

RESEARCH ARTICLE

Validity and reliability of the Patient Health Questionnaire scale (PHQ-9) among university students of Bangladesh

Mahir A. Rahman¹, Tahia Anan Dhira^{2*}, Abdur Razzaque Sarker¹, Jeenat Mehareen³

1 Bangladesh Institute of Development Studies, Dhaka, Bangladesh, **2** Department of Economics, University of Dhaka, Dhaka, Bangladesh, **3** Department of Economics, East West University, Dhaka, Bangladesh

* dhiraanan@du.ac.bd



Abstract

This study investigated the reliability and factorial validity of Patient Health Questionnaire-9 (PHQ-9) in the context of university students in Bangladesh. The research aimed to assess whether the original one-dimensional model or a model containing both somatic and cognitive-emotional factors is appropriate in the case of a sample of university students. A repeated cross-sectional survey design based on convenience sampling was used to collect data from 677 university students from both public and private universities. The factor structure of the PHQ-9 was assessed using confirmatory factor analysis (CFA). Measurement invariances were assessed across gender, type of university, level of education and victim of domestic violence. Its convergent validity was determined by investigating its correlations with Generalized Anxiety Disorder-7 (GAD-7) and Patient Health Questionnaire Anxiety-Depression Scale (PHQ-ADS). Results showed excellent reliability of PHQ-9 as measured by both Cronbach's α and McDonald's ω . CFA suggested that a modified one-factor model where the error variances between item-3 ('sleeping difficulties') and item-6 ('feeling as a failure'), item-6 and item-9 ('suicidal thoughts'), item-4 ('feeling tired') and item-9, item-3 and item-9 were allowed to covary is appropriate for the sample. This model provided high values of comparative fit index (CFI), goodness of fit index (GFI), and Tucker Lewis Index (TLI), low value of standardized root mean square residual (SRMR) and a non-significant root mean square error of approximation (RMSEA) as well as a high Factor Determinacy Score Coefficient. Correlation between PHQ-9 and GAD-7 was 0.751 and 0.934 between PHQ-9 and PHQ-ADS. Finally, the model is strictly invariant across gender and university type. Overall, the study provided support for modified unidimensional structure for PHQ-9 and showed high internal consistency along with good convergent validity.

OPEN ACCESS

Citation: Rahman MA, Dhira TA, Sarker AR, Mehareen J (2022) Validity and reliability of the Patient Health Questionnaire scale (PHQ-9) among university students of Bangladesh. PLoS ONE 17(6): e0269634. <https://doi.org/10.1371/journal.pone.0269634>

Editor: Pedro Vieira da Silva Magalhaes, Universidade Federal do Rio Grande do Sul, BRAZIL

Received: December 28, 2021

Accepted: May 24, 2022

Published: June 8, 2022

Copyright: © 2022 Rahman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting information files](#).

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Depression is one of the most common mental disorders, affecting more than 300 million people globally [1]. At its worst, depression can lead to suicide [2, 3]. Furthermore, depressive disorders caused about 50 million Years Lived with Disability (YLD) globally in 2015 [1].

Symptoms of depression is particularly prevalent among university students [4–6]. A study indicated that approximately 82% of university students in Bangladesh were suffering from mild to severe symptoms of depression, a rate which is higher than those reported in the studies conducted on other general population [6–8]. University students go through significant changes in their emotional (e.g., loneliness, personal autonomy) and physical environment (e.g., transition from college to university) [9]. They also deal with the stress of academic competition, plan for their future careers, and become more involved in family matters as adults [10, 11]. Hence, increased vulnerability of university students to symptoms of depression is not unexpected.

Despite high prevalence and potentially severe consequences of including decreased quality of life, multiple somatic complaints, and increased mortality due to suicide, cardiovascular diseases, stroke, obesity morbidity, etc., depression has been largely ignored as a mental health condition [12, 13]. People suffering from depressive disorders are frequently under-diagnosed and/or misdiagnosed, which hinders effective treatment [14–16]. To overcome these obstacles, a reliable, valid, and rapid screening tool for depressive disorders is needed, especially for university students. The nine-item Patient Health Questionnaire scale (PHQ-9) was developed with the purpose of screening for, and assessing the severity of depressive symptoms in both clinical and research contexts [17]. The initial validation research of PHQ-9 demonstrated strong reliability, as well as criteria, concept, and external validity, in a large primary care sample [17]. Subsequently, the psychometric properties of PHQ-9 have been evaluated in different samples of patients, including other primary care samples [18–20], psychiatric patients [21, 22], specific groups of medical patients [23–25] and also in multicultural populations [26–28]. In academic setting, validation of PHQ-9 scale was performed among university students of Korea [29], China [30], Nigeria [31] etc. In the context of Bangladesh, PHQ-9 has been widely used in various studies as a screening tool for depressive symptoms among adolescents, adults and university students [6, 32–35]. However, while Dhira et al. (2021) have explored the psychometric properties of Generalized Anxiety Disorder-7 (GAD-7) for a sample of university students, to the best of our knowledge, no such study has been conducted for the case of PHQ-9 [36].

While studies have found evidence of adequate internal consistency and moderate to strong correlation with other comorbid disorders (anxiety, worry, etc.), findings regarding the factor structure of PHQ-9 have not been consistent. Symptoms of depression can include both psychological symptoms such as: loss of interest and enjoyment, feelings of low self-worth, difficulty in concentrating and making decisions, as well as physiological symptoms such as: changes in appetite and weight unrelated to diet, changes in patterns of sleeping, increased fatigue, increase in purposeless physical activity (e.g., inability to sit still, pacing, handwringing), slowed movements or speech, etc. [37]. Accordingly, some studies showed support for a unidimensional factor structure (containing only a somatic factor) which aligns with the result of the original validation study [20, 29, 38, 39]. In contrast, others indicated a two-dimensional factor structure [40–42], containing both somatic and cognitive-emotional factors. The lack of consensus regarding the factor structure of PHQ-9 in different contexts warrants a comprehensive validation study consisting of Bangladeshi university students.

Against this backdrop, we conducted this study with the objective of investigating the reliability and factorial validity of PHQ-9 on a sample of university students in Bangladesh. The research aimed to evaluate whether the original one-dimensional model or a model containing both somatic and cognitive-emotional factors is appropriate for university students. We also assessed the convergent validity of PHQ-9 with other relevant measures of mental health conditions, namely Generalized Anxiety Disorder-7 (GAD-7) and Patient Health Questionnaire Anxiety-Depression Scale (PHQ-ADS). Finally, after testing for measurement invariance, we

examined the mean PHQ-9 scores of the students across different demographic and socioeconomic correlates. We expect that the study will contribute to the growing body of literature pertaining to validation studies assessing symptoms of depressive disorders in university students.

Material and methods

Procedure and sampling

A repeated cross-sectional survey was used to collect responses from the university students of Bangladesh. We utilized a snowball sampling strategy in order to capture both public and private university students. Data was collected in two waves: July 18–July 31, 2020 and February 10–February 22, 2021; using the survey Administration software Google Form [43]. To be eligible for the study, the participants had to meet the following criteria: (a) be willing to participate in the study; (b) be enrolled in any public or private university in Bangladesh; (c) have internet access; and (d) be able to read, write, and comprehend the English questionnaire.

Approximately 1.3 million students currently pursue higher education in 47 public and 107 private universities in Bangladesh [44–46]. Considering this population, we calculated the sample size based on the formula:

$$n = \frac{z^2 p(1-p)}{e^2}$$

where, n is the sample size, z is the selected critical value of the desired confidence level, p is the estimated proportion of an attribute that is present in the population, and e is the desired level of precision. Using 5% margin of error, 99% confidence level, and 50% response distribution, the sample size was estimated to be 666.

