Psihologijske teme, 32 (2023), 3, 451-470
Izvorni znanstveni rad UDK: 159.97.072-057.875(85)
https://doi.org/10.31820/pt.32.3.3 616.895.4-057.875(85)

Psychometric Properties of the Patient Health Questionnaire-9 in Peruvian University Students

Susana K. Lingán-Huamán[1], Sergio Dominguez-Lara[2], María Soledad Alarcón-Almeyda[1], Giovanna Barroca de Moura[3, 4], and Tamyres Tomaz Paiva[5]

[1] Universidad San Ignacio de Loyola, Lima, Peru
[2] Universidad Privada Norbert Wiener, Lima, Peru
[3] Universidade Federal da Paraíba, Paraíba, Brazil
[4] Universidade de Coimbra, Coimbra, Portugal
[5] Faculdade de Enfermagem Nova Esperança (FACENE), Paraíba, Brazil

Abstract

Changes that occurred in people's lives due to the COVID-19 pandemic negatively affected the mental health of the universities, where an increase in symptoms of depression was reported since the beginning of the pandemic, so its evaluation is essential in this context. Therefore, the objective of this study was to analyze the psychometric properties of the Patient Health Questionnaire-9 (PHQ-9) in Peruvian university students. A total of 409 university students between 16 and 45 years of age participated (female = 72.9%; M_{age} = 22.15; SD_{age} = 3.75). The PHQ-9 and the Emotional Exhaustion Scale (EES) were used. A series of confirmatory factor analyses were conducted to evaluate structures suggested by previous studies, namely a one factor model, a two correlated factor model and a bi-factor model. We found evidence of a one-dimensional structure, both from factor analysis and item response theory, and a positive association with academic emotional exhaustion. Moreover, construct reliability and score reliability reached high magnitudes. Therefore, the PHQ-9 is an instrument with adequate psychometric properties to be used as a measure of depression in university students in the context of the COVID-19 pandemic.

Keywords: university students, depression, validity, reliability, PHQ-9

Susana K. Lingán-Huamán https://orcid.org/0000-0003-4587-7853
Sergio Dominguez-Lara https://orcid.org/0000-0002-2083-4278
María Soledad Alarcón-Almeyda https://orcid.org/0000-0002-8835-123X
Giovanna Barroca de Moura https://orcid.org/0000-0001-7970-4323
Tamyres Tomaz Paiva https://orcid.org/0000-0001-9415-0963

✉ Sergio Dominguez-Lara, South American Center for Education and Research in Public Health (Universidad Privada Norbert Wiener), Av. Arequipa 440, Lima 01, Peru. E-mail: *sdominguezmpcs@gmail.com*

Introduction

Depression is a public health issue affecting more than 300 million people worldwide. On the American continent, the prevalence of depressive symptomatology is 15% (World Health Organization [WHO], 2017), while in Peru, it is between 14% in urban areas and 12.5% in rural areas (Villarreal-Zegarra et al., 2020). However, according to Mental Health Department of the Ministry of Health of Peru (MINSA, 2022), based on more than 300,000 patients diagnosed with depression during 2021, there is a 12% increase of the incidence of depressive symptoms compared to the pre-pandemic years. Therefore, stressful events, such as those that characterized the COVID-19 pandemic, constitute a risk factor that should be considered since the perception of risk affects people's mental health during a public health crisis (Ding et al., 2020). Specifically, in the case of depression, a meta-analytical study that considered research in the general population during the COVID-19 pandemic found a global prevalence of 33.7% (95% CI [27.5 – 40.6]), and the most affected age group was between 21 and 40 years of age (Salari et al., 2020).

In this scenario, the concern for the mental health of university students is relevant given that it is a group prone to experience high levels of stress, anxiety, and depression since some reports indicate that more than one-third of the university population suffered from certain mental health issues before the pandemic (Zivin et al., 2009). Furthermore, during the pandemic, the situation was even more concerning (Aristovnik et al., 2020). University students present more risk of experiencing mental disorders than other young people of the same age in the general population (Farrer et al., 2016). Similarly, Naser et al. (2020) found a higher prevalence of depression (38%) and anxiety (21%) at Jordan University, in contrast to healthcare professionals and the general population. In Peru, Sánchez-Carlessi et al. (2021) found a 45% prevalence of depression in university students during the COVID-19 pandemic. Thus, it is evident that depressive symptoms are persistent among university students.

The situation described above can be explained by the biopsychosocial model of depression, according to which university students may experience greater uncertainty as they are in a stage of transition to adult life, with greater concern about their personal, academic, and professional future, which could lead to an increase in depressive symptoms (Berry, 2004).

Depressive symptoms can generate negative, immediate, and delayed consequences in the academic, family, and personal environment of university students, including a decrease in academic performance, absence from school, interpersonal and physical health problems, and decline in the quality of life (Farrer et al., 2016; Gao et al., 2020; Stallman, 2010), which could also trigger symptoms of academic burnout (AB; Azzi et al., 2021), as it causes academic emotional exhaustion (AEE). There is a moderate association between depression and AB (r =

.52; Koutsimani et al., 2019), since academic exhaustion is related to negative affectivity, lack of concentration, and lack of interest in tasks, which is related to depressive disorders. In addition, depression has been linked to a higher risk of suicide among university students (Bantjes et al., 2016; Furr et al., 2001).

Despite the high prevalence of depression and its harmful consequences in people's lives, depression is usually infra-diagnosed; therefore, there is no access to timely intervention (Wu & Fang, 2014). In Peru, only 1 in 27 patients with severe depressive disorder manage to receive treatment (Villarreal-Zegarra et al., 2020). Therefore, it is relevant that student guidance departments perform early detection of depression symptoms in the university population to provide appropriate psychological intervention, especially in countries with limited resources for mental health care (Furr et al., 2001).

The detection of mental health problems still lacks biological markers as it is mainly based on the application of questionnaires and scales that evaluate somatic and psychological symptoms from the subjective perception of the subjects (Prata et al., 2014); therefore, valid and reliable instruments for the identification of depressive symptoms and the need for psychological care is required for university students. Although there are currently several questionnaires for the measurement of depressive symptoms, most of them include more than 10 items and are not directly related to the established diagnostic criteria (e.g. Beck Depression Inventory [BDI-II], Center of Epidemiological Study of Depression Scale [CES-D], Depression Anxiety Stress Scale-21 [DASS-21]).

