Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Measurement invariance of the Arabic version of the WHOQOL-BREF among special education teachers

Mohamed Mohamedalhadi Hassan Suliman [1], Bander M. Almohayya [2,*]

[1]Psychology Department, College of Education, University of Hail, Hail, Saudi Arabia
[2]Department of Special Education, College of Education, University of Hail, Hail, Saudi Arabia

ARTICLE INFO

Article history:
Received 28 January 2025
Received in revised form
20 May 2025
Accepted 16 July 2025

Keywords:
Quality of life
Arabic validation
WHOQOL-BREF
Psychometric analysis
Special education

ABSTRACT

This study aimed to validate the Arabic version of the World Health Organization Quality of Life: Brief Version (WHOQOL-BREF) using a convenience sample of 232 special education teachers. The scale's psychometric properties were assessed through exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and measurement invariance testing. EFA supported a four-factor structure explaining 55.87% of the variance, and the scale demonstrated strong internal consistency (Cronbach's alpha = .923; McDonald's omega = .921). Convergent validity was confirmed by a significant positive correlation with the Quality of Life Scale (QOLS), while discriminant validity was supported by a significant negative correlation with the Health and Suffering Scale (HSS; r = –0.539, p = .001). Additionally, the QOLS was negatively correlated with the HSS (r = –0.589, p = .001), reinforcing the validity of the chosen quality-of-life measures. CFA confirmed the four-factor model, and multi-group CFA showed measurement invariance across gender but not across educational levels. Model fit was slightly weaker for the bachelor's group compared to the master's and Ph.D. groups. The study concludes with a discussion of its limitations and implications.

1. Introduction

A line of research that is emphasized in the social sciences, medicine, and economics is the evaluation of quality of life (QoL; Cummins and Lau, 2006), and multidisciplinary research on QoL has attracted a lot of attention in recent years (Kalfoss et al., 2021). These research findings have been used to raise standards and improve outcomes for treatments and interventions for medical conditions, chronic diseases, and disabilities over the last few decades (Burgess and Gutstein, 2007; Barneveld et al., 2014; Oliveira et al., 2016; Sosnowski et al., 2017; Lima-Castro et al., 2021). However, research on QoL is challenging because of the varied conceptions of this construct, the variety of dimensions utilized in its measurement, and all the factors that affect QoL (Lima-Castro et al., 2021).

In an effort to provide a common definition, the World Health Organization (WHO) put forth a definition that is based on cultural and translation research conducted across various cultural boundaries with the goal of combining criteria, defining QoL as follows: "The perception that an individual has of his situation in life within the cultural context and value system in which he lives and in relation to his objectives, expectations, norms, and interests" (WHOQOL Group, 1998). This description highlights the multifaceted, cultural, and subjective nature of QoL. The World Health Organization Quality of Life: Brief Version (WHOQOL-BREF) questionnaire was created using this definition as a criterion. The WHOQOL-BREF is a widely used generic tool for assessing life quality and is notable for its rigorous and meticulous translation and intercultural adaptation procedure, which enables the extraction of trustworthy results that are reflected in its equivalence across 15 different cultures (Bowden and Fox-Rushby, 2003; Crocker et al., 2015; Lima-Castro et al., 2021).

One of the most used general questionnaires for assessing quality of life in both healthy and unwell groups is the WHOQOL-BREF (Krägeloh et al., 2013; Skevington et al., 2004; Kalfoss et al., 2021). For more than two decades, the WHOQOL-BREF has

* Corresponding Author.
Email Address: banderalmohayya2024@gmail.com (B. M. Almohayya)
https://doi.org/10.21833/ijaas.2025.08.012
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-2798-8912

been translated into many languages across the world, with more than 60,000 adults, both healthy and unhealthy, having completed it (Kalfoss et al., 2021). Furthermore, this scale has distinguished itself in comparison to other used generic QoL measures in clinical, case-control, and cross-sectional studies because it allows for data collection on 26 items categorized into four domains: environment, social relationships, physical health, and psychological, which are frequently assessed in clinical fields. The application of the WHOQOL-BREF questionnaire extends beyond the field of public health (Skevington and Epton, 2018), with excellent consistency, reliability, and construct validity demonstrated in cross-cultural studies among both the general population and various patient groups characterized by diseases that result in various disabilities (Skevington et al., 2004; Perera et al., 2018), as well as reports of factor invariance across gender (Perera et al., 2018).

Although the WHOQOL-BREF is still used extensively around the world, a structural evaluation is necessary because of inconsistent findings. Based on exploratory and confirmatory factor analysis, a four-domain model has been suggested (CFA) of a sample of 1,068 healthy and unhealthy individuals in Taiwan (Yao et al., 2002). Similarly, there is evidence supporting the original WHO four-domain model. For example, a recent study conducted in Singapore with 3,400 adults of various ages and with various health problems found well fit of the four domains (Suárez et al., 2018). However, the original WHOQOL-BREF structure has not been replicated in subsequent replication research (Benítez-Borrego et al., 2014; Moreno et al., 2006; Ohaeri et al., 2007; Oliveira et al., 2016).

In summary, given the linguistic and cultural differences that can be identified even when the language is the same (Hambleton et al., 2004), it is essential to make the structural factors of the WHOQOL-BREF questionnaire clear in a variety of contexts (Ohaeri et al., 2007; Perera et al., 2018; Snell et al., 2016). Furthermore, it is important to understand that many different factors, in both the general population and clinical groups, are related to QoL. For example, some previous studies have suggested that improving the QoL of older adults and people with depression and/or chronic illnesses is strongly affected by their personal sense of self-worth (Tavares et al., 2016).

Despite the importance of quality of life as a study topic, the psychometric properties of the WHOQOL-BREF questionnaire in Arabic in the Saudi Arabian context have not been studied before. Therefore, the purpose of this study is to assess the consistency, reliability, and construct validity, measurement invariance, and structural validity of the WHOQOL-BREF using a sample of the Arabic-speaking population.

Even yet, research on the psychometric properties of the WHOQOL-BREF indicates that the scale's validity and reliability are generally acceptable (Naumann and Byrne, 2004; Kalfoss et al., 2021); there has been some variation in the factor structure and item cross-loading reported across validation studies, raising concerns about the generalizability of the factor structure across demographics (Lima-Castro et al., 2021). The findings of some studies contradict the scale's proposed four-factor dimensionality of the scale without prior alterations to the instrument, and on occasion, it is evident that the reliability of the social and environmental domains is poor (Naumann and Byrne, 2004; Kalfoss et al., 2021). Furthermore, measurement invariance across gender, but not age, is reported in one study (Perera et al., 2018), and some studies (Yao and Wu, 2005; Lin et al., 2016), but not others (Theuns et al., 2010), report support for construct validity in this regard, as cited in Kalfoss et al. (2021).

