Examining the construct validity of the Portuguese version of the Patient Health Questionnaire-9 among college students

Sara Monteiro1,2, Ana Bártilo1,2, Ana Torres2,3, Anabela Pereira1,4 & Emília Albuquerque5
1 Department of Education and Psychology, University of Aveiro, Aveiro, Portugal
2 Center for Health Technology and Services Research (CINTESIS), Department of Education and Psychology, University of Aveiro, Aveiro, Portugal
3 ESEnfCVPOA - School of Nursing Portuguese Red Cross Oliveira de Azeméis, Oliveira de Azeméis, Portugal
4 CIDTFF - Didactic Research Center and Technology in the Training of Trainers, Department of Education and Psychology, University of Aveiro, Aveiro, Portugal
5 Portuguese Oncology Institute of Coimbra - Francisco Gentil, E.P.E, Coimbra, Portugal

Abstract: The present study examines a new factor structure and the convergent validity of the Portuguese version of the Patient Health Questionnaire-9 (PHQ-9) with college students. This measure has been used to evaluate depressive symptoms in adults. The total sample included 958 college students. Data were collected from a Web-based survey carried out in schools forming the Polytechnic Institute of Coimbra. Students completed the PHQ-9 and Hospital Anxiety and Depression Scale (HADS) and Brief Symptom Inventory (BSI) subscales as convergent measures. Results pointed to good fit of a second-order factor model with three first-order factors (somatic, cognitive and affective dimensions of depression). Strong positive correlations were found between PHQ-9 scores and HADS depression and BSI depression subscales. Our findings reinforce the PHQ-9 as a valid tool in higher education settings. Future studies should re-examine the dimensional structure of the tool considering its implications for the clinical interpretation of the measure.

Keywords: Depression; PHQ-9; college students; validation.

Examinando a validade de construto da versão Portuguesa do Patient Health Questionnaire-9 entre estudantes universitários: O presente estudo examina uma nova estrutura fatorial e a validade convergente da versão Portuguesa do Patient Health Questionnaire-9 (PHQ-9) usando estudantes universitários. Esta medida é usada para avaliar sintomas depressivos em adultos. A amostra inclui 958 estudantes universitários. Os dados foram recolhidos através de uma pesquisa online conduzida no Instituto Politécnico de Coimbra. Os estudantes completaram o PHQ-9, a Escala de Ansiedade e Depressão Hospitalar (EADH) e subescalas do Inventário de Sintomas Psicopatológicos (BSI) como medidas convergentes. Os resultados apontaram para o bom ajuste de um modelo fatorial de segunda ordem com três fatores de primeira ordem (somático, cognitivo e afetivo). Foram encontradas fortes correlações positivas entre as pontuações do PHQ-9 e as subescalas de depressão. O PHQ-9 mostra ser um instrumento válido para a aplicação no ensino superior. Estudos futuros devem (re)examinar a estrutura fatorial desta escala considerando as suas implicações para a interpretação clínica do instrumento.

Palavras-chave: Depressão; PHQ-9; estudantes universitários; validação.

Depression is a disorder that increasingly affects adolescents and young adults (Stein & Fazel, 2015), causing impairment in inter-personal, occupational and social functioning. In this diagnosis are present feelings of sadness and/or loss of interest leading to a variety of physical and emotional problems (APA, 2014). Beginning university is one of the major transitions that may increase the risk of developing depression (Horgan, Sweeney, Behan & McCarthy, 2016). During this period, new skills are required to cope with new experiences, particularly, for many students, separation from their social network.

Large-scale research using The World Health Organization World Mental Health Surveys showed that 20.3% of college students in 21 countries, including Portugal, presented a 12-month prevalence of some DSM-IV/CIDI disorders (Auerbach et al., 2016) and depression is presented as the most common and incapacitating problem in this setting (Ibrahim, Kelly, Adams & Glazebrook, 2013). Empirical studies have even pointed to an association between depression and academic failure (Hysenbegasi, Hass & Rowland, 2005), alcohol abuse (Horgan et al, 2016) and suicide ideation (Bazzay, Karver & Verona, 2016).

1Correspondence address: Ana Bártilo, Department of Education and Psychology, Campus Universitário de Santiago 3810-193 Aveiro, Portugal. E-mail: anabartilo@ua.pt
Validity of the PHQ-9

2016). It is a serious public health problem and university is an important setting to assess and address mental health. However, mental illness among students often goes unrecognized, and therefore, untreated (Auerbach et al., 2016) and the longer the delay in starting treatment, the more severe the symptoms and their impact on the life domains of the student. Thus, the success of an adjusted and timely intervention on depressive symptoms involves the analysis and availability of reliable screening tools (Stein & Fazel, 2015).

