Validation of the Hamilton Anxiety Rating Scale and State Trait Anxiety Inventory A and B in Arabic among the Lebanese population

Souheil Hallita,b,∗, Chadia Haddadb,c,1, Rabih Hallita, Marwan Akeib,d, Sahar Obeida,b,c,e, Georges Haddadb,c, Michel Soufia, Wael Khansa, Rony Khourya, Nelly Kheirf, Christiane Abi Elias Hallite, Pascale Salamehb,g,h

a Faculty of Medicine and Medical Sciences, Holy Spirit University of Kaslik (USEK), Jounieh, Lebanon
b INSPECT-LB: Institut National de Santé Publique, Épidémiologie Clinique et Toxicologie, Beirut, Lebanon
c Research and Psychology Departments, Psychiatric Hospital of the Cross, Jal Eddib, Lebanon
d School of Pharmacy, Lebanese International University, Beirut, Lebanon
e Faculty of Arts and Sciences, Holy Spirit University of Kaslik (USEK), Jounieh, Lebanon
f Faculty of Pedagogy, Holy Family University, 5534 Batroun, Lebanon
g Faculty of Pharmacy, Lebanese University, Beirut, Lebanon
h Faculty of Medicine, Lebanese University, Beirut, Lebanon

ARTICLE INFO

Keywords:
Anxiety
HAM-A
STAI-A
STAI-B
Arabic
Validation

ABSTRACT

Objectives: Our aim in this study was to translate the Hamilton Anxiety Rating Scale (HAM-A), State Trait Anxiety Inventory (STAI-A and B) scales to Arabic, linguistically validate them for use in a representative sample of the Lebanese population, and to check the reliability of these Arabic versions.

Methods: This study is cross-sectional, conducted between November 2017 and March 2018, which enrolled 1332 community dwelling participants using a proportionate random sample from all Lebanese Mohafazat.

Results: Three factor analyses for the anxiety scales were run over the whole sample (N = 1332). All of the HAM-A, STAI-A and STAI-B items could be extracted from the list. All items from all the scales did not over-correlate to each other (r > 0.9), did not have a low loading on factors (< 0.3) or a low communality (< 0.3). The factor analysis results showed 2 factors for HAM-A (Cronbach alpha = 0.921), 3 factors for STAI-A (Cronbach alpha = 0.928) and STAI-B (Cronbach alpha = 0.898). A significantly high ICC was found between the HAM-A, STAI-A (ICC = 0.709) and STAI-B (0.704). In addition, a significantly high ICC was found between the STAI-A and B scales (ICC = 0.884).

Conclusion: The linguistically validated Arabic versions of these scales can be used to screen for anxiety among the Lebanese populations.

1. Introduction

Anxiety is as a transitory emotional state affected by individual personality traits that varies across culture and situations.1 The intensity of anxiety varies over time and differs from person to other according to stressful life events, situations and cultural background.2 It is considered as a universal pheromone that can be presents across cultures.

The severity of anxiety symptoms could be measured by using different psychometric tools such as the state- trait anxiety inventory (STAI) and the Hamilton Anxiety Rating Scale (HAM-A). These scales are widely and extensively used in research and clinical studies and are available in different versions for adults and children.3–6 The STAI has been validated and translated to several languages and is considered as a standard international tool for measuring anxiety in research.5 A validated Arabic version of the STAI, with an adequate internal consistency reliability, exists among patients attending dental clinics.7 Also, the HAM-A has been proven to have a high reliability and sensitivity for measuring anxiety and has been validated among adults and adolescents as well.8,9

More specifically, the Hamilton Anxiety Rating Scale (HAM-A) is a psychological questionnaire used by clinicians to rate the severity of a patient's anxiety. Anxiety can denote things such as “a mental state … a drive … a response to a particular situation … a personality trait … and
a psychiatric disorder. Nonetheless it was one of the first anxiety rating scales to be published, the HAM-A remains widely used by clinicians. It was originally published by Max Hamilton in 1959. For clinical purposes, only severe or inadequate anxiety is attended to. This scale is considered a “clinical rating” of the range of anxiety, and is proposed for individuals that are “already diagnosed with anxiety neurosis”.