Concerns about whether the data the WHOQOL-BREF generates are well presented by the proposed four-factor structure and whether the WHOQOL-BREF indicates the same structure across various populations persist despite the instrument's widespread use and the evidence demonstrating its psychometric soundness (Kalfoss et al., 2021). Based on research findings by Noerholm et al. (2004), Benítez-Borrego et al. (2014), and Wang et al. (2023), which are all cited by Kalfoss et al. (2021), the four-factor structure fits general populations well. According to Pomeroy et al. (2013) and Rocha et al. (2012), rescoring or removing items may be necessary to achieve appropriate fit indices for the four-factor model. Items may also have significant correlations with different domains; however, they are shown to correlate most strongly with their supposedly intended domain (Kalfoss et al., 2021). Indeed, reports of altered versions of the WHOQOL-BREF's four-factor structure are frequently published. (Pomeroy et al., 2013, as cited in Kalfoss et al., 2021), and the findings of some studies also support a one-factor structure as a solution (Bech, 2001). Similarly, Oliveira et al. (2016) found that the best model fit, along with increased reliability, in a clinical population came from a five-domain model that divided daily activities and capacity to work into domains, denominating the levels of independence possessed by individuals. Because of these differences within and between populations, the psychometric properties of the WHOQOL-BREF need to be continually evaluated to improve QoL instrumentation in various populations (Kalfoss et al., 2021).

Consequently, in this research, we use a study of Arabic special education teachers to assess the psychometric properties of the Arabic WHOQOL-BREF questionnaire. In addition to extending the scope of earlier studies on psychometric properties, our goal is to test for measurement invariance across gender and level of education. The Arabic version of the WHOQOL-BREF has been evaluated using CFA, with a sample of 623 subjects recruited from the general population, and a five-domain model was found to be the best fit (Ohaeri et al., 2007). To the best of our knowledge, convergent validity,

discriminant validity, and measurement invariance have not been established for the Arabic version of the WHOQOL-BREF. An essential step in assessing the construct and structural validity of the scales of a psychometric tool is to test for factorial invariance (Dimitrov, 2010). Consistency of the theoretical structure of the underlying constructs across different groups of people provides evidence of factorial invariance (Dimitrov, 2010). Because measurement invariance has not been established for the Arabic version of the WHOQOL-BREF, we evaluate the structural and factorial invariance of the instrument using a study sample of special education teachers. Teachers who teach the typical population and special education educators have not been recruited or included in any WHOQOL-BREF scale in the Arabic or international contexts. Therefore, we believe that this is the first study to evaluate the psychometric properties of the structure of the WHOQOL-BREF scale and the first to recruit them to measure the levels of quality of life of special education teachers. This study also contributes to the field of psychology by determining the psychometric properties of the Arabic version of the WHOQOL-BREF among special education teachers and by, in fact, using sophisticated analyses; EFA, CFA, and measurement invariance. Our objective is to evaluate the structure of the scale using a specific population to investigate its measurement invariance using advanced analysis.

2. Methods

A quantitative research approach was adopted for this cross-sectional study, and ethics approval was obtained from the Scientific Research Ethics Committee of the University of Hail.

2.1. Participants and procedure

The study participants were selected from among the full-time special education teachers working in the Hail metropolitan area, in Saudi Arabia. A non-probability sampling method was used because it is the easiest for the researcher to access. This was due to the availability and the time given to complete the study. Therefore, an online survey created using Google Forms was disseminated to potential participants via an email containing a hyperlink. Arabic-speaking teachers with special needs who had been enrolled in inclusive classroom programs at general education schools and were teaching in primary schools at the time were sent an invitation and recruited to the study.

Before filling out the study questionnaire, potential participants gave their digital informed consent. The first items of the study questionnaire confirmed whether participants met the inclusion criteria, i.e., being at least 21 years of age and currently teaching. No participant was identified as not meeting the inclusion criteria, and no exclusion criteria were adopted.

2.2. Measures

2.2.1. World Health Organization quality of life questionnaire

The WHOQOL-BREF comprises two distinct items that measure overall QoL and health satisfaction, as well as one item from each of the 24 facets of the WHOQOL-100 (WHOQOL Group, 1998). Together with the overall QoL and health satisfaction metric, these 26 questions form four unique domains of QoL: Physical health, psychological, social relationships, and environmental. Higher scores correspond to a higher QoL. Each item is scored on a Likert scale ranging from 1 to 5, with variable scale answer anchors. “How much do you enjoy life?” is one example, which can be rated using the options that appear below the question. Not at all, a little, a moderate amount, a lot, and an extreme amount are the possible answers. The time frame of the assessed QoL status is the last two weeks (Kalfoss et al., 2021).

2.2.2. Quality of life scale (QOLS)

American psychologist John Flanagan originally developed the QOLS in the United States in the mid-1970s, and it has since been revised for use in populations with chronic illnesses. In developing this tool, approximately 3,000 individuals of a wide range of ages, ethnicities, and cultures from all over the United States of America were asked to share events that meant something to them or that they found fulfilling using the critical incident technique. Considerable effort was made to incorporate seniors, rural dwellers, low-income populations, and individuals from ethnic minorities. Flanagan stated, “The purpose of using the regional samples and diverse groups was not to obtain accurate estimates of frequencies but rather to ensure that differing points of view and types of experience were represented” (Flanagan, 1978). Except for the Cantril ladder (Cantril, 1965), no other QoL instrument in use today was created with as much consideration for individual perspective and diversity as the QOLS. Five conceptual categories of QoL are represented by 15 elements in the original QOLS. For the scaling of the items, Flanagan (1978) utilized two five-point scales in his original structure. The seven answers were also utilized. The five-point importance scale and the seven-point delighted–terrible scale were utilized in subsequent research conducted to modify the QOLS for use among chronically ill American populations (Burckhardt et al., 1989; Burckhardt and Anderson, 2003).

2.2.3. The health and suffering scale (HSS)

The development of the Health and Suffering Scale (HSS) has both an empirical and theoretical basis (Andermo et al., 2018). This 20-item self-report measure uses a semantic visual analog scale

to assess subjective suffering related to health. Based on Erikson et al.'s (1986) hypothesis, perceived suffering is measured on a visual analog scale by marking a point between word pairs that reflect health and suffering, such as "life without meaning – meaningful life" and "lost grip on life–understanding about life." "Barriers to health–Health" and "Unbearable suffering–bearable suffering" are two of the 20 items on the HSS that specifically address the concepts of health and suffering. Initially, the remaining 18 items were associated with five sub-domains of health and suffering: Relationships, personal freedom, meaning, and presence in life. None of the items is reverse scored (Gebhardt et al., 2022).

