BMC Psychology

RESEARCH Open Access

Comparing the psychometric evidence of the Depression, Anxiety, and Stress Scale-21 (DASS-21) between Spanish and Chinese primary schoolteachers: insights from classical test theory and rasch analysis

Xu Wang[1,2], Cui-Hong Cao[1,3], Xiao-Ling Liao[4], Xing-Yong Jiang[5], Mark D. Griffiths[6], I-Hua Chen[7*], Chung-Ying Lin[8,9,10*] and Olga Malas[11]

Abstract

Background Primary schoolteachers play a pivotal role in the education of children, highlighting the importance of addressing their psychological well-being and mental health. While the Depression, Anxiety, and Stress Scale-21 (DASS-21) is a globally recognized instrument for evaluating mental health, its systematic validation among primary schoolteachers, particularly in a cross-national context, remains understudied. The present cross-cultural study evaluated the psychometric properties of DASS-21 among primary schoolteachers in Spain and China to compare the DASS-21 between teachers across Western and Eastern cultures.

Methods The study sample comprised 1,350 Spanish and 2,580 Chinese primary schoolteachers. The DASS-21, alongside the Emotional Exhaustion Scale for Teachers, was used to evaluate its psychometric properties. The construct and concurrent validity of the DASS-21 were examined using Rasch analysis, Confirmatory Factor Analysis (CFA), and Structural Equation Modeling (SEM). Additionally, measurement invariance was tested across two national teacher cohorts.

Results Rasch analysis confirmed the validity of most DASS-21 items among both the Spanish and Chinese teachers, with only the same item from the Anxiety subscale showing inadequate fit in both countries. CFA across both countries initially favored a bifactor model, which was subsequently excluded due to problematic factor loadings. Instead, a single-factor model provided the best fit for Chinese teachers, while a correlated three-factor model was optimal for Spanish teachers. SEM demonstrated the DASS-21's concurrent validity with emotional exhaustion, with consistent findings across both samples. After excluding Differential Item Functioning (DIF) items, strict measurement invariance between Spanish and Chinese teachers was achieved, as verified by multi-group CFA.

*Correspondence:
I-Hua Chen
aholechen@gmail.com
Chung-Ying Lin
cylin36933@gmail.com

Full list of author information is available at the end of the article

Conclusion The DASS-21 exhibits commendable psychometric properties, rendering it a suitable instrument for evaluating the mental illness of primary schoolteachers in both Spain and China.

Keywords Psychometric properties, The DASS-21, Primary schoolteachers, China, Spain, Classical test theory, Rasch analysis

Introduction

Primary schoolteachers are foundational figures in basic education who shape societal fabric through their responsibility of educating children [1, 2]. Public expectations toward teaching have intensified, increasing teacher stress yearly [3]. This occupational stress involves psychological and physiological disturbances within school environments [4], manifesting as negative emotional states from various sources [5] including overwhelming workloads, adaptation challenges to educational reforms, reduced self-efficacy, and strained relationships [6–10]. Primary schoolteachers face more complex stress than higher education counterparts due to their multifaceted role encompassing academic instruction, behavior management, parent communication, and administrative duties [11, 12], underscoring the importance of understanding their mental health needs.

Indeed, the impact of these stressors on teachers' mental health is profound because many experience physiological and psychological consequences such as insomnia, headaches, anhedonia, and self-deprecation [13, 14]. Prolonged high stress can trigger mental health issues such as obsessive-compulsive disorder and depression [15, 16], impairing work efficiency and quality of life [15], and reducing life satisfaction and personal well-being [16]. Importantly, teachers' mental health problems can negatively affect their students' academic performance [17].

The critical role of primary schoolteachers within the educational system highlights the need to focus on their mental health, necessitating a trustworthy and efficacious instrument for assessing their psychological well-being [12]. The Depression, Anxiety, and Stress Scale (DASS), widely recognized among psychological assessments [18], can potentially serve this purpose. Originally developed as a 42-item scale by Lovibond et al. in 1995 [19], the DASS-42 was designed to assess the fundamental symptoms associated with depression and anxiety disorders. Subsequent factor analysis by the same research team identified a third component (i.e., stress), leading to the integration of this dimension into the DASS framework [19]. Consequently, the DASS encapsulates three primary dimensions: *depression*, *anxiety*, and *stress*. A subsequent revision led to the development of a shorter 21-item version, the DASS-21 [19].

The DASS-21, acclaimed for its conciseness, reliability, and extensive applicability, has been extensively utilized in global research and clinical settings, owing to its straightforward interpretability [20]. It has been translated into many languages, including Chinese [21], French [22], and Spanish [23]. The continuous validation of DASS-21 has broadened its application from its initial focus on the general adult population to other populations, including children [24], adolescents [25], university students [26], older people [27], and specific professional groups, such as athletes and teachers [20, 28, 29].

The present study conducted a cross-cultural comparison of psychometric properties of the DASS-21 between Spanish and Chinese primary schoolteachers. The research addresses a significant gap in the literature because this population remains relatively understudied despite their crucial societal role and specific occupational stressors. By validating the DASS-21 in these two distinct educational contexts, the present study sought to establish a foundation for developing culturally appropriate mental health interventions for teachers. The following sections review existing cross-cultural DASS-21 validation studies, contextualize the socioeconomic and mental health challenges faced by Spanish and Chinese teachers, and identify relevant topics and methodological gaps in previous research that the present study aimed to address.

Cross-cultural validation and application of the DASS-21 in global mental health research

Building upon the diverse and global application of the DASS-21 described above, several studies have further extended its use through cross-national comparisons to examine the scale's application across different countries [30–33]. For instance, Scholten et al. [30] selected representative samples of individuals aged 18 years and above in Poland, Russia, United Kingdom, and United States (US) to evaluate the psychometric properties of the DASS-21 among individuals in these nations. The findings indicated that the DASS-21 could meaningfully (without measurement bias) be used to compare differences in the levels of depression, anxiety, and stress between individuals in these countries [30].

Similarly, Bibi et al. [31] conducted an extensive survey to explore the applicability of the DASS-21 among university students in Pakistan and Germany. This study demonstrated that the DASS-21 had good reliability, construct validity, and factorial validity among individuals in both countries. Moreover, the samples from both countries showed a negative association between scale scores of depression, anxiety, and stress and scores on positive

mental health scales, such as scales assessing positive mental health, social support, and life satisfaction [31].

Building upon the cross-cultural utility of the DASS-21, Zanon et al. [32] examined university student samples from a diverse range of countries and regions, including Brazil, Canada, Hong Kong, Romania, Taiwan, Turkey, United Arab Emirates, and the US. Their investigation primarily employed confirmatory factor analyses (CFA) to compare the fit of four distinct DASS-21 factor structures: the unidimensional model, the three-correlated-factors model, the higher-order model, and the bifactor model (comprising a general factor and three specific orthogonal factors). While the bifactor model demonstrated superior fit across all countries studied, the composite scores of the specific factors lacked adequate psychometric properties. Consequently, using subscales for cross-country comparisons may yield biased interpretations because the subscale raw scores were substantially influenced by the general psychological distress factor rather than purely reflecting the three specific constructs (depression, anxiety, and stress). Moreover, the bifactor model was excluded from group comparisons due to convergence failure of the baseline model in the Taiwanese sample [32].

Oei et al. [33] conducted a study to address the lack of systematic examination of the DASS-21's psychometric properties in Asian populations. Consequently, they surveyed over 2000 corporate employees from six Asian countries/regions, namely Malaysia, Indonesia, Singapore, Sri Lanka, Taiwan, and Thailand, to assess the psychometric characteristics of the DASS-21. Their study affirmed that the factor structure among corporate employees in these Asian nations also exhibited a correlated three-factor structure, consistent with findings from Western countries [34].

Despite these valuable contributions, existing cross-cultural DASS-21 research shows considerable variation in factor structure across populations and contexts, with inconsistent findings regarding which model best represents the scale's underlying dimensions. This variability highlights the need for targeted validation among specific occupational groups such as primary schoolteachers, where cultural factors may uniquely influence measurement properties.

The mental health of primary school teachers in Spain and China

In the present study, primary schoolteachers from Spain and China were selected for a cross-national comparison using the DASS-21. The choice of these specific educational populations warrants explanation. As a member of the European Union, Spain is considered a developed country in terms of education on a global scale. Compared to countries such as Bulgaria, Portugal, and Malta, Spanish teachers generally enjoy higher average incomes, minimum wage guarantees, and greater job satisfaction [35]. However, in recent years, Spanish education has faced multiple challenges, including adapting to technological and social changes [36]. Consequently, several studies have explored factors influencing the mental health of Spanish teachers, such as emotional regulation, role stress [37], working conditions [38], self-efficacy, occupational stress, and burnout [39]. During the COVID-19 pandemic, the mental health of Spanish teachers was significantly impacted [40]. Research on the effects of lockdown restrictions revealed that primary and secondary schoolteachers experienced more pronounced negative effects on their quality of life compared to their counterparts in preschool, vocational training, and university education [41]. In Northern Spain, approximately one-third of teachers exhibited mild depressive symptoms, with around half showing moderate levels of anxiety and stress [42].

In contrast, Chinese teachers, despite recent advancements in China's educational system, often earn incomes in the middle to low range compared to other occupations [43], raising concerns about their well-being [44]. Their sense of well-being is relatively unbalanced, with health status below the general population's average [44, 45]. During the COVID-19 outbreak, depression was prevalent among Chinese teachers [46, 47]. Multiple studies have documented that teachers exhibit mental health levels with higher rates of occupational burnout [48, 49].

Given the socio-economic and mental health challenges faced by primary schoolteachers in Spain and China, particularly concerning depression, work stress, and professional burnout, this provides a compelling rationale for cross-cultural validation of the DASS-21. Understanding the instrument's psychometric properties across these distinct educational contexts will enhance assessment accuracy and facilitate meaningful comparisons that account for cultural differences in symptom expression. Such validation is essential for developing targeted interventions addressing teachers' specific needs in each cultural context.

The present study

A literature review identified two key gaps. Firstly, few studies use the DASS-21 to assess primary schoolteachers' psychological well-being, despite its validation across diverse groups. This cohort remains underrepresented, which risks mischaracterizing their mental health. Teachers face distinct stressors—such as classroom management, parental expectations, and educational policy pressures—that differ from other professions, compounded by cultural influences on how distress manifests (e.g., collectivist versus individualist norms). Without

teacher-specific validation, assessments may fail to detect these unique experiences, obscuring the development of tailored interventions to support their well-being and, by extension, student outcomes. An exception is Zhou et al. [50], who tested the DASS-21's factorial validity using CFA among teachers in China and Spain. They employed Maximum Likelihood Estimation with Robust Standard Errors (MLR) for their analysis. However, Li [51] demonstrated through a simulation study that Diagonally Weighted Least Squares (DWLS) produces less biased estimates than MLR when analyzing ordinal data, such as the items in the DASS-21. This finding suggests that DWLS may be a more suitable approach for studies involving diverse teacher populations. Another exception, Katsantonis [52], examined well-being and stress across 15 countries with different scales, finding cultural non-invariance between Anglo-Saxon and East Asian groups. Although not DASS-21-focused, this highlights the need for invariance testing in teacher research to ensure accurate, policy-relevant results.