Thus, the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) is a brief questionnaire that evaluates the levels of the severity of depression; it is widely used in different contexts (Lamela et al., 2020) and has been translated into more than 70 languages (Reich et al., 2018). Its nine items respond to the nine criteria of the diagnosis of depressive disorders, which is based on the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association [APA], 2013). Despite the recent version of DSM (DSM-V), the use of PHQ-9 is still in force since the nine essential diagnostic criteria and the two-week period remain unchanged (Keum et al., 2018).

In primary care, although with some variations at the cut-off points, the sensitivity of PHQ-9 is between .71 and .84, while the specificity is between .90 and .97 (Wittkampf et al., 2007). In a meta-analytical study, a high specificity (> .90) of PHQ-9 was found, although the results of the evaluation of its sensitivity are heterogeneous (between .39 and .73; Manea et al., 2015). In addition to sufficient specificity and sensitivity, the extended use of PHQ-9 can be explained by presenting other advantages, such as free access, brevity, ease of application, qualification, and interpretation (Manea et al., 2012).

Its psychometric properties were evaluated in different populations, including general population samples (Villarreal-Zegarra et al., 2019), primary care samples (González-Blanch et al., 2018), psychiatric patients (Sun et al., 2020), patients with

certain specific medical conditions (Seo & Park, 2015), multicultural populations (Lamela et al., 2020), and in the university context, which demonstrated its usefulness for the detection of depressive symptoms in Bangladesh university students (Rahman et al., 2022), Korea (Kim & Lee, 2019), China (Du et al., 2017), United States (Keum et al., 2018), Japan (Umegaki & Todo, 2017), Lithuania (Pranckeviciene et al., 2022) and Nigeria (Adewuya et al., 2006).

Despite the widespread use of PHQ-9, due to the aforementioned advantages, although previous studies that explored their psychometric properties found adequate levels of reliability ($\alpha > .80$; Du et al., 2017; Keum et al., 2018; Kim & Lee, 2019; Pranckeviciene et al., 2022; Rahman et al., 2022), it was reported that the instrument presents mixed results concerning its dimensionality (Doi et al., 2018; Lamela et al., 2020). Dimensionality refers to the number of latent variables that can be estimated through the data, and it is possible to support the sum scores used to interpret the results of the application of a test with this. Consequently, using a total score from a sum is justified if there is evidence of the one-dimensionality of said latent variable (Zwitser & Maris, 2016).

Hence, the results of the analysis of the factorial structure of PHQ-9 have not been consistent since some studies support the original premise of the instrument (Kroenke et al., 2001) and report a one-dimensional structure (Adewuya et al., 2006; Keum et al., 2018; Kim & Lee, 2019; Pranckeviciene et al., 2022; Rahman et al., 2022; Reich et al., 2018; Sun et al., 2020; Umegaki & All, 2017); while other research found that a two-dimensional structure represented, mainly by psychological symptoms (e.g. loss of interest and joy, change in sleep patterns, etc.), presents better statistical adjustment (Doi et al., 2018; Elhai et al., 2012; Petersen et al., 2015; Zhou et al., 2020).

In a recent systematic review, it was found that 19 studies supported the unifactorial model, while 12 found evidence in favor of the two-factor model, although the data comes from heterogeneous samples both in configuration (general, hospitable, and clinical population) and origin (different countries) (Lamela et al., 2020). However, the studies of González-Blanch et al. (2018), Boothroyd et al. (2019), and Stochl et al. (2022) found that when using different psychometric analysis techniques, one- and two-factor models had adequate fit. In addition, given that the factors were highly correlated, it was suggested to adopt the model of a factor and use a single total sum score as an appropriate summary statistic for the PHQ-9.

In Peru, there are psychometric studies focused on validity evidence based on internal structure with the general population (Villarreal-Zegarra et al., 2019) and in university medical students (Huarcaya-Victoria et al., 2020). Regarding the internal structure of PHQ-9, the one-dimensional model presents the strongest evidence because the interfactorial correlations are close to unity in models with more than one factor (Villarreal-Zegarra et al., 2019). Although another study highlights the relevance of a bifactor model (Huarcaya-Victoria et al., 2020), a limitation was found when modeling the interfactorial correlation into the bifactor model, which would

question its conclusions. Therefore, the study of the internal structure of PHQ-9 in Peruvian university students is still a pending task, both from the classical theory of tests and from the item response theory (IRT) point of view, since there are no Peruvian studies with this methodology.

Therefore, the objective of this study was to analyze the psychometric properties of PHQ-9 in Peruvian university students. The research hypothesis states that PHQ-9 will be unidimensional (H1) considering the two-factor models and the bifactor model will be competing models. Furthermore, by being directly and significantly associated with the AEE (H2), PHQ-9 will obtain adequate magnitudes of construct reliability and scores reliability (H3).

This study is important because a frequent difficulty that researchers and professionals encounter in their work context is related to the lack of availability of adequate instruments that can offer evidence of validity and reliability and are, at the same time, brief. Therefore, this study seeks to satisfy part of that demand and highlight the importance for Peru of carrying out studies aimed at the construction, adaptation, or validation of brief instruments for measuring the different disorders and problems of health, especially depressive disorders, whose prevalence is increasing, especially during the pandemic.

Method

Participants

The sample was obtained using non-probabilistic snowball-type sampling. A total of 409 university students aged 16 – 45 (M_{age} = 22.15; SD_{age} = 3.75) and residents of Metropolitan Lima participated. Of the total sample, 298 (72.9%) were women, and 111 (27.1%) were men; 22.98% of students were in the first two years of their course of studies, while 31.54% and 45.48% were in the middle and final cycles of their careers, respectively. In addition, 33.7% were in the psychology program, 18.75% in economic sciences, and 47.8% in other professional degree programs.

Measures

Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001) is a nine-item self-report that assesses emotional, cognitive, and somatic symptoms of depression. The items have a Likert-type response format with four options (from *Not at all* [0] to *Almost every day* [3]). The version adapted by Villarreal-Zegarra et al. (2019) to the Peruvian general population was used, and this version showed that using standard CFA, a one-dimensional model, presented the best fit (CFI = .936; RMSEA = .089; SRMR = .039) and PHQ-9 reliability was optimal (α = .87; ω = .87).