2.3. Statistical analysis

The means and standard deviations of the data collected on the entire sample (n = 232) and the frequencies and percentages of all the variables were computed using descriptive statistics analysis. To determine construct validity, we used principal component analysis (PCA) and, subsequently, CFA to evaluate the fit of the WHOQOL-BREF domain model. Using PCA with varimax rotation (n = 232), we first investigated the model fit to evaluate the structural validity of the instrument. The suitability of the sampling and the correlation structure were evaluated using Bartlett's test of sphericity and the Kaiser–Meyer–Olkin test. The resulting model was analyzed in terms of its initial eigenvalues, communalities, the cross-loadings of each item, and the total of the squared loadings and variance explained by each factor after rotation. Communalities with a value of less than 0.40, loadings with a value of less than 0.32, and cross-loadings with a value of more than 0.32 were deemed problematic, as recommended by Costello and Osborne (2005).

McDonald's omega and Cronbach's alpha were used to evaluate internal consistency. The Pearson correlations between the WHOQOL-BREF and other QoL measures (QOLS, HSS) were used to evaluate the convergent and discriminant validity of the relationships. In addition, Pearson's correlation coefficient analysis was used to assess the convergent and discriminant validity of the WHOQOL-BREF by analyzing the interrelation between its four dimensions. The WHOQOL-BREF and the QOLS were used to test convergent validity, while the WHOQOL-BREF and the HSS were used to test discriminant validity.

Following the PCA, a CFA of the determined factor structure of the sample (n = 232) was conducted. A variety of techniques were used to assess the goodness of fit: The normed chi-square (χ^2/df), where values below 2, or in more generous recommendations, values below 5, have been proposed as acceptable; the comparative fit index (CFI), where values ≥ .95 indicate a good fit; the root mean square error of approximation (RMSEA), where values ≤ .06 indicate a good fit, and values ≤ .08 may indicate an acceptable fit—especially if the upper limit of the 90% confidence interval falls below this threshold; and finally, the standardized root mean square residual (SRMR), where values 0.08 indicate a good fit (Brown, 2015). Using CFA, were to confirm the results of the EFA, and to compare the model of the scale and its goodness-of-fit to that of competing models.

The internal consistency of the original and proposed domains was evaluated using Cronbach's alpha, McDonald's Omega, and item-domain correlations. Item-domain correlations of 0.30 or higher and a Cronbach's alpha of 0.70 or higher were deemed adequate. As a preliminary test of construct validity, we determined the Pearson correlation (one-tailed) between the QOLS and the HSS, and between the original and the suggested measures.

Based on published research on establishing the measurement invariance of models (Lin et al., 2016), the measurement invariance of the WHOQOL-BREF's scale items across gender and educational level was tested. To check for measurement parameter invariance, we performed hierarchical tests, and to determine whether the population variance matrices were alike, we first conducted a hypothesis test—an RMSEA score below the lowest cutoff would be consistent with the overall invariance of the instrument. We then examined the configuration invariance model (also known as pattern invariance), which in this study does not impose any equality limitations on the model parameters. This condition is required to test for invariance by comparing the configuration invariance model with alternative invariance models based on fit indices. Second, we explored metric invariance, often known as the weak invariance model. The factor loadings are regarded as gender-neutral in this model. This ensures that, for meaningful comparisons, the measures are regarded as being on the same scale across genders. Third, we investigated the strong invariance model. Item intercepts and factor loadings are both subject to gender invariance under this model, which facilitates cross-gender comparisons of the underlying factor. Fourth, we examined the strict invariance model, which requires the invariance of the residual variances, intercepts, and factor loadings.

Evidence of invariance between the less restrictive model (e.g., the configural invariance model) and the more restrictive model (e.g., weak measurement invariance models) was based on recommendations from the literature (Cheung and Rensvold, 2002). If the value of the change in CFI (ΔCFI) is less than or equal to 0.01, we do not reject the invariance hypothesis. The crucial values for the change in the Tucker–Lewis index (ΔTLI) and ΔRMSEA are 0.015 and 0.01, respectively. For every comparison, the chi-square difference test was also performed. To identify any problematic items that may contribute to a misfit with the data, factor loadings of 0.40 and above were used, along with significant p-values, standardized residuals, and modification indices (MIs).

We considered the meaningfulness of including the covariances among the identified items. Therefore, residual covariances among the items were included as an additional parameter based on MI values. Except for the CFA and measurement invariance, we used SPSS (version 27), AMOS (version 25), and JASP for all statistical analyses.

3. Results

3.1. Descriptive statistics and reliability of the WHOQOL-BREF

Table 1 presents the mean scores and standard deviations. A total of 232 participants took part in this study. Of these, 31.9% were female (n = 74) and 68.1% were male (n = 158). In terms of age, 25.9% were between 21 and 30 years old, 37.5% were between 31 and 40, and 36.7% were 41 years or older. Regarding work experience, 50% had between 1 and 10 years, 29.3% had between 11 and 20 years, 15.9% had between 21 and 30 years, and 4.9% had 31 years or more. In terms of education, 77.2% of participants held a bachelor's degree, 15.9% had a master's degree, and 6.9% had a Ph.D. The approximation rate to the maximum score was relatively high at 74.4%, suggesting a generally positive QoL. The reliability of the WHOQOL-BREF was assessed using 24 items. The results showed strong internal consistency, with a Cronbach's alpha of .923 and a McDonald's omega of .921.

3.2. Structural validity

The sample adequacy for the PCA was confirmed using the Kaiser–Meyer–Olkin value (0.917), and the adequacy of the correlation structure was revealed by the Bartlett's test of sphericity, yielding a significant result ($p < .001$). The subject-to-item ratio was 9:1, which is within the recommended sample size requirements (Costello and Osborne, 2005) and greater than the minimum subject-to-item ratio of 5:1 for PCA. Inspection of the eigenvalues revealed four factors with initial eigenvalues greater than one (Fig. 1).

