Cross-Country Validation of the Arabic version of the WHO-5 Well-Being Index in non-clinical young adults from six Arab countries

Feten Fekih-Romdhane (feten.fekih@gmail.com)  
Tunis Al Manar University

Wissal Cherif  
Tunis Al Manar University

Amthal Alhuwailah  
Kuwait University

Mima Fawaz  
Beirut Arab University

Hanaa Ahmed Mohamed Shuwiekh  
Fayoum University

Mai Helmy  
Sultan Qaboos University

Ibrahim Hassan Mohammed Hassan  
South Valley University

Abdallah Y Naser  
Isra University

Btissame Zarrouq  
Laboratory of Epidemiology and Research in Health Sciences

Marianne Chebly  
Holy Spirit University of Kaslik

Yara El-Frenn  
Holy Spirit University of Kaslik

Gabriella Yazbeck  
Holy Spirit University of Kaslik

Gaelle Salameh  
Holy Spirit University of Kaslik

Ayman Hamdan-Mansour  
the university of Jordan

Eqbal Radwan  
Islamic University of Gaza

Abir Hakiri  
Tunis Al Manar University

Sahar Obeld  
Lebanese American University

Majda Cheour  
Tunis Al Manar University

Souheil Hallit  
Holy Spirit University of Kaslik

Research Article

Keywords: Subjective well-being, WHO-5, Arab countries, Arabic, psychometric properties, cross-country validity
Abstract

Background: Subjective well-being (SWB) is a culturally-dependent and context-driven concept; it thus varies widely across and within cultures. Therefore, cross-cultural validation studies are crucial to prove that the well-being measure covers transcultural components of the construct subjective, and can be used for cross-cultural comparison purposes in international multicenter research. In this regard, we aimed to perform a cross-country validation of the Arabic version of the WHO 5-item Well-Being Index (WHO-5) in terms of factor structure, composite reliability, cross-gender measurement invariance, and concurrent validity by calculating Pearson correlation coefficients between the WHO-5 and measures of depression, anxiety, stress, suicidal ideation and insomnia.

Method: We carried-out a cross-sectional, web-based study among a total of 3247 young adults (aged 18-35 years) from six Arab countries (i.e., Tunisia, Lebanon, Egypt, Jordan, Morocco and Kuwait).

Results: We found that WHO-5 mean scores varied significantly across countries, ranging from 32.2 ± 22.72 in Egypt to 44.2 ± 26.84 in Morocco. Confirmatory Factor Analyses showed that the fit of a one-factor model to the data proved to be acceptable in all six countries. In addition, the Arabic WHO-5 yielded high reliability coefficients in samples from each of the six countries (McDonald’s ω = 0.92-0.96) and both genders. Results from multi-group analyses demonstrated that configural, metric, and scalar invariance was supported across gender in the total sample and by country. As for concurrent validity, WHO-5 scores showed a strong significant inverse correlation with depression scores. Negative correlations have also been demonstrated between WHO-5 scores and different symptoms of mental health problems (anxiety, stress, suicidal ideation, insomnia) in our sample.

Conclusion: By verifying the cross-country validity of the Arabic WHO-5, our study opens its wider application to epidemiologically explore SWB among Arabic-speaking community young adults from broad geographic areas.

1. Introduction

As defined by the World Health Organization (WHO), mental health refers to “a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community” [1]. Within this framework, a state of good mental health is not limited to the absence of mental illnesses, but is also described as a state of well-being in all bodily, psychological and social domains [2]. The concept of subjective well-being (SWB) encompasses both negative (e.g., depression, anxiety) and negative aspects (e.g., happiness, satisfaction, contentment) [3, 4]. Although SWB is a complex construct that may have various connotations for different populations and cultures [5], it has universally and consistently proven to be a key outcome and predictor of several major life domains, and to benefit both physical and mental health [6]. Indeed, SWB was found to be closely related to a range of important life domains, including positive development, successful learning [7], high quality social relationships, better academic/work performance, less mental distress and increased resilience in face of stressors [6]. The well-established salutary impact of SWB on health several researchers called for including SWB as a measure of outcome of patient-centered mental healthcare [8], and several countries have already included SWB as a routine assessment to inform government decisions and public policy [9, 10]. As such, particular emphasis was placed in recent years on collecting self-rated SWB data in clinical settings [11], in the general population [12], as well as in research [11, 13], in an attempt to deepen understanding of the SWB concept and its applications.

