Patient Health Questionnaire-9 and Generalized Anxiety Disorders-7: Arabic Version Reliability in Jordan

Mustafa S. Yousuf*, Heather L. Harvey2, Aiman S. Al Sharei1, Khaled A. Albakri1, Yasmeen J. Alabdallat1

Abstract
Background: Depression and anxiety are common mental disorders that are often missed in primary healthcare settings due to the lack of diagnosis criteria. Using valid and reliable easy-to-use instruments can overcome this problem. Aims: To translate and culturally adapt the Patient Health Questionnaire-9 (PHQ9) and the Generalized Anxiety Disorders-7 (GAD7) instruments into Arabic, determine the reliability of these translated versions, and compare them with previous attempts. Methods: Forward and blind back-translation was used to translate the instruments into Arabic. An online version was created and sent to medical school students at the Hashemite University, Jordan. For each instrument, internal consistency reliability and inter-item correlation were calculated. For each item, the mean, standard deviation, item-total correlation, and value of Cronbach’s alpha if the item was deleted were determined. Analysis was performed using IBM SPSS version 25. Results: There were 256 respondents to the translated instruments. The average age was 18.85 and 61% were females. The alpha for the PHQ9 was 0.876, with an average inter-item correlation of 0.444. Removal of any of the nine items decreased the value of alpha. The GAD7 had an alpha of 0.895 with an average inter-item correlation of 0.551. Deletion of any item decreased the internal consistency reliability of the instrument. Conclusion: Using an efficient translation process that was both accurate and culturally sensitive enabled the construction of Arabic versions of the PHQ9 and the GAD7 scales that had high internal consistency reliability and good inter-item correlation.

Keywords: Anxiety, depression, generalized anxiety disorders-7, internal consistency reliability, patient health questionnaire-9, translation.

Introduction
Depression and anxiety are globally the two most common mental health illnesses [1–2]. In the context of the COVID-19 pandemic, the incidence of these disorders increased, which expanded their prevalence [3]. This increase may have occurred due to the lockdown [4] and was evident among health care workers [5–6] and students [7–8]. Because of this, the early and accurate diagnosis of such disorders is essential.

Patients suffering from physical illnesses often refer to primary health care for the initial management of their conditions. Although attempts to integrate mental health into primary health care have been made [9–10], various barriers against such efforts still exist.

*Corresponding author: msysalman@yahoo.com
ORCID: 0000-0001-8991-1270

(J Med J 2022; Vol. 56 (3):250-260)
Received January 15, 2021
Accepted November 25, 2021
One option for overcoming these obstacles is the application of easy-to-use self-reporting instruments in the diagnosis of mental health disorders [12]. In this regard, several valid and reliable instruments have been employed. The nine item Patient Health Questionnaire-9 (PHQ9) is one such instrument addressing depression, and was first introduced in 1999 [13]. Its validity and reliability have been ascertained by several studies [14–16]. This has made the PHQ9 useful in various settings, including primary health care [17]. The Generalized Anxiety Disorder-7 (GAD7) instrument is a seven item self-reported scale designed to measure anxiety, first developed in 2006 [18]. Its validity and reliability have also been shown in different studies [19–21]. Both the PHQ-9 and the GAD-7 instruments are characterized by their brevity and simplicity, and they are widely used in different situations.

Translating such instruments into the native language of the target population would make the understanding of these instruments more feasible. To achieve this, the translation must be effective, taking into consideration both accuracy and cultural context [22]. Because of its great benefits, the PHQ9 has been translated into several languages, including Arabic [23–24], Spanish [25], and Chinese [26], among others, while the GAD7 has been translated into Arabic [23–24], Korean [27], and Finnish [28], for example. In order to promote easy recognition and prevent the negligence of depression and anxiety disorders in primary health care, this study aimed to translate and culturally adapt the PHQ9 and the GAD7 into Arabic, determine the reliability of these translated versions, and compare them with previous attempts.