Due to the rotation, the four-factor solution explains 55.87% of the total variance. The first factor comprises 10 related items (i.e., Items 5, 6, 7, 8, 9, 10, 11, 16, 17, and 26) and accounts for 37.72% of the variance. Six items comprise the second component (i.e., Items 12, 13, 14, 15, 24, and 25), which explains 7.41% of the variation. The third factor comprises six items (i.e., Items 18, 19, 20, 21, 22, and 23), which explain 5.96% of the variance. The final factor comprises two items (i.e., Items 3 and 4), which explain 4.77% of the variance. All items loaded significantly on only their own factors, thus demonstrating good factor discriminability.

Table 1: Characteristics of study participants (n = 232)

Descriptive statistics	Frequency	%
Age		
21–30	60	25.9%
31–40	87	37.5%
41 and older	85	36.7%
Gender		
Male	158	68.1%
Female	74	31.9%
Years of work experience		
1–10 years	116	50.0%
11–20	68	29.3%
21 and more	48	20.6%
Educational level		
Bachelor	179	77.2
Master's and Ph.D.	53	22.8

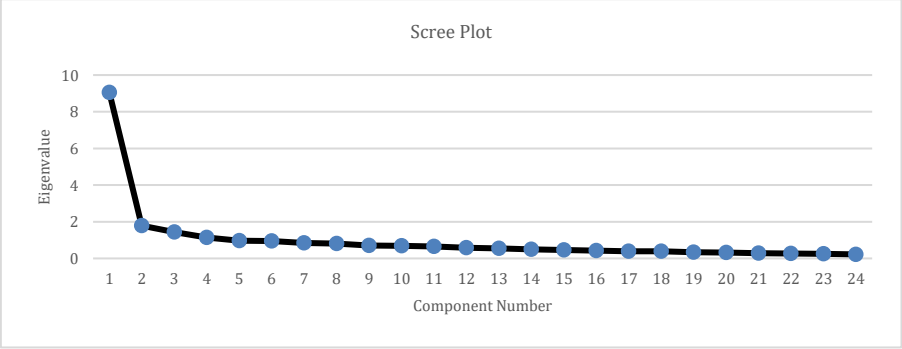

Fig. 1: Eigenvalues of items

The factors that were found to differ from the original domains of the WHOQOL-BREF. However, the factor loadings indicate that three of the items loaded on the first factor of the Arabic version of the WHOQOL-BREF, similar to the physical domain in the original scale (i.e., Items 10, 16, and 17). However, an additional five items loaded on the same first factor of the Arabic version of the WHOQOL-BREF scale: Items 5, 6, 7, 8, 9, and 26. All these new items had a factor loading of > .40 in their respective new domains. Only one item (Item 16) showed low communality (< .40), and all nine items

(Items 5, 6, 7, 8, 9, 10, 11, 17, and 26) had no cross-loadings. Half of the items were originally operationalized under the psychological domain in the original WHOQOL-BREF (i.e., Items 5, 6, 7, 11, and 26), while the other three were originally operationalized under the physical domain (i.e., Items 10, 16, and 17). In contrast to the original scale, only two items (i.e., Items 8 and 9) loaded strongly on the first factor of the Arabic version of the scale (> .60). Nonetheless, the first principal component was identified as the psychological component because the loadings for the psychological items were higher.

However, the second factor included five items of the environment factor of the original scale, and the six new items (i.e., Items 12, 13, 14, 24, and 25) loaded on the second factor of the Arabic version of the WHOQOL-BREF. However, in contrast to the original WHOQOL-BREF, three items (Items 8, 9, and 23) did not load on the Arabic version for this second factor. In addition, Item 15 loaded on this factor, although it belongs to the physical factor in the original scale. However, this is justifiable because this item is structured and related to mobility and getting around physically. Therefore, we retained the item under this third factor and named it environment, as it is named in the original.

For the third factor (i.e., Items 18, 19, 20, 21, 22, and 23), its items loaded similarly to the social relationships factor in the original scale. However, three new items loaded on this third factor of the Arabic WHOQOL-BREF questionnaire were originally operationalized under different factors of the original scale. Therefore, Items 18, 19, and 23 were newly loaded on this factor and seem to pertain to satisfaction with work, self, and living space, three concepts that involve social relationships in some way. We retained the third factor, as it loaded with its three new items.

Finally, and unexpectedly, the fourth factor of the Arabic version of the WHOQOL-BREF included only two items (Items 3 and 4). These two items loaded on the fourth factor were new and included two very high loadings (0.84 and 0.83 for Q3 [Item 3] and Q4 [Item 4], respectively). These two factors relate to physical pain and the need for medical treatment; hence, we named it the physical factor. This was retained, although the factor was limited to loading only two items (Table 2).

The WHOQOL-BREF item factor loadings in PCA were performed using varimax rotation. The factor structure, along with the total squared loadings added and an explanation of the variation of each factor, is presented in Table 2.

Table 2: PCA factor loadings and communalities (h^2) for the Arabic WHOQOL-BREF (n = 232)

Item	Factor 1 (Psychological)	Factor 2 (Environment)	Factor 3 (Social)	Factor 4 (Physical)	Original domain	New domain	Mean	SD	h^2
Item 5	0.757				Psychological	Psychological	3.79	0.817	0.690
Item 6	0.703				Psychological	Psychological	4.11	0.814	0.547
Item 7	0.594				Psychological	Psychological	3.56	0.776	0.520
Item 8	0.673				Environment	Psychological	4.03	0.894	0.600
Item 9	0.618				Environment	Psychological	3.42	0.908	0.526
Item 10	0.567				Physical	Psychological	3.72	0.874	0.536
Item 11	0.623				Psychological	Psychological	4.16	0.884	0.547
Item 16	0.509				Physical	Psychological	3.41	0.989	0.365
Item 17	0.543				Physical	Psychological	3.71	0.838	0.583
Item 26	0.597				Psychological	Psychological	3.06	0.866	0.420
Item 12		0.595			Environment	Environment	3.17	0.898	0.432
Item 13		0.491			Environment	Environment	3.61	0.754	0.367
Item 14		0.632			Environment	Environment	2.97	0.962	0.500
Item 15		0.655			Physical	Environment	4.03	0.837	0.538
Item 24		0.728			Environment	Environment	4.03	0.885	0.587
Item 25		0.774			Environment	Environment	3.06	0.866	0.635
Item 18			0.593		Physical	Social	4.01	0.770	0.628
Item 19			0.519		Psychological	Social	4.19	0.778	0.614
Item 20			0.671		Social	Social	4.04	0.857	0.595
Item 21			0.586		Social	Social	3.68	0.986	0.469
Item 22			0.737		Social	Social	3.70	0.909	0.668
Item 23			0.560		Environment	Social	3.87	0.880	0.514
Item 3				0.849	Physical	Physical	4.25	0.919	0.777
Item 4				0.838	Physical	Physical	3.79	0.817	0.690

h^2 (Communality): Proportion of item variance explained by all extracted factors; SD: Standard deviation

3.2.1. Confirmatory factor analysis (CFA)

No outliers were found in the results when the fourth factor solution from the PCA was cross-validated with the entire sample (n = 232), and all items were checked for conformity with the linearity and multivariate normality assumptions. For the CFA, we evaluated the factor structure identified via PCA in the sample (n = 232). We present the fit indices for each model tested in Table 3. First, we ran the original model as reported in the original study by the WHOQOL Group (1998).

