JERN Journal of Education and Research in Nursing

Original Article

DOI: 10.14744/jern.2025.54782
J Educ Res Nurs. 2025;22(2):1-7

Validity and Reliability of the Turkish Version of the Gender Equality Scale in Nursing Education

Neşe Ataman Bor,[1] Fırat Keskin[2]

[1]Department of Nursing, Hakkari University Faculty of Health Sciences, Hakkari, Türkiye

[2]Department of English Language Teaching, Hakkari University Faculty of Foreign Languages, Hakkari, Türkiye

Abstract

Background: Equality has become an increasingly prominent issue across various sectors of society, with gender inequality emerging as a key topic of discussion in nursing education. A comprehensive evaluation is essential to advancing gender parity in this field.

Aim: This study aimed to assess the validity of the Turkish version of the *Gender Equality Scale in Nursing Education* (GES-NE).

Methods: This methodological, descriptive, and correlational study was conducted between February 10 and April 30 with 408 senior nursing students aged 18 to 36. Data were collected via an online questionnaire, which included a *Socio-Demographic Data Collection Form* and the GES-NE scale. Prior to implementation, the scale underwent language adaptation, expert review, and a pilot study. The primary sample was subsequently assessed using item-total score analysis, Cronbach's alpha, and both confirmatory and exploratory factor analyses.

Results: The scale consisted of four subdimensions and 23 items, accounting for 50% of the total variance. Both confirmatory and exploratory factor analyses showed that all factor loadings were above 0.30. Confirmatory factor analysis revealed that all fit indices exceeded the acceptable threshold of 0.80, with a Root Mean Square Error of Approximation (RMSEA) value of 0.080. The overall Cronbach's alpha coefficient was 0.93, with each subdimension exceeding 0.60.

Conclusion: The Turkish version of the GES-NE scale is a valid and reliable instrument for measuring gender equality in nursing education among Turkish students.

Keywords: Gender discrimination, gender equality, gender role, nursing education, reliability, validity

Introduction

Diversity encompasses a broad range of personal, demographic, and societal characteristics, including, but not limited to, race, ethnicity, gender, sex, age, and gender identity. In contemporary societies, it is imperative to ensure equitable treatment for all individuals and to dismantle artificial barriers, preconceived notions, and biases. In educational environments that value diversity and inclusivity, students, faculty, and administrators increasingly recognize the importance of diversity in achieving success in teaching, learning, research, and practice.[1]

The American Association of Colleges of Nursing (AACN) highlights the importance of diversity, inclusion, and equity in nursing education and the development of a strong nursing workforce capable of delivering high-quality healthcare services.[2] Historically, nursing has been perceived as a profession primarily suited for women, based on the belief that nursing is inherently feminine and therefore inappropriate for men.[3] As a result, research on gender inequality in nursing has largely focused on the experiences of male nurses and nursing students, rather than measuring gender-based disparities or examining the biases and challenges they face.

The literature indicates that gender biases and stereotypes hinder the advancement and inclusivity of the nursing profession. For example, Madlala et al.[4] in 2021 found that male students experienced gender discrimination during midwifery education in clinical practice. They suggest that policy revisions or curriculum enhancements in nursing education could serve as effective strategies to address gender inequality.[4] Similarly, Petges and Sabio[5] im 2020 emphasize that promoting gender equality in nursing education requires the inclusion of diverse genders in mentoring programs, with faculty members playing a key role in supporting this process. Recognizing the importance of gender parity in nursing promotes inclusivity in curricula by encouraging students from diverse ethnic and cultural backgrounds to participate in and successfully complete nursing education programs. Additionally, Green1 suggests that nursing schools should publicly demonstrate their commitment to equity in nursing education while also complying with accreditation renewal requirements.[1] Several studies have also focused on reducing barriers for men entering nursing education and promoting greater male inclusion within the profession.[3,6]

Nevertheless, gender parity concerns in nursing extend beyond issues affecting men. According to the Global Health Workforce Network, established by the World Health Organization and the Center for Gender Equality, significant gender inequality persists within the healthcare industry.[7] The report highlights that gender stereotypes not only discourage men from pursuing nursing as a profession but also limit women's access

Cite this article as: Ataman Bor N, Keskin F. Validity and Reliability of the Turkish Version of the Gender Equality Scale in Nursing Education. J Educ Res Nurs. 2025;22(2):1-7.