As for the State-Trait Anxiety Inventory (STAI), it measures two types of anxiety – state anxiety, or anxiety about an event, and trait anxiety, or anxiety level as a personal characteristic; and used in diagnoses, in both clinical and other medical settings. The advantage of this scale is that it could be applied towards assessing different types of anxiety. This was a new development because all other questionnaires focused on one type of anxiety at the time.

Appropriate, validated and translated scales are needed to explain and predict anxiety in each country due to the influence of linguistic and cultural differences. The use of the validated scale in different languages, to assess the severity of the anxiety symptoms, demands two concepts: (1) maintaining a level of psychometric properties comparable to the original version and (2) adapting the scale to the cultural of the country. The validated and translated scale with a good reliability, will be helpful in clinical settings and research.

Therefore, it is essential to validate the Arabic version of the STAI and HAM-A scales for cultural adaptation among the Lebanese population in order to measure the severity of anxiety. Our aim in this study was to translate the HAM-A, STAI-A and B scales to Arabic, linguistically validate them for use in a representative sample of the Lebanese population, and check the reliability of these Arabic versions.

2. Methods

2.1. Study design and sampling (sample 1)

This study is cross-sectional, conducted between November 2017 and March 2018, which enrolled 1332 community dwelling participants using a proportionate random sample from all Lebanese Mohafazat (Beirut, Mount Lebanon, North, South and Bekaa). Each Mohafaza is divided into Caza (stratum), two villages were randomly selected from the list of villages provided by the Central Agency of Statistics in Lebanon. Patients were randomly selected from each village. All participants above 18 years of age were eligible to participate. Excluded where the patients with psychotic problems, mental retardation, dementia or who refused to fill the questionnaire. Data collection was performed through personal interviews with participants by trained, study independent personnel.

2.2. Ethical approval

The Psychiatric Hospital of the Cross Ethics and Research Committee, in compliance with the Hospital’s Regulatory Research Protocol, approved this study protocol (HPC-009-2018) based on the fact that the autonomy and confidentiality of participants were respected and since it was an observational study, no harm will be prompted to them. The purpose and requirement of the study was informed to each patients. Consent was obtained as written approval on the ethical consent form.

2.3. Minimal sample size calculation

Comrey and Lee suggested that a minimum of 10 observations per variable is necessary in order to avoid computational difficulties. Since the STAI-A and B scales questionnaire contains 40 questions (20 questions for each scale), a minimal sample of 400 patients was needed to conduct an exploratory factor analysis.

2.4. Questionnaire

The questionnaire used during the interviewed was in Arabic, the native language of Lebanon. A trained staff was in charge of collecting the data, via a personal interview with each participant. This person was independent of this study. The first part assessed the socio-demographic characteristics of the included population (age, gender, educational level of the patient and his parents, marital status, socioeconomic level, alcohol consumption, family history of psychiatric diseases). The socioeconomic level, defined as the family monthly income, was divided into 3 categories: low (< 1000 USD, intermediate (1000–2000 USD) and high (2000 USD). Also, we asked the patient if he got a physician diagnosis of anxiety and was taking a medication to treat it. The other parts comprised the different scales used in this study as follows:

2.5. Hamilton anxiety scale (HAM-A)

The HAM-A entails 14 items, each categorized by a series of symptoms, and measures mental agitation and psychological distress, as well as anxiety-related physical complaints. The responses on the scale were measured on a 5-point Likert scale: 0 (symptoms not present), 1 (mild symptoms), 2 (moderate symptoms), 3 (severe symptoms) and 4 (very severe symptoms). The total score was calculated by summation of the 14 items.

2.6. State-trait anxiety inventory (STAI) A and B

It is composed of 40 questions that measure two types of anxiety - state anxiety, or anxiety about an event (STAI-A), and trait anxiety, or anxiety level as a personal characteristic (STAI-B). The answers follow a 4-point Likert scale, with 1 (not at all), 2 (somewhat), 3 (moderately so) and 4 (very much so) for STAI-A and 1 (almost never), 2 (sometimes), 3 (often) and 4 (almost always) for STAI-B. Higher scores are positively correlated with higher levels of anxiety. A permission was obtained from Mind Garden Inc. to use the STAI A and B scales in this study.