The original model, using four factors and 24 items, had an acceptable fit, although it should be noted that the values of the Tucker–Lewis index (TLI) and the CFI failed to meet the criterion set with the following values: 2(df = 246) = 653.562, CFI = .82, TLI = .806, and RMSEA=.85. We then tested our Arabic version of the WHOQOL-BREF scale proposed by the PCA of the sample. In the first run for model fit, the values were as follows: 2(df = 265) = 499.838, CFI = .90, TLI = .89, and RMSEA = .062, SRMR = 0.049. The results of the attempts are presented in Table 3.

Table 3: Fit indices for each model tested for CFA (n = 232)

Model	χ^2	df	χ^2/df	TLI	CFI	RMSEA	GFI	NFI	AIC	CAIC	ECVI
Model 1	653.562	246	2.65	.806	.82	.85	.81	.752	761.562	1001.686	3.297
Model 2	499.838	265	1.88	.89	.90	.062	.85	.81	619.838	886.643	2.683
Model 3	450.8620	242	1.86	.89	.91	.061	.86	.82	566.862	824.773	2.454

Model 1: Original WHOQOL-BREF (24 items, 4 factors); Model 2: Arabic version – initial CFA; Model 3: Arabic version – final CFA with modifications; AIC: Akaike information criterion; CAIC: Consistent Akaike information criterion; CFA: Confirmatory factor analysis: CFI: Comparative fit index; NFI: Normed fit index; RMSEA: Root mean square error of approximation; TLI: Tucker–Lewis index; ECVI: Expected cross-validation index; df: Degrees of freedom

The MIs improved the fit indices as a further attempt, and the second run produced an improvement in the following values: 2(df = 242) = 450.862, CFI = .91, TLI = .89, and RMSEA = .061. We noticed that the AIC and CAIC of the four-factor model were found to decrease with each attempt, and the GFI and NFI increased slightly, indicating relative improvements in fit over the earlier attempt. The model provided a good fit to the data, as seen in Table 3, which presents attempts at modifications of the model fit indices for each of the tested modified models. The model improved and, subsequently, we deemed it a good model fit.

3.3. Convergent and discriminant validity

Based on the Pearson's correlation coefficients, all the original and proposed domains had positive correlations ($p < .001$) with the items on general health and overall QoL, as well as with each other. Discriminant and convergent validity between the two QoL scales was substantiated by positive and medium-sized correlations.

In Table 4, we present the correlations between the WHOQOL-BREF and the other instruments. Convergent validity was assessed by calculating the correlations between the Arabic version of the WHOQOL-BREF and the QoL measures used in this study. The Arabic version of the WHOQOL-BREF demonstrated a significant and positive correlation with the QOLS ($r = 0.810$, $p = .001$). The HSS and WHOQOL-BREF exhibited a negative correlation ($r = −0.539$, $p = .001$), indicating that the Arabic version of the WHOQOL-BREF has discriminant validity. The selection of QoL measures to test discriminant validity in this study was further supported by the QOLS and HSS having a negative correlation ($r = –0.589$, $p = .001$).

3.4. Measurement invariance

Measurement invariance was assessed for gender and educational level. For gender, the configural invariance model fit the data adequately (Table 5). From Table 5, the configural invariance model provides an adequate fit for gender in the data. Next, we compared this configural model with a more constrained measurement invariance (i.e., metric measurement invariance) model. From Table 5, the weak invariance model, the first and most restrictive model, provides a good fit for the data. When comparing the metric invariance model with the configural invariance model, the changes in CFI, TLI, and RMSEA were all acceptable values ($\Delta CFI = 0.003$, $\Delta TLI = 0.01$, $\Delta RMSEA = −0.002$), indicating that the factor score metric was gender-invariant. In simpler terms, the meaning of the items used to estimate the factor loadings was the same for men and women. The fit of the strong invariance model, which was the next restrictive model, with the data is presented in Table 5.

Strong invariance ($\Delta CFI = −0.005$, $\Delta TLI = 0.001$, $\Delta RMSEA = 0.002$) was demonstrated using the second, more restrictive model, which constrained the factor loadings and item intercepts to generate a strong invariance model, indicating that the item intercepts and factor loadings are gender-invariant. Next, by constraining the factor loadings, item intercepts, and residual variances, we examined the final, more restrictive model, which is a strict invariance model. The modified indices of fit ($\Delta CFI = −0.002$, $\Delta TLI = 0.005$, $\Delta RMSEA = −0.002$) were all within the same range as the recommended values. This implies that comparisons of average item scores between males and females are valid.

For the educational level variable, we repeated the same analysis, beginning with the configural invariance model, which fit the data (Table 5). Next, we compared this configural model with the more constrained measurement invariance (i.e., metric measurement invariance) model. However, configural and scalar invariance could not be established for educational level, as the RMSEA values were greater than the threshold value of 0.080 (Table 5); CFI, TLI, and RMSEA values were 0.844, 0.822, and 0.085, respectively. Furthermore, for the metric run, the CFI, TLI, and RMSEA values were 0.842, 0.828, and 0.084, respectively. We determined that the model fit for the bachelor's group was somewhat poorer than those of the master's and Ph.D. groups, notwithstanding the fact that the metric did not make the structure worse. We decided to describe the metric invariance model as having a poorer fit for educational level, notwithstanding the changes in CFI, TLI, and RMSEA between the metric invariance model and the configural invariance model being acceptable values ($\Delta CFI = 0.002$, $\Delta TLI = −0.016$, $\Delta RMSEA = 0.002$).