Emotional Exhaustion Scale (EES; Fontana, 2011) one-dimensionally measures academic emotional exhaustion (AEE) through 10 items scaled in Likert format with five response options ranging from *rarely* (1) to *always* (5). Dominguez-Lara et al. (2021) found adequate psychometric properties with Peruvian university students using CFA; the results showed that the one-dimensional model shows favorable evidence (CFI = .945; WRMR = 1.386; RMSEA = .124) regarding their factor loadings ($\lambda_{average} > .70$) and reliability ($\omega = .91$). With the data from this report, an acceptable reliability was found ($\alpha = .88$).

Procedure

Questionnaires and informed consent forms were submitted using Google forms. The data was collected by sharing the link on social networks, such as Facebook and WhatsApp, between the months of April and May 2020. The link shared with the students presented an informed consent in which the objectives of the study were reported, as well as the characteristics expected of the person answering the questionnaires, indicating the voluntary nature of the study, its risks, and benefits, and that the information provided will be anonymous. The research was approved by the corresponding Faculty Review Board of the Coordination of the Psychology Course of Studies of Universidad San Ignacio de Loyola based on the ethical guidelines within the institutional regulations, the Helsinki Convention for research with human beings, and the College of Psychologists of Peru Code of Ethics (2017). The raw data of this research will be made available by email request.

Data Analysis

As for validity evidence based on internal structure, three measurement models were analyzed: a one-dimensional model, another model of oblique two-factor model, somatic dimension and cognitive–affective dimension; and a bifactor model, which considers the presence of a general factor (GF) that explains the variability of the items even in the presence of two specific factors (SFs).

A confirmatory factorial approach with the weighted least squares mean and variance adjusted extraction method was used. The model was evaluated considering the Tucker–Lewis index (> .90; Hair et al., 2006), comparative fit index (CFI > .90; Hair et al., 2006), upper limit of the CI of the root mean square error of approximation (RMSEA < .10; West et al., 2012), and weighted root mean square residual (WRMR > 1.00; DiStefano et al., 2018). In addition, factor loadings greater than .50 were considered acceptable (Dominguez-Lara, 2018), and interfactorial correlations greater than .90 would be indicators of multicollinearity (Brown, 2015). For the bifactor model, the relevance of interpreting two factors in the presence of a GF was evaluated, for which the magnitude of the overall hierarchical omega (ω_h) and for each dimension (ω_{hs}) was considered. So, if ω_{hs} is greater than .30 and the explained common variance (ECV) is less than .50, the SFs can be interpreted (Gignac &

Kretzschmar, 2017). Additionally, in the one-dimensional model, correlations between residuals that could be misspecifications were progressively modeled (Saris et al., 2009) based on their modification indices (MIs). The calculation was performed with a specific module (Dominguez-Lara & Merino-Soto, 2018). Factor analyses were performed with the Mplus 7 software (Muthén & Muthén, 1998-2015).

Based on the one-dimensional model, the internal structure was analyzed based on the item response theory (IRT) through the Samejima gradual response model, which models item discrimination by making it have different difficulty values (Boateng et al., 2018). We estimated discrimination parameters (a) representing the item's ability to differentiate individuals with different latent trait levels, as well as difficulty parameters (b), which indicates how difficult the item is for the individual to match, but since PHQ-9 refers to the mental health context, location was used instead of difficulty (Yang & Kao, 2014).

Regarding discrimination, values between 0.10 and 0.34 are considered low, between 0.65 and 1.34 are considered moderate, between 1.35 and 1.69 are considered high, and greater than 1.70 are regarded as very high (Baker, 2001). With regard to the location, it can be negative or positive and suggests that participants may present milder (theta close to −3) or more severe (theta close to +3) symptoms with respect to depression.

In addition, the fit indices of each item to the proposed one-dimensional model were analyzed. Therefore, the higher the infit, the more likely it is that the item contains residuals and, thus, does not represent the latent trait, so values between 0.5 and 1.5 are considered acceptable (Linacre & Wrigh, 1994). All procedures related to the analysis from IRT were performed with the *mirt* package (Chalmers, 2012).

In relation to the reliability of the scores, the alpha coefficient was considered (> .70), and with respect to the reliability of the construct, the omega coefficient was used (> .80). Finally, a correction was implemented in both coefficients due to the presence of correlated residuals. In all cases, for calculated and corrected coefficients, the CI of the coefficient was calculated using the bootstrap method.

Finally, regarding the evidence of convergence validity due to its relationship with other variables, the association between the measure of depression (PHQ-9) and that of AEE was explored with the Pearson correlation coefficient, which was interpreted under a size approach effect, where a value between .20 and .50 is considered low, between .50 and .80, moderate, and greater than .80, high (Ferguson, 2009).

Results

Evidence of Validity Based on Internal Structure

In general, the three models obtained favorable fit indices (Table 1), as well as acceptable factor loadings. However, in the case of the two-factor model, a high interfactorial correlation (.89) is observed, which suggests an overlapping of factors; additionally, the complementary indices to the bifactor model (third model) indicate the relevance of interpreting a single factor (ECV = .86; PUC = .50; ω_h = .89) given that the SFs do not explain enough variance ($\omega HS_{cognitive\text{-}emotional}$ = .039; $\omega HS_{somatic}$ = .16). For this reason, the subsequent analyses (misspecifications, IRT, and reliability) were performed based on the one-dimensional model.

Table 1

Factor Parameters of the PHQ-9 Measurement Models

	One-dimensional model	Oblique two-factor model		Bifactor model		
		F1	F2	GF	F1	F2
Cognitive–affective symptoms						
1. Little interest or pleasure in doing things	.79	.80		.82	-.17	
2. Feeling down, depressed, or hopeless	.85	.86		.88	-.13	
6. Feeling bad about yourself or feeling like you are a failure or have failed yourself or your family	.82	.83		.81	.22	
7. Difficulty in concentrating on things, such as reading the newspaper or watching television	.68	.69		.66	.28	
8. Moving or speaking so slowly that other people may have noticed, or the opposite: being so fidgety or restless that you have been moving around a lot more than usual	.58	.59		.55	.32	
9. Thoughts that you would be better off dead or that you wanted to hurt yourself in some way	.79	.79		.75	.43	
Somatic symptoms						
3. Trouble in sleeping or staying asleep or sleeping too much	.63		.66	.58		.37
4. Feeling tired or low on energy	.79		.85	.76		.35
5. Poor appetite or overeating	.67		.71	.64		.26
CFI	.97	.98		.99		
TLI	.96	.97		.98		
RMSEA	.09	.08		.07		
90% CI RMSEA	.08 – .11	.07 – .10		.05 – .09		
WRMR	1.06	0.95		0.68		

Note. GF = General Factor; F1 = Cognitive–affective symptoms; F2 = Somatic symptoms.

