44	0.45		0.50	0.50
Item16	0.48	0.49	0.47		

Confirmatory factor analysis

CFA results for the one-factor-16-item scale showed that all factor loadings were statistically significant and

moderately robust, except for item 13, which had a mild negative loading. The Chi-square test indicated an imperfect model fit. The CFI was 0.854 and TLI 0.831, suggesting moderate fit, while RMSEA (0.073) and

SRMR (0.067) fell within acceptable limits (<0.08). Though the above indices support a moderately acceptable fit, the model does not meet more stringent criteria, suggesting the 16-item structure may not optimally represent burnout. Removing item 13 produced

similar fit indices, indicating its exclusion does not significantly affect the scale reliability. The 15-item model retained structural stability and comparable performance, aligning with the standardized coefficients (Table 5).

Table 5: Standardized coefficients* and fit indices from confirmatory factor analysis.

Items	One-factor model		One-factor model- subscales			Two-factor model	
	16 items	15 items	Exhaustion	Disengagement	Disengagement-removing item 13	16 items-two dimensions	15 items- two dimensions
item2	0.46	0.46	0.49			0.46	0.46
item4	0.48	0.48	0.50			0.48	0.47
item5	0.66	0.66	0.67			0.65	0.66
item8	0.67	0.67	0.65			0.67	0.67
item10	0.66	0.66	0.67			0.66	0.66
item12	0.64	0.64	0.63			0.64	0.63
item14	0.62	0.63	0.61			0.62	0.62
item16	0.42	0.42	0.39			0.42	0.42
item1	0.36	0.36		0.29	0.29	0.36	0.36
item3	0.47	0.47		0.59	0.59	0.45	0.45
item6	0.65	0.65		0.64	0.64	0.63	0.63
item7	0.31	0.31		0.32	0.32	0.30	0.30
item9	0.44	0.45		0.44	0.44	0.43	0.43
item11	0.58	0.58		0.51	0.50	0.58	0.58
item13	-0.15			-0.10		-0.15	
item15	0.40	0.40		0.35	0.35	0.39	0.40
Fit indices							
χ^2	244.4	227.7	30.58	55.36	51.55	242.2	226
df	104	90	20	20	14	103	89
P value	<0.001	<0.001	0.061	<0.001	<0.001	<0.001	<0.001
CFI	0.854	0.856	0.976	0.819	0.81	0.855	0.857
TLI	0.831	0.832	0.967	0.747	0.714	0.831	0.831
RMSEA	0.073	0.078	0.046	0.083	0.103	0.073	0.078
SRMR	0.067	0.068	0.04	0.065	0.072	0.067	0.068
CD	0.871	0.87	0.816	0.687	0.686	0.848	0.85

*Standardized coefficients ≥ 0.40 are considered cautiously acceptable.

The single-factor analysis confirmed the high reliability of the exhaustion subscale, and the strong factor loadings for exhaustion in the two-factor model further supported this. The estimated covariance between exhaustion and disengagement was 1.04, indicating perfect correlation. The results suggest that a two-factor model did not fit better than the one-factor 16-item model, indicating that separating exhaustion and disengagement may not improve the explanatory power of the scale in this context.

DISCUSSION

The study translated and validated the OLBI into Malayalam, addressing the absence of a tool to measure burnout among Malayalam speakers, and demonstrated a culturally appropriate translation with excellent content validity indices, temporal stability, and acceptable

internal consistency. Factor analysis indicated some issues with item 13, but the fit indices were moderately acceptable. However, there was no clear evidence to support the original two-factor model over a one-factor model while considering the Malayalam version.

The translation and back-translation of OLBI followed a rigorous process to ensure that it is conceptually meaningful rather than a literal translation. Challenges included difficulty in finding equivalent Malayalam words, variations in translation by different individuals, and partial use of Google Translate. However, these differences contributed to refining and improving the translated tool. Similar issues related to uncertainty in wording during translation were noted in the Chinese OLBI translation study, where active discussions helped resolve them and improve the translated version.¹⁵ We observed an excellent content validity in our study, which

was also in line with the Chinese translation.¹⁵ Expert feedback and cognitive interviews confirmed that the Malayalam version retained the intended meaning and was well understood, supporting good content and face validity.