One of the well-known, free-to-use and most widely used scales for assessing SWB is the WHO 5-item Well-Being Index (WHO-5; [11, 14]). The WHO-5 allows for a simple, brief self-report evaluation of the SWB construct over a two-week period using five positively worded items scored on a five-point scale. All items focus on positive health statements [11], and measure a global hedonic dimension of SWB [15]. The WHO-5 has evidenced good psychometric qualities in a unidimensional structure, with high internal consistency, and high convergent associations with other well-being measures (e.g., [16]). Since its development, the WHO-5 has gained global popularity and was translated into more than 30 languages [1, 3], predominantly in high-income Western and Asia-Pacific settings. The different linguistic versions of the WHO-5 include Icelandic [17], Swedish [18], Spanish [19], Polish [20], Italian [21], Romanian [22], Danish [23], Sinhala [24], Brazilian Portuguese [25], Farsi [26], Turkish [27], Malay [28], Thai [29], Taiwanese [30], Bangla [31], Japanese [32], Korean [33], Chinese [34], Swahili Kenyan [35]. All these versions confirmed the robustness of the WHO-5 and its utility in different research settings and across different geographical contexts [11]. Additionally, we could find two previous validations of the WHO-5 into the Arabic language. The first one was performed among a relatively small, gender-disproportionate (N
different ways. First, we propose to investigate, for the first time, the cross-cultural validity of the WHO-5 across different Arab countries and cultural backgrounds. In this paper, we aimed to contribute to the literature on SWB in Arab countries and to provide evidence for the implementation of evidence-informed interventions to improve Arab people's well-being.

Well-being in the Arab world

People from Arab countries have been struggling over the past years with a high burden of mental health problems. Indeed, mental disorders rates exceeded the expected values in Arab Eastern Mediterranean countries, generating steadily increasing and higher than globally burden levels [51]. This burden is expected to be on the rise due to the unstable political, economic, and social climate in the Arab region (e.g., [52, 53]); and mental health will likely pose major challenges and strains on the already-fragile resources in the coming years [54]. Despite these alarming predictions, mental health care systems in Arab countries continue to be centralized, hospital-based, mainly focused on secondary care and diseases treatment, thus neglecting the crucial role that SWB may play in alleviating mental health issues and promoting adaptive psychological outcomes [55]. Such strategies are inappropriate and ineffective for dealing with the highly challenging conditions and deteriorating mental health that most of the Arab general populations are facing. Therefore, using contextual and culturally sensitive prevention approaches focused on SWB becomes urgently needed in Arab countries.

Recently, a growing attention has been directed to the positive psychology field, and first local research initiatives aiming at promoting SWB have begun to emerge [56]. However, emerging studies are in no way comparable with non-Arab research in this field, both in terms of quality and quantity [56]. In addition, experimental research on SWB in the Arab region is still in its infancy and suffers from major methodological flaws [56]. We could find only little information available on SWB and a very few studies using the WHO-5 among Arab people in specific populations (e.g., Youth in Jordan [57], Saudi women [58], Emirati and other Arabic speaking adults [59], aid workers exposed to cumulative trauma in Palestine [60]). One of the main factors that hampers advances in mental health research and access to evidence-informed care in Arab countries is the lack of valid and reliable measurement tools [61]. Providing psychometrically sound measures of the SWB construct could aid in designing and implementing evidence-informed interventions to improve Arab people's well-being.

Rationale of the present study

The vast majority of previous validation and adaptation studies of the WHO-5 were performed in Western countries with individualistic backgrounds. However, SWB is a culturally-dependent and context-driven concept; it thus varies widely across-and within cultures, based on geographical situations [62, 63]. For instance, some findings indicated that individuals from collectivist cultures tend to exhibit lower ratings as compared to those from individualistic cultures, which may result in distinct levels of functioning of the WHO-5 items and overall measure [64]. Despite these data, the cross-cultural validity of well-being scales is still an unexplored question [65]. Some previous studies have investigated the cross-cultural validity of the WHO-5. For example, Carrozzino et al. [66] investigated the validity of the WHO-5 in a sample of 3762 adults from five European (i.e., Italy, Poland, Denmark) and non-European (i.e., China, Japan) countries. Sischka et al. [67] demonstrated that the WHO-5 is psychometrically appropriate and cross-culturally applicable in different nationally representative samples of individuals (N = 43,469) across 35 European countries. Another study also found that the WHO-5 showed good validity and reliability across Spain, Chile and Norway in nurses who worked during the COVID-19 pandemic [42]. Cross-cultural validation studies are crucial to prove that the measure covers transcultural components of the construct subjective, and can be used for cross-cultural comparison purposes in international multicenter research.