Initially, a misspecification associated with the correlation between the residuals of the items 6 (*Feeling bad about yourself or feeling like you are a failure or have failed yourself or your family*) and 9 (*Thoughts that you would be better off dead or that you wanted to hurt yourself in some way*) (MI = 24.29, Expected Parameter Change [EPC] = .49) was detected, and when modeling said correlation, the fit improved (CFI = .98; RMSEA = .08, 90% CI [.06 – .10]; WRMR = 0.94) and a low magnitude correlation was observed between the residuals of the mentioned items ($\phi_{6.9}$ = .40). Subsequently, a misspecification associated with the correlation between the residuals of items 1 (*Little interest or pleasure in doing things*) and 2 (*Feeling down, depressed, or hopeless*) (MI = 13.35, EPC = .38) was also found, and after implementing it, the fit improved (CFI = .98; RMSEA = .08, 90% CI [.06 – .09]; WRMR = 0.84) and a low magnitude correlation was also observed ($\phi_{6.9}$ = .307). Finally, the correlation between the residuals of items 2 (*Feeling down, depressed, or hopeless*) and 6 (*Feeling bad about yourself or feeling like you are a failure or have failed yourself or your family*) (MI = 10.53, EPC = .34) was modeled, and after implementing it, the fit improved (CFI = .98; RMSEA = .07, 90% CI [.05 – .09]; WRMR = 0.81) and a low magnitude correlation was also observed ($\phi_{6.9}$ = .28).

Regarding the results based on IRT, the difficulty parameter indicates that the more the individual experiences depressive symptoms, the more likely it is for them to choose the option that indicates more frequency or intensity of symptoms. Items with negative values in b1 indicate how easy it is for this category to be selected by individuals who do not have any depressive symptoms; for example, checking the option *Not at all* (0), in items 1 to 5 (*Feeling down, depressed, or hopeless*). A crucial threshold parameter is b3 as it shows how necessary it is for the individual to mark the maximum category, *to almost every day* (3). That is, only individuals who feel some type of discomfort or some symptom of depression will be likely to choose the most severe response options compared to a patient with a lower level of depression. Figure 1 shows the ability (or depression) of an item to accurately estimate scores in theta. The optimally informative items will have a large location and broad category coverage (as indicated by the location parameters) about theta, and the occupation area of each item is also shown in Table 2. Items 2, 6, and 9 have a longer information curve (Figure 1), with theta ranging between 0.90 and 1.00, and located on the positive side of the latent trait among people with thetas located at levels 2 and 3 (as shown in Figure 2). This indicates that the higher the information, the greater the accuracy of the item in the PHQ-9. Thus, items 3, 5, and 8 present a flatter information curve (Figure 1), with thetas ranging between 0.30 and 0.50 and located in the latent trait at levels 0 and 2 (as shown in figure 2); that is, these items contribute little to the accuracy of the PHQ-9.

Regarding discrimination, the items that most differentiated the participants were also items 2 (a = 2.84) and 6 (a = 2.37), which presented high levels, although in general the other items presented moderate to high discrimination. Finally,

regarding the fit indexes of each item to the one-dimensional model, all items presented favorable values for the construction of the latent trait. Thus, this analysis based on IRT shows that the nine items of PHQ-9 can discriminate individuals in a wide or narrow range of theta and, therefore, adequately measure the construct. Parameter infit means inlier-sensitive or information-weighted fit, ranging from 0.78 to 0.97 (Table 2) and is considered acceptable according to Linacre and Wright (1994). This indicates that people were sensitive to each item in PHQ-9.

Table 2

Discrimination, Difficulty, Proportion of Information, and Infit

	a	b1	b2	b3	Information Ratio	Infit
1. Little interest or pleasure in doing things	2.26	-0.49	1.03	1.79	.97	0.88
2. Feeling down, depressed, or hopeless	2.84	-0.16	1.13	2.06	.97	0.78
3. Trouble in sleeping or staying asleep or sleeping too much	1.38	-0.97	0.64	1.47	.91	0.97
4. Feeling tired or low on energy	2.29	-0.82	0.87	1.56	.98	0.88
5. Poor appetite or overeating	1.51	-0.63	0.80	1.63	.93	0.93
6. Feeling bad about yourself or feeling like you are a failure or have failed yourself or your family	2.37	0.08	1.13	1.82	.96	0.83
7. Difficulty in concentrating on things, such as reading the newspaper or watching television	1.57	0.21	1.37	2.64	.97	0.93
8. Moving or speaking so slowly that other people may have noticed, or the opposite: being so fidgety or restless that you have been moving around a lot more than usual	1.14	0.37	1.93	3.53	.97	0.96
9. Thoughts that you would be better off dead or that you wanted to hurt yourself in some way	2.00	0.82	2.00	2.35	.93	0.92

Figure 1

Information Curve per Item

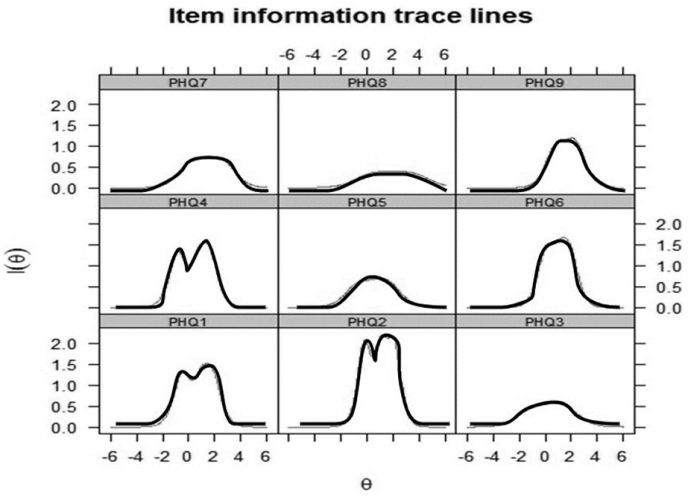
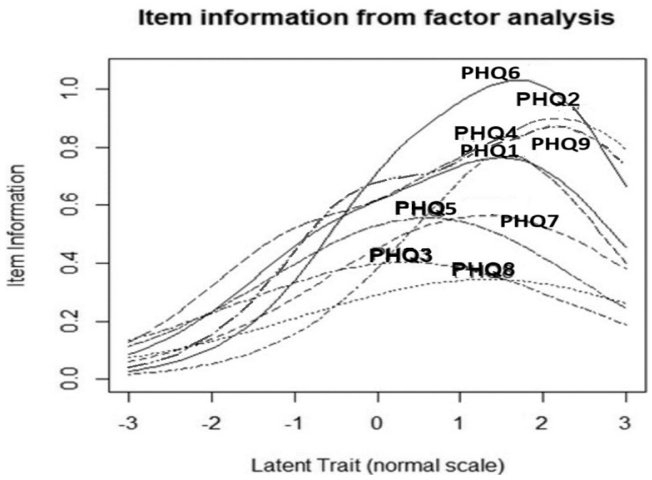