Secondly, a key methodological gap in the literature is the heavy reliance on CFA to assess DASS-21 invariance across countries and ethnic groups [26, 53], with little focus on alternatives like the Rasch measurement model. Proposed by Georg Rasch [54], this item response model aligns abilities and item difficulties on an interval scale, offering advantages over CFA, such as flexibility in scale assumptions [55, 56]. Salzberger et al. [57] compared the Rasch model to multigroup CFA using the Technophobia Scale, finding differences in scale bandwidth and item locations, highlighting its value for cross-national studies. While the Rasch model has validated the DASS-21 among non-teacher groups [58, 59], its application to teachers, particularly across nations, remains underexplored, limiting comprehensive invariance testing.

To address these gaps, the present study conducted a comprehensive analysis using Classical Test Theory (e.g., CFA) and the Rasch model. To further validate the DASS-21, the study explored its association with emotional exhaustion, a core burnout dimension defined by fatigue from depleted emotional resources [60]. Characterized by diminished feelings, care, trust, interest, and spirit [61], emotional exhaustion reflects the individual stress component of burnout and aligns with psychological distress, as outlined in the Stressor–Strain–Outcome (SSO) model [62]. Prior studies with teacher samples have consistently shown strong correlations between emotional exhaustion and DASS-21 factors—depression, anxiety, and stress [63–65]. Accordingly, the present study assessed the DASS-21's concurrent validity by examining its relationship with emotional exhaustion. Measurement invariance was also tested using multi-group CFA and differential item functioning (DIF) analysis.

Method

Procedure and participants

In early 2022, amid the COVID-19 pandemic and the dominance of the Omicron variant, a survey targeted primary schoolteachers in Spain and China. In Spain, a strong vaccination program mitigated Omicron's challenges, allowing primary schools to sustain in-person teaching [66]. Conversely, China's strict zero-tolerance policy triggered immediate lockdowns in affected areas, severely restricting face-to-face instruction [67]. The study received ethical clearance differently in each country: in Spain, the Institutional Review Board (CEIM IRJV) deemed it exempt, while in China, the Jiangxi Psychological Consultant Association granted approval.

For Spanish participants, the sampling was non-probabilistic, and the instruments were administered online. Recruitment involved emailing a survey link to all primary schools in the Autonomous Community of Catalonia, Spain, under the jurisdiction of the Departament d'Educació (the Catalan government's education department). The school directors were asked to forward this email to their teachers. Participation was entirely voluntary for both school directors and teachers, and there was no compensation. The first page of the survey informed participants about the study's objectives, the anonymity and confidentiality of their data, and their option to withdraw at any time. Participants provided digital informed consent before completing the survey, thereby acknowledging their acceptance of the study terms. The inclusion criterion was being an active primary education teacher without a diagnosis of mental health illness at the time of the survey. A total of 1,350 Spanish primary schoolteachers participated in this part of the study.

For the Chinese participants, the survey was part of a larger project aimed at monitoring teachers' mental health during the pandemic, focusing particularly on the effects of emergency online teaching in a central Chinese city. After a strict lockdown and a period of online teaching, the survey was conducted as schools were reopening. In collaboration with the local government's education department and using purposive sampling, the survey targeted teachers involved in emergency online teaching. The inclusion criteria included being employed in a primary school, registration in the project, and providing online consent, with no mental illness reported. The online survey platform was designed to ensure complete responses. Out of 2,764 surveys distributed via email, 2,580 completed surveys were received.

Table 1 presents the demographic characteristics of the participants from Spain and China. The data showed that the average age of Spanish participants was higher, with a mean age of 42.86 years (standard deviation = 10.18), compared to their Chinese counterparts, who had a mean age of 34.72 years (standard deviation = 9.95). Regarding

Table 1 The characteristics of the participants

Source	Spanish	China
Data collection period	Jan-Feb 2022	Jan. 2022
Sample size	1350	2580
Gender (Female); *n* (%)	1151 (85.25)	1962 (76.04)
Age; Mean (SD)	42.86 (10.18)	34.72 (9.95)

gender composition, 85.26% of Spanish teachers were female ($n = 1{,}151$), and 76.05% of Chinese teachers were female ($n = 1{,}962$).

To ensure adequate sample size for psychometric examinations (including CFA and Rasch analysis) on the DASS-21, a power analysis using the semPower package in R targeted 80% power to detect good model fit (RMSEA ≤ 0.05, $\alpha = 0.05$) for its correlated three-factor structure, indicating a minimum of 200–250 participants for CFA [68]. For Rasch analysis of the DASS-21, which uses a four-category response format (0–3), a sample size of 250–500 participants is commonly used to ensure stable item calibrations, with 250–300 being sufficient for general purposes, and 300–400 per group considered ideal for robust cross-cultural comparisons, such as between Spanish and Chinese teachers, particularly when detecting DIF of 0.5 logits [58, 69]. The samples of 1,350 Spanish and 2,580 Chinese primary schoolteachers exceeded these requirements, ensuring robustness for both analyses, accounting for issues such as non-normality, and supporting cross-cultural comparisons, including DIF analyses.

Instruments

The present study utilized the Spanish and Chinese versions of the DASS-21, along with a scale to assess emotional exhaustion, also administered in the respective languages. The following sections provide a detailed description of these instruments.

Depression, Anxiety, and Stress Scale (DASS-21)

As aforementioned, the DASS-21 consists of 21 items, with seven items assessing each of the three subscales: depression (e.g., *"Life was meaningless"*), anxiety (e.g., *"Experienced trembling"*), and stress (e.g., *"Intolerant of anything"*). Items are rated on a scale from 0 (*Did not apply to me at all*) to 3 (*Applied to me very much, or most of the time*) with total scores in each subscale ranging from 0 to 21. Higher scores on the Depression, Anxiety, and Stress subscales indicate greater levels of depression, anxiety, and stress, respectively. The total score indicates the overall level of psychological distress. In the present study, the Spanish version of the DASS-21 [70] was administered to Spanish participants. It demonstrated strong internal consistency with McDonald's omega coefficients of 0.879, 0.931, and 0.895 for the Anxiety, Depression, and Stress subscales respectively, and 0.957 for the overall scale [70]. For the Chinese cohort, the Chinese version of the DASS-21 [71] was utilized, a version widely used across various populations in Mainland China, including minors [72], university students [73], and teachers [20].

According to the suggested cut-off scores indicative of clinical depression, anxiety, and stress in the DASS-21 are 5 or more, 4 or more, and 8 or more, respectively [74]. These criteria were adhered to in the present study for both Spanish and Chinese participants. In the present study, the McDonald's omega coefficients for general psychological distress and the three factors within the DASS-21 ranged from 0.88 to 0.97 across two countries, indicating good internal consistency.

Emotional Exhaustion Scale for Teachers (EEST)

In the present study for the Spanish sample, the EEST developed by Aluja et al. [75] was used to assess emotional exhaustion. As an adapted version of the Emotional Exhaustion subscale from the Maslach Burnout Inventory [76], this scale is specifically tailored for schoolteachers. It consists of six items (e.g., *"I feel used up at the end of the workday"*), each rated on a seven-point Likert scale that ranges from 0 (*never*) to 6 (*every day*). The total score ranges from 0 to 36, and higher scores indicate greater emotional exhaustion. The scale demonstrated very good internal consistency among the Spanish sample in the present study, with a McDonald's omega coefficient of 0.877.

Chinese Primary and Secondary School Teachers' Job Burnout Questionnaire (CTJBO)

In the present study for the Chinese sample, the CTJBO was used to assess emotional exhaustion [77]. More specifically, the Emotional Exhaustion subscale from the CTJBO based on the Maslach Burnout Inventory for schoolteachers [78] was used. This instrument is specifically designed to assess job burnout among primary and secondary schoolteachers within the Chinese educational context. The CTJBO is inclusive of culturally relevant content and wording. The Emotional Exhaustion subscale of the CTJBO comprises eight items (e.g., *"I feel like I'm overdrawn"* and *"I feel that my teaching work has exhausted my emotions"*) rated on a seven-point Likert scale, ranging from 0 (*strongly disagree*) to 6 (*strongly agree*). The total score ranges from 0 to 48, and higher scores indicate greater emotional exhaustion. This subscale has demonstrated a unidimensional structure with excellent factorial validity and high reliability in the Chinese teaching context, as evidenced by recent studies [79, 80]. The scale demonstrated excellent internal consistency among the Chinese sample in the present study, with a McDonald's omega coefficient of 0.954.

Data analysis

In the present study, all analyses were conducted using jamovi 2.6.44, software based on the R package [81]. The analysis began with descriptive statistics to summarize the data characteristics, followed by the application of Pearson's correlation coefficient. This approach was adopted due to the relatively large sample size, which mitigates concerns about deviations from normality, as supported by Bishara and Hittner [82], who demonstrated the robustness of Pearson's correlation among large samples even if they have non-normal data distributions. Subsequently, an individual item assessment within the DASS-21 using Rasch analysis was conducted. CFA was then employed to evaluate the overall scale measurement. Structural equation modeling (SEM) was further used to examine the association between DASS-21 and emotional exhaustion, to evaluate concurrent validity. Finally, the examination of measurement invariance between participants from China and Spain was conducted. The following sections detail these procedures.

Rasch analysis

In conducting the Rasch analysis, the guidelines established by Tennant and Conaghan [54] were followed. The initial step involved selecting the most appropriate model, choosing between the Partial Credit Model (PCM) and the Andrich Rating Scale Model (RSM). Unidimensionality was then assessed, applying the criterion that the eigenvalue of the first component in the principal component analysis (PCA) of residuals should be less than 2.00 [83]. To verify that response categories increased monotonically, a person-item map was utilized. Item validity was evaluated using information-weighted (INFIT) and outlier-sensitive (OUTFIT) fit statistics, with acceptable ranges for valid item responses set between 0.50 and 1.50 [84]. Additionally, item and person reliability and separation indices were assessed separately. Following Boone et al. [85], item reliability was evaluated, with a threshold of >0.90 indicating good reliability; item separation was assessed, with a threshold of >3 indicating that the person sample is sufficiently large to confirm the item difficulty hierarchy (construct validity); for person separation, a person separation index of 1.50 indicates acceptable separation, 2.00 indicates good separation, and 3.00 indicates excellent separation; and for person reliability coefficients, values of 0.90 distinguish 3 or 4 levels, 0.80 distinguishes 2 or 3 levels, and 0.50 distinguishes 1 or 2 levels.