Although people from different Arab countries share similarities (including the language, geography, collectivist identity, religion, a young age structure [11, 40]), diversities do also exist. Large cross-cultural studies have shown that the way Arab people view and behave towards mental health issues is not uniform, and appears to be largely shaped by the local context of each Arab country [68, 69]. Taking into consideration these cultural disparities, it is necessary to examine whether the WHO-5 measures the SWB construct accurately in different Arab countries and cultural backgrounds. In this paper, we aimed to contribute to the literature on SWB in different ways. First, we propose to investigate, for the first time, the cross-cultural validity of the WHO-5 across different Arab
countries, to ensure its suitability to capture and provide reliable information on the SWB construct in different Arab contexts. Second, as the two previous validations were conducted in Arab Middle East countries, we intended to expand our investigation to an Arab region and countries (i.e., North Africa, Tunisia and Morocco) that have not been subject of previous validation studies of the WHO-5. Third, we sought to examine important psychometric properties that have not been previously examined, such as measurement invariance across genders. Indeed, gender differences in SWB are culturally-determined, as they may be substantially affected by social norms and adherence to traditional gender roles [70]. However, variations across genders may also be largely driven by methodological factors [71]. It is therefore required to verify that the WHO-5 invariantly measures the SWB factor across gender groups. Fourth, we aspired to include larger samples of participants than have been used in the past in order to provide stronger and more reliable results.

The objective of the present study was to perform a cross-country validation of the Arabic version of the WHO-5. In particular, we aimed to explore its (1) factor structure and composite reliability by country, (2) cross-gender measurement invariance, and (3) concurrent validity by calculating Pearson correlation coefficients between the WHO-5 and measures of depression, anxiety, stress, suicidal ideation and insomnia. We hypothesized that the Arabic version of the WHO-5 will show a single-factor structure and a satisfactory composite reliability in all samples from different countries, and will be invariant across gender groups. We also expected that concurrent validity of the Arabic WHO-5 will be supported through significant negative correlations with depression and other psychopathology measures.

2. Methods

2.1. Study design and Participants

This was a multi-country, web-based, cross-sectional study performed from February to June 2022. Several researchers from different institutions of the 22 Arab countries were invited to collaborate in our multinational project, and to join our team as co-investigators and co-authors. Researchers from six Arab countries accepted our invitation: Tunisia, Lebanon, Egypt, Jordan, Morocco and Kuwait. Arabic-speaking individuals from the general population, aged between 18 and 35 years, and residing in an Arab country during the study period, were considered eligible to participate. This age range was chosen to guarantee a homogeneous sampling and eliminate any differences resulting from age. Indeed, young adults aged 18–35 years are found to display a worse health profile than both adolescents and those in their late 30s [72]. The committee on Improving the Health, Safety, and Well-Being of Young Adults (convened by the National Research Council and the Institute of Medicine) concluded in their report that young adulthood is developmentally “of critical nature” within the life course [73]. Accordingly, the committee recommended that “outcomes are measured specifically for young adults”, and that young adults be treated as a distinct subpopulation in programming, planning, policy, and research [73]. Following these recommendations, we aimed to test the psychometric properties of the WHO-5 exclusively in young adults and a relatively narrow age range.

All participants fulfilling these criteria were sampled using convenience sampling technique and were invited to respond to a uniform anonymous web-based questionnaire through social media platforms (including Instagram, Facebook, and WhatsApp). Participants were also asked to forward the link to other eligible people they might know, using the snowball technique [74]. Snowballing techniques and online recruitment of non-help-seeking participants are typically adopted for research in this area (see, e.g., [75–77]). The questionnaire was administered using the free online survey tool provided by google forms. The study information and answering instructions were provided online via text; participants were asked to read them and give their informed consent before filling out the survey. Participants were free to accept or decline participation; no fee was given to any participant whatsoever. The study was performed following the Declaration of Helsinki for human research. The research protocol has been approved by the Psychiatric Hospital of the Cross ethics committee, Lebanon (Ref: HPC-012-2022). All collaborators who collected data were asked to follow the ethical guidelines of their Institutional Review Board (IRB), acting either on the ethical approval received from their local IRBs or that of the Principal Investigators.