The nine-item Patient Health Questionnaire (PHQ-9) is a self-report tool used in screening and tracking depressive symptoms in adults in several countries such as Nigeria (Adewuya, Ola & Afolabi, 2006), Spain (Diez-Quevedo, Rangil, Sanchez-Planell, Kroenke, Spitzer, 2001), Canada (Liu & Wang, 2015) and Taiwan and China (Liu et al., 2011; Zhang et al., 2013). It was originally designed for primary care (Kroenke, Spitzer & Williams, 2001), but in the last decade, its psychometric properties have been examined with other clinical (e.g. Hinz et al., 2016) and non-clinical samples (e.g. Liu & Wang, 2015). More specifically, studies developed in Nigeria (Adewuya et al., 2006) and China (Zhang et al., 2013) indicated the usefulness of the PHQ-9 also during the transition to adulthood, presenting good reliability and validity in higher education settings.

Currently, there are valid measures to assess depression in Portugal, namely HADS–Hospital Anxiety and Depression Scale (Pais-Ribeiro et al., 2007; Zigmond & Snaith, 1983) and BDI–Beck Depression Inventory–II (Beck, Steer & Brown, 1996; Martins, Coelho, Ramos & Barros, 2000). The PHQ-9 stands out from other measures because it is a solution based on DSM-IV major depression criteria (APA, 2002) that were kept in DSM-5 (APA, 2014), reflecting, for this reason, the presence of major or minor depression. This argument supported the need to bring this measure to Portugal, and a preliminary study, adapting the English version of the PHQ-9, indicated the potential validity of the tool to assess depression in Portuguese university students (Monteiro, Torres, Pereira, Albuquerque & Morgadinho, 2013). Good internal consistency and moderate to strong convergent validity with external scales to assess depression were reported with this version of the measure (Monteiro et al., 2013), but its factor structure was not explored and the full work was not published. Based on this, the present study evaluates the validity of this Portuguese version of the PHQ-9, also recently used among breast cancer patients (Torres, Monteiro, Pereira & Albuquerque, 2016).

Through the collection of a new sample of college students, our aim was to perform a comprehensive psychometric validation of the PHQ-9. Controversial findings are found in relation to the PHQ-9 structure using confirmatory analyses. The majority of studies confirmed a one-factor structure (“depression severity”) (Cameron, Crawford, Lawton & Reid, 2008; Huang, Chung, Kroenke, Delucchi & Spitzer, 2006; Yu, Tam, Wong, Lam & Stewart, 2012). In turn, a factor model, originally proposed by Krause, Bombardier and Carter (2008), involving somatic and cognitive/affective dimensions of depression as two independent latent factors, has also been confirmed mainly in clinical settings (e.g. Chilcot et al., 2013). However, to our knowledge, no published study has hypothesized a second-order factor model to account for the relations among somatic, affective and cognitive symptoms in clinical or non-clinical samples.

Based on this, our study proposed a new framework to examine the factor validity of the PHQ-9 among college students. We consider an alternative model comprising the three dimensions of depression as first-order factors loading on a second-order factor, that is, depression severity. More specifically, the main objective was to compare the one-factor and two-factor solutions with an alternative factor structure. The proposed conceptual model considers the relations between somatic, cognitive and affective dimensions of depression and may contribute to guiding therapeutic intervention. Additionally, the convergent validity of the Portuguese version of the PHQ-9 was also examined considering external scales which assess depression. Other mental health indicators such as anxiety, somatic symptoms (Haug, Myklebust, & Dahl, 2004) and interpersonal sensitivity (Wilhelm, Boyce, & Brownhill, 2004) have been related to depression. Thus, we hypothesized that there would be a moderate to strong positive correlations between these indicators and the depressive symptoms assessed by the PHQ-9.

**METHOD**

**Participants and Procedures**

The total sample included 958 Portuguese students over 18 years old [71.6% women and 28.4% men, with an average age of 24.03 (SD=7.12)] recruited through a web-based survey of college students’ mental health and health risk behaviors. The link to the survey was divulged in different institutions of the Polytechnic Institute of Coimbra. We provided electronic informed consent and asked if participants would agree to participate. Most college students were single (87.6%) and 59.6% were studying away from home. Regarding education, 74.6% were undergraduates, 16.5% were master students, and 9.0%
Ph.D. students. Participants attended different training areas, namely health (33.0%), education of applied social sciences and arts (26.7%), accountancy, administration, management and marketing (23.8%), engineering (13.0%) and technology, science and agricultural engineering (3.4%). Of the students, 11.4% take medication regularly for various clinical reasons including chronic diseases such as asthma and diabetes.