The questionnaire was circulated among two public and three private university students. Students from these universities were most likely to have access to a suitable internet connection and also use English as a mode of learning. Therefore, it was convenient for us to reach them through social media platforms while keeping the questionnaire in its original form. The questionnaire (Google Form link) was initially shared with faculty members of those selected universities, and they were asked to distribute the questionnaire in their respective classrooms either via e-mail or through any course material sharing platform that they were using for communication. We also asked the faculty members to encourage the students to pass on the survey link among their classmates. The final collection of data had a sample of 677 participants studying at different levels of university who responded anonymously to a structured questionnaire which included questions regarding socio-demographic information as well as the items of Patient Health Questionnaire (PHQ-9).

Description of the instruments

The PHQ-9 is a self-administered version of the Primary Care Evaluation of Mental Disorders (PRIME-MD) diagnostic instrument for common mental health disorders, which is used to determine severity of initial symptoms of depression, and also to monitor symptom changes and treatment effects over time [47]. Participants are asked how often they have encountered symptoms of depression such as: hopelessness, trouble concentrating, etc. during the last two weeks. Response options for each item range from 0 to 3 on a 4-point Likert-scale (0 = not at all, 1 = several days, 2 = more than half the days and 3 = nearly every day). Adding the scores of all nine items provide the PHQ-9 total score differing from 0 to 27. Several validation studies have detected cut-points of ≥ 5 , ≥ 10 , ≥ 15 and ≥ 20 based on receiver operating

characteristics analyses for PHQ-9, standing for mild, moderate, moderately severe and severe depression levels, respectively [48–51].

We have also used the GAD-7 and PHQ-ADS scale to test for convergent validity of PHQ-9. The self-administered seven-item instrument GAD-7 is used as a screening tool to assess the presence and severity of GAD [52–54]. In the assessment, participants are asked how often during the last two weeks they have encountered anxiety symptoms like feeling nervous, trouble relaxing, etc. The range of the scale's response options and calculation of total score is similar to PHQ-9. [S1 Table](#) describes the items and scores of PHQ-9 and GAD-7 questionnaires.

Table 1. Socio-demographic characteristics.

Variables	Categories	N	% in the sample
Age	18–22 years	550	81.24
	23–27 years	127	18.76
Gender	Male	348	51.40
	Female	329	48.60
Education Level	First year	309	45.64
	Second year	98	14.48
	Third year	131	19.35
	Fourth year	95	14.03
	Masters	44	6.50
University Type	Public University	440	65.19
	Private University	235	34.81
Marital Status	Married	8	1.19
	Single	661	97.93
	Others	6	0.89
Student employment	Yes	102	15.07
	No	575	84.93
Family monthly income	<25,000 BDT	154	22.75
	25,000–54,999 BDT	250	36.93
	55,000–99,999 BDT	174	25.70
	> = 1,00,000 BDT	99	14.62
Principal Income Source	Government Service Holder	189	27.92
	Agricultural wage labor	35	5.17
	Organized Trade/Business	173	25.55
	Pension/ Rent	87	12.85
	Private Service Holder	181	26.74
	Others	12	1.77
Joint Family	No	536	79.17
	Yes	141	20.83
Family Size	< = 4 members	340	50.22
	>4 members	337	49.78
Majority of time spent	Alone	191	28.21
	With family	409	60.41
	With friends	66	9.75
	With pets	11	1.62
Domestic violence	Yes	96	14.18
	No	581	85.82
Victim of domestic violence	Yes	65	67.71
	No	31	32.29

<https://doi.org/10.1371/journal.pone.0269634.t001>

Lastly, the PHQ-ADS is a composite measure that assesses the overall burden of anxiety and depressive symptoms (mental distress) by combining the sum of the PHQ-9 and GAD-7 scores [55]. Thus, the scale can range from 0 to 48, with higher scores indicating higher levels of depression and anxiety symptomatology. Cut points of 10, 20, and 30 on the PHQ-ADS can be considered as thresholds of mild, moderate, and severe distress symptoms, respectively.

Statistical analysis

Characteristics of the items were examined by exploring item mean score and item-intercorrelations (Table 2, S2 Table). While Cronbach's α is widely used as a measure of internal consistency and reliability, there are several issues involved. For example, holding the average inter-item correlation constant, α increases if the number of items increase [56]. If the number of items is sufficiently large, α could be large even though the intercorrelation between the items is generally quite small [57, 58]. Therefore, α does not directly measure internal consistency or homogeneity of item responses. Furthermore, Cronbach's α relies on tau-equivalent measurement model, a measurement model that requires a number of assumptions to be met for the estimate to accurately reflect the data's true reliability [59–61]. Therefore, we use both Cronbach's α and McDonald's ω to measure internal consistency and reliability [61] (Table 2).

For applicability purpose, Bartlett Test of Sphericity and Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was assessed [62]. To analyze construct validity of PHQ-9, confirmatory factor analysis (CFA) was performed with structural equation model (SEM) [63, 64]. Using CFA, we tested four alternative models for the structure of the PHQ-9 to understand the dimensionality of the scale [29]. Model 1 is the original one-factor model suggested by Kroenke et al. [17]. Model 2 is a two-factor model suggested by Krause et al. [65], where the items 'sleeping difficulties', 'fatigue' and 'appetite change' get loaded on a somatic factor. Model 3 is another two-factor model derived by Richardson and Richard [66], where the somatic factor also includes 'concentration difficulties' and 'retardation/slowed down speech and movement' along with the three items suggested in Model 2. Model 4 is the modified one factor model suggested by inspection of the modification indices in our study.

Next, these models were compared using several model fit indices and their criteria, including (i) the chi-square (χ^2) and its degrees of freedom (df), (ii) root mean square error of approximation (RMSEA) and its 90% confidence interval, (iii) comparative fit index (CFI), (iv) goodness of fit index (GFI), (v) Tucker Lewis Index (TLI) and (vi) standardized root mean square residual (SRMR) (Table 3). RMSEA values of less than or equal to 0.05 represents close fit, while values between 0.05 to 0.08 are considered acceptable fit [67, 68]. GFI values greater than 0.9 indicate good fit [69]. CFI [70] and TLI [71] are incremental fit indices and values of greater than or equal to 0.95 of these indices indicate very good fit [72] and values of 0.90 or above are considered acceptable fit [73]. SRMR values up to 0.05 indicate close-fit, while values between 0.05 to 0.10 suggest acceptable fit [73]. We also tested factor score determinacy coefficient to evaluate the goodness of fit of the models. According to Gorsuch (1983) this coefficient should be ≥ 0.90 if the factor score is to be used as a substitute for the factor itself [74].

In order to utilize the PHQ-9 for meaningful comparisons in depressive symptoms across different socioeconomic groups, we tested whether measurement invariance holds across these groups [75]. Thus, we carried out multiple-group confirmatory factor analysis based on the unidimensional modified model between two gender groups (male, female), two types of university students (public, private), five different years of education levels in the university and whether the respondent was a victim of domestic violence to investigate if PHQ-9 assesses the same construct across these groups and that observed differences in PHQ-9 scores among these groups reflect true group differences in depressive symptoms. The first one of the four

increasingly constrained CFA models had all parameters free (configural invariance). The second one took equal loadings (weak invariance), while the third model required equal loadings and intercepts (strong invariance). Lastly, the fourth one is the most constrained model with equal loadings, residuals, and intercepts (strict invariance). The essential criterion for comparing models with additional constraints were the change in CFA and RMSEA. $\Delta\text{CFI} < 0.01$ and $\Delta\text{RMSEA} < 0.015$ support for measurement invariance [76].

To assess convergent validity of the PHQ-9, the association between PHQ-9 and the Generalized Anxiety Disorder (GAD-7) and Patient Health Questionnaire Anxiety and Depression Scale (PHQ-ADS) were examined using Pearson's correlation (r) and its significance. Mean scores of PHQ-9 index across sample characteristics for which measurement invariance holds were also studied using t test and analysis of variance (ANOVA) (Table 4).