Figure 2

Information Curve Relative to the Factor



Regarding reliability, the magnitudes were acceptable both in the alpha coefficient (α = .91, 95% CI [.89 – .92]) and in the omega coefficient (ω = .92, 95% CI [.90 – .93]), although after correcting for the presence of correlated residuals, the omega coefficient obtained acceptable magnitudes ($\omega_{corrected}$ = .89, 95% [CI .87 –

.91]), but the alpha coefficient, which decreased in the presence of these parameters ($\alpha_{corrected}$ = .70, 95% CI [.68 – .73]), did not.

Finally, the correlation with AEE was moderate (r = .62).

Discussion

The changes caused by the COVID-19 pandemic had a negative impact on the mental health of people, especially university students, who reported an increase in depressive symptoms since the start of the pandemic in different parts of the world (Naser et al., 2020). Therefore, it is important to have valid and reliable evaluation instruments that are able to identify symptoms related to the mental health of young people in a timely manner. To that regard, in the present study, we sought to evaluate the psychometric properties of PHQ-9 in a sample of Peruvian university students by analyzing the internal structure, evidence of concurrent validity, and reliability.

The results of the analysis of the internal structure, carried out through the CFA, show that the items of PHQ-9 are grouped in a one-dimensional structure, which means that all the items constitute adequate indicators to measure depression as the only latent trait, which confirms what was found in previous studies (Adewuya et al., 2006; Keum et al., 2018; Kim & Lee, 2019; Pranckeviciene et al., 2022; Rahman et al., 2022; Reich et al., 2018; Sun et al., 2020; Umegaki & Todo, 2017; Villarreal-Zegarra et al., 2019) and is consistent with the original version, designed as a one-dimensional instrument that considers the nine main criteria on which diagnosing a major depressive disorder is based, according to the DSM-IV (Kroenke et al., 2001). However, the findings are contrary to what was concluded in the study by Huarcaya-Victoria et al. (2020), which pointed out that the bifactor model has advantages over the one-dimensional and two-factor models. Notwithstanding, it is necessary to specify that this study has some limitations and methodological errors when modeling the interfactorial correlation and not considering complementary indices of the bifactor model for their interpretation, which could call their conclusions into question. Additionally, the bifactor model aims to model the influence of a GF and other SFs to conclude on the one-dimensionality of the instrument and to be subsequently evaluated. It is not constituted as a measurement model for practical purposes or professional use. On the contrary, although there are studies in the scientific literature in favor of the two-factor models (Lamela et al., 2020), the results of this study show a high interfactorial correlation (> .85), indicating redundant factors, which was also found in other studies (Boothroyd et al., 2019; González-Blanch et al., 2017; Stochl et al., 2020; Villarreal-Zegarra et al., 2019). In a complementary way, based on the IRT, all the items of PHQ-9 presented an adequate fit to the proposed one-dimensional model and, as well as discrimination between moderate and high, items 2 (*Feeling down, depressed, or hopeless*) and 6 (*Feeling bad about yourself or feeling like you are a failure or have failed yourself or your*

family) stand out for their discriminative power (Yang & Kao, 2014). This information could be relevant to briefly assess depression in the university context because, along with career, disappointments and bad feelings about their own performance could affect students' perspective about their future. So, with respect to its internal structure, the evidence suggests that using PHQ-9 from the perspective of the two factors would not provide any useful information, since they are overlapping factors, and the one-dimensional model has greater empirical support.

It is necessary to mention that an element linked to the analysis of the internal structure is the modeling of correlation between residuals. This practice is frequent in many studies, but its use and objectives are questionable (Dominguez-Lara, 2019), since, if it is used to increase the fit indices and appears to be a good measurement model, its implementation could distort the quality of the model. However, it is acceptable if their presence helps the user to understand the nature of these associations in order to improve the metric quality of the scale (Pérez-Fuentes et al., 2020). In the case of the present study, the correlations between residuals that were found could be explained because the items involved share the type of symptom assessed, i.e., cognitive-emotional, which was indicated as an independent factor by other studies, but which does not explain enough variance to establish itself as such. However, the correlations between residuals of various items may vary depending on the participating sample, so they should be routinely explored after performing a large evaluation.

Regarding the evidence of convergence validity, based on the relationship with other variables, a positive and moderate correlation was found between depression and emotional exhaustion. This is consistent with the scientific literature, both in the pre-pandemic COVID-19 context (Koutsimani et al., 2019) and in the pandemic scenario (Azzi et al., 2021). In fact, the magnitude of the correlation found in this study is similar to that found in a meta-analytical study that analyzes the link between the two constructs (Koutsimani et al., 2019), as well as coincides with another report where elevated levels of depression are associated with high scores in AB in the context of lockdown in university students (Azzi et al., 2021).

On the contrary, regarding reliability, the alpha and omega coefficients reached adequate values. Therefore, it is suggested that PHQ-9 is a reliable instrument, as demonstrated in previous research (Boothroyd et al., 2019; Du et al., 2017; Keum et al., 2018; Kim & Lee, 2019; Pranckeviciene et al., 2022; Rahman et al., 2022; Villarreal-Zegarra et al., 2019). However, regarding the influence of the correlation between residuals on reliability, a greater impact was observed on the alpha coefficient than on the omega coefficient, which is usually not reported despite being relevant and could change the perception of the metric quality of the scale (e.g., Wisting et al., 2021). As mentioned above, although the pattern of correlations between residuals is specific to each sample, it is necessary to correct the reliability coefficients, as the measurement could contain more error than allowed and, thus, inaccurate results would be provided.