**Materials and Methods**

Two instruments were used, the 9-item PHQ9 deals with depression [13] and the 7-item GAD7 deals with anxiety [18]. Each instrument is a Likert scale with four possible scores: 0 (not at all), 1 (several days), 2 (more than half the days), and 3 (nearly every day). The original layout of the instruments was maintained in the translated versions. The process of translation generally used the guidelines proposed by Sousa and Rojjanasrirat [29], and the following steps were employed: (1) forward translation from English to Arabic by a member of the research team fluent in both languages; (2) the translated instrument was discussed by members of the team fluent in Arabic and a preliminary version was agreed upon; (3) the Arabic version was back translated into English by a team member fluent in both languages and not involved in the first two steps; (4) the back-translated version of the instrument was compared with the original English version by a team member fluent in English; and, (5) suggestions made in the previous step were discussed and appropriate modifications made before the finalized version was decided upon. The face validity of the instrument was judged by a member of the team specialized in psychology. An online version of the finalized translation was created using Google Forms. The link to the online version, along with a consent form, was sent to students enrolled in the first three years of medical school at the Faculty of Medicine, Hashemite University, Jordan, through social media groups. Responses were collected during November, 2020.

To determine the minimum sample size, an online sample size calculator [30] based on the formula proposed by Bonett [31] was used.
The values of the formula variables were chosen as: minimum acceptable alpha, 0.70; expected alpha, 0.80; significance level, 0.05; power, 0.8; number of items, nine (the maximum of our two instruments); and, dropout rate, 10%. As such, the minimum sample size was calculated to be 123.

The collected responses were tabulated using Microsoft Excel. For each instrument, the internal consistency reliability (Cronbach’s alpha) and inter-item correlation were calculated. For each item, the mean, standard deviation, item-total correlation, and value of Cronbach’s alpha if the item was deleted were determined. Two factors were also considered: the gender of the respondent and their age. Analysis of the data was performed using IBM SPSS version 25.

Results
The number of respondents was 256, of whom 156 (61%) were females. Of the respondents, 89 (35%) were first years, 141 (55%) were second years, and 25 (10%) were third years. The age of the participants ranged from 17–22, with a mean of 18.85 and SD of 0.78.

Primary Health Questionnaire-9
There were 252 valid responses for this instrument. Cronbach’s alpha value was 0.876. The statistics of the nine items of this instrument are shown in Table 1. The removal of any of the nine items decreased the value of Cronbach’s alpha. Item seven (trouble concentrating) had the greatest impact on the results: it had the highest correlation to the total score (0.720) and its deletion led to the largest decrease in the value of alpha (0.852). Item three (sleep problems), on the other hand, had the least impact with the least item-total correlation (0.521) and the smallest effect on alpha (0.871).

Table 2 shows the inter-item correlation matrix of the PHQ9 items. The greatest correlation (0.605) was found between item one (little interest or pleasure in doing things) and item two (feeling down or hopeless). The least correlation (0.304) was between items nine (thoughts of hurting oneself) and three (problems sleeping). The average inter-item correlation was 0.444.

General Anxiety Disorders-7
The value of Cronbach’s alpha was 0.895 from 253 valid responses. Table 3 shows the statistics of the seven items of this instrument. Item two (not being able to stop or control worrying) had the greatest item-total correlation (0.761) and its deletion decreased Cronbach’s alpha the most (0.872). Item seven (being fearful something awful might happen) had the least item-total correlation (0.619) and alpha would decrease the least if this item were omitted (0.889). Deletion of any of the seven items decreased the value of Cronbach’s alpha.

The inter-item correlation matrix of the GAD7 items is shown in Table 4. The highest correlation (0.703) was found between item two (on continual worrying) and item one (on feeling nervous/anxious). The smallest correlation (0.430) was between becoming irritable (item six) and having trouble relaxing (item four). The average inter-item correlation was 0.551.