Data cleaning, validation, and all statistical analyses were performed using Stata/IC 16.1 (StataCorp, College Station, TX, USA) and R studio, with the packages 'lavaan', 'semTools', and 'psych'.

Ethical considerations

Ethical permission for data collection was taken from respective faculty and department heads of the universities where the questionnaire was distributed. All participants gave their informed consent to anonymously (unidentified to the authors) participate in the study. In the consent form, participants were provided with information concerning the purpose, procedure and nature of the study, the option to take part as well as the right to revoke their data at any point of the study. The research is approved by the Department of Economics, East West University and procedures of this study complied with the provisions of the Declaration of Helsinki (1989) regarding research on human participants.

Results

Table 1 shows the distributions of the key socio-demographic variables. Males made up 51.40% of the 677 participants, while females made up 48.60%, with public university students accounting for over a third of the sample (65.19%). Distribution of PHQ-9 items are represented in S3 Table.

Item characteristics

Item characteristics are summarized in Table 2. The highest reported score is on item-3 'Trouble falling asleep or sleeping too much' (1.58 ± 1.13) and the lowest reported score is on item-9 'Thoughts of dying, or hurting self' (0.64 ± 0.99). Correlation between the items were significant and moderate enough (ranging from 0.10–0.51) to justify conducting factor analysis [77, 78] (S2 Table).

Reliability

The value of the reliability coefficient Cronbach's α and McDonalds' ω for the overall PHQ-9 scale are 0.824 and 0.86 respectively, which is greater than the recommended value of 0.80, indicating excellent reliability [79]. (Table 2).

Construct validity

Construct validity of the scale was tested with confirmatory factor analysis. Applicability of factor analysis was tested using KMO and Bartlett Test of Sphericity. The KMO coefficient is 0.879 surpasses the recommended value of 0.6, while Bartlett Test of Sphericity is found

Table 2. Characteristics of items and total PHQ-9 scale.

PHQ-9 Items	Mean (95% CI)	SD	Factor Loadings	Cronbach's α	McDonald's ω
1. Little interest or pleasure in doing things?	1.27 (1.19–1.34)	0.96	0.276	0.824	0.86
2. Feeling down, depressed, or hopeless?	1.44 (1.36–1.53)	1.08	0.749		
3. Trouble falling or staying asleep, or sleeping too much?	1.58 (1.50–1.67)	1.13	0.653		
4. Feeling tired or having little energy?	1.39 (1.31–1.47)	1.06	0.743		
5. Poor appetite or overeating?	1.03 (0.95–1.11)	1.08	0.673		
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down?	1.31 (1.22–1.40)	1.16	0.730		
7. Trouble concentrating on things, such as reading the newspaper or watching television?	1.28 (1.19–1.37)	1.16	0.663		
8. Moving or speaking so slowly that other people could have noticed?	0.78 (0.71–0.85)	0.99	0.660		
9. Thoughts that you would be better off dead, or of hurting yourself in some way?	0.64 (0.57–0.72)	0.99	0.588		
PHQ-9 Total Score	10.74 (10.27–11.20)	6.20	—		

CI = Confidence Interval; SD = Standard Deviation.

<https://doi.org/10.1371/journal.pone.0269634.t002>

statistically significant ($\chi^2 = 1689.151$, $df = 36$, $p < 0.001$), indicating the suitability of performing factor analysis on this sample [79]. All the items of PHQ-9 have statistically significant loadings ($p < 0.001$). Therefore, all seven items of the measure are important to interpret (Table 2).

We performed CFA on four alternative models. CFA does not satisfy adequate fit criteria for the original one-factor model (Model 1) (Table 3). While the value of SRMR is less than 0.05, the chi-square value is significant at $p < 0.001$ suggesting poor fit. Besides, chi-square provides inflated value when sample size is large and does not work well where sample size is small, and the underlying distribution may be non-normal [80]. Moreover, CFI, GFI and TLI values are less than 0.950, again indicating poor fit. RMSEA value is comparatively higher and statistically significant ($p < 0.001$) suggesting unacceptable fit (Table 3).

Goodness of fit indices for both of the two-factor models also do not meet the required cut off values. Consequently, a modification was conducted to improve the values of goodness of

Table 3. Goodness of fit indices for the PHQ-9 item factor models (N = 677).

Model	k	χ^2	df	CFI	GFI	TLI	RMSEA (90% CI)	SRMR
Model 1	9	153.823***	27	0.924	0.910	0.898	0.083*** (0.071–0.096)	0.046
Model 2	9	113.820***	26	0.947	0.933	0.927	0.071** (0.058–0.084)	0.040
Model 3	9	108.974***	26	0.950	0.936	0.931	0.069** (0.056–0.082)	0.040
Model 4	9	74.740***	23	0.969	0.956	0.951	0.058 (0.043–0.073)	0.033

*** $p < 0.001$;

** $p < 0.05$;

* $p < 0.1$

k = number of items; df = degrees of freedom; CFI = comparative fit index; GFI = goodness of fit index; TLI = Tucker-Lewis index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean residual

Model 1: Originally validated one-factor model

Model 2: Two-factor model of Krause et al. (2008) with item 3, 4 and 5 loaded on one somatic factor and the other six items loaded on an affective factor

Model 3: Two-factor model of Richardson and Richard with item 3,4,5,7,8 loaded on the somatic factor and the others on affective factor

Model 4: Modified one-factor model.

<https://doi.org/10.1371/journal.pone.0269634.t003>

fit indices for the original one-factor model and the error variances between item-3 ('sleeping difficulties') and item-6 ('feeling as a failure'), item-6 and item-9 ('suicidal thoughts'), item-4 ('feeling tired') and item-9, item-3 and item-9 were combined to construct Model 4. Modified one-factor model provides non-significant and the lowest RMSEA, the highest and acceptable values of CFI, GFI and TLI which are all greater than the required cut off score 0.950 [72] and the lowest value of SRMR (0.033) [81]. Furthermore, Factor Determinacy Coefficient suggested that the modified unidimensional model is the most well-defined for this sample (S4 Table).

Convergent validity

Convergent validity of the PHQ-9 was determined by its Pearson's correlations with other measures used in the study. Scores of the PHQ-9 scale were highly and positively correlated with the scores of GAD-7 and PHQ-ADS. Correlation between PHQ-9 and GAD-7 is 0.751 and between PHQ-9 and PHQ-ADS is 0.934. Both the correlations are statistically significant ($p < 0.001$) (S2 Table).

Thus, we conclude that the modified one factor model is the best fit to the data for our sample. All factor loadings and error covariances are statistically significant ($p < 0.001$), suggesting that the indicator variables are significantly related to their respective factor. Confirmatory factor analysis path diagram is represented in Fig 1 and the fit indices are shown in Table 3.

Measurement invariance

Configural, weak and strong invariance models for both gender and university type had statistically insignificant Chi-square, suggesting strong invariance. The strict invariance model for gender and university type were significant at 5% level, according to Chi-square difference. However, as it is highly perceptible to the sample size and minor mis-specification can result in substantial Chi-square difference [82], we emphasize on differences of CFI and RMSEA. All models with progressively stronger constraints exhibited $\Delta CFI < 0.01$ and $\Delta RMSEA < 0.015$, for gender and type of university, suggesting strict invariance (S5 Table). However, victim of domestic violence in the family and level of education are weakly invariant as the Chi-square differences were significant and also the ΔCFI was not less than 0.01 for strong invariant model. These results lend sufficient support for a comparison of mean PHQ-9 scores across these socioeconomic groups of our sample.