Corresponding author: Neşe Ataman Bor
E-mail: neseatamanbor@hakkari.edu.tr

Received: March 31, 2024
Accepted: April 10, 2025
Publication Date: June 01, 2025

to leadership and senior roles in the field. Brandford and Brandford-Stevenson[8] in 2021 note that while women comprise 70% of the global health workforce, they hold only 25% of senior positions, reflecting a stark disparity. Therefore, a comprehensive examination of gender equality in nursing education is crucial to identifying and addressing the challenges faced by women in the profession.

Previous research has primarily focused on assessing the gender-based disparities experienced by male nursing students. For instance, O'Lynn in 2004 developed the *Inventory of Men's Friendship in Nursing Programs* (IMFNP) to measure the gender-related challenges perceived by male nursing students during their educational journey.[3] However, a more holistic understanding is necessary to develop effective strategies for promoting gender equality in nursing education. Accordingly, this study aimed to translate and adapt Cho et al.'s[9] in 2022 *Gender Equality Scale in Nursing Education* (GES-NE) from English to Turkish.

Research Questions of the Study

1. Is the *Gender Equality Scale in Nursing Education* a valid measurement tool?
2. Is the *Gender Equality Scale in Nursing Education* a reliable measurement tool?

Materials and Methods

Study Design

A methodological-descriptive-correlational design was employed in this study.

Universe and Sampling

This study utilized a methodological approach to evaluate the validity and reliability of the *Gender Equality Scale in Nursing Education*, adapted for use in Turkish. The target population included senior nursing students aged 18 and older in Türkiye. Convenience sampling was used to recruit participants from this population. A total of 710 senior nursing students who met the inclusion criteria were invited to participate and were encouraged to share the link with their peers.

According to the literature, sample size adequacy is classified as inadequate for values up to 200, fair for up to 300, good for up to 500, very good for up to 700, and excellent for values exceeding 1000.[10–12] The study sample consisted of 408 senior nursing students who voluntarily consented to participate between February 10 and April 30, 2023. A pilot study was conducted with a group of 20 senior nursing students. For the pilot application, data were collected from the first 20 volunteer students who accessed the study link shared via social media. These participants were excluded from the main sample to avoid any potential influence on the scale outcomes. Thus, the final sample included 408 senior nursing students. Sampling criteria required participants to be senior nursing students aged 18 or older, in good physical and mental health, and to have voluntarily agreed to participate. The participation rate was recorded at 57.4%.

Data Collection Tools

Data were collected using the Socio-Demographic Data Collection Form and the Gender Equality in Nursing Education Scale (GES-NE).

Socio-Demographic Data Collection Form

The form, developed by the researchers based on a review of the relevant literature, includes seven questions related to student's age, gender, education level, and employment status.[9]

Gender Equality in Nursing Education

The five-point Likert-type GES-NE scale, consisting of 23 items across four sub-dimensions, was developed by Cho et al.[9] The scale provides five response options, ranging from "1=strongly agree" to "5=strongly disagree." Item factor loadings range from 0.30 to 0.87. The four sub-dimensions: *gender personal experience of inequality, gender role perception, gender discrimination*, and *gender biases*, together account for 50% of the total variance.

The first sub-dimension, *gender personal experience of inequality*, consists of eight items (16, 15, 22, 17, 18, 26, 20, 14). The second sub-dimension, *gender role perception*, includes five items (3, 4, 5, 1, 2). The third sub-dimension, *gender discrimination*, comprises seven items (12, 8, 13, 24, 25, 9, 23). The fourth sub-dimension, *gender biases*, contains three items (7, 6, 19). The sub-dimensions and items in the Turkish version are consistent with those of the original scale.

Cronbach's alpha was used to assess the reliability of the scale. The overall Cronbach's α value was 0.93, with sub-dimension values ranging from 0.73 to 0.90.

Data Collection

Data were collected online due to the nationwide shift to distance education following the significant earthquake in Şanlıurfa on February 6, 2023. The researchers disseminated a link containing details about the study via social media platforms such as Instagram, WhatsApp, and Facebook.

Procedure

The Turkish version of the scale was developed in accordance with the guidelines of the International Test Commission (ITC).[13] Written permission was obtained from the scale's original developer to conduct the research. The adaptation process involved both back-translation and group translation methods. Five translators, proficient in both languages, translated the scale from English to Turkish after receiving the necessary approval from the scale's developer. The draft Turkish version, reflecting the most accurate Turkish equivalents, was then back-translated into English by three professionals (two academic nurses and one linguist), who had no prior exposure to the original scale. Following the back-translation, the researchers made the necessary revisions, and the final Turkish version of the scale was prepared. The scale was then submitted to experts for evaluation in terms of language, cultural equivalence, and content validity.[10,14]

According to the literature, consultation with at least three experts is recommended for such evaluations.[14] In this study, expert opinions were obtained from 11 faculty members from different universities, all specializing in gender equality research within the public health nursing departments of three higher education institutions.