2.2.1. The World Health Organization 5-item Well-Being Index (WHO-5). This instrument was developed in 1998 and has been translated into thirty different languages. WHO-5 consists of 5 items and assesses subjective psychological well-being. Each item is scored on a 5 point Likert scale with 5 = all of time to 0 = none of time. The total score is therefore ranging from 0 = absence of well-being to 25 = maximum well-being [78]. Raw scores are then multiplied by 4 to obtain a percentage score ranging from 0 (worst) to 100 (best). The Arabic version of this instrument was validated in Lebanon among elderly people [36].
2.2.2. Columbia Suicide Rating Scale. This scale is composed of 5 items, rated as no/yes type of answer. Higher scores indicate higher suicidal ideation. This scale has been validated in Lebanon among adolescents [79] and adults [80] (ω = .79).

2.2.3. Insomnia Severity Index. Validated in Lebanon [81], this scale is composed of 7 items, rated on a 4-point Likert scale. Higher scores reflect more severe insomnia (ω = .82).

2.2.4. Depression, Anxiety, and Stress Scale 8 items (DASS-8). Validated in Arabic [82], this scale is composed of 8 items rated on a 4-point Likert scale, which yields three scores: depression (ω = .91), anxiety (ω = .90) and stress (ω = .73).

2.2.5. Demographics. Participants were asked to provide their demographic details consisting of age, sex, and education level.

2.3. Analytic Strategy

2.3.1. Confirmatory factor analysis (CFA). There were no missing responses in the dataset. We used data from the total sample to conduct a CFA using the SPSS AMOS v.29 software. As a rule of thumb, simulation studies show that with normally distributed indicator variables and no missing data, a reasonable sample size for a simple confirmatory factor analysis model is about N = 150 [83], which was exceeded in our sample. Our intention was to test the original model of the WHO-5 scale (i.e., one-factor model). Parameter estimates were obtained using the maximum likelihood method and fit indices. For this purpose, the normed model chi-square (χ²/df), the Steiger-Lind root mean square error of approximation (RMSEA), the Tucker-Lewis Index (TLI) and the comparative fit index (CFI). Values ≤ 5 for χ²/df, and ≤ .08 for RMSEA, and .90 for CFI and TLI indicate good fit of the model to the data [84]. Additionally, evidence of convergent validity was assessed in this subsample using the Fornell-Larcker criterion, with average variance extracted (AVE) values of ≥ .50 considered adequate [85]. The absence of multicollinearity was verified through tolerance values > .2 and variance inflation factor (VIF) values < 5. Multivariate normality was not verified at first (Bollen-Stine bootstrap p = .002); therefore, we performed a non-parametric bootstrapping procedure (available in AMOS).

2.3.2. Gender invariance. To examine gender invariance of WHO-5 scores, we conducted multi-group CFA [86] using the total sample. Measurement invariance was assessed at the configural, metric, and scalar levels [87]. We accepted ΔCFI ≤ .010 and ΔRMSEA ≤ .015 or ΔSRMR ≤ .010 (.030 for factorial invariance) as evidence of invariance [86].

2.3.3. Further analyses. Composite reliability in both subsamples was assessed using McDonald’s ω, with values greater than .70 reflecting adequate composite reliability [88], because of known problems with the use of Cronbach’s α [89]. The WHO-5 total score was considered normally distributed since the skewness and kurtosis values varied between ± 1 [90]. To assess concurrent and divergent validity, we examined bivariate correlations between WHO-5 and the CSRS, ISI and DASS-8 scores using the Pearson test. Based on Cohen (1992) [91], values ≤ .10 were considered weak, ~ .30 were considered moderate, and ~ .50 were considered strong correlations.