PHQ-9 scores across socio-demographic characteristics

Table 4 shows the responsiveness of PHQ-9 index over the values of the variables those are important for our study and have been used to test measurement invariance. From the table, we can see students who are female, studying in public university, and faced domestic violence in the family have significantly stronger symptoms of depression. We also found significantly higher PHQ-9 score for students who are enrolled in third and fourth year of undergraduate studies.

Discussion

PHQ-9 has been used to detect symptoms of depressive disorders across diverse populations, beyond its original application in primary-care settings. However, a paucity of studies conducted on vulnerable groups such as university students necessitates a contribution to the existing gap in the literature. In this context, our study examined the psychometric properties of the PHQ-9 on a sample of university students in Bangladesh, using CFA.

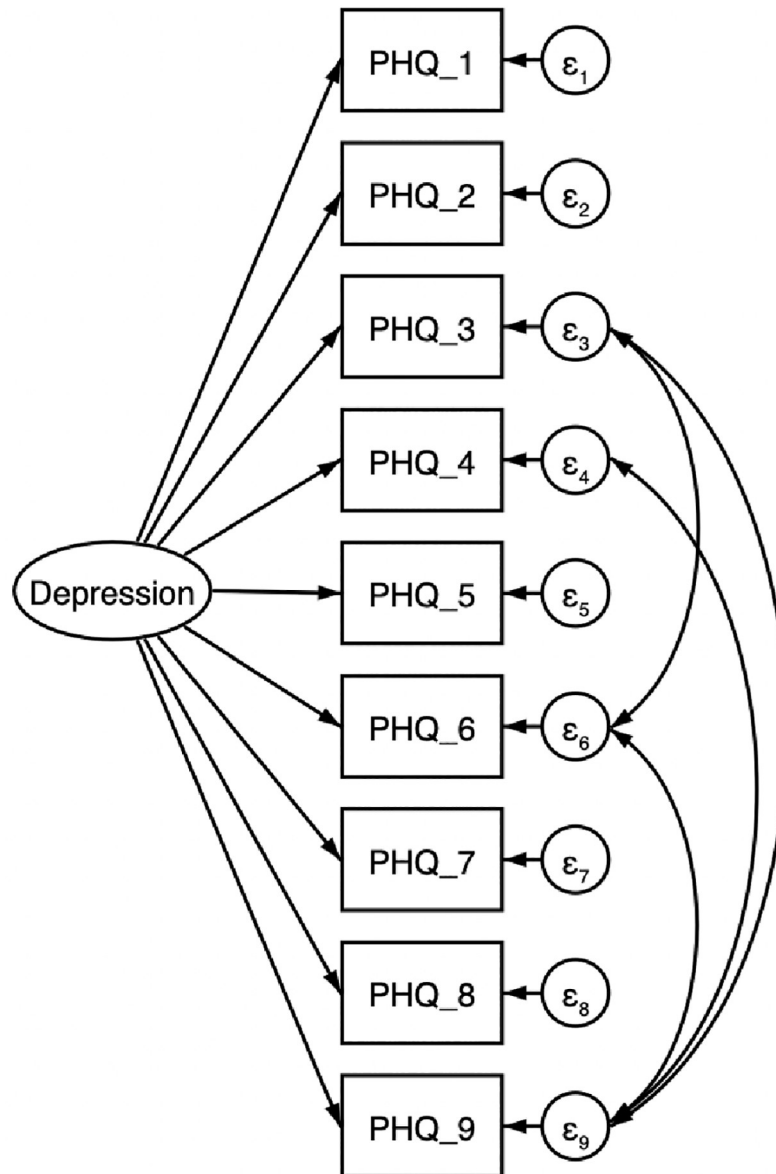


Fig 1. Confirmatory factor analysis path diagram for modified one-factor model of PHQ-9 factors. (All factor loadings and error covariances were significant at $p < 0.001$).

<https://doi.org/10.1371/journal.pone.0269634.g001>

Internal consistency of the scale was excellent in our sample, reflected by the overall Cronbach's α of 0.824. To conclude regardless of tau-equivalence assumption, we have also computed McDonald's ω [61], which is 0.86, ensuring reliability of the scale in our sample. The original validation study conducted on 3000 primary care patients and 3000 ob-gyn patients found excellent internal reliability of PHQ-9, with a Cronbach's α of 0.89 and 0.86 respectively [17]. Later, clinical studies in Chile [83], England [41], Germany [39], United States [84], Iran [85], and China [64] in addition to non-clinical studies conducted in Ghana [86], Hong Kong [87], Philippines [88] revealed the evidence of good internal consistency of PHQ-9 scale across different populations through excellent Cronbach's α coefficient. Specifically, studies

Table 4. Association of PHQ-9 score with socio-demographic characteristics (N = 677).

Variables	Categories	PHQ-9 score, Mean (SD)	t/F	P-value
Gender	Male	9.61 (6.01)	4.922	0.000
	Female	11.92 (6.17)		
Type of University	Public University	11.58 (6.25)	-5.068	0.000
	Private University	9.09 (5.74)		
Level of Education	First year	9.74 (5.95)	5.46	0.000
	Second year	10.48 (6.12)		
	Third year	12.38 (6.47)		
	Fourth year	11.99 (6.41)		
	Masters	10.68 (5.45)		
Domestic violence in family	Yes	12.82 (6.71)	-3.594	0.000
	No	10.39 (6.04)		

*Group differences were performed using t test and analysis of variance. Significant group differences are printed in bold (P < 0.05)

SD = Standard Deviation.

<https://doi.org/10.1371/journal.pone.0269634.t004>

conducted on university students from Korea (Cronbach's $\alpha = 0.83$) [29], Nigeria (Cronbach's $\alpha = 0.85$) [31] and China (Cronbach's $\alpha = 0.80$) [30] found excellent internal consistency of PHQ-9.

We tested the convergent validity of PHQ-9 with two other scales, GAD-7 and PHQ-ADS. Correlation coefficient between PHQ-9 and GAD-7 and, between PHQ-9 and PHQ-ADS were significant and greater than 0.75, suggesting satisfaction of convergent validity (S2 Table). Previous studies have observed the comorbidity of PHQ-9 with anxiety disorders [89–91]. The study on Korean university students has also found good convergent validity of PHQ-9 with GAD-7, with a correlation coefficient of 0.68 [92]. Other studies have also found strong evidence of convergent validity of PHQ-9 with similar psychometric instruments in different settings [29, 31, 47, 93–98]. Together, these findings suggest the reliability and validity of applying the PHQ-9 scale as a measure of symptoms of depression in the context of university students in different countries.

The unidimensional model showed a marginal fit to our context. The original model was therefore revised using the examination of modification indices. Dependency of the error variances between item-3 ('sleeping difficulties') and 6 ('feeling as a failure'), 6 and 9 ('suicidal thoughts'), 4 ('feeling tired') and 9 and, 3 and 9 upgraded the fitness of the model. Our modified one-factor model was partially similar to that of Maroufizadeh et al. (2019) for patients with infertility [85] where they found covariance between Item 1 and Item 2 as well as between Item 7 and Item 8, and Item 6 and Item 9. Furthermore, Beards et al. (2016) found evidence for covariance between item 7 (concentration difficulty) and item 8 (motor slowing/restlessness) in case of a two-factor model [22]. On the other hand, Kim & Lee (2019) found support for a one-factor model in the case of Korean university students [29]. A number of studies have also suggested two-factor models with a cognitive and a physical latent factor [40, 65, 66, 99, 100]. However, the modified one-factor model provided the best values of all the goodness of fit indices as well as in terms of factor score determinacy index (S4 Table). This may be due to our non-clinical sample consisting of university students. In contrast, the studies concluding a somatic and a physical factor for PHQ-9 are mainly conducted with clinical populations such as patients with infertility, spinal cord injury etc. [85, 63].