The experts were provided with both the English and Turkish versions of the scale and were asked to assess the appropriateness of each item. The evaluators were provided with both the source (English) and target (Turkish) versions of the scale and were asked to rate the appropriateness of each item using a 4-point scale: 1 = minimal adjustments needed, 2 = some adjustments needed, 3 = appropriate, and 4 = highly appropriate. The ratings were analyzed using the Davis Content Validity Index (CVI), and both the item-level CVI (I-CVI) and the scale-level CVI (S-CVI) were calculated. For the overall scale, the I-CVI ranged from 0.99 to 1.00, while the S-CVI was 0.99. I-CVI and S-CVI values above 0.80 are considered sufficient to establish content validity. Following expert approval of both the Turkish and English versions, a preliminary study was conducted with 20 senior nursing students. The pilot results indicated that the questions were clear and no negative feedback was received. Data were collected via an online survey distributed through Google Forms. Based on the pilot findings, no changes were made to the Turkish version of the scale, and the researchers proceeded to administer it to the main sample. The 20 who participated in the pilot study were excluded from the main sample to prevent any potential influence on the results. No personal data or email addresses were collected from participants during the online survey administration.

Invariance of the Scale Over Time

The stability of the scale over time was assessed using the test-retest reliability method. Nursing students completed the GES-NE scale twice, with a two-week interval between administrations. The test-retest reliability coefficient was determined to be 0.94.

Data Analysis

Data analysis was conducted using the Statistical Package for the Social Sciences version 24.0 (IBM SPSS Corp., Armonk, NY, USA) and Analysis of Moment Structures version 25 (Amos Development Corporation, Chicago, IL, USA).[15] Descriptive statistics, including percentages, means, standard deviations, and frequencies, were used to summarize the data. Normality was assessed using skewness and kurtosis values, along with normality tests. For language adaptation, the translation and back-translation method was employed. Validity analysis included content validity, surface validity, and construct validity. For content validity, both the Content Validity Index and Content Validity Ratio were calculated using the Davis

Table 1. Descriptive characteristics of nursing students (n=408)

Characteristic	n	%	M (SD)
Gender			
Male	176	43.1	
Female	232	56.9	
Marital status			
Single	352	86.3	
Married	56	13.7	
Presence of male faculty members			
Yes	400	98.0	
No	8	2.0	
Views opposite sex as brother/sister			
Yes	264	64.7	
No	144	35.3	
Received instruction from male instructors			
Yes	374	94.1	
No	24	5.9	
Proportion of male students in class			
<10%	88	21.6	
>10%–20%	252	61.8	
>20%–30%	60	14.7	
>30%	8	2.0	
Satisfaction with nursing (VAS 1–10)			5.68±3.16
Determination to become a nurse (VAS 1–10)			5.96±2.98
Total scale score			81.05±17.24

n: Number, %: Percentage, M: Mean, SD: Standard deviation, VAS: Visual analogue scale.

technique. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used to assess construct validity. For reliability analysis, item-total correlations, Cronbach's alpha, and split-half methods were employed. A p-value of <0.05 was considered the threshold for statistical significance.

Ethics Committee Approval

Written permission to translate the GES-NE scale into Turkish was obtained via email from the scale's original author. The study received ethical approval from Hakkari University Rectorate Scientific Research and Publication Ethics Committee (Date: 23.01.2023, Number: 2023/14-1).

The study's objective was explained to nursing students through a consent form accessible via a Google Form link. The study included senior nursing students who voluntarily provided informed consent to participate. The research was conducted in accordance with the principles of the Declaration of Helsinki.

Results

Among the nursing students, 56.9%, were identified as female, while a significant majority, 86.3%, reported being single. Additionally, 61.8% indicated that the percentage of male students enrolled in their nursing programs ranged from 10% to 20%. A smaller proportion, 5.9%, reported that their schools had no male faculty members. The data show that the majority of nursing students, specifically 98%, acknowledged the presence of male faculty members at their institutions. Additionally, 64.7% of the students reported perceiving individuals of the opposite sex as siblings. The average total score on the scale was found to be 81.05±17.24 (Table 1).

Validity Analysis of the Scale

Content Validity

Content validity was assessed based on responses from 11 experts who were provided with the form developed for the study. The Content Validity Index was calculated using the Davis technique.[16] In this study, the item-level Content Validity Index ranged from 0.99 to 1.00, while the scale-level Content Validity Index was 0.99.