3. Results

3.1. Participants

The total sample consisted of 3247 participants, with a mean age of 23.36 ± 4.62 years. The majority of the participants were females (71.6%), single (75.9%) and a university level of education (79.5%). The details of the sample by country are summarized in supplementary material (Table S1).

3.2. Confirmatory Factor Analysis of the Arabic WHO-5

CFA indicated that fit of the one-factor model of the Arabic WHO-5 was acceptable: χ² = 223.44, df = 5 (p < .001), RMSEA = .116 (90% CI .103, .129), SRMR = .017, CFI = .985, TLI = .970. When adding a correlation between residuals of items 1 and 4 (after showing a high modification index), the results improved more as follows: χ² = 57.71, df = 4 (p < .001), RMSEA = .064 (90% CI .050, .080), SRMR = .009, CFI = .996, TLI = .991. The standardised estimates of factor loadings (Fig. 1) and the AVE value (~ 1.03) were all excellent. The same analysis was conducted per country and showed adequate results as well (Table 1).
### 3.4. Composite reliability

Composite reliability of the WHO-5 scores was adequate in the total sample ($\omega = .94$), in men ($\omega = .95$), and women ($\omega = .94$), and in each country as follows: Tunisia ($\omega = .96$), Lebanon ($\omega = .95$), Kuwait ($\omega = .94$), Egypt ($\omega = .92$), Jordan ($\omega = .93$) and Morocco ($\omega = .94$).

### 3.3. Measurement Invariance

As reported in Table 2, indices suggested that configural, metric, and scalar invariance was supported across gender in the total sample and by country (Table 3). The results showed that a significantly higher mean WHO-5 score was found in males compared to females ($10.30 \pm 6.65$ vs $8.73 \pm 6.17$; $t = 6.37; p < .001$). Moreover, the highest mean well-being score was found in Morocco compared to all other countries (Table 3); the Bonferroni post-hoc analysis showed a significant difference between Tunisia and Lebanon ($p < .001$), Tunisia and Kuwait ($p < .001$), Tunisia and Morocco ($p < .001$), Lebanon and Egypt ($p < .001$), Kuwait and Egypt ($p < .001$), Egypt and Morocco ($p < .001$), and Jordan and Morocco ($p = .027$).

#### Table 1

<table>
<thead>
<tr>
<th>Country</th>
<th>SRMR</th>
<th>$\chi^2$</th>
<th>df</th>
<th>TLI</th>
<th>CFI</th>
<th>RMSEA [90% CI]</th>
<th>Standardized loading factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tunisia</strong></td>
<td>.011</td>
<td>34.73</td>
<td>4</td>
<td>.980</td>
<td>.992</td>
<td>.112 [.079-.147]</td>
<td>Item 1 .88</td>
</tr>
<tr>
<td><strong>Lebanon</strong></td>
<td>.008</td>
<td>13.63</td>
<td>4</td>
<td>.992</td>
<td>.997</td>
<td>.063 [.028-.101]</td>
<td>Item 2 .92</td>
</tr>
<tr>
<td><strong>Kuwait</strong></td>
<td>.013</td>
<td>20.55</td>
<td>4</td>
<td>.985</td>
<td>.994</td>
<td>.079 [.047-.114]</td>
<td>Item 3 .93</td>
</tr>
<tr>
<td><strong>Egypt</strong></td>
<td>.010</td>
<td>13.40</td>
<td>4</td>
<td>.992</td>
<td>.997</td>
<td>.051 [.023-.083]</td>
<td>Item 4 .94</td>
</tr>
<tr>
<td><strong>Jordan</strong></td>
<td>.016</td>
<td>15.82</td>
<td>4</td>
<td>.971</td>
<td>.989</td>
<td>.107 [.055-.165]</td>
<td>Item 5 .93</td>
</tr>
<tr>
<td><strong>Morocco</strong></td>
<td>.007</td>
<td>2.47</td>
<td>4</td>
<td>1.004</td>
<td>1.000</td>
<td>.001 [.001-.085]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 1: Invariance by gender</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model Comparison</strong></td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Configural vs metric</td>
</tr>
<tr>
<td>Metric vs scalar</td>
</tr>
</tbody>
</table>

Note. CFI = Comparative fit index; RMSEA = Steiger-Lind root mean square error of approximation; SRMR = Standardised root mean square residual.
Table 3
Comparison between countries in terms of the WHO-5 scores.