2.7. Forward translation into Arabic

A single bilingual translator, Arabic native fluent in English, aware of the concepts of the three anxiety scales translated the English versions of the scales into Arabic. An expert committee, composed of health care professionals (psychiatrists and psychologists), a language professional and the original translator, had reviewed and revised the translated questionnaire in order to check for idiomatic and conceptual equivalence of the Arabic translated version.

2.8. Back translation into English

A native English speaker translator person, fluent in Arabic, had back translated the Arabic versions of the three scales into the English language. The translator was unaware with the concepts of the anxiety scales and the original English versions. The expert committee compared the back-translated English version of the questionnaire with the original English one, in order to check for inconsistencies and to solve any discrepancies between the versions. All ambiguities disappeared after repeating the process of forward-back translation.

2.9. Sample 2

We conducted another cross-sectional study in May 2018 on a sample of Lebanese patients attending psychology clinic that enrolled 155 participants. To ensure the validity of the results, the scales was tested on another sample (sample 2), independent from the first one. Patients filled a questionnaire through a face-to-face interview.
2.10. Statistical analyses

The SPSS software version 23 was used for data analysis. A descriptive analysis were done using the counts and percentages for categorical variables and mean and standard deviation for continuous measures. The anxiety scales’ validity was tested in two different methods. First, construct validity was determined by an exploratory factor analysis, using the principal components analysis technique (Sample 1): The Bartlett’s test of sphericity and Kaiser-Meyer-Olkin measure of sampling adequacy were confirmed to be adequate. The number of factors retained corresponded to Eigenvalues higher than one. Items with factor loading > 0.4 were considered as belonging to a factor. The Cronbach’s alpha values and the Intraclass correlation coefficient were used to evaluate the internal consistency of the scales. An ICC > 0.7 indicate a good reproducibility.24

Second, a confirmatory factor analysis was carried out in Sample 2 using the maximum likelihood method for discrepancy function to assess the structure of the instrument. We also reported several goodness-of-fit indicators: the relative chi square (x²/df), the Root Mean Square Error of Approximation (RMSEA), the Goodness of Fit Index (GFI) and the Adjusted Goodness of Fit Index (AGFI). The value of x² divided by the degrees of freedom (x²/df) has a low sensitivity to sample size and the Adjusted Goodness of Fit Index (AGFI). The value of x² divided by the degrees of freedom (x²/df) has a low sensitivity to sample size and may be used as an index of goodness of fit (cut-off values: < 2–5). The RMSEA tests the fit of the model to the covariance matrix. As a guideline, values of < 0.05 indicate a close fit and values below 0.1 an acceptable fit. The GFI and AGFI are chi-square-based calculations independent of degrees of freedom. The recommended thresholds for acceptable values are ≥0.9. The Receiver-Operating Characteristics (ROC) curve was sketched in order to determine the cutoff point of whether the presence of anxiety was predicted or not with a great percentage. The sensitivity and specificity were calculated to assess criterion validity.

3. Results

The sociodemographic characteristics of the participants are summarized in Table 1. The results showed that the mean age of the participants was 28.08 ± 16.79 years, with 63.9% females. The majority (75.7%) had a university level of education, unemployed (68.8%), single (72.6%), with a low monthly income (< 1000 USD) (73.5%). Only 5.2% of the participants had a family history of psychiatric illnesses, whereas 11.8% had a history of medical illness.

3.1. Factor analysis

Three factor analyses for the anxiety scales were run over the whole sample (N = 1332). All of the HAM-A, STAI-A and STAI-B items could be extracted from the list. All items from all the scales did not over-correlate to each other (r > 0.9), did not have a low loading on factors (< 0.3) or low communality (a communality is the extent to which an item correlates with all other items) (< 0.3). The number of loading factors, KMO and Bartlett’s test of sphericity values and the Cronbach alpha values for the HAM-A scale are summarized in Table 2. For STAI-A, scale’s items yielded a total of 3 factors, explaining 56.04% of the total variance (KMO = 0.953; Bartlett’s test of sphericity p < 0.001). Loading factors ranged between 0.467 for item “I am relaxed” to 0.795 for item “I feel steady”. Moreover, the Cronbach’s alphas were recorded as follows: factor 1 = 0.852; factor 2 = 0.855; factor 3 = 0.780; total scale = 0.928. For STAI-B, scale’s items yielded a total of 3 factors, explaining 58.27% of the total variance (KMO = 0.940; Bartlett’s test of sphericity p < 0.001). Loading factors ranged between 0.450 for item “I feel that difficulties are piling up so that I cannot overcome them” to 0.898 for item “I wish I could be as happy as others seem to be”. Moreover, the Cronbach’s alphas were recorded as follows: factor 1 = 0.875; factor 2 = 0.839; factor 3 = 0.742; total scale = 0.898.