Construct Validity

The construct validity of the scale was evaluated using both Exploratory Factor Analysis and Confirmatory Factor Analysis. Following the EFA, the promax rotation method and principal axis factoring were applied to determine the construct validity of the scale, as the data were normally distributed. Bartlett's test yielded a χ^2 value of 13,651.996, the Kaiser-Meyer-Olkin coefficient was 0.857, and p<0.01. The analysis revealed that the scale consists of four sub-dimensions. These four dimensions collectively explained 58.13% of the total variance:

The first sub-dimension item factor loads varied from 0.73 to 0.80. The items in the second sub-dimension had factor loads ranging from 0.56 to 0.79. The third sub-dimension item factor loads varied from 0.51 to 0.73. The fourth sub-dimension item factor loads range from 0.50 to 0.67 (Table 2).

In the four-factor model, items 1 through 9 loaded most heavily onto the first factor.

The four sub-dimensions identified through the Exploratory Factor Analysis were further evaluated using CFA. Goodness-of-fit indices used in the CFA included the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Chi-square/degrees of freedom ratio (CMIN/DF), Incremental Fit Index (IFI), and the Root Mean Square Error of Approximation (RMSEA). The fit indices from the CFA were as follows: χ^2=472.864, df=224, χ^2/df=2.111, RMSEA=0.055, CFI=0.90, and IFI=0.91 (Table 3). In the first sub-dimension, which measures *personal experience of gender inequality*, the factor loadings ranged from 0.70 to 0.91. In the second sub-dimension, which measures *perception of gender roles*, factor loadings ranged from 0.58 to 0.88. In the third sub-dimension, which measures *gender discrimination*, factor loadings ranged from 0.64 to 0.98. Finally, in the fourth sub-dimension, which measures *gender biases*, factor loadings ranged from 0.84 to 0.91 (Figure 1).

Results of the Scale's Reliability Analysis

The Cronbach's coefficients for the overall scale and each sub-dimension were as follows:

Overall scale 0.96, personal experience of gender inequality 0.94, perception of gender roles 0.83, gender discrimination 0.95, and gender biases 0.91.

Cronbach alpha values for the first and second halves of the split-half analysis were determined to be 0.93 and 0.94, respectively. The Spearman-Brown coefficient was found to be .90, the Guttman split-half coefficient to be .89, and the two halves' correlation coefficient to be 802. The analysis yielded the following results for Hotelling: T2=479.44, F=20.66, and p=0.000. The inter-item correlation ranged from 0.09 to 0.69 (Table 4).

The study found a range of correlations between the scale items and the scale total score, with values ranging from 0.48 to 0.83. Table 5 presents the correlation coefficients between the total score of each subdimension and its corresponding subdimension. Table 5 displays the correlation coefficients for each subdimension. The first subdimension has a range of 0.71–0.96, the second subdimension has a range of 0.74–0.89, the third subdimension has a range of 0.87–0.96, and the fourth subdimension has a range of 0.87–0.90.

Lower and Upper Group Item Analysis

In scale development studies, lower and upper-group item analysis is used to determine the discriminatory power of items.[10,12] To evaluate item distinctiveness in the GES-NE scale, the scores of 408 nursing students were ranked from highest to lowest. The mean total scores of the upper and lower 27% groups (n=204 each) were compared using an independent sample t-test. The t-test revealed that the difference between the upper and lower group scores was statistically significant (p<0.05). Based on this test, all 23 items on the scale were found to be distinctive. These results indicate that the scale has strong discriminatory power, effectively measures the intended construct, and can clearly differentiate the upper and lower 27% of respondents (Table 6).

Discussion

A minimum CVI of 0.80 is required to establish content validity.[17] In this study, the scale-level Content Validity Index was 0.99, while the item-level Content Validity Index ranged from 0.99 to 1.00. These values exceed the threshold of 0.80. The high I-CVI and S-CVI values in the current study indicate that content validity was achieved, expert

DOI: 10.14744/jern.2025.54782
J Educ Res Nurs. 2025;22(2):1-7

Table 2. Factor loadings for the four-factor structure of the Turkish version of the gender equality scale in nursing education (n=408)