Confirmatory factor analysis and structural equation modeling

With respect to CFA, factorial validity was assessed by testing multiple factor structures: a single-factor (unidimensional) model, a correlated two-factor model (comprising a depression factor and a combined anxiety and stress factor), a correlated three-factor model, and a bifactor structure. The correlated two-factor model was adopted due to strong evidence of high correlations between anxiety and stress, observed among children [86] and hospital workers ($r>0.93$) [87]. This model reflects their shared traits, such as physiological arousal and negative affect, distinguishing them from depression, which is more aligned with cognitive features. Acceptable model fit was determined using established criteria: a Comparative Fit Index (CFI) and Non-Normed Fit Index (NNFI) exceeding 0.90, a Root Mean Square Error of Approximation (RMSEA) below 0.06, and a Standardized Root Mean Square Residual (SRMR) below 0.08 [88]. DWLS estimation was employed, as it is particularly suited for ordinal data such as the DASS-21 items [51].

Within the bifactor model, indices of the explained variance by the general factor, including Explained Common Variance (ECV), Percentage of Uncontaminated Correlations (PUC), and Omega Hierarchical (ωH), were calculated. ECV represents the ratio of variance explained by the general factor to the total variance explained by the model. PUC indicates the proportion of uncontaminated correlations to unique correlations. The ωH for the general factor reflects the percentage of variance in total scores attributed to the general factor. If the general factor ωH is >0.80, total scores can be considered unidimensional. If both ECV and PUC are >0.70, the relative bias is slight, and the common variance can be considered essentially unidimensional [89, 90].

Moreover, SEM was used to evaluate the concurrent validity of the DASS-21 by testing its association with emotional exhaustion. Gender and age were included as control variables based on empirical evidence from the preliminary analyses. More specifically, significant gender differences were observed among Spanish teachers, with female teachers reporting higher levels across all three dimensions (depression, anxiety, and stress) compared to male teachers (t-values ranging from 3.52 to 5.61, all $p<0.001$). Age showed significant negative correlations with anxiety (r = -0.07, $p=0.01$) and stress (r = -0.11, $p<0.001$) among Spanish teachers. Although these demographic variables did not show significant associations with the three DASS-21 factors in the Chinese sample, they were retained as control variables in the SEM models for consistency and to account for their demonstrated effects in the initial analyses. The model fit was first evaluated using the same criteria as in the aforementioned CFA, and then scrutinized the path coefficient with emotional exhaustion.

Measurement invariance in the DASS-21 among teachers from Spain and China

Finally, to establish a basis for comparing primary schoolteachers from two countries, DIF of Rasch analysis and multi-group CFA were conducted. DIF assessed item difficulty differences between groups [85], with a value above 0.50 considered substantial [91]. Regarding measurement invariance in CFA, several model comparisons were made in nested models of baseline model-configural model, factor-loading constrained model, and factor-loading plus item-intercept constrained model. Indices including ΔCFI, ΔRMSEA, and ΔSRMR were then used to decide if measurement invariance was supported. Measurement invariance is supported if ΔCFI > -0.01, ΔRMSEA < 0.015, and ΔSRMR < 0.03 (for factor loading) or < 0.01 (for item intercept), following the criteria proposed by Chen [92].

Results

Descriptive statistics and pearson correlation

Table 2 outlines the mean observed scores of participants from both countries as assessed by the DASS-21, detailing the correlations with the overall scale and its respective subscales. Additionally, Table 2 integrates the correlations that relate these scores to the variable of emotional exhaustion. The findings indicated that Spanish educators exhibited higher levels of depression and stress compared to their Chinese counterparts. In contrast, anxiety levels appeared to be comparable in both groups.

Pearson correlation coefficients showed significant positive correlations between the three emotional disorders among teachers from both countries (Spain: r ranged from 0.76 to 0.78, all $p < 0.001$, large effect size; China: r ranged from 0.89 to 0.91, all $p < 0.001$, large effect size). Notably, within the Chinese teacher cohort, these correlations reached approximately 0.90, suggesting that depression, anxiety, and stress are difficult to differentiate as distinct constructs in this population.

According to the cutoffs specified in the DASS manual [74], among the Spanish teachers, 34.7%, 30.4%, 38.2%, and 43.4% exhibited clinical symptoms of psychological distress, depression, anxiety, and stress, respectively, at a mild level or above. In contrast, the corresponding percentages for the Chinese teachers experiencing mild or higher levels of these symptoms were 30.2%, 30.6%, 40.7%, and 19.0% for psychological distress, depression, anxiety, and stress, respectively. The chi-square analyses showed significant differences in the proportion of psychological distress and stress between Spanish and Chinese teachers ($\chi^2 = 8.31$, $p = 0.004$ and $\chi^2 = 266.17$, $p < 0.001$, respectively; see Table S1). No significant differences were observed for depression or anxiety.

Table 2 Descriptive statistics and pearson correlations for Spanish and Chinese primary schoolteachers

	Spanish primary schoolteachers ($n = 1,350$)							Chinese primary schoolteachers ($n = 2,580$)						
	ω	*Mean*	*SD*	1	2	3	4	ω	*Mean*	*SD*	1	2	3	4
1. DASS-21 (range:0–63)	0.95	14.69	12.79	1.00				0.97	11.41	12.30	1.00			
2. Depression subscale (range: 0–21)	0.91	3.72	4.51	0.92***	1.00			0.92	3.25	4.17	0.96***	1.00		
3. Anxiety subscale (range: 0–21)	0.88	3.78	4.33	0.91***	0.78***	1.00		0.91	3.80	4.21	0.97***	0.91***	1.00	
4. Stress subscale (range:0–21)	0.90	7.18	5.06	0.93***	0.77***	0.76***	1.00	0.91	4.37	4.34	0.97***	0.89***	0.91***	1.00
5. Emotional exhaustion (range:0–36 for Spanish teachers; 0–48 for Chinese teachers)	0.88	13.92	8.59	0.69***	0.63***	0.58***	0.68***	0.95	20.12	12.47	0.52***	0.47***	0.50***	0.55***

Item validity with rasch analysis

In the Rasch analysis, first, the assumption of dimensionality was tested by subjecting all 21 items of the DASS-21 to a comprehensive Rasch analysis. The eigenvalues of the first component in the PCA of the standardized residuals after Rasch analysis were 2.26 for the Spanish sample and 2.64 for the Chinese sample, exceeding the recommended threshold of 2.0, and therefore supporting the multidimensional structure of the DASS-21 and confirming that analyzing the three subscales (Depression, Anxiety, and Stress) separately was appropriate. Following this, a likelihood ratio test was employed to determine the most suitable model for each subscale among participants from the Spanish and Chinese samples, revealing that the PCM was more suitable for most cases, except for the Anxiety subscale for Spanish teachers, where the RSM was appropriate; all subsequent Rasch analyses were conducted using these respective models.

For each subscale, the assumption of unidimensionality was confirmed, with eigenvalues of the first component in the PCA of standardized residuals after Rasch analysis all being less than 2.00. Person-item maps demonstrated proper monotonic category functioning across both Spanish and Chinese groups (Figures S1 and S2). The INFIT and OUTFIT analysis showed that only Item 2 of the Anxiety subscale fell outside the acceptable range, while all other items displayed satisfactory validity (Table 3). This problematic item was further investigated through complementary CTT approaches, which confirmed its psychometric weakness: item-rest correlations for Item 2 were notably lower (Spanish: 0.47; Chinese: 0.55) compared to other items (all above 0.60). Similarly, EFA with a forced one-factor solution showed substantially lower factor loadings for Item 2 (Spanish: 0.50; Chinese: 0.57) compared to other items in the Anxiety subscale (Spanish: 0.69–0.80; Chinese: 0.74–0.83). This triangulation of evidence from both Rasch analysis and CTT methods provides robust validation for identifying the problematic performance of Item 2.

The person reliability coefficients for the Depression, Anxiety, and Stress subscales of the DASS-21 were 0.77, 0.68, and 0.85 for Spanish teachers, and 0.79, 0.77, and 0.79 for Chinese teachers, respectively. The Stress subscale reached or neared the threshold for good reliability (>0.80), with a value of 0.85 in the Spanish sample (exceeding the threshold) and 0.79 in the Chinese sample (falling just below it), suggesting it can generally distinguish 2–3 levels of the trait. The Depression and Anxiety subscales were slightly below this criterion in most cases, with values ranging from 0.77 to 0.79, except for the Spanish Anxiety subscale (0.68), which exhibited notably lower reliability. Corresponding person separation indices were 1.83, 1.46, and 2.38 for Spanish teachers, and 1.94, 1.83, and 1.94 for Chinese teachers. The Stress subscale exceeded the threshold for good separation (>2.00) in the Spanish sample (2.38) but was slightly below it in the Chinese sample (1.94), while the Depression and Anxiety subscales were close to this level, ranging from 1.83 to 1.94, except for the Spanish Anxiety subscale (1.46), which fell below the minimum acceptable

Table 3 Fit statistics of rasch analysis for items

Source	Spanish teachers				Chinese teachers			
	Difficulty	Point-biserial	Infit MnSq	Outfit MnSq	Difficulty	Point-biserial	Infit MnSq	Outfit MnSq
Subscale of Depression								
3. Couldn't experience positive feeling	0.26	0.78	0.94	0.92	-0.75	0.81	1.30	1.31
5. Difficult to work up the initiative to do things	-0.46	0.74	1.48	1.44	0.07	0.80	1.16	1.15
10. Had nothing to look forward to	-0.16	0.81	0.79	0.70	-0.65	0.83	1.02	1.03
13. Felt down-hearted and blue	-1.61	0.86	0.92	0.91	-0.41	0.84	0.90	0.91
16. Unable to become enthusiastic	0.30	0.79	0.81	0.68	0.11	0.83	0.81	0.81
17. Not worth much as a person	0.35	0.72	1.02	1.15	0.96	0.71	0.87	0.59
21. Life was meaningless	1.33	0.61	1.04	0.95	0.67	0.75	0.91	0.89
Subscale of Anxiety								
2. Dryness of my mouth	-1.48	0.72	1.47	1.62	-1.38	0.74	1.63	1.68
4. Experienced breathing difficulty	-0.11	0.74	0.98	0.90	0.05	0.79	0.94	0.89
7. Experienced trembling	1.09	0.63	0.95	0.79	0.72	0.73	0.91	0.75
9. Worried about situations of panic and making a fool of myself	-0.25	0.73	1.00	1.04	-0.60	0.80	1.10	1.12
15. Feeling of close to panic	0.83	0.68	0.84	0.74	0.44	0.79	0.72	0.64
19. Aware of the action of my heart	-0.44	0.78	0.85	0.81	0.22	0.79	0.82	0.82
20. Felt scared without any good reason	0.36	0.71	0.87	0.83	0.55	0.76	0.78	0.69
Subscale of Stress								
1. Hard to wind down	-0.77	0.82	0.88	0.87	-0.20	0.71	1.50	1.49
6. Tended to over-react	0.98	0.69	1.21	1.22	0.55	0.75	0.94	0.89
8. Using a lot of nervous energy	-0.99	0.79	1.11	1.13	-0.98	0.8	1.13	1.15
11. Getting agitated	-0.85	0.81	0.93	0.92	0.31	0.81	0.68	0.63
12. Difficult to relax	-0.92	0.85	0.72	0.71	-0.01	0.83	0.67	0.64
14. Intolerant of anything	1.98	0.63	1.19	1.24	0.18	0.76	1.08	1.10
18. Felt rather touchy	0.56	0.76	0.99	0.95	0.15	0.77	0.96	0.96

Values in gray shadows indicate that they were outside the acceptable range for INFIT and OUTFIT MnSq (0.5 to 1.5).

separation of 1.50. Therefore, while most subscales demonstrated reliability and separation values approaching acceptable standards, the Spanish Anxiety subscale stood out as a concern due to its weaker performance.