Measurements

PHQ-9. The PHQ-9 (Monteiro et al., 2013; Kroenke et al., 2001) is a self-report measure consisting of nine items. Participants were instructed to report symptoms experienced during the two weeks prior to completing the questionnaire on a 4-point scale. Scores for each item in the PHQ-9 range from 0 (not at all), to 1 (several days), 2 (more than half of the days) and 3 (nearly every day), while summed scores range from 0 to 27. Higher scores were associated with more severe depression. As depression severity categories, we considered the recommendations of Kroenke et al. (2001) namely: minimal 0-4, mild 5-9, 10-14 moderate, moderately severe 15-19 and severe 20-27. The global scale Cronbach’s α was .88.

HADS. The HADS (Pais-Ribeiro et al., 2007; Zigmond & Snaith, 1983) was used as a convergent measure of the PHQ-9. HADS consists of two subscales, one measuring anxiety (HADS-A), and one measuring depression (HADS-D), both with seven items, which are scored separately. Each item is answered using a 4-point (0–3) response scale, so possible scores range from 0 to 21 for anxiety, and from 0 to 21 for depression. In the present study, Cronbach α values were .83 for anxiety and .79 for depression.

Brief Symptom Inventory (BSI). The BSI (Canavarro, 1999; Derogatis, 1982) is a self-report measure developed to assess general psychiatric symptomatology. This tool was also used to examine convergent validity of the PHQ-9. It contains 53 items rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely). This instrument assesses nine primary symptom dimensions and three global indices of distress, namely the general index of symptoms (GIS). In this study, we only used depression, somatization, interpersonal sensitivity and anxiety subscales. Here, the Cronbach alpha reliability ranged from .83 (somatization) to .89 (depression) for primary symptom dimensions. For the GIS, the Cronbach α value was .97.

Statistical analyses

Confirmatory factor analysis (CFA) was conducted through SPSS AMOS, version 24, using asymptotically distribution-free as the estimator. To assess overall model fit we evaluated the parameters recommended by Hu and Bentler (1999): (i) chi-square statistic ($\chi^2$); (ii) comparative fit index (CFI); (iii) root-mean-square error of approximation (RMSEA); and (iv) standardized root-mean-square residual (SRMR). Convergent validity was assessed by exploring Spearman’s rank correlations between the PHQ-9 scores and the HADS and BSI domains using Statistical Package for Social Sciences software, version 24.

RESULTS

Comparing Alternative Models

Three alternative models were tested using CFA. We estimated the unidimensional model, a model with two factors involving somatic and cognitive/affective dimensions and a second-order factor model with three first-order factors (somatic items: 3, 4 and 5; affective items: 1, 2, 6 and 9; and cognitive items: 7 and 8). Chi-square values were significant, rejecting the models. However, we know that this statistic is sensitive to sample size. Therefore, other fit indexes were analyzed following the cut-off criteria: CFI ≥.95, RMSEA<.06 and SRMR <.08 (Hu & Bentler, 1999) considered as indicators of the model’s good fit. The results confirmed the poor fit of the one-factor solution. This original unidimensional structure did not pass in the majority of the recommended evaluation criteria, as shown in Figure 1.
Validity of the PHQ-9

Figure 1. One-factor model.

Given that the suggested modifications were not conceptually justified, this model was not re-specified again. In turn, the two-factor structure proposed by Krause, Bombardier and Carter (2008) was substantially better than the one-factor model (CFI=.816, RMSEA=.056, SRMR=.043). The standardized factor loadings ranged from .721 to .787 for somatic dimension and ranged from .489 to .814 for cognitive/affective dimension. In this framework, the RMSEA and SRMR were within the cut-off criterion showing that model parameter estimates fit the population covariance matrix. However, the second-order factor model had slightly better fit indexes (see Figure 2). All standardized factor loadings of the item parcels were statistically significant showing that question items were good indicators for each first-order latent factor (p<.001). All the first-order factors loaded significantly on the second-order factor (depression severity). Additionally, the three latent factors showed satisfactory internal consistency (α=.767, α=.810 and α=.653 for somatic, affective and cognitive factors, respectively).
Figure 2. Proposed conceptual model: Second-order factor structure.

Convergent Validity
The Portuguese version of the PHQ-9 was compared to other self-report measures, as shown in Table 1. As hypothesized, the moderate to strong positive correlations between the PHQ-9 global score and the HADS-A and HADS-D subscales were found. Regarding the BSI symptoms, the PHQ-9 index correlated most strongly with the BSI depression subscale, but also presented significant positive associations with all of the other domains. The GIS presented a strong positive correlation with the global score of the PHQ-9. The somatic, affective and cognitive factors of the PHQ-9 were also positively related to the HADS and BSI dimensions. The BSI somatization subscale showed a stronger correlation with the somatic dimension of depression.
Table 1. Convergent validity: Spearman’s correlations between the PHQ-9, HADS-D, HADS-A and the BSI dimensions.