Discussion
In this study, the internal consistency reliability of an Arabic version of the PHQ9 and GAD7 instruments was evaluated in a sample of medical students. Acceptable values
of alpha typically range from 0.70–0.95; however, an alpha >0.90 may indicate that some items are asking about the same concept [32]. The team which originally introduced this instrument tested its reliability on two samples and found them to be 0.86 and 0.89 [14]. Another study found Cronbach’s alpha of the English version of the PHQ9 to be 0.803. This instrument also had good internal consistency reliability after being translated into different languages. A Spanish version had an alpha of 0.87 [25], while the Chinese version of the PHQ9 recorded an alpha of 0.86 in one study [33] and a very high value of 0.938 in another [26]. In this study, the Arabic version of the PHQ9, similarly, showed a high internal consistency with Cronbach’s alpha being 0.876. Al Hadi et al. [24] screened university students with an Arabic version of the PHQ9 and found the alpha to be 0.857, which was lower than the value in this study. Moreover, the average inter-item correlation calculated in this study (0.444) was also slightly higher than the 0.400 recorded in the study by Al Hadi et al. [24]. However, the alpha in this study was slightly lower than the 0.88 calculated by Sawaya et al. [23], who used another Arabic version to screen psychiatric outpatients. Despite the differences, which could have been due to the target population or the size of the sample, all these alpha values can be considered good.

The removal of any item led to a decrease in alpha. Item seven (trouble concentrating) contributed the most to the reliability of the PHQ9. This was item two (feeling down) in the Chinese version [34] and in an Arabic version [24]. With the ongoing COVID-19 pandemic and the reliance on the relatively new approach of online education, our sample of medical students seemed to have so much on their minds that they were unable to concentrate. This probably affected all the aspects they were asked about in the instrument. The mean inter-item correlation for item seven was the highest in this study (0.509), while item three (sleep problems) affected the overall alpha the least. In the aforementioned Chinese and Arabic versions, it was item nine (better dead) that affected alpha the least and, in the Arabic version used by Al Hadi et al. [24], omitting item nine slightly increased alpha.

The Arabic version of the GAD7 scale in this study also demonstrated a high internal consistency reliability, with a Cronbach’s alpha of 0.895. This was reported as 0.92 for the English version [18] and 0.93 for a Korean version [27]. One Arabic version [23] reported a similarly high alpha of 0.95, whereas another [24] reported a lower value of 0.763. In this study, item two (worrying) affected reliability the most. This could also be a reflection of the worrisome situation in which our sample of students found themselves.

The high reliability of both the PHQ9 and GAD7 scales in different languages (including our results) is an indication of how, even after translation, these scales maintain their internal consistency, possibly because of how the original English versions are constructed. The items in these scales are written in an easy-to-understand manner that does not depend on scientific or clinical terms, thus making the translation and cultural adaptation easier. To give an example, the words ‘worry’, ‘nervous’, and ‘anxious’ could all be translated into the same or different Arabic words, depending on the context of their sentence. The items in the two scales are stated so as to make the context clear, which facilitated the way the items were translated.
The translation and validity of this study were rigorous. To the best of our knowledge, this was the first attempt at translating these instruments into Arabic in Jordan. The results obtained in this research were similar to those obtained in previous Arabic translation attempts [23–24]. Compared to our results, the Cronbach’s alpha value obtained by Sawaya et al. [23] was higher, which may have been due to the focused sample used (psychiatric outpatients), while our sample was medical students. Moreover, the result for the GAD7 was >0.9, which is considered unfavorable. The alpha values in the study by Al Hadi et al. [24] were lower than our results (for the GAD7, it was much lower), and this may have been due to the translation being performed by independent certified translators. As such, a more literal rather than contextual translation may have been produced which could have caused some confusion for the participants. In addition, the omission of one item in [24] increased the value of alpha. Although not explained by the authors, this could have resulted from the participants’ misunderstanding of the translated question. The sample obtained in our study was more than double the minimum size required for such research, meaning that the sample was more than sufficient to instill confidence in the results that could have led to a type one error. However, the Arabic version constructed in our study still needs to be validated (only face validation was performed by a member of the team). It is possible that we may need to draw samples from different populations, especially high-risk populations. The results need to be compared to other scales that measure the same constructs and to diagnoses made by specialists.

Conclusions
An accurate, effective and culturally adapted translation process led to the construction of Arabic versions of the PHQ9 and the GAD7 scales, with high internal consistency reliability and good inter-item correlation.

Ethical approval
The Institutional Review Board of the institution in which this research was conducted approved this study. All participants signed a written informed consent prior to this study. Their participation was voluntary, and they had full autonomy to leave the study at any time without explanation.