In contrast, item performance was notably stronger across both samples, with item reliability coefficients of 0.99 for all assessed subscales, surpassing the 0.90 threshold for excellent reliability. Item separation indices were also robust, ranging from 10.70 (Depression), 13.48 (Anxiety), to 20.26 (Stress) in the Spanish sample, and from 12.06 (Depression) to 15.22 (Anxiety) in the Chinese sample, with a Stress item separation of 11.30. These high item separation indices far exceeded the threshold of 3, indicating exceptional measurement precision and construct stability across both cultural contexts.

Construct validity with confirmatory factor analysis and structural equation modeling

Table 4 presents the model fit results for various factor structures of the DASS-21 among primary school teachers in Spain and China, encompassing the single-factor, correlated two-factor (depression and combined anxiety-stress), correlated three-factor, and bifactor models. The findings show that the single-factor, correlated two-factor, and correlated three-factor models all exhibited acceptable fit within the teacher populations of both countries. However, applying the parsimony principle, no significant χ^2 difference was observed among the single-factor, correlated two-factor, and correlated three-factor models for Chinese teachers, indicating the single-factor model as preferable due to its simplicity. In contrast, the correlated three-factor model demonstrated superior fit for Spanish teachers, evidenced by a significant χ^2 difference relative to the other two models. Consequently, the standardized factor loadings for the selected single-factor (Chinese teachers) and correlated three-factor (Spanish teachers) models are presented in Table 5, demonstrating strong convergent validity with most values exceeding 0.70. Although the bifactor model yielded the best fit across all structures, its specific factor loadings were deemed unreasonable (e.g., negative values, see Table S2), leading to its exclusion as a suitable representation of the DASS-21 structure in these populations.

Although the bifactor model was not supported in either country, ECV and ωH based on its factor loadings were calculated (see Table S2) to explore the presence of a general factor within the DASS-21. The results showed that general factor loadings ranged from 0.55 to 0.82 for Spanish teachers and from 0.62 to 0.86 for Chinese teachers, with an ECV of 0.82 (Spain) and 0.92 (China), and ωH of 0.91 (Spain) and 0.96 (China). Given a PUC of 0.70 for both countries, these findings suggest that the common variance of the scale is essentially unidimensional, albeit with a slight bias. Consequently, while the correlated three-factor structure was better supported among Spanish teachers, the scale scores should primarily be interpreted as reflecting a general factor of overall psychological distress. Moreover, the ωH for the specific factors of depression, anxiety, and stress were 0.15, 0.17, and 0.12 for Spanish teachers, respectively, and notably lower for Chinese teachers at 0.05, 0.01, and 0.02, respectively, reinforcing the model selection. The correlated three-factor structure was better supported for Spanish teachers, while the unidimensional structure was preferred for Chinese teachers.

In assessing the concurrent validity of the DASS-21 with respect to emotional exhaustion, this association was examined by adopting a correlated three-factor structure for Spanish teachers and a unidimensional factor structure for Chinese teachers within an SEM framework. The results indicated acceptable model fit for participants from both countries. For Spanish teachers (see Fig. 1), the fit indices were: χ^2 (364) = 635.44, RMSEA = 0.024, SRMR = 0.040, CFI = 0.996, and NNFI = 0.996; for Chinese teachers (see Fig. 2), the indices were: χ^2 (430) = 1889.42, RMSEA = 0.037, SRMR = 0.051, CFI = 0.990, and NNFI = 0.990. These values satisfied established criteria for model validity, permitting further

Table 4 Results of CFA model fit

	χ^2 (df)	CFI	NNFI	RMSEA; 90% CI	SRMR
Spain					
One-factor structure	548.29 (189)	0.991	0.991	0.038; 0.034–0.041	0.055
Two-factor structure	468.11(188)	0.993	0.993	0.033; 0.029–0.037	0.050
Three-factor structure	342.64 (186)	0.996	0.996	0.025; 0.021–0.029	0.043
Bifactor model	205.65 (168)	0.999	0.999	0.013; 0.005–0.019	0.033
China					
One-factor structure	519.96 (189)	0.996	0.995	0.026; 0.023–0.029	0.045
Two-factor structure	516.36 (188)	0.996	0.995	0.026; 0.023–0.029	0.045
Three-factor structure	514.34 (186)	0.996	0.995	0.026; 0.023–0.029	0.044
Bifactor model	435.44 (168)	0.997	0.996	0.025; 0.022–0.028	0.040

CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CI = confidence interval

Table 5 Factor loadings for the One-Factor and correlated Three-Factor models of the DASS-21

	One-factor structure		Three-factor structure					
	Spanish	Chinese	Spanish			Chinese		
	Distress		Dep	Anx	Stress	Dep	Anx	Stress
Item 3	0.77	0.76	0.81			0.76		
Item 5	0.67	0.76	0.70			0.76		
Item 10	0.75	0.81	0.79			0.82		
Item 13	0.82	0.86	0.86			0.87		
Item 16	0.74	0.83	0.79			0.83		
Item 17	0.68	0.71	0.73			0.72		
Item 21	0.58	0.74	0.62			0.75		
Item 2	0.54	0.61		0.57			0.60	
Item 4	0.66	0.78		0.71			0.78	
Item 7	0.59	0.76		0.63			0.75	
Item 9	0.71	0.78		0.75			0.78	
Item 15	0.73	0.86		0.77			0.86	
Item 19	0.72	0.80		0.77			0.80	
Item 20	0.69	0.81		0.74			0.80	
Item 1	0.75	0.62			0.79			0.62
Item 6	0.64	0.79			0.67			0.79
Item 8	0.69	0.71			0.72			0.71
Item 11	0.75	0.86			0.79			0.86
Item 12	0.78	0.84			0.82			0.84
Item 14	0.67	0.72			0.69			0.72
Item 18	0.72	0.78			0.75			0.78

Dep = Depression subscale; Anx = Anxiety subscale; Stress = Stress subscale

analysis of the path coefficients. Notably, the correlated three-factor structure for Chinese teachers was explored to align with the Spanish model, given its acceptable fit among Chinese participants. However, excessively high correlations among depression, anxiety, and stress raised concerns about potential multicollinearity. This resulted in unreasonable coefficient estimates (exceeding 1.0) and variance inflation factor (VIF) values approaching 10, rendering this model unsuitable for Chinese teachers.

Subsequent analysis showed significant standardized coefficients for the associations between the DASS-21 factors and emotional exhaustion. For Spanish teachers, the standardized coefficients were 0.26 ($t = 3.64$, $p < 0.001$) for depression and 0.68 ($t = 11.00$, $p < 0.001$) for stress, with no significant association for anxiety; for Chinese teachers, the standardized coefficient for the psychological distress factor was 0.54 ($t = 68.97$, $p < 0.001$). Collectively, these findings provide robust support for the concurrent validity of the DASS-21 in relation to emotional exhaustion across these two populations.

Testing measurement invariance between Spanish and Chinese primary schoolteachers

In the final analysis phase, measurement invariance of the DASS-21 was examined across Spanish and Chinese primary school teachers. Utilizing DIF within the framework of Rasch analysis (see Table 3), the findings indicated that most items in the Depression subscale exhibited substantial DIF, with the exceptions of Items 10 and 16, which showed no significant functional discrepancies across teachers in either country. In the Anxiety subscale, Item 19 displayed pronounced DIF, while in the Stress subscale, significant DIF was observed for Items 1, 11, 12, and 14.

Moreover, the analysis examined whether individual DIF items influenced model-level invariance across the two countries. The nested CFA models were conducted using the correlated three-factor model because DIF was identified through individual Rasch analyses on the three subscales. Additionally, to enable comparison across the two countries, the same structure was required, and the correlated three-factor model demonstrated no significant difference in fit compared to other models for the Chinese sample while providing the best fit for the Spanish sample, facilitating the examination of measurement invariance. When comparing nested models, a significant deterioration in model fit was observed upon constraining factor loadings to be equal across the two countries, indicating a failure to achieve weak invariance, as evidenced by changes in model fit indices (see Table 6): ΔCFI = -0.010, ΔRMSEA = 0.021, and ΔSRMR = 0.021.