<table>
<thead>
<tr>
<th>Country</th>
<th>Mean ± SD</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>15.96</td>
<td>&lt; .001</td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>33.6 ± 25.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>39.8 ± 25.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuwait</td>
<td>40.4 ± 27.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>32.2 ± 22.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jordan</td>
<td>36.84 ± 24.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>44.2 ± 26.84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers in bold indicate significant p values.

3.4. Concurrent validity (total sample)

As for concurrent validity, WHO-5 scores showed a moderate significant inverse correlation with DASS depression (r = −.28; p < .001), anxiety (r = −.29; p < .001), and stress (r = −.27; p < .001) subscales scores, suicidal ideation (r = −.16; p < .001) and insomnia severity (r = −.37; p < .001).

4. Discussion

This study is the first to explore the cross-country validity of the WHO-5 among young adults across six Arab countries (i.e., Tunisia, Lebanon, Egypt, Jordan, Morocco and Kuwait) of the Middle East and North Africa (MENA) region. Findings supported that all 5 items were loaded into a single underlying factor in all six countries, which showed adequate reliability (McDonald’s \( \omega = .94 \)), as well as good convergent and divergent validity. Measurement invariance across gender groups was also evidenced at the configural, metric, and scalar levels. Overall, these findings suggest that the Arabic WHO-5 measures the originally intended SWB construct in the specific context conditions of Arab general populations. By verifying the cross-country validity of the Arabic WHO-5, our study opens its wider application to epidemiologically explore SWB among Arabic-speaking community young adults from broad geographic areas.

We found that WHO-5 mean scores varied significantly across countries, ranging from 32.2 ± 22.72 in Egypt to 44.2 ± 26.84 in Morocco. Despite these wide variations, WHO-5 scores reported in all six Arab countries are much lower than those observed among the adult general population in other international studies (e.g., 56 in Sri Lanka [24], 64.74 in Iceland [17], 73.37 in southern Brazil [25]), but comparable to that found in a Middle Eastern country (i.e., 35.8 in Iranian people [26]). These high mean scores should sound warning bells for clinicians, researchers and policy-makers working in Arab settings; and further highlights that local culturally-sensitive strategies are urgently needed to address well-being issues among Arab young adults.

The construct validity of the WHO-5 was examined using CFA, which is consistently advocated by validation researchers as a crucial step in scale validation [92, 93]. Indeed, unlike exploratory factor analysis (EFA), it imposes meaningful constraints in evaluating a measure validity [17]. Although the WHO-5 has been extensively validated in dozens of languages and countries, scant WHO-5 assessments used CFA [34, 94], and several validation studies only relied on EFA [18, 32, 38, 95–97]. Analyses showed that the fit of a one-factor model to the data proved to be acceptable in each of the six countries, thus replicating the factor structure of the original WHO-5 [15, 16, 98], and that obtained in other linguistic versions using CFA (e.g., Swahili Kenyan [35], Malay [28], Icelandic [17], Sinhala [24], Chinese [34], and Arabic [37]). Our results suggest that SWB of non-clinical young adults in Arab countries may be explained by a unique underlying trait likewise young adults of other countries, and further support the applicability of the WHO-5 as a unidimensional measure of SWB within Arab settings and contexts. Furthermore, to assess composite reliability in our sample, McDonald’s \( \omega \) coefficients were used as they have proven to provide more realistic estimates of a measure’s reliability than Cronbach’s Alpha [99]. Findings revealed that the Arabic WHO-5 yielded high reliability coefficients in the total sample and both genders, which is in line with previous international studies on other translations of the WHO-5 that mostly relied on Cronbach’s alpha.
coefficients (e.g., alpha values of 0.79–0.91 in Italy [21], 0.86–0.88 in Kenya [35], 0.87 in Poland [20], 0.88 in Romania [22], 0.83 in Sweden [18], 0.91 in Iran [26], 0.81 in Turkey [27], 0.91 in Malaysia [28], 0.75 in Bangladesh [31], 0.83 in Brazil [25], 0.82–0.87 in Iceland [17], 0.81–0.85 in China [34], 0.91 in Saudi Arabia [37]).