Funding sources
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Conflicting Interests
The authors declare that there is no conflict of interest.

Author contributions
All the authors made a substantial contribution to the research conception and design, data collection and analysis, and writing and revising of the manuscript. All the authors approved the final draft of this article.
Table 1: Statistics of the nine items of the Patient Health Questionnaire-9

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Item-total correlation</th>
<th>Cronbach’s Alpha if item is deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.81</td>
<td>0.967</td>
<td>0.651</td>
<td>0.860</td>
</tr>
<tr>
<td>2</td>
<td>1.65</td>
<td>1.066</td>
<td>0.709</td>
<td>0.854</td>
</tr>
<tr>
<td>3</td>
<td>1.65</td>
<td>1.149</td>
<td><strong>0.521</strong></td>
<td><strong>0.871</strong></td>
</tr>
<tr>
<td>4</td>
<td>1.89</td>
<td>0.941</td>
<td>0.677</td>
<td>0.858</td>
</tr>
<tr>
<td>5</td>
<td>1.69</td>
<td>1.081</td>
<td>0.582</td>
<td>0.865</td>
</tr>
<tr>
<td>6</td>
<td>1.29</td>
<td>1.221</td>
<td>0.638</td>
<td>0.860</td>
</tr>
<tr>
<td>7</td>
<td>1.50</td>
<td>1.141</td>
<td><strong>0.720</strong></td>
<td><strong>0.852</strong></td>
</tr>
<tr>
<td>8</td>
<td>0.91</td>
<td>1.111</td>
<td>0.542</td>
<td>0.869</td>
</tr>
<tr>
<td>9</td>
<td>0.76</td>
<td>1.129</td>
<td>0.538</td>
<td>0.869</td>
</tr>
</tbody>
</table>
Table 2: Inter-item correlation matrix of the Patient Health Questionnaire-9

<table>
<thead>
<tr>
<th></th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
<th>Item 8</th>
<th>Item 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td>0.605</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td>0.410</td>
<td>0.406</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td>0.542</td>
<td>0.543</td>
<td>0.459</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>0.382</td>
<td>0.482</td>
<td>0.311</td>
<td>0.559</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td>0.479</td>
<td>0.528</td>
<td>0.389</td>
<td>0.419</td>
<td>0.415</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 7</td>
<td>0.528</td>
<td>0.550</td>
<td>0.451</td>
<td>0.596</td>
<td>0.538</td>
<td>0.514</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 8</td>
<td>0.374</td>
<td>0.432</td>
<td>0.316</td>
<td>0.387</td>
<td>0.349</td>
<td>0.436</td>
<td>0.478</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Item 9</td>
<td>0.404</td>
<td>0.475</td>
<td>0.304</td>
<td>0.355</td>
<td>0.332</td>
<td>0.459</td>
<td>0.421</td>
<td>0.364</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 3: Statistics of the seven items of the Generalized Anxiety Disorders-7

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Item-total correlation</th>
<th>Cronbach’s Alpha if item is deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Feeling nervous, anxious, or on edge</td>
<td>1.80</td>
<td>0.964</td>
<td>0.754</td>
<td>0.873</td>
</tr>
<tr>
<td>2 Not being able to stop or control worrying</td>
<td>1.26</td>
<td>1.014</td>
<td>0.761</td>
<td>0.872</td>
</tr>
<tr>
<td>3 Worrying too much about different things</td>
<td>1.58</td>
<td>0.958</td>
<td>0.745</td>
<td>0.874</td>
</tr>
<tr>
<td>4 Trouble relaxing</td>
<td>1.44</td>
<td>1.047</td>
<td>0.688</td>
<td>0.881</td>
</tr>
<tr>
<td>5 Being so restless that it's hard to sit still</td>
<td>1.06</td>
<td>0.990</td>
<td>0.686</td>
<td>0.881</td>
</tr>
<tr>
<td>6 Becoming easily annoyed or irritable</td>
<td>1.58</td>
<td>0.991</td>
<td>0.621</td>
<td>0.881</td>
</tr>
<tr>
<td>7 Feeling afraid as if something awful might happen</td>
<td>1.29</td>
<td>1.035</td>
<td>0.619</td>
<td>0.889</td>
</tr>
</tbody>
</table>
Table 4: Inter-item correlation matrix of the Generalized Anxiety Disorders-7