Items	Factor loadings			
	Personal experience of gender inequality	Perception of gender roles	Gender discrimination	Gender prejudices/ bias in the classroom
I1	0.806			
I2	0.804			
I3	0.802			
I4	0.790			
I5	0.769			
I6	0.761			
I7	0.753			
I8	0.737			
I9		0.792		
I10		0.657		
I11		0.566		
I12		0.645		
I13		0.593		
I14			0.552	
I15			0.512	
I16			0.645	
I17			0.593	
I18			0.650	
I19			0.739	
I20			0.727	
I21				0.673
I22				0.645
I23				0.508
Explained variance (%)	37.45	8.41	7.42	4.85
Total explained variance (%)	58.13			
Eigenvalue	13.37	1.94	1.71	1.12
Kaiser-Meyer-Olkin	857			
Barlett's Test (χ^2)	13,651.996			
p	0.000			

Table 3. Model fit indices of the gender equality scale in nursing education

Model	χ^2	CFA[a]	χ^2/DF	RMSEA[b]	GFI[c]	CFI[d]	IFI[e]
Four-factor structure	472.864	224	2.111	0.055	0.95	0.90	0.91

[a]: Confirmatory factor analysis, [b]: Root mean square error of approximation, [c]: Goodness of fit index, [d]: Comparative fit index; [e]: Incremental fit index. DF: Degree of freedom.

consensus was reached, and the scale effectively measured the intended construct. The content validity results were consistent with those reported in the original study.[9]

There are several techniques for assessing construct validity, with factor analysis being the most commonly used method.[11,14,17] Factor analysis can be conducted using two approaches: Exploratory Factor Analysis and Confirmatory Factor Analysis. To conduct factor analysis effectively, the sample size should be sufficiently large, ideally five to ten times the number of items on the scale.[17,18] Initially, sample adequacy should be assessed using the Kaiser-Meyer-Olkin (KMO) test, followed by Bartlett's test to evaluate the suitability of the data for factor analysis. The KMO test determines whether the sample is adequate, while Bartlett's test assesses the appropriateness of the data structure for factor analysis.[10,14] KMO values are interpreted as follows: Below 0.50=Poor, 0.50–0.59=Fair, 0.60–0.69=Good, 0.70–0.79= Very good and 0.80–1.00 = Excellent.[17,18]

If the KMO value is close to 1 and greater than 0.60, the data are considered suitable for factor analysis.[19] In this study, the EFA results indicated a Kaiser-Meyer-Olkin coefficient of 0.857 and a Bartlett's test χ^2 value of 13,651.996 ($p<0.01$), confirming that the sample size and data were appropriate for factor analysis.[20] The sample size and dataset used in this study were similar to those of the original research.[9] Additionally, the scale consists of four factors, with items numbered 1 to 9 showing the strongest association with the the first factor.

Variance rates between 40% and 60% are generally considered acceptable when determining the number of components.[17] In this study, the four-factor structure explained 50% of the total variance. In comparison, the initial study reported that the scale accounted for 60.7% of the variance.[9]

According to the EFA results, the factor loadings of the scale items were ≥0.30.18 The factor loadings for the four sub-dimensions ranged from 0.30 to 0.87 based on the results of the EFA. According to the literature, items with factor loadings below 0.30 should be excluded from the scale. The minimum acceptable factor loading for determining an item's inclusion in a factor is 0.30 or higher.[17,10,12] In their study, Cho et al.[9] in 2022 reported factor loadings for the four subscales ranging from 0.58 to 0.85. The similarity between the current findings and those of the original study supports the scale's robust factor structure.

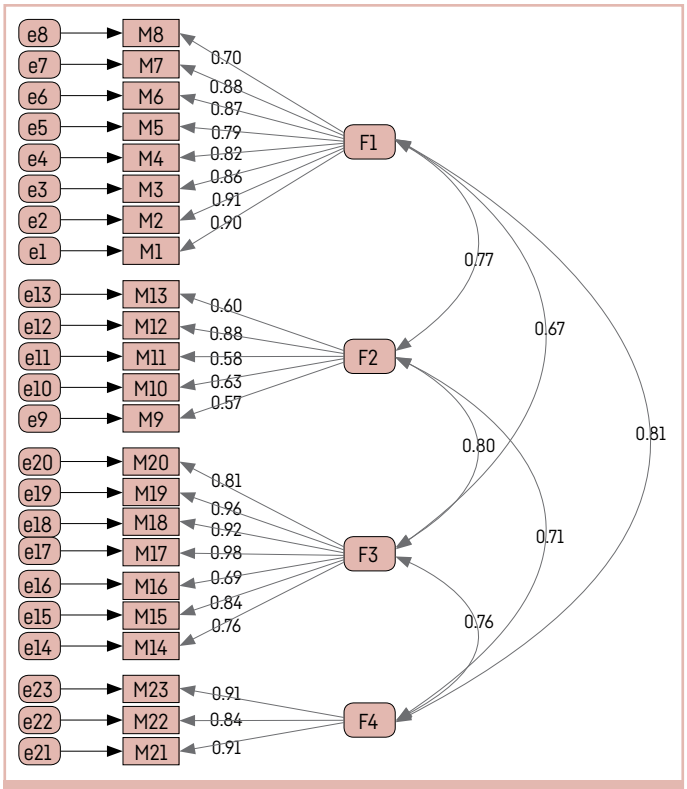

Figure 1. Confirmatory factor analysis (CFA) results of the Gender Equality Scale in Nursing Education (n=408).