Table 4: Correlations among the WHOQOL-BREF, with scales included in the analyses (n = 232)

Measures	HSS	QOLS	WHOQOL-BREF
HSS	1	−0.589**	−0.539**
QOLS	−0.589**	1	0.810**
WHOQOL-BREF	−0.539**	0.810**	1

HSS: Health and suffering scale; QOLS: Quality of life scale; WHOQOL-BREF: WHO quality of life – Brief version; **: p-value < .01

Table 5: Measurement invariance of the Arabic version of WHOQOL-BREF

Measurement invariance	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$(Δdf)	ΔCFI	ΔTLI	ΔRMSEA	ΔSRMR
Gender										
Configural	776.738 (444)	0.856	0.836	0.080 (0.071,0.090)	0.066	-	-	-	-	-
Metric	789.752 (463)	0.859	0.846	0.078(0.069,0.087)	0.071	13.014(19)	0.003	0.01	0.005	0.003
Scalar	824.807 (486)	0.854	0.848	0.078(0.068,0.087)	0.073	35.055 (23)	-0.005	0.001	0.002	0.003
Strict	851.854(509)	0.852	0.853	0.076 (0.067,0.085)	0.076	27.047 (23)	-0.002	0.005	-0.002	0.003
Educational level										
Configural	813.504 (444)	0.844	0.822	0.085(0.076,0.094)	0.063	-	-	-	-	-
Metric	842.044 (463)	0.842	0.828	0.084(0.075,0.093)	0.073	28.54 (19)	-0.002	-0.016	0.002	0.004

CFI: Comparative fit index; TLI: Tucker–Lewis index; RMSEA: Root mean square error of approximation; SRMR: Standardized root mean square residual; Δ: Change between models

This finding indicates that the factor score metric was not invariant among the educational level groups. In other words, the items utilized for estimating the factor loadings are different for the master's and Ph.D. groups than for the bachelor's group.

4. Discussion

Based on our inclusive study sample of teachers, the results of this study demonstrate that the WHOQOL-BREF is a measure with adequate psychometric properties for use in Arabic cultural settings. The WHOQOL-BREF demonstrates a good degree of internal consistency, and the four proposed QoL domains fit together adequately. Although a four-factor solution was found to best fit the data, the factor analysis did not group the same as the original instrument (Table 2). The major difference between the two versions is that the psychological domain in the WHOQOL-BREF with major differences in the other three factors.

In this study, we aimed to assess the construct validity, internal consistency, and structural validity of the WHOQOL-BREF using a sample of special education teachers. A four-factor model encompassing the physical, psychological, social, and environmental domains of QoL was broadly supported by all the analyses performed in this study. Nevertheless, the PCA results indicate a different item distribution among these domains. The new model structure outperformed the original domain structure on all indices of fit, notwithstanding the CFA results indicating a good fit of the original domain structure. This implies that the proposed item distribution is more appropriate for this sample of special education teachers than the item distribution in the original domain structure. Based on the general pattern in the findings from additional comparisons between the original and proposed domain structures, the physical, psychological, and social domains appear to operate somewhat better in the proposed structure.

The convergent validity of the WHOQOL-BREF questionnaire is demonstrated by the significant positive correlation found between its scale domains and broad QOLS measurements. Furthermore, there was a negative correlation between the HSS and the WHOQOL-BREF, and a positive correlation between all four of the WHOQOL-BREF domains. Based on our literature review, numerous global research studies have confirmed the convergent and discriminant validity of the WHOQOL-BREF (WHOQOL Group, 1998; Skevington et al., 2004; Kalfoss et al., 2021), which corresponds with our findings.

The measurement invariance of the structure of the Arabic WHOQOL-BREF across gender (men vs. women) and educational level was assessed using multiple-group CFA models. We evaluated the full factorial invariance of the item responses across gender and educational levels. Our study findings support the acceptable validity and internal consistency (or reliability) of the Arabic WHOQOL-BREF scale. The data demonstrate the gender-group invariance of the WHOQOL-BREF. Nevertheless, scalar invariance could not be proven for the educational level. The model fit was slightly poorer when comparing the bachelor's educational level group to the master's and Ph.D. groups. This means that the items of the scale were understood differently across different groups of participants who have different levels of education. This may be due to the small sample size in the master's and Ph.D. groups. It may also be explained by the diverse educational backgrounds within each of these two groups.

The findings of further investigation indicate that the factor structure of the WHOQOL-BREF might differ from the first factor structure proposed. Languages and cultures other than Spanish-speaking populations in Costa Rica, Peru, Mexico, Cuba, Paraguay, Argentina, Colombia, Spain, and Chile also exhibit variations from the original four-domain model (Benítez-Borrego et al., 2014; Lima-Castro et al., 2021). This has also been observed with populations in African countries (Oliveira et al., 2016; Ohaeri et al., 2007) and Portuguese-speaking populations (Moreno et al., 2006). The observed variations in our findings may be explained by the fact that the study sample was composed of working-age, reasonably healthy adults, as previously found and reported by Moreno et al. (2006) and Suárez et al. (2018). It is probable that the original scale structure has a good fit with a variety of samples with wide ranges in age and health status because the WHOQOL-BREF was initially developed and validated using large general populations (Suárez et al., 2018). This could account for some of the reported findings.

This study provides new insights into gender-related measurement invariance of the WHOQOL-BREF. We examined the measurement and structural invariance of the WHOQOL-BREF using a sample of special education teachers from the Hail

metropolitan area in Saudi Arabia. At the first two levels of analysis, the four-factor model demonstrated metric invariance (also called weak measurement invariance) and scalar invariance (also known as strong measurement invariance) across gender. These findings are consistent with those reported in previous studies (Yao and Wu, 2005; Lin et al., 2016). Furthermore, measurement invariance across gender was established, which also aligns with the findings of other recent studies (Kalfoss et al., 2021). However, the results of our study do not establish the presence of measurement invariance across educational levels, which contrasts with the findings of Kalfoss et al. (2021).

We chose to include covariance between residual items from other factors as well as those from the same factor in this study. These changes to the proposed WHOQOL-BREF model were made after we had obtained adequate theoretical support and were based on the MI values in the output of Jeffreys's Amazing Statistics Program. Items from the same factor were implicated in two residual covariances: Q20, which asked about satisfaction with one's personal relationships (factor: Social), and Q22, which asked about satisfaction with the help received from friends (factor: Social), were in question here. Given that the two items are derived from the same factor, the covariance between the residuals for both items was reasonable. Furthermore, because friendships are a type of relationship, it is possible to connect these two items based on their common social backgrounds. To obtain a better-fitting model, these residual covariances were included in the models after substantive meaningfulness was taken into account. This is not unexpected. These factors can make significant substantive sense, particularly in social psychology research; therefore, they should be incorporated into the model (Kueh et al., 2018).