The practical implications of the present study relate to evidence that PHQ-9 is an instrument with adequate psychometric properties for use as a measure of depression in college students in the context of COVID-19 pandemic. It will allow for a timely identification of students who require psychological counseling and can be used as a tool to develop programs for the promotion and primary care of mental health in the university community. PHQ-9 can also be used to explore how depression is related to other functional and dysfunctional behaviors in the context of a pandemic and, thereby, produces new knowledge about the emotional and cognitive responses of university students in this new scenario.

Despite the benefits, it is necessary to identify the limitations of the study. First, although the findings obtained are consistent with previous literature, the characteristics of the sample and the sampling procedures used in this study limit the ability to generalize the results obtained for the Peruvian university population. Therefore, it is necessary to conduct research with nationally representative samples to confirm the results. Second, the reliability of PHQ-9 considered only the internal consistency perspective; therefore, the conclusions regarding the measurement's temporary stability are pending.

On the contrary, while the exclusive use of self-reporting measures for data collection within a study could introduce biases caused by factors that affect the recall and self-perception capacity of participants' own subjective states, in this work, it can be justified, because the central objective was to learn the dimensionality of the scale, and the method of evaluation was online during the first year of the pandemic due to the restrictions of mobilization and meetings already discussed at the beginning of the work.

However, it would be useful for future research to consider other validity strategies based on external criteria, such as clinical diagnosis or psychological interview, to analyze the sensitivity and specificity of PHQ-9 in university students. Thus, cross-cultural studies could be useful among students from different countries or between general and clinical Peruvian samples for cut-offs for professional contexts.

References

Adewuya, A. O., Ola, B. A., & Afolabi, O. O. (2006). Validity of the Patient Health Questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. *Journal of Affective Disorders*, *96*(1-2), 89–93. https://doi.org/10.1016/j.jad.2006.05.021

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (DSM-5). American Psychiatric Publishing, Inc.

Aristovnik, A., Keržič, D., Ravšelj, D., Tomaževič, N., & Umek, L. (2020). Impacts of the COVID-19 pandemic on life of higher education students: A global perspective. *Sustainability*, *12*(20), 8438. https://doi.org/10.3390/su12208438

Azzi, D. V., Melo, J., Neto, A. de A. C., Castelo, P. M., Andrade, E. F., & Pereira, L. J. (2022). Quality of life, physical activity and burnout syndrome during online learning period in Brazilian university students during the COVID-19 pandemic: A cluster analysis. *Psychology, Health & Medicine*, *27*(2), 466–480. https://doi.org/10.1080/13548506.2021.1944656

Baker, F. B. (2001). *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation.

Bantjes, J. R., Kagee, A., McGowan, T., & Steel, H. (2016). Symptoms of posttraumatic stress, depression, and anxiety as predictors of suicidal ideation among South African university students. *Journal of American College Health*, *64*(6), 429–437. https://doi.org/10.1080/07448481.2016.1178120

Berry, D. (2004). The relationship between depression and emerging adulthood. *Advances in Nursing Science*, *27*(1), 53–69. https://doi.org/10.1097/00012272-200401000-00007

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00149

Boothroyd, L., Dagnan, D., & Muncer, S. (2019). PHQ-9: One factor or two? *Psychiatry Research*, *271*, 532–534. https://doi.org/10.1016/j.psychres.2018.12.048

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Cehua Alvarez, E. A., Virú-Flores, H., Alburqueque-Melgarejo, J., Roque-Quezada, J. C., Guerra Valencia, J., Gonzales Matta, G. A., Gonzales Menéndez, M. J. M., & Laván Quiroz, D. A. (2022). Validation of the Patient Health Questionnaire-9 (PHQ-9) in human medicine interns at a reference university in Peru during the COVID-19 pandemic. *Revista de La Facultad de Medicina Humana*, *22*(3), 540–546. https://doi.org/10.25176/RFMH.v21i1.3179

Chalmers, R. P. (2012). mirt : A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Colegio de Psicólogos del Perú. (2017). *Código de ética y deontología [Code of ethics and deontology]*. https://www.cpsp.pe/documentos/marco_legal/codigo_de_etica_y_deontologia.pdf

Ding, Y., Xu, J., Huang, S., Li, P., Lu, C., & Xie, S. (2020). Risk perception and depression in public health crises: Evidence from the COVID-19 crisis in China. *International Journal of Environmental Research and Public Health*, *17*(16), 5728. https://doi.org/10.3390/ijerph17165728

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 453–466. https://doi.org/10.1080/10705511.2017.1390394

Doi, S., Ito, M., Takebayashi, Y., Muramatsu, K., & Horikoshi, M. (2018). Factorial validity and invariance of the Patient Health Questionnaire (PHQ)-9 among clinical and non-clinical populations. *Plos One*, *13*(7), e0199235. https://doi.org/10.1371/journal.pone.0199235

Dominguez-Lara, S. (2018). Propuesta de puntos de corte para cargas factoriales: Una perspectiva de fiabilidad de constructo [Proposal for cut-offs for factor loadings: A construct reliability perspective]. *Enfermería Clínica*, *28*(6), 401–402. https://doi.org/10.1016/j.enfcli.2018.06.002

Dominguez-Lara, S. (2019). Correlación entre residuales en análisis factorial confirmatorio: Una breve guía para su uso e interpretación [Correlation between residuals in confirmatory factor analysis: A brief guide to its use and interpretation]. *Interacciones: Revista de Avances En Psicología*, *5*(3), e207. https://doi.org/10.24016/2019.v5n3.207

Dominguez-Lara, S., Alarcón-Parco, D., Fernández-Arata, M., Tamayo-Agudelo, W., Bernal-Vargas, L., & Campos-Uscanga, Y. (2021). Invarianza factorial de una medida de burnout académico entre estudiantes universitarios de Perú, México y Colombia [Factorial invariance of a measure of academic burnout among university students in Peru, Mexico, and Colombia]. *Revista Argentina de Ciencias Del Comportamiento*, *13*(3), 70–86. https://doi.org/10.32348/1852.4206.v13.n3.25946

Dominguez-Lara, S., & Merino-Soto, C. (2018). Evaluación de las malas especificaciones en modelos de ecuaciones estructurales [Evaluation of misspecification in structural equation models]. *Revista Argentina de Ciencias del Comportamiento*, *10*(2), 19–24.