Subsequently, items were excluded if they showed a difficulty contrast exceeding 0.5 from the DIF analysis. This led to the removal of Items 3, 13, 17, and 21 from the

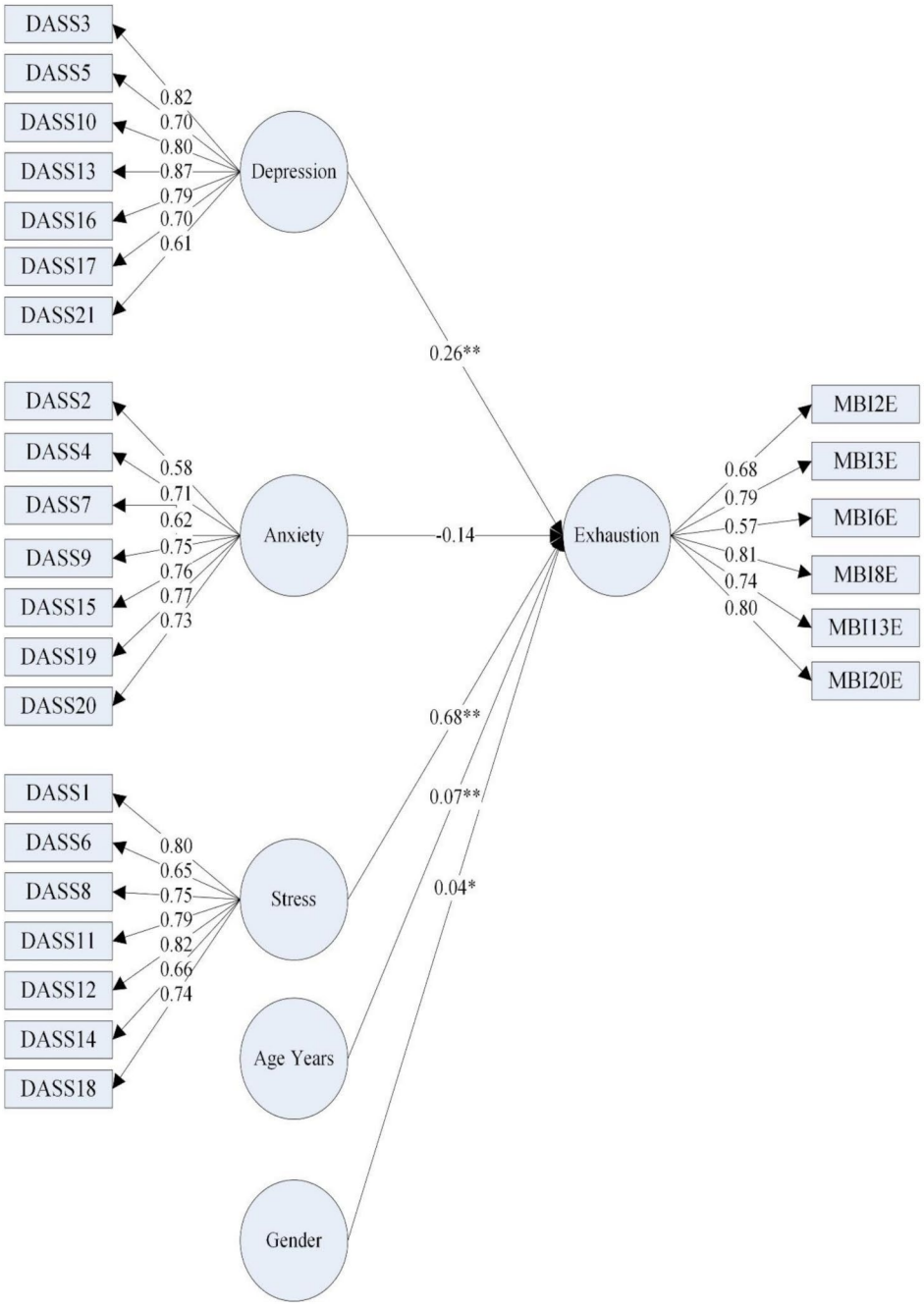

Fig. 1 Structural equation model of the relationship between Depression, Anxiety, and Stress Scale (DASS), exhaustion, age, and gender in a Spanish sample. $^{**}p < 0.01$, $^{*}p < 0.05$

Depression subscale; Item 19 from the Anxiety subscale; and Items 11, 12, and 14 from the Stress subscale. Due to the reduction of the Depression subscale to only two items after this exclusion, Item 5 was retained because its DIF (0.53) was marginally deviated from 0.50 threshold. The post-exclusion results demonstrated that the correlated three-factor model of the scale could attain strict measurement equivalence, characterized by passing tests for both factor loading and item intercept constraints equally, in the sample of primary schoolteachers from both countries (see Table 6).

Discussion

The present study assessed the psychometric properties of the DASS-21 among primary teachers in Spain and China to support future international research in specific demographics. Results from Rasch analysis indicated generally satisfactory item validity for most DASS-21

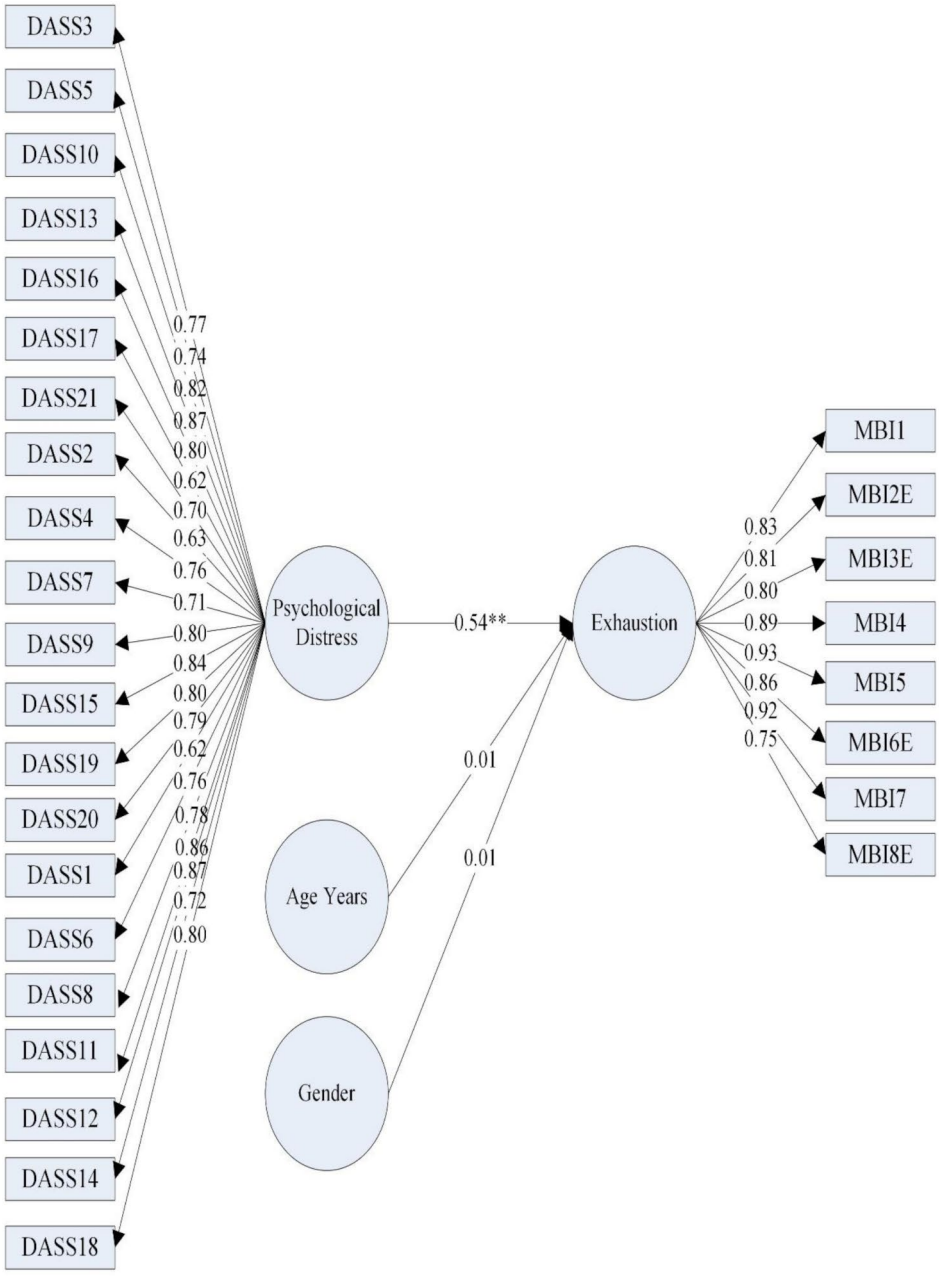

Fig. 2 Structural equation model of the relationship between Depression, Anxiety, and Stress Scale (DASS), exhaustion, age, and gender in a Chinese sample. $^{**}p < 0.01$, $^{*}p < 0.05$

items among both the Spanish and Chinese teacher cohorts, except for one item in the Anxiety subscale, which was also verified through a complementary CTT examination. CFA indicated that the bifactor model, despite its superior statistical fit, was excluded due to problematic negative loadings. For the Chinese cohort, the single-factor model was selected because intercorrelations between the DASS-21 factors exceeded 0.90, with an ECV of 0.92 and an ωH of 0.96, reflecting a dominant general distress factor. Conversely, the correlated three-factor model was chosen for the Spanish cohort, offering better fit and theoretical alignment, with an ECV of 0.82 and an ωH of 0.91, yet retaining distinct depression, anxiety, and stress factors due to meaningful unique variance (ωH ranging from 0.12 to 0.17). SEM confirmed concurrent validity, showing strong associations between DASS-21 factors and emotional exhaustion among both cohorts. DIF testing showed cross-cultural differences, and excluding affected items achieved strict measurement invariance.

Table 6 Measurement invariance testing: changes in fit indices between Spain and China

Model	χ^2	df	CFI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	ΔCFI	ΔRMSEA	ΔSRMR
Full items										
Configural Invariance	856.98	372	0.996	0.026	0.044	-	-	-	-	-
Loading constrained	2101.27	390	0.986	0.047	0.065	1244.29	18	-0.010	0.021	0.021
Thresholds and loadings constrained	3085.74	408	0.978	0.058	0.073	984.47	18	-0.008	0.011	0.008
Items after removed [a]										
Configural Invariance	319.97	124	0.995	0.029	0.039	-	-	-	-	-
Loading constrained	629.61	134	0.989	0.043	0.054	309.64	10	-0.006	0.014	0.015
Thresholds and loadings constrained	787.37	144	0.985	0.048	0.057	157.76	10	-0.004	0.005	0.003

Notes: CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; a: The model labeled "Items after removed" includes the final set of retained items (Items 1, 2, 4, 5, 6, 7, 8, 9, 10, 15, 16, 18, and 20) following psychometric evaluation

The Rasch analysis and CFA, supplemented by bifactor indices such as ECV and ωH, yielded seemingly contradictory findings on the DASS-21's dimensionality due to their distinct methodological lenses, with Rasch emphasizing item-level properties and CFA focusing on overall factor structure. Rasch analysis often detects multidimensionality because it flags deviations from a single latent trait through high eigenvalues in PCA of residuals (e.g., >2.0), reflecting the DASS-21's design to assess three related constructs—depression, anxiety, and stress. In contrast, CFA with a bifactor model highlights a dominant general distress factor, supported by high ECV and ωH values (e.g., 0.92 and 0.96), suggesting that the scale can be interpreted unidimensionally at the factor level despite item-level complexity. This interplay indicates that while Rasch underscores the DASS-21's theoretical multidimensionality, CFA supports its practical use as a unidimensional measure when the general factor predominates, offering a dual perspective that enriches its application in research and clinical settings across diverse populations.

Rasch model analysis of item validity in DASS-21

In the present study's Rasch analysis of the DASS-21, critical insights were garnered, particularly within the Anxiety subscale for Spanish and Chinese primary schoolteachers. Item 2 ('*dryness of my mouth*') in the Anxiety subscale demonstrated an outfit of 1.62 for Spanish teachers and INFIT and OUTFIT of 1.63 and 1.68, respectively, for Chinese teachers, exceeding the acceptable range and indicating diminished item validity. This finding is consistent with the results of Shea et al. [58] and Cao et al. [20]. The elevated outfit for Item 2 suggests a divergence from the constructs assessed by other items, because a high outfit represents a deviation from the construct assessed by other items [93]. It is hypothesized that this item might be capturing symptoms related to stress, because physiological responses such as mouth dryness can manifest under high stress [94].