We are aware that this study has some limitations. We are aware that self-reported survey data is impacted by response bias, and that this might lower the accuracy of the information provided by participants. Furthermore, based on impression management theories, individuals may respond to survey questions in a way that positively impacts the persona they would like to project. Therefore, response bias was considered while collecting data for this study. Consequently, we consistently encouraged and reminded participants to be truthful in responding to any questions about their personal lives and physical and mental well-being. In addition, we informed all participants that their responses would remain anonymous and be treated as confidential. Another limitation is that the findings of this investigation may be limited to special education teachers and may not be generalizable to other dissimilar populations. However, we are convinced that if a scale this specific to a healthy, educated, working-class population shows promising results, which the Arabic WHOQOL-BREF does, then our findings further establish the validity and reliability of the WHOQOL-BREF as a regional psychometric scale.

List of abbreviations

AIC	Akaike information criterion
CAIC	Consistent Akaike information criterion
CFA	Confirmatory factor analysis
CFI	Comparative fit index
df	Degrees of freedom
ECVI	Expected cross-validation index
EFA	Exploratory factor analysis
GFI	Goodness of fit index
HSS	Health and Suffering Scale
KMO	Kaiser–Meyer–Olkin (test)
MIs	Modification indexes
NFI	Normed fit index
PCA	Principal component analysis
QOLS	Quality of Life Scale
QoL	Quality of life
RMSEA	Root mean square error of approximation
SRMR	Standardized root mean square residual
TLI	Tucker–Lewis index
WHO	World Health Organization
WHOQOL-BREF	World Health Organization Quality of Life: Brief Version

Compliance with ethical standards

Ethical considerations

This study was approved by the Scientific Research Ethics Committee of the University of Hail. All participants were informed about the nature and purpose of the research and provided digital informed consent prior to participation. Participation was voluntary, and confidentiality and anonymity of responses were ensured throughout the study. The research complied with the ethical standards of the Declaration of Helsinki.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Andermo S, Sundberg T, Falkenberg T, Nordberg JH, and Arman M (2018). Measuring patients' health and suffering–the first stages of instrument development. Scandinavian Journal of Caring Sciences, 32(4): 1322-1331. https://doi.org/10.1111/scs.12577 **PMid:29855064**

Barneveld PS, Swaab H, Fagel S, Van Engeland H, and de Sonneville LM (2014). Quality of life: A case-controlled long-term follow-up study, comparing young high-functioning adults with autism spectrum disorders with adults with other psychiatric disorders diagnosed in childhood. Comprehensive Psychiatry, 55(2): 302-310. https://doi.org/10.1016/j.comppsych.2013.08.001 **PMid:24290884**

Bech VNP (2001). The WHO quality of life (WHOQOL) questionnaire: Danish validation study. Nordic Journal of Psychiatry, 55(4): 229-235.

https://doi.org/10.1080/080394801681019075
PMid:11839112

Benítez-Borrego S, Guàrdia-Olmos J, and Urzúa-Morales A (2014). Factorial structural analysis of the Spanish version of WHOQOL-BREF: An exploratory structural equation model study. Quality of Life Research, 23: 2205-2212.
https://doi.org/10.1007/s11136-014-0663-2
PMid:24609387

Bowden A and Fox-Rushby JA (2003). A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. Social Science and Medicine, 57(7): 1289-1306.
https://doi.org/10.1016/S0277-9536(02)00503-8
PMid:12899911

Brown TA (2015). Confirmatory factor analysis for applied research. Guilford Publications, New York, USA.

Burckhardt CS and Anderson KL (2003). The quality of life scale (QOLS): Reliability, validity, and utilization. Health and Quality of Life Outcomes, 1: 60.
https://doi.org/10.1186/1477-7525-1-60
PMid:14613562 PMCid:PMC269997

Burckhardt CS, Woods SL, Schultz AA, and Ziebarth DM (1989). Quality of life of adults with chronic illness: A psychometric study. Research in Nursing and Health, 12(6): 347-354.
https://doi.org/10.1002/nur.4770120604 **PMid:2602575**

Burgess AF and Gutstein SE (2007). Quality of life for people with autism: Raising the standard for evaluating successful outcomes. Child and Adolescent Mental Health, 12(2): 80-86.
https://doi.org/10.1111/j.1475-3588.2006.00432.x
PMid:32811109

Cantril H (1965). The pattern of human concerns. Rutgers University Press, New Brunswick, USA.

Cheung GW and Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2): 233-255.
https://doi.org/10.1207/S15328007SEM0902_5

Costello AB and Osborne J (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment, Research, and Evaluation, 10(1): 7.

Crocker TF, Smith JK, and Skevington SM (2015). Family and professionals underestimate quality of life across diverse cultures and health conditions: Systematic review. Journal of Clinical Epidemiology, 68(5): 584-595.
https://doi.org/10.1016/j.jclinepi.2014.12.007
PMid:25662007

Cummins RA and Lau AL (2006). Using health and subjective wellbeing for quality of life measurement: A review. Social Policy Review 18: 165-192.
https://doi.org/10.56687/9781847429186-012

Dimitrov DM (2010). Testing for factorial invariance in the context of construct validation. Measurement and Evaluation in Counseling and Development, 43(2): 121-149.
https://doi.org/10.1177/0748175610373459

Erikson EH, Erikson JM, and Kivnick HQ (1986). Vital involvement in old age. W. W. Norton & Company, New York, USA.

Flanagan JC (1978). A research approach to improving our quality of life. American Psychologist, 33(2): 138-147.
https://doi.org/10.1037//0003-066X.33.2.138

Gebhardt A, Langius-Eklöf A, Andermo S, and Arman M (2022). The health and suffering scale: Item reduction, reliability and validity among women undergoing rehabilitation for exhaustion and long-lasting pain. Nursing Open, 9(6): 2781-2792.
https://doi.org/10.1002/nop2.980
PMid:34216090 PMCid:PMC9584466

Hambleton RK, Merenda PF, and Spielberger CD (2004). Adapting educational and psychological tests for cross-cultural assessment. Psychology Press, New York, USA.
https://doi.org/10.4324/9781410611758

Kalfoss MH, Reidunsdatter RJ, Klöckner CA, and Nilsen M (2021). Validation of the WHOQOL-Bref: psychometric properties and normative data for the Norwegian general population. Health and Quality of Life Outcomes, 19: 13.
https://doi.org/10.1186/s12955-020-01656-x
PMid:33413455 PMCid:PMC7792093

Krägeloh CU, Kersten P, Rex Billington D, Hsu PHC, Shepherd D, Landon J, and Feng XJ (2013). Validation of the WHOQOL-BREF quality of life questionnaire for general use in New Zealand: Confirmatory factor analysis and Rasch analysis. Quality of Life Research, 22: 1451-1457.
https://doi.org/10.1007/s11136-012-0265-9
PMid:22983781