<table>
<thead>
<tr>
<th>Items</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genitourinary symptoms</td>
<td>12</td>
<td>0.881</td>
</tr>
<tr>
<td>Cardiovascular symptoms</td>
<td>9</td>
<td>0.821</td>
</tr>
<tr>
<td>Autonomic symptoms</td>
<td>13</td>
<td>0.800</td>
</tr>
<tr>
<td>Respiratory symptoms</td>
<td>10</td>
<td>0.774</td>
</tr>
<tr>
<td>Gastrointestinal symptoms</td>
<td>11</td>
<td>0.721</td>
</tr>
<tr>
<td>Somatic (sensory)</td>
<td>8</td>
<td>0.667</td>
</tr>
<tr>
<td>Behavior at interview</td>
<td>14</td>
<td>0.643</td>
</tr>
<tr>
<td>Somatic (muscular)</td>
<td>5</td>
<td>0.551</td>
</tr>
<tr>
<td>Tension</td>
<td>2</td>
<td>0.910</td>
</tr>
<tr>
<td>Anxious mood</td>
<td>1</td>
<td>0.903</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>6</td>
<td>0.746</td>
</tr>
<tr>
<td>Insomnia</td>
<td>4</td>
<td>0.692</td>
</tr>
<tr>
<td>Fears</td>
<td>3</td>
<td>0.529</td>
</tr>
<tr>
<td>Intellectual</td>
<td>5</td>
<td>0.486</td>
</tr>
</tbody>
</table>

Factor 1 = Somatic items; Factor 2 = Psych/Psychological items.
Cronbach’s alphas: factor 1 = 0.898; factor 2 = 0.853; total scale = 0.921.
Percentage of variance explained: 58.51%.
Kaiser-Meyer-Olkin (KMO) = 0.947.
Bartlett’s test of sphericity = p < 0.001.
3.2. Sample 2

A confirmatory factor analysis was run on sample 2, using the structure obtained in Sample 1. The following results were obtained: for the HAMA scale, the Maximum Likelihood Chi-Square = 152.509 and Degrees of Freedom = 76, which gave an x²/df = 2.01. For non-centrality fit indices, the Steiger-Lind RMSEA was 0.077 [0.057–0.096]. Moreover, the Joreskog GFI equaled 0.98 and AGFI equaled 0.93.

For the STAI-A scale the Maximum Likelihood Chi-Square = 405.995 and Degrees of Freedom = 167, which gave an x²/df = 2.43. For non-centrality fit indices, the Steiger-Lind RMSEA was 0.077 [0.057–0.096]. Moreover, the Joreskog GFI equaled 0.98 and AGFI equaled 0.93.

For the STAI-B scale the Maximum Likelihood Chi-Square = 344.374 and Degrees of Freedom = 165, which gave an x²/df = 2.08. For non-centrality fit indices, the Steiger-Lind RMSEA was 0.087 [0.074–0.099]. Moreover, the Joreskog GFI equaled 0.806 and AGFI equaled 0.753.

3.3. Intraclass correlation coefficient between insomnia scales

A significantly high ICC was found between the HAM-A, STAI-A (ICC = 0.709) and STAI-B (0.704). In addition, a significantly high ICC was found between the STAI-A and B scales (ICC = 0.884) (Table 3).

3.4. ROC curves

The receiver operating characteristic (ROC) curve of the anxiety score as calculated by the HAM-A scale, comparing patients with a physician diagnosis of anxiety to healthy ones, showed that the area under the curve was high = 0.795 [0.760–0.830] (P < 0.001); at value = 11.50, the sensitivity was 76.3% and the specificity was 71.1% (Fig. 1).

The ROC curve of the anxiety score as calculated by the STAI-A scale, comparing patients with a physician diagnosis of anxiety to healthy ones, showed that the area under the curve was high = 0.771 [0.732–0.810] (P < 0.001); at value = 42.50, the sensitivity was 71.3% and the specificity was 69.1% (Fig. 2).