Additionally, Rasch reliability analysis showed further concerns: person reliability coefficients for the Anxiety subscale were 0.68 for Spanish teachers and 0.77 for Chinese teachers, meeting the minimum threshold of 0.50 for basic reliability but falling below the 0.80 threshold for good reliability, indicating a constrained ability to distinguish nuanced levels of anxiety in these cohorts. Corresponding person separation indices were 1.46 (Spanish) and 1.83 (Chinese), with the Spanish value failing to meet the minimum acceptable separation of 1.50, further underscoring the subscale's limited precision in this group. At a practical level, while the Anxiety subscale can still broadly differentiate between low and high anxiety levels (e.g., identifying teachers with minimal versus elevated anxiety), its inability to reliably detect finer gradations of severity—particularly among Spanish teachers—may reduce its utility for precise clinical screening or tailored interventions requiring detailed anxiety profiling. In contrast, item reliability was robust at 0.99 across both cohorts, exceeding the 0.90 threshold for excellent reliability, suggesting that the items themselves consistently assess the intended constructs despite the person-level reliability challenges.

DASS-21: DIF in cross-national comparisons among primary schoolteachers

In order to comprehensively assess measurement invariance in the DASS-21 among primary schoolteachers in Spain and China, both Rasch analysis (DIF) and multiple-group CFA were employed. The results showed substantial DIF contrasts, exceeding 0.5 in several items. These included Items 3, 13, 17, and 21 in the Depression subscale, Item 19 in the Anxiety subscale, and Items 11, 12, and 14 in the Stress subscale. Following the removal of these items, the multi-group CFA results confirmed that the remaining DASS-21 items met the criteria for strict measurement invariance. It is important to note that the removal of specific DASS-21 items does not necessarily implicate their lack of validity but rather their unsuitability for direct comparison between the two countries. This highlights the importance of context-sensitive item selection in cross-cultural psychological assessment tools such as the DASS-21. Consequently, the present study

provides valuable insights into the nuances of cross-national comparative psychological research.

These findings resonate with previous cross-national comparative studies. More specifically, Oei et al. [33] identified significant cross-loadings for Items 8, 11, and 12 of the DASS-21 among samples from Asian countries, a phenomenon not observed in Western populations. They attributed this discrepancy to cultural differences, concluding that these items may not be appropriate for use in Asian contexts. Similarly, Bibi et al. [31] conducted a measurement invariance analysis of the DASS-21 among university students in Pakistan and Germany. Their analysis indicated that Items 2, 9, and 13 from the depression and Anxiety subscales, along with Items 6, 8, 12, and 14 from the Stress subscale, were inappropriate for cross-group comparisons between the two cohorts.

The present study further corroborates these findings, showing that Item 13 from the Depression subscale, and Items 11, 12, and 14 from the Stress subscale, exhibited significant DIF and are not suitable for comparing emotional disorders among teachers in Spain and China. Collectively, the present study's findings align with prior investigations, emphasizing the imperative of cautious DASS-21 application across diverse national and cultural backdrops. This highlights the criticality of accounting for cultural variances and ensuring measurement invariance in the realm of cross-national research endeavors.

Factor structure and correlation among emotional disorders within DASS-21 among teachers in two countries

The present study's findings diverged from prior research on the DASS-21's factor structure [26, 95, 96], which often identified the bifactor model as the best representation compared to single-factor or correlated three-factor models. Despite the bifactor model demonstrating the best statistical fit in the present study, it was excluded due to problematic factor loadings on the specific factors, including negative values and notably low loadings, which compromised its interpretability and theoretical coherence. For the Chinese cohort, the single-factor model was preferred due to very high intercorrelations between DASS-21 factors, indicating a lack of discriminant validity and supporting a unidimensional structure, further reinforced by bifactor indices that highlighted a dominant general distress factor with minimal contribution from specific factors. Conversely, the correlated three-factor model was selected for the Spanish cohort because it provided a better fit and aligned with the scale's theoretical framework distinguishing depression, anxiety, and stress, while bifactor indices suggested a strong general factor alongside meaningful unique variance from the specific factors.

High correlations between the three DASS-21 factors were observed in both groups, with the Chinese cohort showing particularly strong overlaps, indicating substantial similarity in the constructs of depression, anxiety, and stress. While strong correlations are common in teacher-focused studies, such as those by Ishak et al. [97] during the COVID-19 pandemic in Malaysia and Lacomba-Trejo et al. [98] among Chilean teachers, the notably higher correlations in the Chinese cohort are less typical but have been noted in other Chinese teacher populations [20, 99]. This suggests that the DASS-21 may have limitations in differentiating between the three mood states (depression, anxiety, and stress) when used with Chinese teachers, a pattern also observed in studies involving primary school students where strong correlations among DASS-21 factors indicate challenges in distinguishing between depression and anxiety [100].

This pattern may reflect cultural differences in emotional expression between Eastern and Western contexts [101]. Traditional Chinese culture often encourages implicit emotional articulation [102–104], potentially blurring distinctions between psychological and somatic symptoms, a phenomenon also observed among Chinese adolescents [105]. Consequently, when conducting cross-cultural comparisons of teachers' mental health, interpreting DASS-21 scores as indicators of overall emotional distress, rather than specific emotional disorders, may be more appropriate, particularly among populations where cultural practices influence emotional differentiation [106].

Limitations

While the present study had several significant findings, it is not without limitations. Firstly, in terms of sample selection, the sample primarily consisted of participants from specific regions or specific types of schools, and this non-random sampling may not fully represent the general population of teachers in each country. Additionally, Spanish teachers were notably older than their Chinese counterparts. This discrepancy in age matching among participants may stem from differences in demographic factors between the two countries. This discrepancy in age could potentially impact the outcomes of the present study, particularly when comparing the differences in DASS scores between teachers from the two countries.

Secondly, in terms of criterion validity, the present study opted for emotional exhaustion as the criterion variable. Although emotional exhaustion is one of the most frequently discussed correlates of negative emotions within the teaching community, the study did not include other potential criterion variables that might be associated with the DASS among this population. To enhance the assessment of the DASS-21's effectiveness in assessing teachers' emotional states, future research could consider exploring a more diversified set of criterion variables. For instance, the inclination to leave the

profession could be employed as a criterion variable, with the DASS-21 elucidating teachers' intentions to resign. Such a research design would offer more comprehensive and multi-faceted validity evidence, thereby enriching the understanding of emotional measurement among teachers.

The study also utilized two distinct measures of emotional exhaustion. The differences between these measurements may have influenced the correlation coefficients when emotional exhaustion is used as the criterion variable in relation to the DASS-21 scores. Lastly, it is important to acknowledge that the study primarily relied on self-reported data to assess teachers' emotional states. Therefore, participants' responses may be influenced by social desirability. In the context of this study, teachers may feel compelled to report emotions that conform to societal norms, potentially affecting the accuracy of the collected data.

Conclusion

In the present study, the DASS-21 demonstrated commendable psychometric properties and practical value in assessing the psychological health of primary schoolteachers in both Spain and China. The integration of Rasch model analysis and CFA results indicated satisfactory validity across most items of the DASS-21. However, due to variations in item difficulties across the DASS-21 items between the two countries, direct comparison of total scores using the full set of items is not advisable; instead, the influence of items exhibiting DIF must be considered to ensure accurate cross-cultural comparisons. Moreover, it is not recommended to directly apply the interpretations of individual subscale scores (such as those of the Anxiety subscale) to the Chinese teacher population. The scale provides researchers and clinical professionals with a potent tool for understanding and intervening in teachers' psychological health issues, thereby enhancing teachers' work efficacy and quality of life. Nonetheless, future research should further explore the applicability and consistency of the DASS-21 across different cultural backgrounds and professional groups to enhance the scale's widespread application and interpretability.

Abbreviations

DASS-21	Depression Anxiety and Stress scale 21
CFA	Confirmatory factor analysis
SEM	Structural equation modeling
DIF	Differential item functioning
EFA	Exploratory factor analysis
SSO	Stressor–Strain–Outcome
CTJBO	Chinese Teachers' Job Burnout Questionnaire
PCM	Partial credit Rasch model
RSM	Rating scale model
PCA	Principal component analysis
INFIT	Information-weighted fit statistic
OUTFIT	Outlier-sensitive fit statistics
CFI	Comparative fit index
NNFI	Non-normed fit index
RMSEA	Root mean square error of approximation
SRMR	Standardized root mean square residual
ECV	Explained common variance
PUC	Percentage of uncontaminated correlations
ωH	Omega hierarchical
DIF	Differential item functioning

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40359-025-02728-7.

Supplementary Material 1

Acknowledgements

We thank all the participants who took part in the study.

Author contributions

XW made contributions to conceptualization, formal analysis, and writing - original draft preparation. CC made contributions to writing - review and editing, resources. XL made contributions to the methodology, formal analysis, and resources. XJ made contributions to the investigation, and data curation. MDG made contributions to writing - review and editing. IC made contributions to the conceptualization, methodology, validation, investigation, writing - original draft preparation, writing - review and editing, supervision, and funding acquisition. CL made contributions to the validation, visualization, and project administration. OM made contributions to resources, methodology, supervision, and project administration. All authors reviewed the manuscript.