Aside from the issues discussed above, the mean PHQ-9 scores across different sample characteristics were compared with similar analyses from existing literature on university

students. As a prerequisite, we tested measurement invariance of the scale across gender, type of university, victim of domestic violence and level of education. Strict invariances were observed for gender, that is consistent with the outcomes of some previous research [20, 38, 101]. In contrast, other studies find no or weak measurement invariance across gender [102]. In the context of Bangladesh, examining measurement invariance across public and private university students is also important [5]. Additionally, level of education and victim of domestic violence were found weakly invariant in our study. Previous research primarily based on adult population found level of education to be strictly invariant [20, 75]. Similar results were also obtained for domestic violence [103]. However, the nature of non-clinical student sample could be the reason behind the absence of strong invariance across education level and domestic violence in our study.

As measurement invariance was established, it is meaningful to discuss differences in PHQ-9 scores across the specific sociodemographic groups of university students. In terms of gender, our results show that higher PHQ-9 scores were associated with female students, in line with the findings of other studies [104, 105]. In case of level of education, we observe that students enrolled in higher level of their undergraduate study have significantly higher PHQ-9 scores. Advanced undergraduate students often need to deal with factors such as failure in love affairs, lack of self-confidence, job and financial insecurity and familial problems. All these factors might contribute to low self-esteem which is associated with increasing depression [106]. The results obtained from our sample also show that the PHQ-9 scores are significantly higher for students from public universities. As public university students in Bangladesh mostly come from a poorer socio-economic background compared to private university students, they have an additional pressure of finding jobs just after or even during their study. As a result, fear of delayed completion of degree and uncertainty of jobs are likely to be contributing factors to the high score [5]. Our results also indicate that students who witnessed domestic violence in the family suffer more from depression compared to those who did not [107, 108].

The study has several limitations which should be considered before interpreting the results. First, the data collected through web-based platform captured a relatively homogeneous sample of students characterized by high literacy and easy internet access, potentially resulting in selection bias. Hence, the findings of this study cannot be generalized across other populations such as older adults, adolescents, patients, etc. The nature of non-clinical student sample could be the reason behind the absence of strong invariance across education level and domestic violence. Second, self-reported mental health metrics such as PHQ-9 might be unduly affected by reporting bias [109, 110]. Lastly, sensitivity and specificity of PHQ-9 for university students should be explored in future research.

Conclusion

This is the first study to evaluate the psychometric properties of PHQ-9 in university students of Bangladesh and hence contributes to minimize a major gap in the literature. The study adds to the growing evidence of PHQ-9 as a concise, simply administered self-reported questionnaire. The results also provide support for a modified unidimensional structure of PHQ-9 and show high internal consistency as well as good convergent validity for the sample. Such successful validation of PHQ-9 scale in the context of university students of Bangladesh will allow early diagnosis and treatment, thus helping the policy makers and public health authorities to take necessary and timely interventions to combat the prevalence of such disorders.

Supporting information

S1 Table. PHQ-9 and GAD-7 items and scores.

(DOCX)

S2 Table. Pearson's correlation coefficients (r) between PHQ-9 items and with other questionnaires, (n = 677).

(DOCX)

S3 Table. Distribution of PHQ-9 items.

(DOCX)

S4 Table. Factor determinacy coefficients.

(DOCX)

S5 Table. Multigroup—CFA: Fit measures of the invariance test.

(DOCX)

S1 Dataset. PHQ-9_dataset.

(DTA)

Acknowledgments

We would like to thank all the participants who voluntarily offered their time to complete the questionnaire.

Author Contributions

Conceptualization: Mahir A. Rahman, Tahia Anan Dhira, Jeenat Mehareen.

Data curation: Mahir A. Rahman, Jeenat Mehareen.

Formal analysis: Mahir A. Rahman, Tahia Anan Dhira.

Investigation: Mahir A. Rahman, Tahia Anan Dhira.

Methodology: Mahir A. Rahman, Tahia Anan Dhira.

Supervision: Mahir A. Rahman, Tahia Anan Dhira, Abdur Razzaque Sarker, Jeenat Mehareen.

Validation: Mahir A. Rahman, Tahia Anan Dhira.

Visualization: Mahir A. Rahman, Tahia Anan Dhira.

Writing – original draft: Mahir A. Rahman, Tahia Anan Dhira.

Writing – review & editing: Mahir A. Rahman, Tahia Anan Dhira, Abdur Razzaque Sarker, Jeenat Mehareen.

References

1. World Health Organization. Depression and Other Common Mental Disorders Global Health Estimates. 2017.
2. Kalin NH. Insights Into Suicide and Depression. *American Journal of Psychiatry*. 2020; 177. <https://doi.org/10.1176/appi.ajp.2020.20081207> PMID: 32998546
3. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet*. 2007; 370. [https://doi.org/10.1016/S0140-6736\(07\)61414-7](https://doi.org/10.1016/S0140-6736(07)61414-7) PMID: 17826169

4. Islam MA, Barna SD, Raihan H, Khan MNA, Hossain MT. Depression and anxiety among university students during the COVID-19 pandemic in Bangladesh: A web-based cross-sectional survey. *PLOS ONE*. 2020; 15. <https://doi.org/10.1371/journal.pone.0238162> PMID: 32845928
5. Mehareen J, Rahman MA, Dhira TA, Sarker AR. Prevalence and socio-demographic correlates of depression, anxiety, and co-morbidity during COVID-19: A cross-sectional study among public and private university students of Bangladesh. *Journal of Affective Disorders Reports*. 2021;5. <https://doi.org/10.1016/j.jadr.2021.100179>
6. Mridha MK, Hossain MM, Khan MSA, Hanif AAM, Hasan M, Mitra D, et al. Prevalence and associated factors of depression among adolescent boys and girls in Bangladesh: findings from a nationwide survey. *BMJ Open*. 2021; 11. <https://doi.org/10.1136/bmjopen-2020-038954> PMID: 33455924
7. Faisal RA, Jobe MC, Ahmed O, Sharker T. Mental Health Status, Anxiety, and Depression Levels of Bangladeshi University Students During the COVID-19 Pandemic. *International Journal of Mental Health and Addiction*. 2021. <https://doi.org/10.1007/s11469-020-00458-y> PMID: 33424514
8. National Mental Health Survey in Bangladesh. NIMH Fact Sheet. 2019.
9. Smith L, Jacob L, Yakkundi A, McDermott D, Armstrong NC, Barnett Y, et al. Correlates of symptoms of anxiety and depression and mental wellbeing associated with COVID-19: a cross-sectional study of UK-based respondents. *Psychiatry Research*. 2020; 291. <https://doi.org/10.1016/j.psychres.2020.113138> PMID: 32562931
10. Bernal-Morales B, Rodríguez-Landa JF, Pulido-Criollo F. Impact of Anxiety and Depression Symptoms on Scholar Performance in High School and University Students. *A Fresh Look at Anxiety Disorders*. InTech; 2015. <https://doi.org/10.5772/60711>
11. Arusha AR, Biswas RK. Prevalence of stress, anxiety and depression due to examination in Bangladeshi youths: A pilot study. *Children and Youth Services Review*. 2020; 116. <https://doi.org/10.1016/j.childyouth.2020.105254> PMID: 32834273
12. Penninx BW, Milaneschi Y, Lamers F, Vogelzangs N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC Medicine*. 2013; 11. <https://doi.org/10.1186/1741-7015-11-129> PMID: 23672628
13. Rogers MA, Greene MT, Young VB, Saint S, Langa KM, Kao JY, et al. Depression, antidepressant medications, and risk of Clostridium difficile infection. *BMC Medicine*. 2013; 11. <https://doi.org/10.1186/1741-7015-11-121> PMID: 23647647
14. Wu Z, Fang Y. Comorbidity of depressive and anxiety disorders: challenges in diagnosis and assessment. *Shanghai Archives of Psychiatry*. 2014; 26: 227–231. PMID: 25317009
15. Ibrahim AW, Jidda MS, Wakil MA, Rabbebe IB, Omeiza AB, Yusuph H, et al. Prevalence, correlates and under-diagnosis of clinical depression among adults on highly active antiretroviral therapy in a Tertiary Health Institution in northeastern Nigeria. *Journal of Public Health in Africa*. 2014; 5. <https://doi.org/10.4081/jphia.2014.342> PMID: 28299127
16. Blackburn TP. Depressive disorders: Treatment failures and poor prognosis over the last 50 years. *Pharmacology Research & Perspectives*. 2019; 7. <https://doi.org/10.1002/prp2.472> PMID: 31065377
17. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. *Journal of General Internal Medicine*. 2001; 16. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x> PMID: 11556941
18. Liu S-I, Yeh Z-T, Huang H-C, Sun F-J, Tjung J-J, Hwang L-C, et al. Validation of Patient Health Questionnaire for depression screening among primary care patients in Taiwan. *Comprehensive Psychiatry*. 2011; 52. <https://doi.org/10.1016/j.comppsy.2010.04.013> PMID: 21111406
19. Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, et al. Validation of PHQ-2 and PHQ-9 to Screen for Major Depression in the Primary Care Population. *The Annals of Family Medicine*. 2010; 8. <https://doi.org/10.1370/afm.1139> PMID: 20644190
20. González-Blanch C, Medrano LA, Muñoz-Navarro R, Ruiz-Rodríguez P, Moriana JA, Limonero JT, et al. Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PLOS ONE*. 2018; 13. <https://doi.org/10.1371/journal.pone.0193356> PMID: 29474410
21. Sun Y, Fu Z, Bo Q, Mao Z, Ma X, Wang C. The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. *BMC Psychiatry*. 2020; 20. <https://doi.org/10.1186/s12888-020-02885-6> PMID: 32993604
22. Beard C, Hsu KJ, Rifkin LS, Busch AB, Björgvinsson T. Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*. 2016; 193. <https://doi.org/10.1016/j.jad.2015.12.075> PMID: 26774513
23. Seo J-G, Park S-P. Validation of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with migraine. *The Journal of Headache and Pain*. 2015; 16. <https://doi.org/10.1186/s10194-015-0552-2> PMID: 26174509