The ROC curve of the anxiety score as calculated by the STAI-B scale, comparing patients with a physician diagnosis of anxiety to healthy ones, showed that the area under the curve was high = 0.779 [0.741–0.817] (P < 0.001); at value = 42.50, the sensitivity was 73.1% and the specificity was 71.6% (Fig. 3).

3.5. Face validity

When comparing the scales scores between patients with and without physician diagnosis of anxiety, the results showed that significantly higher means HAM-A, STAI-A and B scores were found in patients with a physician diagnosis of anxiety compared to those without the diagnosis (p < 0.001 for the three scales) (Fig. 4).

4. Discussion

The Arabic versions of the HAM-A, STAI-A and B scales were validated in the current study for use among the Lebanese population. The results showed a high reliability and validity of these scales that provide an initial evidence that could be used as a screening instrument for anxiety in Lebanon. The translation process was conducted by independent translators, the point that decreases the subjectivity of the
scales’ adaptation. A comparison of the fit of the form obtained through the confirmatory factor analysis revealed that the 2 factors for HAM-A, 3 factors for STAI-A and 4 factors for STAI-B had high fit indices.

4.1. Validation of the scale

The construct validity of the STAI-A and B scales were adequate because items converged over three factors for STAI-A and 4 factors for STAI-B, with adequate factor loadings for all items. The internal consistency of both scales was similar to that obtained in other translated versions for state (Cronbach alpha of 0.93 for state and 0.92 for trait scales in Greece26 and 0.93 to 0.95 for state and 0.91 to 0.93 for trait anxiety in a group of Lebanese and American students27), but lower to that obtained in. The results of the Arabic version of the STAI validity and reliability, had been similar to the psychometric properties reported by other studies.26,28,29

Factor analysis indicated that a three-factor for STAI-A and a four-factor for STAI-B provide the best explaining criteria of anxiety enclosing the state and traits subscales. Different models were detected in other studies; Abdullatif27 had found two and three factors for STAI-A and STAI-B for American and Lebanese samples. In a Korean study,28 factor analysis had revealed two factors, while a single factor and 2 factor models was revealed by the confirmatory factor analysis. These variations might be due to cultural differences between countries; in fact, the experience and expression of emotions differ from person to other due to the influence of cultural background.13 The physiology of the illness syndromes and social context depend on the cultural beliefs that varied across person, thus anxiety disorders are strongly related to racial, ethnic and cultural factors.20

The construct validity of the HAM-A scale was also adequate because items converged over two factors (Somatic and Psychic/Psychologic), with adequate factor loadings for all items, similar to the results obtained from previous research in adults.11,31 The internal consistency of the HAM-A scale was excellent (0.921), suggesting that this scale’s items are able to screen for anxiety among the Lebanese population in an excellent way.

4.2. Validity

The construct validity of the Arabic version of these scales was addressed by calculating the sensitivity and specificity of the scale. This study is the first to our knowledge to assess this property during the validation process of these scales. We obtained good areas under the curves, sensitivities and specificities for all three scales. These good results may be due to the fact that we compared used the physician diagnosis variable to compare between individuals with anxiety and those without.

4.3. Limitations

The current study has several limitations. An information bias could have occurred since the patients might not have understood the questions well or over/underestimated the answers to some questions. A selection bias is possible because of the refusal rate. The levels of experiential and expressive anxiety are gender different thus our results could be affected by the unequal number of males and females. In addition, the majority of the participants were young according to the mean age, therefore, future studies should be targeted towards older
adults. Despite these limitations, the sample size is acceptable and the study results can be extrapolated to the whole population because of the study design used.

5. Conclusion

Since language barriers are an important obstacle to proper medical communication and management, translation of the HAM-A, STAI-A and B scales to Arabic and validation of the Arabic forms was essential. The linguistically validated Arabic versions of these scales can be used to screen for anxiety among the Lebanese populations. Further studies are needed in other Arabic-speaking countries to confirm our findings and extrapolate the use of the Arabic versions of these scales.

Funding sources

None.

Declaration of competing interest

The authors have nothing to disclose.

Acknowledgments

The authors would like to thank Mrs Timi Most for her support and help towards the publication of this paper.

References