Funding

This research received funding from the Key Project of Guangxi Higher Education Teaching Reform Project in 2022: Mining of Curriculum Ideological and Political Elements and Teaching Practice for the Major of Mathematics and Applied Mathematics under the Background of Excellent Teachers 2.0 (Project No. 2022JGZ165).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (IRB) of the Jiangxi Psychological Consultant Association (IRB ref: JXSXL-2020-J013). The electronically informed consent was obtained from all of the participants, where the purpose of the study, researcher's affiliation, and privacy guarantee were explained.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Faculty of Education, Qufu Normal University, Qufu 273165, China
[2]Yellow River Delta Research Center of Green Low-carbon and High-quality Development, School of Economics Management, Humanities and Law, Shandong Institute of Petroleum and Chemical Technology, Dongying 257061, China
[3]School of Foreign languages, Shandong Women's University, Jinan 250300, China
[4]Faculty of Education, Jiangxi Science and Technology Normal University, Nanchang 330038, China
[5]Yangan Primary School of Qionglai City, Qionglai 611535, China
[6]International Gaming Research Unit, Psychology Department, Nottingham Trent University, Nottingham, UK

[7]Chinese Academy of Education Big Data, Qufu Normal University, Qufu 273165, China
[8]Institute of Allied Health Sciences, College of Medicine, National Cheng Kung University, Tainan 701401, Taiwan, ROC
[9]Biostatistics Consulting Center, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan 701401, ROC
[10]School of Nursing, College of Nursing, Kaohsiung Medical University, Kaohsiung 807378, Taiwan, ROC
[11]Department of Psychology, Sociology and Social Work Faculty of Education, Psychology and Social Work, University of Lleida, Lleida, Spain

Received: 3 April 2024 / Accepted: 10 April 2025
Published online: 29 April 2025

References

1. Fuller FF, Bown OH. Becoming a teacher. Teach Coll Rec. 1975;76(6):25–52.
2. Eklund G, Sundqvist C, Lindell M, Toppinen H. A study of Finnish primary school teachers' experiences of their role and competences by implementing the three-tiered support. Eur J Spec Needs Educ. 2021;36(5):729–42.
3. Yang R, You X, Zhang Y, Lian L, Feng W. Teachers' mental health becoming worse: the case of China. Int J Educ Dev. 2019;70:102077.
4. Chirico F. Religious belief and mental health in Lay and consecrated Italian teachers. J Relig Health. 2017;56(3):839–51.
5. Putwain DW, von der Embse NP. Teacher self-efficacy moderates the relations between imposed pressure from imposed curriculum changes and teacher stress. Educ Psychol. 2019;39(1):51–64.
6. Ballet K, Kelchtermans G. Struggling with workload. Primary teachers' experience of intensification. Teach Teach Dduc. 2009;25(8):1150–7.
7. Klassen RM, Chiu MM. Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. J Educ Psychol. 2010;102(3):741–56.
8. Perryman J, Ball S, Maguire M, Braun A. Life in the pressure cooker–School league tables and english and mathematics teachers' responses to accountability in a results-driven era. Br J Educ Stud. 2011;59(2):179–95.
9. Dicke T, Parker PD, Marsh HW, Kunter M, Schmeck A, Leutner D. Self-efficacy in classroom management, classroom disturbances, and emotional exhaustion: A moderated mediation analysis of teacher candidates. J Educ Psychol. 2014;106(2):569–83.
10. Skaalvik EM, Skaalvik S. Teacher job satisfaction and motivation to leave the teaching profession: relations with school context, feeling of belonging, and emotional exhaustion. Teach Teach Educ. 2011;27(6):1029–38.
11. Martínez-Monteagudo MC, Inglés CJ, Granados L, Aparisi D, García-Fernández JM. Trait emotional intelligence profiles, burnout, anxiety, depression, and stress in secondary education teachers. Pers Indiv Differ. 2019;142:53–61.
12. Stapleton P, Garby S, Sabot D. Psychological distress and coping styles in teachers: A preliminary study. Aust J Educ. 2020;64(2):127–46.
13. Collie RJ, Mansfield CF. Teacher and school stress profiles: A multilevel examination and associations with work-related outcomes. Teach Teach Educ. 2022;116:103759.
14. Desouky D, Allam H. Occupational stress, anxiety and depression among Egyptian teachers. J Epidemiol Glob Health. 2017;7(3):191–8.
15. Phillips AC, Carroll D, Der G. Negative life events and symptoms of depression and anxiety: stress causation and/or stress generation. Anxiety Stress Copin. 2015;28(4):357–71.
16. Hamama L, Ronen T, Shachar K, Rosenbaum M. Links between stress, positive and negative affect, and life satisfaction among teachers in special education schools. J Happiness Stud. 2013;14:731–51.
17. von der Embse N, Ryan SV, Gibbs T, Mankin A. Teacher stress interventions: a systematic review. Psychol Schools. 2019;56(8):1328–43.
18. Crawford JR, Henry JD. The Depression Anxiety Stress Scales (DASS): normative data and latent structure in a large non-clinical sample. Br J Clin Psychol. 2003;42(2):111–31.
19. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck depression and anxiety inventories. Behav Res Ther. 1995;33(3):335–43.
20. Cao C-H, Liao X-L, Jiang X-Y, Li X-D, Chen I-H, Lin C-Y. Psychometric evaluation of the Depression, Anxiety, andStress Scale-21 (DASS-21) among Chinese primary and middle school teachers. BMC Psychol. 2023;11(1):209.
21. Gong X, Xie X-y, Xu R, Luo Y-j. Psychometric properties of the Chinese versions of DASS-21 in Chinese college students. Chin J Clin Psychol. 2010;18(4):443–6.
22. Ramasawmy S. Validation of the French Depression Anxiety Stress Scales (DASS-21) and predictors of depression in an adolescent Mauritian population. Aix-Marseille; 2015.
23. Daza P, Novy DM, Stanley MA, Averill P. The Depression Anxiety Stress Scale-21: Spanish translation and validation with a Hispanic sample. J Psychopathol Behav Assess. 2002;24:195–205.
24. Szabo M, Lovibond PF. Development and psychometric properties of the DASS-Youth (DASS-Y): an extension of the depression anxiety stress scales (DASS) to adolescents and children. Front Psychol. 2022;13:766890.
25. Szabó M. The short version of the Depression Anxiety Stress Scales (DASS-21): factor structure in a young adolescent sample. J Adolescence. 2010;33(1):1–8.
26. Lee B, Kim YE. Validity of the Depression, Anxiety, and Stress Scale (DASS-21) in a sample of Korean university students. Curr Psychol. 2022;41(6):3937–46.
27. Thapa DK, Visentin DC, Kornhaber R, Cleary M. Prevalence and factors associated with depression, anxiety, and stress symptoms among older adults: a cross-sectional population-based study. Nurs Health Sci. 2020;22(4):1139–52.
28. Vaughan RS, Edwards EJ, MacIntyre TE. Mental health measurement in a post Covid-19 world: psychometric properties and invariance of the DASS-21 in athletes and non-athletes. Front Psychol. 2020;11:590559.
29. Belcher J, Wuthrich VM, Lowe C. Teachers use of fear appeals: association with student and teacher mental health. Br J Educ Psychol. 2022;92(2):610–26.
30. Scholten S, Velten J, Bieda A, Zhang XC, Margraf J. Testing measurement invariance of the Depression, Anxiety, and Stress Scales (DASS-21) across four countries. Psychol Assess. 2017;29(11):1376–90.
31. Bibi A, Lin M, Zhang XC, Margraf J. Psychometric properties and measurement invariance of Depression, Anxiety and Stress Scales (DASS-21) across cultures. Int J Psychol. 2020;55(6):916–25.
32. Zanon C, Brenner RE, Baptista MN, Vogel DL, Rubin M, Al-Darmaki FR, Gonçalves M, Heath PJ, Liao H-Y, Mackenzie CS. Examining the dimensionality, reliability, and invariance of the Depression, Anxiety, and Stress Scale–21 (DASS-21) across eight countries. Assessment. 2021;28(6):1531–44.
33. Oei TP, Sawang S, Goh YW, Mukhtar F. Using the Depression Anxiety Stress Scale 21 (DASS-21) across cultures. Int J Psychol. 2013;48(6):1018–29.
34. Norton PJ. Depression Anxiety and Stress Scales (DASS-21): psychometric analysis across four Racial groups. Anxiety Stress Copin. 2007;20(3):253–65.
35. Zakariya YF, Bjørkestøl K, Nilsen HK. Teacher job satisfaction across 38 countries and economies: an alignment optimization approach to a cross-cultural mean comparison. Int J Educ Res. 2020;101:101573.
36. Sancho-Gil JM, Sánchez-Valero J-A, Domingo-Coscollola M. Based insights on initial teacher education in Spain. Eur J Teach Educ. 2017;40(3):310–25.
37. Mérida-López S, Extremera N, Rey L. Emotion-regulation ability, role stress and teachers' mental health. Occup Med. 2017;67(7):540–5.
38. Pascual E, Pérez-Jover V, Mirambell E, Ivanez G, Terol M. Job conditions, coping and wellness/health outcomes in Spanish secondary school teachers. Psychol Health. 2003;18(4):511–21.
39. Betoret FD. Stressors, self-efficacy, coping resources, and burnout among secondary school teachers in Spain. Educ Psychol. 2006;26(4):519–39.
40. Gómez-Domínguez V, Gómez-Domínguez T, Navarro-Mateu D, Giménez-Espert MC. The influence of COVID-19 and psychosocial risks on burnout and psychosomatic health problems in non-university teachers in Spain during the peak of the pandemic regressions vs. FsQCA Sustain. 2022;14(20):13030.
41. Garvey AM, Jimeno García I, Fernández CM. Which teachers were most affected in terms of mental health during the strict COVID-19 lockdown in Spain? Teachers Teach. 2023;29(7–8):883–907.
42. Santamaría MD, Mondragon NI, Santxo NB, Ozamiz-Etxebarria N. Teacher stress, anxiety and depression at the beginning of the academic year during the COVID-19 pandemic. Glob Ment Health. 2021;8:e14.
43. Liu S, Onwuegbuzie AJ. Chinese teachers' work stress and their turnover intention. Int J Educ Res. 2012;53:160–70.
44. Tang CSK, Au WT, Schwarzer R, Schmitz G. Mental health outcomes of job stress among Chinese teachers: role of stress resource factors and burnout. J Organ Behav. 2001;22(8):887–901.
45. Yang X, Ge C, Hu B, Chi T, Wang L. Relationship between quality of life and occupational stress among teachers. Public Health. 2009;123(11):750–5.
46. Zhou J, Yuan X, Huang H, Li Y, Yu H, Chen X, Luo J. The prevalence and correlative factors of depression among Chinese teachers during the COVID-19 outbreak. Front Psychiatry. 2021;12:644276.
47. Yi J, Chen I-H, Lin C-Y, Li C-C, Liao X-L, Wei Z-H, Gamble JH. The effect of primary and middle school teachers' problematic internet use and fear of