Du, N., Yu, K., Ye, Y., & Chen, S. (2017). Validity study of Patient Health Questionnaire-9 items for Internet screening in depression among Chinese university students. *Asia-Pacific Psychiatry*, *9*(3), e12266. https://doi.org/10.1111/appy.12266

Elhai, J. D., Contractor, A. A., Tamburrino, M., Fine, T. H., Prescott, M. R., Shirley, E., Chan, P. K., Slembarski, R., Liberzon, I., Galea, S., & Calabrese, J. R. (2012). The factor structure of major depression symptoms: A test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Research*, *199*(3), 169–173. https://doi.org/10.1016/j.psychres.2012.05.018

Farrer, L. M., Gulliver, A., Bennett, K., Fassnacht, D. B., & Griffiths, K. M. (2016). Demographic and psychosocial predictors of major depression and generalised anxiety disorder in Australian university students. *BMC Psychiatry*, *16*(1), 241. https://doi.org/10.1186/s12888-016-0961-z

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538. https://doi.org/10.1037/a0015808

Furr, S. R., Westefeld, J. S., McConnell, G. N., & Jenkins, J. M. (2001). Suicide and depression among college students: A decade later. *Professional Psychology: Research and Practice*, *32*(1), 97–100. https://doi.org/10.1037/0735-7028.32.1.97

Gao, L., Xie, Y., Jia, C., & Wang, W. (2020). Prevalence of depression among Chinese university students: A systematic review and meta-analysis. *Scientific Reports*, *10*(1), 15897. https://doi.org/10.1038/s41598-020-72998-1

Gignac, G. E., & Kretzschmar, A. (2017). Evaluating dimensional distinctness with correlated-factor models: Limitations and suggestions. *Intelligence*, *62*, 138–147. https://doi.org/10.1016/j.intell.2017.04.001

González-Blanch, C., Medrano, L. A., Muñoz-Navarro, R., Ruíz-Rodríguez, P., Moriana, J. A., Limonero, J. T., Schmitz, F., & Cano-Vindel, A. (2018). Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *Plos One*, *13*(2), e0193356. https://doi.org/10.1371/journal.pone.0193356

Hair, J. F., Black, W. C., Babin, B., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Prentice-Hall.

Huarcaya-Victoria, J., De-Lama-Morán, R., Quiros, M., Bazán, J., López, K., & Lora, D. (2020). Propiedades psicométricas del Patient Health Questionnaire (PHQ-9) en estudiantes de medicina en Lima, Perú [Psychometric properties of the Patient Health Questionnaire (PHQ-9) in medical students in Lima, Peru]. *Revista de Neuro-Psiquiatria*, *83*(2), 72–78. https://doi.org/10.20453/rnp.v83i2.3749

Keum, B. T., Miller, M. J., & Inkelas, K. K. (2018). Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse U.S. college students. *Psychological Assessment*, *30*(8), 1096–1106. https://doi.org/10.1037/pas0000550

Kim, Y. E., & Lee, B. (2019). The psychometric properties of the Patient Health Questionnaire-9 in a sample of Korean university students. *Psychiatry Investigation*, *16*(12), 904–910. https://doi.org/10.30773/pi.2019.0226

Koutsimani, P., Montgomery, A., & Georganta, K. (2019). The relationship between burnout, depression, and anxiety: A systematic review and meta-analysis. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00284

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lamela, D., Soreira, C., Matos, P., & Morais, A. (2020). 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*, *276*, 220–233. https://doi.org/10.1016/j.jad.2020.06.066

Linacre, J. M., & Wright, B. D. (1994). Reasonable mean square fit values. *Rasch Measurement Transactions, 8*(2), 370.

Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *Canadian Medical Association Journal*, *184*(3), E191–E196. https://doi.org/10.1503/cmaj.110829

Manea, L., Gilbody, S., & McMillan, D. (2015). A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. *General Hospital Psychiatry*, *37*(1), 67–75. https://doi.org/10.1016/j.genhosppsych.2014.09.009

Ministerio de Salud del Perú. (2020, January). *Minsa: Más de 300 mil casos de depresión fueron atendidos durante el 2021 [Minsa: More than 300 thousand cases of depression were treated during 2021]*. https://www.gob.pe/institucion/minsa/noticias/575899-minsa-mas-de-300-mil-casos-de-depresion-fueron-atendidos-durante-el-2021

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's guide* (7th ed.). Muthén & Muthén.

Naser, A. Y., Dahmash, E. Z., Al-Rousan, R., Alwafi, H., Alrawashdeh, H. M., Ghoul, I., Abidine, A., Bokhary, M. A., AL-Hadithi, H. T., Ali, D., Abuthawabeh, R., Abdelwahab, G. M., Alhartani, Y. J., Al Muhaisen, H., Dagash, A., & Alyami, H. S. (2020). Mental health status of the general population, healthcare professionals, and university students during 2019 coronavirus disease outbreak in Jordan: A cross-sectional study. *Brain and Behavior*, *10*(8). https://doi.org/10.1002/brb3.1730

Petersen, J. J., Paulitsch, M. A., Hartig, J., Mergenthal, K., Gerlach, F. M., & Gensichen, J. (2015). 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*, *170*, 138–142. https://doi.org/10.1016/j.jad.2014.08.053

Pranckeviciene, A., Saudargiene, A., Gecaite-Stonciene, J., Liaugaudaite, V., Griskova-Bulanova, I., Simkute, D., Naginiene, R., Dainauskas, L. L., Ceidaite, G., & Burkauskas, J. (2022). Validation of the Patient Health Questionnaire-9 and the generalized anxiety disorder-7 in Lithuanian student sample. *Plos One*, *17*(1), e0263027. https://doi.org/10.1371/journal.pone.0263027

Prata, D., Mechelli, A., & Kapur, S. (2014). Clinically meaningful biomarkers for psychosis: A systematic and quantitative review. *Neuroscience & Biobehavioral Reviews*, *45*, 134–141. https://doi.org/10.1016/j.neubiorev.2014.05.010

Rahman, M. A., Dhira, T. A., Sarker, A. R., & 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

Reich, H., Rief, W., Brähler, E., & Mewes, R. (2018). Cross-cultural validation of the German and Turkish versions of the PHQ-9: An IRT approach. *BMC Psychology*, *6*(1), 26. https://doi.org/10.1186/s40359-018-0238-z

Salari, N., Hosseinian-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, S., Mohammadi, M., Rasoulpoor, S., & Khaledi-Paveh, B. (2020). Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: A systematic review and meta-analysis. *Globalization and Health*, *16*(1), 57. https://doi.org/10.1186/s12992-020-00589-w