In our study, the Malayalam version of the OLBI showed good test-retest reliability. Although the ICC for disengagement was slightly lower than that for the exhaustion dimension, ICCs for both the overall scale and its dimensions were above acceptable thresholds, indicating temporal stability. Minor variations in specific items, such as item number 1, could be attributed to changes in workload, personal experiences, and other factors that happened over the test and retest interval. By contrast, the test-retest reliability of the original English version of OLBI reported only a moderate correlation for the exhaustion dimension and a poor correlation for the disengagement dimension between the two time points.¹⁶ However, in comparison with the validation of the student version of OLBI, the ICCs were slightly lower in the present study.¹⁷

The internal consistency was satisfactory in our study, with a Cronbach's alpha of 0.84 for the overall scale, higher than the Malay version ($\alpha=0.80$) and lower than the Tamil ($\alpha=0.93$) and Chinese versions ($\alpha=0.905$) of OLBI.^{15,18,19} The alpha for the disengagement dimension was lower, consistent with findings from the validation of the OLBI in the Malay language, which also reported a borderline acceptable value (α -disengagement =0.74).¹⁸ Inconsistency in alpha values of disengagement dimensions was also noted in a study among Nigerian doctors, where the authors identified a borderline alpha value for the disengagement dimension using the OLBI English version.²⁰ Even though the internal consistency was at an acceptable level (>0.7), the original English version of OLBI still reported variations in Cronbach's alpha values across occupational groups for both dimensions.¹⁶

While the overall Cronbach's alpha indicated good internal consistency, a high alpha does not necessarily mean that the scale measures a single underlying construct.¹² Therefore, to assess the dimensionality and validate the structure of the translated OLBI, we conducted exploratory factor analysis (EFA) followed by confirmatory factor analysis (CFA). The factorial validity of the original English version of the scale supported a two-dimensional structure as initially proposed by the authors during development.¹⁶ However, in the present study, EFA showed that the 16-item scale has a dominating single dimensionality, suggesting its ability to measure the latent variable burnout. In addition, the subscales, especially the exhaustion dimension, have a strong single dimensionality, showing the ability of the subscale. The disengagement dimension remains a dominant factor, although not as strongly as the original scale or exhaustion. Item 13 also showed a low impact on

the construct, as observed in the Cronbach's alpha analysis.

CFA showed similar results for both the single-factor and two-factor models, suggesting that the two-factor model may not be distinct enough, or that a single dimension is sufficient for explaining the observed relationship between the variables. A high correlation between the two subscales also indicates that they are not distinct constructs, and a single factor could be a better explanation for the latent construct. However, from the analysis, we also found that the exhaustion dimension exists as a strong and reliable measure, and the total scale is, in fact, impacted by this dimension.

However, the high correlation between the two subscales does not specify the direction of influence, whether exhaustion causes disengagement, vice versa, or if both are affected by a third factor. The borderline acceptability of the dimension of disengagement is consistent in reliability analysis, EFA, and CFA. The low impact of item 13 is consistent in all the analyses, suggesting the need for further item refinement or cultural adaptation in this dimension. The results indicate that the overall structure and reliability of the scale remain stable even without item 13, which also shows consistency. Hence, in the present study, EFA indicated a one-dimensional model, and CFA provided no clear evidence favouring the two-dimensional model over the one-dimensional model.

On the other hand, our results are more consistent with the factorial validity of the original English version as compared to the Malay translation of OLBI, where the scale demonstrated poor fit in both one-dimensional and two-dimensional models, leading to the stepwise removal of seven items from the translated scale to achieve acceptable fit. Item 13, which showed a low impact in our study, was similarly identified in the validation of the Malay version of OLBI as well. The Malay OLBI established a two-dimensional model comprising nine items, demonstrating good validity and model fit.¹⁸ In contrast, the Tamil and Chinese translations of OLBI showed the best fit for the two-dimensional structure, with all items retained, comparable to the original English translation.^{15,19} However, a notable observation from our study was that the exhaustion dimension demonstrated high fit indices and strong factor loadings in the factorial analysis, indicating its ability to reliably measure exhaustion independently, along with good internal consistency.

In conclusion, the differences observed in the present study may be attributed to cultural variations, linguistic modifications, or differences in the occupational context of teachers in Kerala, which underscores the importance of conducting a localized validation of assessment tools before using them in different cultural settings.

A rigorous translation and validation process, which ensured conceptual equivalence between the English and

Malayalam versions of the OLBI, is the key strength of the study. However, the reliability and validity of the instrument were specifically established among secondary school teachers in a district in central Kerala. Hence, further validation is recommended to confirm its contextual relevance and measurement accuracy in other populations and settings.

CONCLUSION

The validated Malayalam version of the Oldenburg burnout inventory (OLBI) offers a standardised and culturally adapted tool for assessing burnout among Malayalam-speaking populations. Validated for use among secondary school teachers, the tool demonstrated strong internal consistency, test-retest reliability, and a near-stable one-factor structure over the originally conceptualised two-factor model. These findings support its suitability for assessing burnout in educational settings. While it may apply to other occupational groups, further validation is required due to differences in job demands and work environments. This study serves as a foundation for future research that can support the wider use of the tool in assessing burnout and informing interventions to improve occupational well-being in Kerala and beyond.