<table>
<thead>
<tr>
<th>PHQ-9</th>
<th>HADS Anxiety</th>
<th>HADS Depression</th>
<th>BSI Depression</th>
<th>BSI Anxiety</th>
<th>BSI Somatization</th>
<th>BSI Interpersonal</th>
<th>GIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>.680</td>
<td>.653</td>
<td>.766</td>
<td>.694</td>
<td>.633</td>
<td>.651</td>
<td>.811</td>
</tr>
<tr>
<td>Somatic dimension</td>
<td>.595</td>
<td>.516</td>
<td>.583</td>
<td>.598</td>
<td>.575</td>
<td>.497</td>
<td>.675</td>
</tr>
<tr>
<td>Affective dimension</td>
<td>.592</td>
<td>.643</td>
<td>.800</td>
<td>.614</td>
<td>.528</td>
<td>.657</td>
<td>.748</td>
</tr>
<tr>
<td>Cognitive dimension</td>
<td>.556</td>
<td>.506</td>
<td>.541</td>
<td>.566</td>
<td>.542</td>
<td>.508</td>
<td>.622</td>
</tr>
</tbody>
</table>

Note: All correlations were statistically significant at p<.001.

DISCUSSION AND CONCLUSION

Our data using 958 students showed that the Portuguese version of the PHQ-9 is a valid tool to assess depression in this period of transition to adulthood. Our comprehensive approach confirmed the strong construct validity of this version. An acceptable fit of the second-order factor model was demonstrated.

Firstly, it is important to underline that analysis using CFA procedures has failed in validation studies with this tool. Most analysis were based on exploratory factor analysis and, for this reason, the fit of alternative factor structures was not tested. In this research, we showed that the one-factor solution was not adequate for the Portuguese version of the PHQ-9 administered to college students. However, the two-factor model proposed by Krause, Bombardier and Carter (2008) and the conceptual model including three first-order factors loading on a second-order factor, named depression severity, presented a good fit. This last model was a little better and items were good indicators for each first-order latent factor. The standardized factor loadings were higher considering the cognitive and affective symptoms as independent latent factors. Thus, our findings suggest the importance of clinicians or researchers considering the relation of the somatic, affective and cognitive dimensions as consistent indicators of depression severity, to allow better conceptual understanding of the PHQ-9 as an instrument. This view of the broad clinical domains reported by college students may lead to better clinical judgment than a single severity score derived from the PHQ-9 total scale.

Concerning convergent validity, the data obtained were in line with previous studies. The relationship between the Portuguese versions of the PHQ-9 and HADS-D subscales is not surprising. Several studies point to a strong correlation of these measures (e.g. Cameron, Crawford, Lawton & Reid, 2008), but with significant differences between their severity categories (e.g. Hansson, Chotai, Nordstöm & Bodlund, 2009). As hypothesized, the HADS-A subscale also presented moderate positive correlation with PHQ-9 total score and their factors. The association between the anxiety and depression constructs is consensually recognized and can justify these findings. More specifically, the studies by Cameron et al. (2008) also demonstrated a close relation between depression and anxiety using the same measures. The correlation between the PHQ-9 index and BSI scores reinforced the convergent validity of the measure, since PHQ-9 presented a strong association with a BSI depression subscale. Other BSI dimensions such as anxiety, somatization and interpersonal sensitivity showed moderate positive correlations with the PHQ-9 scores. Among them, the anxiety domain showed the highest moderate convergence with the PHQ-9 dimensions.

Overall, this study proposes the European Portuguese version of the PHQ-9 as a valid and available tool for depression among students. Our comprehensive approach to validation of PHQ-9 goes beyond other validation studies developed with Western (Adewuya et al., 2006) or non-Western (Zhang et al., 2013) samples of undergraduate students, which mainly focus on the measure’s operating characteristics. This study instigates a return to examining the dimensional structure of the measure in non-clinical communities. An assessment based on a multidimensional structure may allow the identification of the main symptoms of college students with depression and proper treatment. Nevertheless, our conclusions should be interpreted cautiously since they are limited by methodological issues such as data collection being predominantly based on the Internet, increasing bias; geographical limitation in divulging the study; and lastly, the underrepresentation of male college students in the sample. Future studies should include a more representative sample to examine the factor structure of the PHQ-9 using the proposed conceptual model. Moreover, the invariance of this dimensional structure may be examined across men and women college students.
REFERENCES


Copyright © 2019 Associação Portuguesa de Psicologia