24. van Steenberg-Weijenburg KM, de Vroeghe L, Ploeger RR, Brals JW, Vloedveld MG, Veneman TF, et al. Validation of the PHQ-9 as a screening instrument for depression in diabetes patients in specialized outpatient clinics. *BMC Health Services Research*. 2010;10.
25. Monahan PO, Shacham E, Reece M, Kroenke K, Ong'or WO, Omollo O, et al. Validity/Reliability of PHQ-9 and PHQ-2 Depression Scales Among Adults Living with HIV/AIDS in Western Kenya. *Journal of General Internal Medicine*. 2009; 24. <https://doi.org/10.1007/s11606-008-0846-z> PMID: 19031037
26. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry*. 2014; 36. <https://doi.org/10.1016/j.genhosppsych.2014.05.021> PMID: 25023953
27. Lamela D, Soreira C, Matos P, Morais A. Systematic review of the factor structure and measurement invariance of the patient health questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings. *Journal of Affective Disorders*. 2020; 276. <https://doi.org/10.1016/j.jad.2020.06.066> PMID: 32697702
28. Kohrt BA, Luitel NP, Acharya P, Jordans MJD. Detection of depression in low resource settings: validation of the Patient Health Questionnaire (PHQ-9) and cultural concepts of distress in Nepal. *BMC Psychiatry*. 2016; 16. PMID: 26951403
29. Kim YE, Lee B. The Psychometric Properties of the Patient Health Questionnaire-9 in a Sample of Korean University Students. *Psychiatry Investigation*. 2019; 16. <https://doi.org/10.30773/pi.2019.0226> PMID: 31870089
30. Du N, Yu K, Ye Y, Chen S. Validity study of Patient Health Questionnaire-9 items for Internet screening in depression among Chinese university students. *Asia-Pacific Psychiatry*. 2017; 9. <https://doi.org/10.1111/appy.12266> PMID: 28856843
31. Adewuya AO, Ola BA, Afolabi OO. Validity of the patient health questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. *Journal of Affective Disorders*. 2006; 96. <https://doi.org/10.1016/j.jad.2006.05.021> PMID: 16857265
32. Roy T, Lloyd CE, Parvin M, Mohiuddin KGB, Rahman M. Prevalence of co-morbid depression in outpatients with type 2 diabetes mellitus in Bangladesh. *BMC Psychiatry*. 2012; 12. <https://doi.org/10.1186/1471-244X-12-123> PMID: 22909306
33. Kundu S, Bakchi J, Al Banna MH, Sayeed A, Hasan MT, Abid MT, et al. Depressive symptoms associated with loneliness and physical activities among graduate university students in Bangladesh: findings from a cross-sectional pilot study. *Heliyon*. 2021; 7. <https://doi.org/10.1016/j.heliyon.2021.e06401> PMID: 33748473
34. Islam MS, Rahman ME, Moonajilin MS, van Os J. Prevalence of depression, anxiety and associated factors among school going adolescents in Bangladesh: Findings from a cross-sectional study. *PLOS ONE*. 2021; 16. <https://doi.org/10.1371/journal.pone.0247898> PMID: 33793610
35. Khan MA. Prevalence and Correlates of Depressive Symptoms among Bangladeshi Young Adults due to COVID-19 Outbreak. *Journal of Clinical and Experimental Investigations*. 2021; 12. <https://doi.org/10.29333/jcei/9766>
36. Dhira TA, Rahman MA, Sarker AR, Mehareen J. Validity and reliability of the Generalized Anxiety Disorder-7 (GAD-7) among university students of Bangladesh. *PLOS ONE*. 2021; 16: e0261590. <https://doi.org/10.1371/journal.pone.0261590> PMID: 34914811
37. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association; 2013. <https://doi.org/10.1176/appi.books.9780890425596>
38. Villarreal-Zegarra D, Copez-Lonzoy A, Bernabé-Ortiz A, Melendez-Torres GJ, Bazo-Alvarez JC. Valid group comparisons can be made with the Patient Health Questionnaire (PHQ-9): A measurement invariance study across groups by demographic characteristics. *PLOS ONE*. 2019; 14. <https://doi.org/10.1371/journal.pone.0221717> PMID: 31498796
39. Kocalevent R-D, Hinz A, Brähler E. Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry*. 2013; 35. <https://doi.org/10.1016/j.genhosppsych.2013.04.006> PMID: 23664569
40. Doi S, Ito M, Takebayashi Y, Muramatsu K, Horikoshi M. Factorial validity and invariance of the Patient Health Questionnaire (PHQ)-9 among clinical and non-clinical populations. *PLOS ONE*. 2018; 13. <https://doi.org/10.1371/journal.pone.0199235> PMID: 30024876
41. Boothroyd L, Dagnan D, Muncer S. PHQ-9: One factor or two? *Psychiatry Research*. 2019; 271. <https://doi.org/10.1016/j.psychres.2018.12.048> PMID: 30553100
42. Elhai JD, Contractor AA, Tamburrino M, Fine TH, Prescott MR, Shirley E, et al. The factor structure of major depression symptoms: A test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Research*. 2012; 199. <https://doi.org/10.1016/j.psychres.2012.05.018> PMID: 22698261