CFA is another technique used in scale adaptation research to ascertain the structure of the measure.[10,14] The goodness-of-fit indices in CFA indicate how well the proposed structure represents the observed data. These indices partially determine whether the model is acceptable or should be rejected. The most commonly used goodness-of-fit indices include CMIN/DF, Goodness-of-Fit Index (GFI), CFI, Normed Fit Index (NFI), IFI, and RMSEA.[14] It is generally expected that the CMIN/DF value should be less than 5, the CFI value greater than 0.85, and the GFI, NFI, and TLI values greater than 0.80.[10] The interpretation of Root Mean Square Error of Approximation values is as follows: ≤0.05 indicates a good fit; 0.05–0.08 indicates an adequate fit; 0.08–0.10 indicates an acceptable fit; and >0.10 indicates a poor fit.[21,22] In this study, the CFA results were: CFI=0.91, RMSEA=0.055, GFI=0.95, CFI=0.90, and IFI=0.91. The χ^2/df ratio was found to be 2.111, and RMSEA <0.85. According to the literature, a χ^2/df value less than 5 and an RMSEA value below 0.08 are considered indicators of good model fit.[11,14,19] The CFA results in this study meet these criteria. Cho et al.[9] in 2022 confirmed that the final model in their study demonstrated satisfactory goodness of fit. The findings of the present investigation are consistent with those of the original study.

There are several methods to assess the reliability of a measurement tool.[23] In this study, reliability was evaluated using internal consistency coefficient, split-half, and item-total score correlation analyses.[18] The internal consistency coefficient, known as Cronbach's α, measures the degree of internal consistency and is used to assess reliability. Higher values indicate stronger coherence among items. The scale is considered reliable when Cronbach's alpha falls within the range of 0.80 to 1.00.[18,23] In this study, the overall Cronbach's α for the scale was 0.963, and all sub-dimensions had values above 0.80. These results demonstrated the high reliability of the measure, with a reliability value that exceeded that of the initial study (0.824).[9]

To verify the explanatory power of individual items in relation to the overall score, item analysis should be conducted as part of the reliability assessment process.[11,12] Correlation coefficients between the items and the total score should exceed 0.20.[11,23] In this study, the adjusted item-total correlation coefficients ranged from 0.48 to 0.83, all of which were above 0.20 and met the necessary criteria. Item-total correlation data were not available in the original study, so a direct comparison could not be made.[9]

Split-half analysis is a technique used to assess reliability. According to the literature, Guttman and Spearman-Brown split-half coefficients should exceed 0.80.[18] In this study, both the Spearman-Brown and Guttman split-half coefficients were greater than 0.80, demonstrating the strong reliability of the scale's items and overall structure. Our results could not be compared to the findings of the original study, as split-half analysis was not performed in that research.[9] The results of this study demonstrate that the scale is reliable and that the items are consistent with the theoretical framework.

One of the recommended methods for assessing the reliability and validity of scales is the 27% upper-lower group comparison.[10,12,14] In this study, a statistically significant difference was found between the average scores of nursing students in the upper 27% group and those in the lower 27% group. These results indicate that the scale has strong discriminating power, adequately measures the intended construct, and can effectively distinguish between the upper and lower 27% performance groups. Overall, the findings demonstrate that this scale is both valid and reliable and can be used to assess gender equality in nursing education.

Limitations

The study has several limitations despite its many strengths. First, the use of a convenience sampling method may affect the generalizability of the findings. The data were collected based on self-reports from nursing students, and it is possible that some students provided inaccurate responses, which may have introduced response bias. This could have impacted the reliability of the results. Additionally, collecting data through an online platform may be considered another limitation of the study.