Covid-19 on psychological need thwarting of online teaching and psychological distress. Healthcare. 2021;9(9):1199.
48. Wang Y, Ramos A, Wu H, Liu L, Yang X, Wang J, Wang L. Relationship between occupational stress and burnout among Chinese teachers: a cross-sectional survey in Liaoning, China. Int Arch Occup Environ Health. 2015;88:589–97.
49. Zhang L, Zhao J, Xiao H, Zheng H, Xiao Y, Chen M, Chen D. Mental health and burnout in primary and secondary school teachers in the remote mountain areas of Guangdong Province in the People's Republic of China. Neuropsych Dis Treat. 2014;10:123–30.
50. Zhou XH, Shen ZZ, Cao CH, Liao XL, Jiang XY, Griffiths MD, Chen I-H, Lin C-Y, Malas O. Psychometric evaluation of DASS versions among Spanish and Chinese teachers using exploratory structural equation modeling (ESEM). A cta Psychol. 2024;251:104626.
51. Li CH. The performance of ML, DWLS, and ULS Estimation with robust corrections in structural equation models with ordinal variables. Psychol Methods. 2016;21(3):369–87.
52. Katsantonis IG. Teachers' self-efficacy, perceived administrative support and positive attitude toward students: their effect on coping with job-related stress. Hell J Psychol. 2020;17(1):1–14.
53. Kakemam E, Navvabi E, Albelbeisi AH, Saeedikia F, Rouhi A, Majidi S. Psychometric properties of the Persian version of Depression Anxiety Stress Scale-21 items (DASS-21) in a sample of health professionals: a cross-sectional study. BMC Health Serv Res. 2022;22(1):111.
54. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthrit Care Res. 2007;57(8):1358–62.
55. Wongpakaran N, Wongpakaran T, Pinyopornpanish M, Simcharoen S, Suradom C, Varnado P, Kuntawong P. Development and validation of a 6-item revised UCLA Loneliness Scale (RULS-6) using Rasch analysis. Br J Health Psychol. 2020;25(2):233–56.
56. Sbeglia GC, Nehm RH. Do you see what I-SEA? A Rasch analysis of the psychometric properties of the Inventory of Student Evolution Acceptance. Sci Educ. 2019;103(2):287–316.
57. Salzberger T, Sinkovics RR. Reconsidering the problem of data equivalence in international marketing research: contrasting approaches based on CFA and the Rasch model for measurement. Int Market Rev. 2006;23(4):390–417.
58. Shea TL, Tennant A, Pallant JF. Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). BMC Psychiatry. 2009;9:1–10.
59. Medvedev ON, Krägeloh CU, Titkova EA, Siegert RJ. Rasch analysis and ordinal-to-interval conversion tables for the depression, anxiety and stress scale. J Health Psychol. 2020;25(10–11):1374–83.
60. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory. Scarecrow Education; 1997.
61. Maslach C. Understanding burnout: definitional issues in analyzing a complex phenomenon. Job stress and burnout. In: Paine WS, editor, Job stress and burnout. Beverly Hills, CA; 1982.pp. 29–40.
62. Koeske GF, Koeske RD. A preliminary test of a stress-strain-outcome model for reconceptualizing the burnout phenomenon. J Soc Serv Res. 1993;17(3–4):107–35.
63. Agyapong B, Obuobi-Donkor G, Burback L, Wei Y. Stress, burnout, anxiety and depression among teachers: A scoping review. Int J Environ Res Public Health. 2022;19(17):10706.
64. Ozoemena EL, Agbaje OS, Ogundu L, Ononuju AH, Umoke PCI, Iweama CN, Kato GU, Isabu AC, Obute AJ. Psychological distress, burnout, and coping strategies among Nigerian primary school teachers: a school-based cross-sectional study. BMC Public Health. 2021;21(1):2327.
65. Yin H, Huang S, Chen G. The relationships between teachers' emotional labor and their burnout and satisfaction: a meta-analytic review. Educ Res Rev. 2019;28:100283.
66. Herrero-Diez MT, Salado-Valdivieso MI, Carbajal-Domínguez S, Allué-Tango M, Villa-Caballero JC, Berbel-Hernández C. Impact of vaccination against SARS-CoV-2 on the incidence of infection in school settings. Vacunas (English Edition). 2023;24(3):203–9.
67. Shangguan Z, Wang MY. China's community-based crisis management model for COVID-19: a zero-tolerance approach. Front Public Health. 2022;10:880479.
68. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1(2):130–49.
69. Linacre ET, Estimating. US class A pan evaporation from few climate data. WATER INT. 1994;19(1):5–14.
70. Chan RC, Xu T, Huang J, Wang Y, Zhao Q, Shum DH, O'Gorman J, Potangaroa R. Extending the utility of the Depression Anxiety Stress Scale by examining its psychometric properties in Chinese settings. Psychiat Res. 2012;200(2–3):879–83.
71. Chen I-H, Chen C-Y, Liao X-I, Chen X-m, Zheng X, Tsai Y-C, Lin C-Y, Griffiths MD, Pakpour AH. Psychometric properties of the Depression, Anxiety, and Stress Scale (DASS-21) among different Chinese populations: a cross-sectional and longitudinal analysis. Acta Psychol. 2023;240:104042.
72. Cao C-H, Dang C-Y, Zheng X, Chen W-G, Chen I-H, Gamble JH. The psychometric properties of the DASS-21 and its association with problematic internet use among Chinese college freshmen. Healthcare. 2023;11(5):700.
73. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. 2nd ed. Sydney, Australia: Psychology Foundation of Australia; 1996.
74. Aluja A, Blanch A, García LF. Dimensionality of the Maslach Burnout Inventory in school teachers. Eur J Psychol Assess. 2005;21(1):67–76.
75. Maslach C, Jackson SE. MBI: Maslach Burnout Inventory; manual research edition. Palo Alto, CA: Consulting Psychologists; 1986.
76. Wu X-c, Qi Y-j, Yu R-r. Zang W-w. Revision of Chinese primary and secondary school teachers' job burnout questionnaire. Chin J Clin Psychol. 2016;24(5):856–60.
77. Maslach C, Jackson SE. The measurement of experienced burnout. J Organ Behav. 1981;2(2):99–113.
78. Chen I-H, Chen X-M, Liao X-L, Zhao K-Y, Wei Z-H, Lin C-Y, Gamble JH. Evaluating the immediate and delayed effects of psychological need thwarting of online teaching on Chinese primary and middle school teachers' psychological well-being. Front Psychol. 2022;13:943449.
79. Chen L, Gamble JH, Chen I-H, Lee Z-H, Fu Q-L. Psychometric properties and measurement invariance for a Chinese version of a psychological need thwarting scale for teachers. Int J Environ Res Public Health. 2020;17(9):3247.
80. Şahin M, Aybek E. Jamovi: an easy to use statistical software for the social scientists. INT J ASSESS TOOLS E. 2019;6(4):670–92.
81. Bishara AJ, Hittner JB. Testing the significance of a correlation with nonnormal data: comparison of Pearson, spearman, transformation, and resampling approaches. Psychol Methods. 2012;17(3):399–417.
82. Linacre J. Rasch measurement computer program user's guide. Beaverton, Oregon: Winsteps; 2016.
83. Jafari P, Bagheri Z, Safe M. Item and response-category functioning of the Persian version of the KIDSCREEN-27: Rasch partial credit model. Health Qual Life Out. 2012;10:1–6.
84. Boone WJ, Staver JR, Yale MS. Rasch analysis in the human sciences. Springer; 2013.
85. Szabó M, Lovibond PF. Anxiety, depression, and tension/stress in children. J Psychopathol Behav. 2006;28:192–202.
86. Jiang LC, Yan YJ, Jin ZS, Hu ML, Wang L, Song Y, Li NN, Su J, Wu DX, Xiao T. The depression anxiety stress Scale-21 in Chinese hospital workers: reliability, latent structure, and measurement invariance across genders. Front Psychol. 2020;11:247.
87. Brown TA. Confirmatory factor analysis for applied research. New York: Guilford; 2014.
88. Rodriguez A, Reise SP, Haviland MG. Applying bifactor statistical indices in the evaluation of psychological measures. J Pers Assess. 2016;98(3):223–37.
89. Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: calculating and interpreting statistical indices. Psychol Methods. 2016;21(2):137–50.
90. Shih CL, Wang WC. Differential item functioning detection using the multiple indicators, multiple causes method with a pure short anchor. APPL PSYCH MEAS. 2009;33(3):184–99.
91. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Model. 2007;14(3):464–504.
92. Tang WK, Wong E, Chiu HF, Lum C, Ungvari GS. The geriatric depression scale should be shortened: results of Rasch analysis. Int J Geriatr Psych. 2005;20(8):783–9.
93. Abetz LM, Savage NW. Burning mouth syndrome and psychological disorders. Aust Dent J. 2009;54(2):84–93.
94. Osman A, Wong JL, Bagge CL, Freedenthal S, Gutierrez PM, Lozano G. The Depression Anxiety Stress Scales—21 (DASS-21): further examination of dimensions, scale reliability, and correlates. J Clin Psychol. 2012;68(12):1322–38.
95. Alfonsson S, Wallin E, Maathz P. Factor structure and validity of the Depression, Anxiety and Stress Scale-21 in Swedish translation. J Psychiatr Ment HLT. 2017;24(2–3):154–62.
96. Ishak AR, Adnan NA, Aziz MY, Nazli SN, Mualif SA, Ishar SM, Suaidi NA, Aziz MYA. The impact of the Covid-19 pandemic on depression, anxiety, and stress

among teachers in Malaysia: a cross-sectional study. Malays J Med Health Sci. 2022;18(8):43–9.
97. Lacomba-Trejo L, Schoeps K, Valero-Moreno S, Del Rosario C, Montoya-Castilla I. Teachers' response to stress, anxiety and depression during COVID-19 lockdown: what have we learned from the pandemic? J School Health. 2022;92(9):864–72.
98. Cheng X, Ma Y, Li J, Cai Y, Li L, Zhang J. Mindfulness and psychological distress in kindergarten teachers: the mediating role of emotional intelligence. Int J Environ Res Public Health. 2020;17(21):8212.
99. Cao C-h, Liao X-l, Gamble JH, Li L-l, Jiang X-Y, Li X-D, Griffiths MD, Chen I-H, Lin C-Y. Evaluating the psychometric properties of the Chinese Depression Anxiety Stress Scale for Youth (DASS-Y) and DASS-21. Child Adol Psych Ment. 2023;17(1):106.
100. Chen I-H, Gamble JH, Lin C-Y. Peer victimization's impact on adolescent school belonging, truancy, and life satisfaction: a cross-cohort international comparison. Curr Psychol. 2023;42(2):1402–19.
101. Soto JA, Perez CR, Kim Y-H, Lee EA, Minnick MR. Is expressive suppression always associated with poorer psychological functioning? A cross-cultural comparison between European Americans and Hong Kong Chinese. Emotion. 2011;11(6):1450–5.
102. Bond MH. Emotions and their expression in Chinese culture. J Nonverbal Behav. 1993;17(4):245–62.
103. Markus HR, Kitayama S. Culture and the self: implications for cognition, emotion, and motivation. Psychol Rev. 1991;98(2):224–53.
104. Wang M, Armour C, Wu Y, Ren F, Zhu X, Yao S. Factor structure of the CES-D and measurement invariance across gender in Mainland Chinese adolescents. J Clin Psychol. 2013;69(9):966–79.
105. Matsumoto D, Kupperbusch C. Idiocentric and allocentric differences in emotional expression, experience, and the coherence between expression and experience. Asian J Soc Psychol. 2001;4(2):113–31.
106. Malas O, Tolsá MD. Depression, Anxiety and Stress Scales (DASS-21): Factor structure, reliability, invariance and validity of the Catalan Version. Anu Psic/UB J Psychol. 2022;52(3):228–40.

Publisher's note