Kueh YC, Abdullah N, Kuan G, Morris T, and Naing NN (2018). Testing measurement and factor structure invariance of the physical activity and leisure motivation scale for youth across gender. Frontiers in Psychology, 9: 1096.
https://doi.org/10.3389/fpsyg.2018.01096
PMid:30018580 PMCid:PMC6037986

Lima-Castro S, Arias-Medina P, Bueno-Pacheco A, Peña-Contreras E, Aguilar-Sizer M, and Cabrera-Vélez M (2021). Factor structure, measurement invariance and psychometric properties of the quality of life scale WHOQOL-BREF in the Ecuadorian context. Psicologia: Reflexão e Crítica, 34: 29.
https://doi.org/10.1186/s41155-021-00194-9
PMid:34622355 PMCid:PMC8497657

Lin CY, Li YP, Lin SI, and Chen CH (2016). Measurement equivalence across gender and education in the WHOQOL-BREF for community-dwelling elderly Taiwanese. International Psychogeriatrics, 28(8): 1375-1382.
https://doi.org/10.1017/S1041610216000594
PMid:27097756

Moreno AB, Faerstein E, Werneck GL, Lopes CS, and Chor D (2006). Psychometric properties of the World Health Organization abbreviated instrument for quality of life assessment in the Pró-Saúde study. Cadernos de Saude Publica, 22: 2585-2597.
https://doi.org/10.1590/S0102-311X2006001200009
PMid:17096038

Naumann VJ and Byrne GJ (2004). WHOQOL-BREF as a measure of quality of life in older patients with depression. International Psychogeriatrics, 16(2): 159-173.
https://doi.org/10.1017/S1041610204000109
PMid:15318762

Noerholm V, Groenvold M, Watt T, Bjorner JB, Rasmussen NA, and Bech P (2004). Quality of life in the Danish general population–normative data and validity of WHOQOL-BREF using Rasch and item response theory models. Quality of Life Research, 13: 531-540.
https://doi.org/10.1023/B:QURE.0000018485.05372.d6
PMid:15085925

Ohaeri JU, Awadalla AW, El-Abassi AHM, and Jacob A (2007). Confirmatory factor analytical study of the WHOQOL-BREF: Experience with Sudanese general population and psychiatric samples. BMC Medical Research Methodology, 7: 37.
https://doi.org/10.1186/1471-2288-7-37
PMid:17678527 PMCid:PMC1950095

Oliveira SE, Carvalho H, and Esteves F (2016). Toward an understanding of the quality of life construct: Validity and reliability of the WHOQOL-BREF in a psychiatric sample. Psychiatry Research, 244: 37-44.
https://doi.org/10.1016/j.psychres.2016.07.007
PMid:27455149

Perera HN, Izadikhah Z, O'Connor P, and McIlveen P (2018). Resolving dimensionality problems with WHOQOL-BREF item responses. Assessment, 25(8): 1014-1025.

https://doi.org/10.1177/1073191116678925
PMid:27872348

Pomeroy IM, Tennant A, and Young CA (2013). Rasch analysis of the WHOQOL-BREF in post polio syndrome. Journal of Rehabilitation Medicine, 45(9): 873-880.
https://doi.org/10.2340/16501977-1186 **PMid:23979562**

Rocha NS, Power MJ, Bushnell DM, and Fleck MP (2012). Cross-cultural evaluation of the WHOQOL-BREF domains in primary care depressed patients using Rasch analysis. Medical Decision Making, 32(1): 41-55.
https://doi.org/10.1177/0272989X11415112
PMid:21825270

Skevington SM and Epton T (2018). How will the sustainable development goals deliver changes in well-being? A systematic review and meta-analysis to investigate whether WHOQOL-BREF scores respond to change. BMJ Global Health, 3: e000609.
https://doi.org/10.1136/bmjgh-2017-000609
PMid:29379649 PMCid:PMC5759710

Skevington SM, Lotfy M, and O'Connell KA (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL Group. Quality of life Research, 13: 299-310.
https://doi.org/10.1023/B:QURE.0000018486.91360.00
PMid:15085902

Snell DL, Siegert RJ, Surgenor LJ, Dunn JA, and Hooper GJ (2016). Evaluating quality of life outcomes following joint replacement: Psychometric evaluation of a short form of the WHOQOL-BREF. Quality of Life Research, 25: 51-61.
https://doi.org/10.1007/s11136-015-1044-1
PMid:26068734

Sosnowski R, Kulpa M, Ziętalewicz U, Wolski JK, Nowakowski R, Bakuła R, and Demkow T (2017). Basic issues concerning health-related quality of life. Central European Journal of Urology, 70(2): 206-211.
https://doi.org/10.5173/ceju.2017.923

Suárez L, Tay B, and Abdullah F (2018). Psychometric properties of the World Health Organization WHOQOL-BREF quality of life assessment in Singapore. Quality of Life Research, 27: 2945-2952.
https://doi.org/10.1007/s11136-018-1947-8
PMid:30046975

Tavares DMDS, Matias TGC, Ferreira PCS, Pegorari MS, Nascimento JS, and Paiva MMD (2016). Quality of life and self-esteem among the elderly in the community. Ciencia and Saude Coletiva, 21: 3557-3564.
https://doi.org/10.1590/1413-812320152111.03032016
PMid:27828588

Theuns P, Hofmans J, Mazaheri M, Van Acker F, and Bernheim JL (2010). Cross-national comparability of the WHOQOL-BREF: A measurement invariance approach. Quality of Life Research, 19: 219-224.
https://doi.org/10.1007/s11136-009-9577-9
PMid:20087777

Wang D, Sun L, Shao Y, Zhang X, Maguire P, and Hu Y (2023). Research and evaluation of a cyberchondria severity scale in a Chinese context. Psychology Research and Behavior Management, 16: 4417-4429.
https://doi.org/10.2147/PRBM.S431470
PMid:37936970 PMCid:PMC10626044

WHOQOL Group (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychological Medicine, 28(3): 551-558.
https://doi.org/10.1017/S0033291798006667
PMid:9626712

Yao G and Wu CH (2005). Factorial invariance of the WHOQOL-BREF among disease groups. Quality of Life Research, 14: 1881-1888.
https://doi.org/10.1007/s11136-005-3867-7
PMid:16155775

Yao G, Chung CW, Yu CF, and Wang JD (2002). Development and verification of validity and reliability of the WHOQOL-BREF Taiwan version. Journal of the Formosan Medical Association, 101(5): 342-351.