Sánchez-Carlessi, H. H., Yarlequé-Chocas, L. A., Javier-Alva, L., Nuñez-LLacuachaqui, E. R., Arenas-Iparraguirre, C., Matalinares-Calvet, M. L., Gutierrez-Santayana, E., Egoavil-Medina, I., Solis-Quispe, J., & Fernandez-Figueroa, C. (2021). Indicadores de ansiedad, depresión, somatización y evitación experiencial en estudiantes universitarios del Perú en cuarentena por COVID-19 [Indicators of anxiety, depression, somatization,

and experiential avoidance in Peruvian university students quarantined by COVID-19.]. *Revista de la Facultad de Medicina Humana*, *21*(2), 346–353. http://dx.doi.org/10.25176/rfmh.v21i1.3179

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 561–582. https://doi.org/10.1080/10705510903203433

Seo, J.-G., & Park, S.-P. (2015). Validation of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with migraine. *The Journal of Headache and Pain*, *16*(1), 65. https://doi.org/10.1186/s10194-015-0552-2

Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, *45*(4), 249–257. https://doi.org/10.1080/00050067.2010.482109

Stochl, J., Fried, E. I., Fritz, J., Croudace, T. J., Russo, D. A., Knight, C., Jones, P. B., & Perez, J. (2022). On dimensionality, measurement invariance, and suitability of sum scores for the PHQ-9 and the GAD-7. *Assessment*, *29*(3), 355–366. https://doi.org/10.1177/1073191120976863

Sun, Y., Fu, Z., Bo, Q., Mao, Z., Ma, X., & Wang, C. (2020). The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. *BMC Psychiatry*, *20*(1), 474. https://doi.org/10.1186/s12888-020-02885-6

Umegaki, Y., & Todo, N. (2017). Psychometric properties of the Japanese CES–D, SDS, and PHQ–9 depression scales in university students. *Psychological Assessment*, *29*(3), 354–359. https://doi.org/10.1037/pas0000351

Villarreal-Zegarra, D., Cabrera-Alva, M., Carrillo-Larco, R. M., & Bernabe-Ortiz, A. (2020). Trends in the prevalence and treatment of depressive symptoms in Peru: A population-based study. *BMJ Open*, *10*(7), e036777. https://doi.org/10.1136/bmjopen-2020-036777

Villarreal-Zegarra, D., Copez-Lonzoy, A., Bernabé-Ortiz, A., Melendez-Torres, G. J., & Bazo-Alvarez, J. C. (2019). Valid group comparisons can be made with the Patient Health Questionnaire (PHQ-9): A measurement invariance study across groups by demographic characteristics. *Plos One*, *14*(9), e0221717. https://doi.org/10.1371/journal.pone.0221717

Wisting, L., Johnson, S. U., Bulik, C. M., Andreassen, O. A., Rø, Ø., & Bang, L. (2021). Psychometric properties of the Norwegian version of the Patient Health Questionnaire-9 (PHQ-9) in a large female sample of adults with and without eating disorders. *BMC Psychiatry*, *21*(1), 6. https://doi.org/10.1186/s12888-020-03013-0

Wittkampf, K. A., Naeije, L., Schene, A. H., Huyser, J., & van Weert, H. C. (2007). Diagnostic accuracy of the mood module of the Patient Health Questionnaire: A systematic review. *General Hospital Psychiatry*, *29*(5), 388–395. https://doi.org/10.1016/j.genhosppsych.2007.06.004

World Health Organization. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf?s

Wu, Z., & Fang, Y. (2014). Comorbidity of depressive and anxiety disorders: Challenges in diagnosis and assessment. *Shanghai Archives of Psychiatry*, *26*(4), 227. https://doi.org/10.3969/j.issn.1002-0829.2014.04.006

Zhou, Y., Xu, J., & Rief, W. (2020). Are comparisons of mental disorders between Chinese and German students possible? An examination of measurement invariance for the PHQ-15, PHQ-9 and GAD-7. *BMC Psychiatry*, *20*(1), 480. https://doi.org/10.1186/s12888-020-02859-8

Zivin, K., Eisenberg, D., Gollust, S. E., & Golberstein, E. (2009). Persistence of mental health problems and needs in a college student population. *Journal of Affective Disorders*, *117*(3), 180–185. https://doi.org/10.1016/j.jad.2009.01.001

Zwitser, R. J., & Maris, G. (2016). Ordering individuals with sum scores: The introduction of the nonparametric Rasch model. *Psychometrika*, *81*(1), 39–59. https://doi.org/10.1007/s11336-015-9481-x

Psihometrijske karakteristike Upitnika o zdravlju pacijenta – 9 (PHQ-9) kod peruanskih sveučilišnih studenata

Sažetak

Promjene koje su se u ljudskim životima dogodile zbog pandemije uzrokovane bolešću COVID-19 negativno su utjecale na mentalno zdravlje sveučilišnih studenata kod kojih je od početka pandemije evidentiran porast simptoma depresije te je u tome kontekstu važno provesti njihovu procjenu. Stoga je cilj ovoga istraživanja bio analizirati psihometrijske značajke Upitnika o zdravlju pacijenta – 9 (PHQ-9) kod peruanskih sveučilišnih studenata. U istraživanju je sudjelovalo 409 sveučilišnih studenata u dobi od 16 do 45 godina (ženski = 72.9 %; prosječna dob = 22.15; SD_{dob} = 3.75). Korišteni su upitnik PHQ-9 i Skala emocionalnoga iscrpljenja (EES). Proveden je niz konfirmatornih faktorskih analiza da bi se procijenile strukture predložene u prethodnim istraživanjima: jednodimenzionalni model, model s dvama koreliranim faktorima i bifaktorski model. Pronašli smo dokaze o jednodimenzionalnoj strukturi, kako na temelju rezultata faktorske analize, tako i s aspekta teorije odgovora na zadatke, te pozitivnu povezanost s akademskim emocionalnim iscrpljenjem. Osim toga, pouzdanost rezultata i konstrukta dosegnule su visoke vrijednosti. Stoga je upitnik PHQ-9 instrument s prikladnim psihometrijskim značajkama za upotrebu pri mjerenju depresije kod sveučilišnih studenata u kontekstu pandemije uzrokovane bolešću COVID-19.

Ključne riječi: sveučilišni studenti, depresija, valjanost, pouzdanost, PHQ-9

Primljeno: 21. 9. 2022.