Conclusion

Gender equality is one of the least explored domains in nursing, despite being one of the essential components of modern nursing education. There exists scant empirical literature on the evaluation of gender equity within nursing education in Türkiye, and even on a global scale, there are only a few studies that apply 'scale' criteria to assess inequality in nursing education. With this gap in mind, we sought to validate and assess reliability for the Gender Equality Scale for Nursing Education (GES-NE) adapted to Turkish. The results supported that the Turkish version of GES-NE has satisfactory levels of internal consistency and structural validity with four sub-dimensions and 23 items. These four sub-dimensions explained 58.3% of the overall

Table 4. Reliability analysis results of the gender equality scale in nursing education (n=408)

Dimension	Cronbach's α	Split-half analysis				
		First half Cronbach's α	Second half Cronbach's α	Spearman-brown coefficient	Guttman split-half coefficient	Correlation between halves
Total scale	0.96	0.93	0.94	0.90	0.89	0.80
Personal experience of gender inequality	0.94					
Perception of gender roles	0.83					
Gender discrimination	0.95					
Gender biases	0.91					

DOI: 10.14744/jern.2025.54782
J Educ Res Nurs. 2025;22(2):1-7

Table 5. Item-total score and sub-dimension score correlations

Item	Corrected item-subdimension total score correlations				
	Personal experience of gender inequality	Perception of gender roles	Gender discrimination	Gender prejudices/ bias in the classroom	Corrected item-total correlation
I1	0.937				0.806
I2	0.960				0.810
I3	0.947				0.781
I4	0.936				0.724
I5	0.925				0.742
I6	0.959				0.803
I7	0.923				0.820
I8	0.711				0.691
I9		0.749			0.506
I10		0.780			0.556
I11		0.750			0.489
I12		0.895			0.819
I13		0.783			0.503
I14			0.914		0.788
I15			0.888		0.837
I16			0.911		0.771
I17			0.962		0.782
I18			0.942		0.707
I19			0.944		0.760
I20			0.870		0.751
I21				0.858	0.770
I22				0.871	0.740
I23				0.905	0.776

Table 6. t-test of upper and lower group means for the overall scale (n=408)

	Item	Upper 27% (n=204)		Lower 27% (n=204)		Test	
		Mean	SD	Mean	SD	t	p
Personal experience of gender inequality	I1	3.8	1.01	1	0.00	29.168	0.000
	I2	3.73	1.16	1.22	0.78	18.856	0.000
	I3	3.15	1.47	1.07	0.26	14.656	0.000
	I4	4.09	1.09	1.89	1.38	13.162	0.000
	I5	3.87	1.05	1.75	1.36	13.025	0.000
	I6	3.15	1.47	1.27	0.59	12.491	0.000
	I7	3.74	1.2	1.38	1	15.796	0.000
	I8	3.31	1.27	1.56	1.06	11.055	0.000
Perception of gender roles	I9	3.3	1.39	1.89	1.38	7.55	0.000
	I10	3.66	1.23	2.02	1.69	8.258	0.000
	I11	3.37	1.6	2.98	1.53	1.845	0.066
	I12	3.15	1.06	1.07	0.26	19.932	0.000
	I13	3.73	1.58	2.05	1.4	8.328	0.000
Gender discrimination	I14	3.09	1.44	1.07	0.26	14.49	0.000
	I15	2.95	1.44	1.29	0.81	10.539	0.000
	I16	3.09	1.44	1.22	0.78	12.011	0.000
	I17	2.59	1.5	1	0	11.054	0.000
	I18	2.51	1.6	1	0	9.919	0.000
	I19	2.51	1.41	1.07	0.26	10.562	0.000
	I20	2.8	1.43	1.18	0.51	11.225	0.000
Gender prejudices/bias in the classroom	I21	3.31	1.27	1.42	0.73	13.509	0.000
	I22	3.09	1.39	1.42	0.91	10.585	0.000
	I23	3.16	1.4	1.31	0.83	11.917	0.000

SD: Standard deviation.

variance, fulfilling the vital psychometric standards. Therefore, the conclusions of this research proved that the Turkish GES-NE is a valid and reliable tool for evaluating gender equity in nursing education in Türkiye. Moreover, this tool will guide efforts aimed at evaluating organizational and educational initiatives undertaken to promote gender equality in nursing education.

Taking these considerations into account, we recommend that the nursing education programs in Türkiye incorporate the GES-NE into their assessment practices for systematic evaluation at all levels. This would facilitate efforts to achieve gender equality at the institutional level within nursing programs. Moreover, other longitudinal studies employing the GES-NE over extended periods would provide greater clarity and focus on the assessment of gender-equality initiatives over time. Lastly, cross-case studies among varying levels of education in Türkiye may further illuminate the context factors that determine gender equality and strengthen educational opportunities and policies using gender lenses.