ACKNOWLEDGEMENTS

Authors acknowledge sincere gratitude to Dr Evangelia Demerouti, the author of the Oldenburg burnout inventory, for permitting us to translate and validate the OLBI to Malayalam for this research. The authors would like to thank Dr Mini R. Nair, Dr Anjitha Krishna, Sabeley Jose C., Dr Gopika M. G., Ramya P. M., Dr Neethu P. S., Dr Arsha VR, Aravind, Ajayan P., and Ajayan Thenmala for their support in translation, back translation, and initial validation.

Supplementary material

The validated Malayalam version of the tool is provided as supplementary files. Available at: <https://drive.google.com/drive/folders/1uc52Gbn5wLAD0ipvMLNWjSuVUGtxahpQ?usp=sharing>.

Funding: No funding sources

Conflict of interest: None declared

Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

- World Health Organization. Mental health. 2022. Available from: <https://www.who.int/health-topics/mental-health>. Accessed on 8 June 2024.
- World Health Organization. Mental health. 2022. Available from: <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>. Accessed on 8 October 2024.
- ICD-11 for Mortality and Morbidity Statistics. ICD-11 for Mortality and Morbidity Statistics. 2024. Available from: <https://icd.who.int/browse/2024-01/mms/en#129180281>. Accessed on 10 August 2024.
- Freudenberger HJ. Staff burn-out. *J Soc Issues*. 1974;30(1):159-65.
- Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol*. 2001;52:397-422.
- Maslach C, Jackson SE. The measurement of experienced burnout. *J Organ Behav*. 1981;2(2):99-113.
- Demerouti E, Bakker AB, Nachreiner F, Schaufeli WB. The job demands-resources model of burnout. *J Appl Psychol*. 2001;86(3):499-512.
- Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986;35(6):382-5.
- Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health*. 2007;30(4):459-67.
- Waltz CF, Strickland OL, Lenz ER. Measurement in Nursing and Health Research. New York, NY: Springer Publishing Company; 2016
- Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-63.
- Acock AC. Discovering structural equation modeling using Stata. Revised edition. College Station, Tex: StataCorp LP; 2013:306.
- Nunnally JC, Bernstein IH. Psychometric Theory. 3rd ed. New York, NY: McGraw-Hill; 1994.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equat Model Multidiscipl J*. 1999;6(1):1-55.
- Xu H, Yuan Y, Gong W, Zhang J, Liu X, Zhu P, et al. Reliability and validity of the Chinese version of Oldenburg Burnout Inventory for Chinese nurses. *Nurs Open*. 2021;9(1):320-8.
- Halbesleben JRB, Demerouti E. The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. *Work Stress*. 2005;19(3):208-20.
- Smith A, Ellison J, Bogardus J, Gleeson P. Reliability and validity of the student version of the oldenburg burnout inventory in physical therapist students. *J Phys Ther Educ*. 2022;36(3):205.
- Mahadi NF, Chin RWA, Chua YY, Chu MN, Wong MS, Yusoff MSB, et al. Malay language translation and validation of the Oldenburg Burnout Inventory Measuring Burnout. *Educ Med J*. 2018;10:27.
- Subburaj A, Vijayadurai J. Translation, validation and psychometric properties of Tamil version of Oldenburg burnout inventory (OLBI). *Procedia Soc Behav Sci*. 2016;219:724-31.
- Ogunsuji O, Ogundipe H, Adebayo O, Oladehin T, Oiwoh S, Obafemi O, et al. Internal reliability and

validity of Copenhagen burnout inventory and Oldenburg burnout inventory compared with Maslach burnout inventory among Nigerian resident doctors: a pilot study. *Dubai Med J*. 2022;5(2):89-95.

21. Davis LL. Instrument review: getting the most from a panel of experts. *Appl Nurs Res*. 1992;5(4):194-7.
22. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489-97.
23. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic*. 1981;86(2):127-37.

24. Fleiss JL, Levin BA, Paik MC. Statistical methods for rates and proportions. 3rd ed. Hoboken, N.J: Wiley-Interscience; 2003:1.

Cite this article as: Sivadas R, Thulaseedharan JV. Cross-cultural adaptation and psychometric evaluation of the Malayalam version of the Oldenburg burnout inventory among secondary school teachers. *Int J Community Med Public Health* 2026;13:4353-62.