Another relevant contribution of this study was to examine measurement invariance of the WHO-5 scores across gender. Results from multi-group analyses demonstrated that configural, metric, and scalar invariance was supported across gender in the total sample and by country. Evidence of invariance across gender groups have also been reported in other validation studies and different linguistic contexts (e.g., Icelandic national and patient samples [17], Sri Lankan people from the general population [24]). Providing empirical evidence that the Arabic WHO-5 is invariant between males and females implies that mean gender differences in SWB are not attributable to group-level variations in understanding or responding to items, but to real differences in the construct level [100]. We therefore suggest that the Arabic WHO-5 can be used reliably to compare mean differences between gender groups. In this regard, we found that males displayed higher WHO-5 scores than females, which is in accordance with previous studies (e.g., [17, 18, 21, 22, 30]).

As for concurrent validity, WHO-5 scores showed a strong significant inverse correlation with DASS depression sub-scores, which is consistent with several previous validations using various depression measures (e.g., [17, 20, 24, 26, 29, 101]). In the same line, the WHO-5 has consistently shown high sensitivity and specificity in detecting depression, and was extensively applied as a screening tool for this condition [11]. Negative correlations have also been demonstrated between WHO-5 scores and different symptoms of mental health problems (anxiety, stress, suicidal ideation, insomnia) in our sample. These findings also concur with prior research comparing the correlations between the WHO-5 and different measures of mental health problems (e.g., anxiety [17, 32], stress [17, 31], sleep problems [18]). In sum, the current results strongly support the validity of the Arabic WHO-5, and offer additional confirmation that it serves the purposes for which it was originally developed.

**Study limitations and research perspectives**

Despite its valuable contribution to the field of SWB, the present study has some limitations that need to be addressed in future research. First, we did not use a structured clinical interview against which the results from the self-report measure can be validated, which prevented us from assessing specificity and sensitivity of the WHO-5 as a depression screening-tool. To address this limitation, future studies need to have an external criterion measure. Second, our data gathered at a single point in time, which precluded us from testing the stability of the Arabic WHO-5 over time. Therefore, additional validation studies need to examine the test–retest reliability of the scale. Third, the use of a community sample of young adults may undermine the generalizability of our findings to clinical populations. Fourth, we adopted an online survey and a convenience sampling, which both mostly attracted highly-educated and female participants and limited the representativeness of our sample. Finally, the six countries involved in our study are lower-to-middle income countries, and cannot be considered representative of all Arab populations and the Arab world. Young adults from low-income Arab countries (such as Yemen and Syria) may have different SWB, and should be subject of future validation studies.

**5. Conclusion**

In summary, findings indicate that the WHO-5 in its Arabic version has a unidimensional structure among Arabic-speaking young adults across six Arab countries, high internal consistency, good concurrent validity, and measurement invariance across gender. This study contributes to the available amount of research in the field of positive psychology and well-being by providing a brief, valid and reliable Arabic version of the WHO-5 that can be used cross-culturally with a variety of Arabic-speaking young adult populations for screening and research purposes.

**Declarations**

**Ethics Approval and Consent to Participate**

Ethics approval for this study was obtained from the Psychiatric Hospital of the Cross ethics committee, Lebanon (Ref: HPC-012-2022). Written informed consent was obtained from all subjects; the online submission of the soft copy was considered equivalent to receiving a written informed consent. All methods were performed in accordance with the relevant guidelines and regulations.

Consent for publication: Not applicable.
Availability of data and materials: All data generated or analyzed during this study are not publicly available due the restrictions from the ethics committee, but are available upon a reasonable request from the corresponding authors (FFR & SH).

Competing interests: The authors have nothing to disclose.

Funding: None.

Author contributions:

FFR and SH designed the study; FFR and SH wrote the paper; SH carried out the analysis and interpreted the results; WC, AA, MF, HAMS, MH, IHMH, AYN, BZ, MC, YE-F, GY, GS, AH-M, ER, AH, and SO were involved in the data collection; MC reviewed the paper; all authors read and approved the final manuscript.

Acknowledgements: The authors would like to thank all participants.

References


89. McNeish D: Thanks coefficient alpha, we'll take it from here. *Psychological methods* 2018, 23(3):412.


**Figures**
Figure 1

Standardized Factor Loadings Derived from the Confirmatory Factor Analysis of the Arabic WHO-5 in the total sample.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterial.docx