<table>
<thead>
<tr>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.703</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.650</td>
<td>0.661</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.578</td>
<td>0.565</td>
<td>0.599</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.533</td>
<td>0.584</td>
<td>0.551</td>
<td>0.631</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.554</td>
<td>0.548</td>
<td>0.478</td>
<td>0.430</td>
<td>0.512</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>0.523</td>
<td>0.513</td>
<td>0.565</td>
<td>0.467</td>
<td>0.447</td>
<td>0.469</td>
<td>1.000</td>
</tr>
</tbody>
</table>

References


   https://doi.org/10.1016/S0140-6736(18)32279-7

   https://doi.org/https://doi.org/10.1002/da.23077


   https://doi.org/10.1016/j.bbi.2020.05.026

   https://doi.org/10.1007/s10900-020-00921-w

   https://doi.org/10.1017/S003329172000392X


استبيان صحة المريض - (9) واضطرابات القلق العامة - (7): دراسة موثوقية النسخة العربية في الأردن

مصطفى سعد يوسف1، هندي لي هارفي2، خالد الشرع، خالد البكري1، ياسمين العبد اللات

قسم العلوم الطبية الأساسية، كلية الطب، الجامعة الهاشمية، صندوق بريد 330127، الزرقاء 13133، الأردن

جامعة لايمستون، غافني، كارولاين الجنوبية، الولايات المتحدة الأمريكية

الملخص

الخلفية: الاكتئاب والقلق من الاضطرابات النفسية الشائعة، والتي غالبًا ما يتم إغفالها في أماكن الرعاية الصحية الأولية بسبب نقص معايير التشخيص، وإن استخدام أدوات سهلة التطبيق ذات موثوقية وصالحية عالية يمكن أن يؤدي إلى التغلب على هذه المشكلة.

الأهداف: ترجمة استبيان صحة المريض - (9) (PHQ9) وأداة اضطرابات القلق العام - (7) (GAD) إلى اللغة العربية وتعديلها بما يناسب الثقافة المحلية، وتحديد موثوقية هذه النسخ المترجمة ومقارنتها بالمحاولات السابقة.

الأساليب: تم ترجمة الأدوات من الإنجليزية إلى العربية -ترجمة أمماية- ثم ترجمت الأدوات العربية المترجمة عكسيا من العربية إلى الإنجليزية، لمقارنة النسخة الجاهزة مع النسخة الإنجليزية الأصلية، وتم إنشاء نسخة على الإنترنت وإرسالها إلى طلبة كلية الطب في الجامعة الهاشمية، الأردن، وتم حساب موثوقية الإتساق الداخلي والارتباط بين العناصر لكل أدوات، والتنبؤ، وتحديد المتوازن، والانحراف المعياري، وانتشار كل عنصر على_use

التكنولوجيا: كان هناك (256) مستجيبًا للأدوات المترجمة، وكان متوسط أعمارهم (18.85) عامًا (61.6%) من المشتركين كانوا من الإناث، و كان لدى (9) (PHQ9) ألفا قدره (0.876) بمتوسط ارتباط بين العناصر يبلغ (0.444)، وآخذ إرادة أي من العناصر التسعة إلى تقليل قيمة ألفا، وكان لدى (7) (GAD) ألفا قدره (0.895) بمتوسط ارتباط بين العناصر يبلغ (0.551)، وأدى حذف أي عنصر إلى تقليل موثوقية الإتساق الداخلي للاداة.

الخلاصة: أدى استخدام عملية ترجمة فعالة ودقيقة وحساسية تقليفًا إلى إنشاء نسخ عربية من مقاييس (9) (PHQ9) و (7) (GAD) تتمتع بدرجة عالية من موثوقية الإتساق الداخلي، والارتباط الجيد بين العناصر.

الكلمات الدالة: القلق، الكآبة، استبيان صحة المريض - (9)، اضطرابات القلق العام - (7)، موثوقية الإتساق الداخلي، ترجمة.