43. Rayhan RU, Zheng Y, Uddin E, Timbol C, Adewuyi O, Baraniuk JN. Administer and Collect Medical Questionnaires with Google Documents: A Simple, Safe, and Free System. *Applied Medical Informatics Original Research*. 2013. www.docs.google.com.
44. LightCastle Analytics Wing. Tertiary Education in Bangladesh: A Sector in Need of Reform. DATABDCO. 2019.
45. University Grants Commission of Bangladesh. List of Public Universities. 2020 [cited 2 Feb 2021]. <http://www.ugc-universities.gov.bd/>.
46. University Grants Commission of Bangladesh. List Of Private Universities. 2020 [cited 2 Feb 2021]. <http://www.ugc-universities.gov.bd/private-universities>.
47. Giannopoulou I, Efstathiou V, Triantafyllou G, Korkoliakou P, Douzenis A. Adding stress to the stressed: Senior high school students' mental health amidst the COVID-19 nationwide lockdown in Greece. *Psychiatry Research*. 2021; 295. <https://doi.org/10.1016/j.psychres.2020.113560> PMID: 33187723
48. Pinto-Meza A, Serrano-Blanco A, Peñarrubia MT, Blanco E, Haro JM. Assessing depression in primary care with the PHQ-9: Can it be carried out over the telephone? *Journal of General Internal Medicine*. 2005; 20. <https://doi.org/10.1111/j.1525-1497.2005.0144.x> PMID: 16050884
49. Bhana A, Rathod SD, Selohilwe O, Kathree T, Petersen I. The validity of the Patient Health Questionnaire for screening depression in chronic care patients in primary health care in South Africa. *BMC Psychiatry*. 2015; 15. <https://doi.org/10.1186/s12888-015-0503-0> PMID: 26001915
50. Han C, Jo SA, Kwak J-H, Pae C-U, Steffens D, Jo I, et al. Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Ansan Geriatric study. *Comprehensive Psychiatry*. 2008; 49. <https://doi.org/10.1016/j.comppsy.2007.08.006> PMID: 18243897
51. Phelan E, Williams B, Meeker K, Bonn K, Frederick J, LoGerfo J, et al. A study of the diagnostic accuracy of the PHQ-9 in primary care elderly. *BMC Family Practice*. 2010; 11. <https://doi.org/10.1186/1471-2296-11-63> PMID: 20807445
52. Swinson RP. The GAD-7 scale was accurate for diagnosing generalised anxiety disorder. *Evidence-Based Medicine*. 2006; 11. <https://doi.org/10.1136/ebm.11.6.184> PMID: 17213178
53. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine*. 2007; 146. <https://doi.org/10.7326/0003-4819-146-5-200703060-00004>
54. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder. *Archives of Internal Medicine*. 2006; 166. <https://doi.org/10.1001/archinte.166.10.1092> PMID: 16717171
55. Kroenke K, Wu J, Yu Z, Bair MJ, Kean J, Stump T, et al. Patient Health Questionnaire Anxiety and Depression Scale. *Psychosomatic Medicine*. 2016; 78. <https://doi.org/10.1097/PSY.0000000000000322> PMID: 27187854
56. Ramsey JB. Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society Series B (Methodological)*. 1969; 31: 350–371. Available: <http://www.jstor.org/stable/2984219>.
57. Barbaranelli C, Lee CS, Vellone E, Riegel B. The Problem With Cronbach's Alpha. *Nursing Research*. 2015; 64: 140–145. <https://doi.org/10.1097/NNR.0000000000000079> PMID: 25738626
58. Bonett DG. Sample Size Requirements for Testing and Estimating Coefficient Alpha. *Journal of Educational and Behavioral Statistics*. 2002; 27: 335–340. <https://doi.org/10.3102/10769986027004335>
59. Graham JM. Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability. *Educational and Psychological Measurement*. 2006; 66: 930–944. <https://doi.org/10.1177/0013164406288165>
60. Raykov T. Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement*. 1997; 21: 173–184. <https://doi.org/10.1177/01466216970212006>
61. Hayes AF, Coutts JJ. Use Omega Rather than Cronbach's Alpha for Estimating Reliability. *But ... Communication Methods and Measures*. 2020; 14: 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
62. Kaiser HF. A second generation little jiffy. *Psychometrika*. 1970; 35: 401–415. <https://doi.org/10.1007/BF02291817>
63. Krause JS, Saunders LL, Bombardier C, Kalpakjian C. Confirmatory Factor Analysis of the Patient Health Questionnaire-9: A Study of the Participants From the Spinal Cord Injury Model Systems. *PM&R*. 2011; 3: 533–540. <https://doi.org/10.1016/j.pmrj.2011.03.003> PMID: 21665166
64. Xiong N, Fritzsche K, Wei J, Hong X, Leonhart R, Zhao X, et al. Validation of patient health questionnaire (PHQ) for major depression in Chinese outpatients with multiple somatic symptoms: A multicenter cross-sectional study. *Journal of Affective Disorders*. 2015; 174: 636–643. <https://doi.org/10.1016/j.jad.2014.12.042> PMID: 25576931

65. Krause JS, Bombardier C, Carter RE. Assessment of depressive symptoms during inpatient rehabilitation for spinal cord injury: Is there an underlying somatic factor when using the PHQ? *Rehabilitation Psychology*. 2008; 53. <https://doi.org/10.1037/a0013354>
66. Richardson EJ, Richards JS. Factor structure of the PHQ-9 screen for depression across time since injury among persons with spinal cord injury. *Rehabilitation Psychology*. 2008; 53. <https://doi.org/10.1037/0090-5550.53.2.243>
67. Steiger JH. Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*. 1990; 25. https://doi.org/10.1207/s15327906mbr2502_4 PMID: 26794479
68. Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*. 1992; 21. <https://doi.org/10.1177/0049124192021002005>
69. Fan X, Sivo SA. Sensitivity of Fit Indices to Model Misspecification and Model Types. *Multivariate Behavioral Research*. 2007; 42. <https://doi.org/10.1080/00273170701382864>
70. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*. 1980; 88. <https://doi.org/10.1037/0033-2909.88.3.588>
71. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973; 38. <https://doi.org/10.1007/BF02291170>
72. Stegmann R by G. Review of *A Beginner's Guide to Structural Equation Modeling (4th ed.)*, by Randall E. Schumacker & Richard G. Lomax. *Structural Equation Modeling: A Multidisciplinary Journal*. 2017; 24. <https://doi.org/10.1080/10705511.2017.1280798>
73. Pituch KA, Stevens J (James P. Applied multivariate statistics for the social sciences: analyses with SAS and IBM's SPSS. Routledge; 2016.
74. Gorsuch RL. *Factor Analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1983.
75. Patel JS, Oh Y, Rand KL, Wu W, Cyders MA, Kroenke K, et al. Measurement invariance of the patient health questionnaire-9 (PHQ-9) depression screener in U.S. adults across sex, race/ethnicity, and education level: NHANES 2005–2016. *Depression and Anxiety*. 2019; 36: 813–823. <https://doi.org/10.1002/da.22940> PMID: 31356710
76. Chen Feinian, Curran PJ, Bollen KA, Kirby J, Paxton P. An Empirical Evaluation of the Use of Fixed Cutoff Points in RMSEA Test Statistic in Structural Equation Models. *Sociological Methods & Research*. 2008; 36: 462–494. <https://doi.org/10.1177/0049124108314720> PMID: 19756246
77. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*. 1986;54. <https://doi.org/10.1111/j.1467-6494.1986.tb00391.x>
78. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. *Psychological Assessment*. 1995; 7. <https://doi.org/10.1037/1040-3590.7.3.309>
79. Thorndike RM. Book Review: *Psychometric Theory (3rd ed.)* by Jum Nunnally and Ira Bernstein New York: McGraw-Hill, 1994, xxiv + 752 pp. *Applied Psychological Measurement*. 1995; 19. <https://doi.org/10.1177/014662169501900308>
80. Brown TA. *Confirmatory factor analysis for applied research*. Guilford publications; 2015.
81. Pituch KA, Stevens JP. *Applied Multivariate Statistics for the Social Sciences*. Routledge; 2015. <https://doi.org/10.4324/9781315814919>
82. Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*. 2002; 9: 233–255. https://doi.org/10.1207/S15328007SEM0902_5
83. Aslan J, Cova F, Saldivia S, Bustos C, Inostroza C, Rincón P, et al. Psychometric Properties of the Patient Health Questionnaire-9 in Elderly Chilean Primary Care Users. *Frontiers in Psychiatry*. 2020; 11. <https://doi.org/10.3389/fpsy.2020.555011> PMID: 33312135
84. Krause JS, Reed KS, McArdle JJ. Factor Structure and Predictive Validity of Somatic and Nonsomatic Symptoms From the Patient Health Questionnaire-9: A Longitudinal Study After Spinal Cord Injury. *Archives of Physical Medicine and Rehabilitation*. 2010; 91. <https://doi.org/10.1016/j.apmr.2010.04.015> PMID: 20684902
85. Maroufizadeh S, Omani-Samani R, Almasi-Hashiani A, Amini P, Sepidarkish M. The reliability and validity of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with infertility. *Reproductive Health*. 2019; 16. <https://doi.org/10.1186/s12978-019-0802-x> PMID: 31500644
86. Anum A, Adjorlolo S, Kugbey N. Depressive symptomatology in adolescents in Ghana: Examination of psychometric properties of the Patient Health Questionnaire-9. *Journal of Affective Disorders*. 2019; 256. <https://doi.org/10.1016/j.jad.2019.06.007> PMID: 31181377
87. Yu X, Tam WWS, Wong PTK, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. *Comprehensive Psychiatry*. 2012; 53. <https://doi.org/10.1016/j.comppsy.2010.11.002> PMID: 21193179

88. Garabiles MR, Lao CK, Yip P, Chan EWW, Mordeno I, Hall BJ. Psychometric Validation of PHQ-9 and GAD-7 in Filipino Migrant Domestic Workers in Macao (SAR), China. *Journal of Personality Assessment*. 2020;102.
89. Hettema JM. The nosologic relationship between generalized anxiety disorder and major depression. *Depression and Anxiety*. 2008; 25. <https://doi.org/10.1002/da.20491> PMID: 18412057
90. Schoevers RA. Depression and Generalized Anxiety Disorder: Co-Occurrence and Longitudinal Patterns in Elderly Patients. *American Journal of Geriatric Psychiatry*. 2005; 13. <https://doi.org/10.1176/appi.ajgp.13.1.31> PMID: 15653938
91. Shamsuddin K, Fadzil F, Ismail WSW, Shah SA, Omar K, Muhammad NA, et al. Correlates of depression, anxiety and stress among Malaysian university students. *Asian Journal of Psychiatry*. 2013; 6. <https://doi.org/10.1016/j.ajp.2013.01.014> PMID: 23810140
92. Lee B, Kim YE. The psychometric properties of the Generalized Anxiety Disorder scale (GAD-7) among Korean university students. *Psychiatry and Clinical Psychopharmacology*. 2019; 29. PMID: 31870089
93. Granillo MT. Structure and Function of the Patient Health Questionnaire-9 Among Latina and Non-Latina White Female College Students. *Journal of the Society for Social Work and Research*. 2012; 3. <https://doi.org/10.5243/jsswr.2012.6>
94. Al-Busaidi Z, Bhargava K, Al-Ismaily A, Al-Lawati H, Al-Kindi R, Al-Shafae M, et al. Prevalence of Depressive Symptoms among University Students in Oman. *Oman Medical Journal*. 2011. <https://doi.org/10.5001/omj.2011.58> PMID: 22043426
95. Flesch BD, Houvéssou GM, Munhoz TN, Fassa AG. Major depressive episode among university students in Southern Brazil. *Revista de Saúde Pública*. 2020; 54. <https://doi.org/10.11606/s1518-8787.2020054001540> PMID: 32022140
96. Ihab Hafez Elsayy W, Aly Reda Sherif A, Salah El Din Attia M, Ahmed El- Nimr N. Depression among medical students in Alexandria, Egypt. *African Health Sciences*. 2020; 20. <https://doi.org/10.4314/ahs.v20i3.47> PMID: 33402990
97. Olum R, Nakwagala FN, Odokonyero R. Prevalence and Factors Associated with Depression among Medical Students at Makerere University, Uganda. *Advances in Medical Education and Practice*. 2020; Volume 11. <https://doi.org/10.2147/AMEP.S278841> PMID: 33209071
98. Sylla M, Vogel AC, Bah AK, Tassiou NR, Barry SD, Djibo BA, et al. Prevalence, severity, and associations of depression in people with epilepsy in Guinea: A single-center study. *Epilepsy & Behavior*. 2020; 113. <https://doi.org/10.1016/j.yebeh.2020.107475> PMID: 33189054
99. Chilcot J, Rayner L, Lee W, Price A, Goodwin L, Monroe B, et al. The factor structure of the PHQ-9 in palliative care. *Journal of Psychosomatic Research*. 2013; 75. <https://doi.org/10.1016/j.jpsychores.2012.12.012> PMID: 23751240
100. Petersen JJ, Paulitsch MA, Hartig J, Mergenthal K, Gerlach FM, Gensichen J. Factor structure and measurement invariance of the Patient Health Questionnaire-9 for female and male primary care patients with major depression in Germany. *Journal of Affective Disorders*. 2015; 170. <https://doi.org/10.1016/j.jad.2014.08.053> PMID: 25240840
101. Keum BT, Miller MJ, Inkelas KK. Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse U.S. college students. *Psychological Assessment*. 2018; 30: 1096–1106. <https://doi.org/10.1037/pas0000550> PMID: 29565614
102. Galenkamp H, Stronks K, Snijder MB, Derks EM. Measurement invariance testing of the PHQ-9 in a multi-ethnic population in Europe: the HELIUS study. *BMC Psychiatry*. 2017; 17: 349. <https://doi.org/10.1186/s12888-017-1506-9> PMID: 29065874
103. Hegarty K. Domestic violence: the hidden epidemic associated with mental illness. *British Journal of Psychiatry*. 2011; 198: 169–170. <https://doi.org/10.1192/bjp.bp.110.083758> PMID: 21357872
104. Burdzovic Andreas J, Brunborg GS. Depressive Symptomatology among Norwegian Adolescent Boys and Girls: The Patient Health Questionnaire-9 (PHQ-9) Psychometric Properties and Correlates. *Frontiers in Psychology*. 2017; 8. <https://doi.org/10.3389/fpsyg.2017.00887> PMID: 28642720
105. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*. 2017; 143. <https://doi.org/10.1037/bul0000102> PMID: 28447828
106. Nguyen DT, Wright EP, Dedding C, Pham TT, Bunders J. Low Self-Esteem and Its Association With Anxiety, Depression, and Suicidal Ideation in Vietnamese Secondary School Students: A Cross-Sectional Study. *Frontiers in Psychiatry*. 2019; 10. <https://doi.org/10.3389/fpsyg.2019.00698> PMID: 31611825
107. Howard LM, Trevillion K, Agnew-Davies R. Domestic violence and mental health. *International Review of Psychiatry*. 2010; 22. <https://doi.org/10.3109/09540261.2010.512283> PMID: 21047164

108. Matud MP. The Psychological Impact of Domestic Violence on Spanish Women1. *Journal of Applied Social Psychology*. 2005; 35. <https://doi.org/10.1111/j.1559-1816.2005.tb02104.x>
109. Brown S, Harris MN, Srivastava P, Taylor K. Mental Health and Reporting Bias: Analysis of the GHQ-12. 2018. www.iza.org.
110. Heiervang E, Goodman R. Advantages and limitations of web-based surveys: evidence from a child mental health survey. *Social Psychiatry and Psychiatric Epidemiology*. 2011; 46. <https://doi.org/10.1007/s00127-009-0171-9> PMID: 19921078