Ethics Committee Approval: The study was approved by the Hakkari University Rectorate Scientific Research and Publication Ethics Committee (Approval Number: 2023/14-1, Date: 23.01.2023).

Informed Consent: Written informed consent was obtained from the participants.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – N.A.B., F.K.; Design – N.A.B., F.K.; Supervision – N.A.B., F.K.; Resource – N.A.B., F.K.; Materials – N.A.B., F.K.; Data Collection and/or Processing – N.A.B., F.K.; Analysis and/or Interpretation – N.A.B., F.K.; Literature Review – N.A.B., F.K..; Writing – N.A.B., F.K.; Critical Review – N.A.B., F.K.

Conflict of Interests: The authors have no conflicts of interest to declare.

Funding: The authors declared that this study received no financial support.

References

1. Green C. Equity and diversity in nursing education. Teach Learn Nurs. 2020;15(4):280–283. [CrossRef]
2. American Association of Colleges of Nursing. Diversity, equity, and inclusion in academic nursing. Accessed April 14, 2025. https://www.aacnnursing.org/News-Information/Position-Statements-White-Papers/Diversity
3. Teresa-Morales C, Rodríguez-Pérez M, Araujo-Hernández M, Feria-Ramírez C. Current stereotypes associated with nursing and nursing professionals: An integrative review. Int J Environ Res Public Health. 2022;19(13):7640. [CrossRef]
4. Madlala ST, Ngxongo TS, Sibiya MN. Perceptions of student accoucheurs regarding gender inequality in midwifery training at Free State maternal healthcare institutions. Curationis. 2021;44(1):e1-e8. [CrossRef]
5. Petges N, Sabio C. Perceptions of male students in a baccalaureate nursing program: A qualitative study. Nurse Educ Pract. 2020;48:102872. [CrossRef]
6. Pai HC, Liu NY, Hsu WY, Hung CA, Wu PL. Testing and validation of the gender-friendliness barriers in nursing programs questionnaire: A confirmatory factor analysis. 2019. Accessed April 14, 2025. https://www.sigmarepository.org/cgi/viewcontent.cgi?article=6116&context=inrc
7. Boniol M, McIsaac M, Xu L, et al. Gender equity in the health workforce: Analysis of 104 countries. Geneva: World Health Organization. 2019. Accessed April 14, 2025. https://iris.who.int/handle/10665/311314
8. Brandford A, Brandford-Stevenson A. Going up!: Exploring the phenomenon of the glass escalator in nursing. Nurs Adm Q. 2021;45(4):295–301. [CrossRef]
9. Cho S, Kwon SH, Jang SJ. Validity and reliability of the gender equity scale in nursing education. Nurs Health Sci. 2022;24(2):447–457. [CrossRef]
10. Kartal M, Bardakçı S. SPSS ve AMOS uygulamalı örneklerle güvenirlik ve geçerlik analizleri. Ankara: Akademisyen Kitabevi. 2018;1–192. [CrossRef]
11. Özdamar K. Eğitim, sağlık ve davranış bilimlerinde ölçek ve test geliştirme yapısal eşitlik modellemesi. Eskişehir: Nisan Yayınevi. 2016;6–286.
12. Seçer İ. Psikolojik test geliştirme ve uyarlama süreci: SPSS ve LISREL uygulamaları. Ankara: Anı Yayıncılık. 2018;10–168.
13. Gregoire J. ITC guidelines for translating and adapting tests. Int J Testing. 2018;18(2):101–134. [CrossRef]
14. Karagöz Y. SPSS 23 ve AMOS 23 uygulamalı istatistiksel analizler. Ankara, Turkey: Nobel Akademik Yayıncılık. 2016;878–940.
15. IBM Corp. IBM SPSS Statistics for Windows, version 24.0. Armonk, NY: IBM Corp. 2016. [CrossRef]
16. Davis LL. Instrument review: Getting the most from a panel of experts. Appl Nurs Res. 1992;5(4):194–197. [CrossRef]
17. DeVellis RF, Thorpe CT. Scale development: Theory and applications. Thousand Oaks, CA: Sage Publications. 2021;31–59.
18. Tabachnick BG, Fidell LS. Çok değişkenli istatistiklerin kullanımı. Baloğlu M, translator. Ankara: Nobel Academic Publishing. 2015;612–760.
19. Johnson BR, Christensen L. Educational research: Quantitative, qualitative, and mixed approaches. 3rd ed. Thousand Oaks, CA: SAGE Publication. 2014;190–222.
20. Çokluk Ö, Şekercioğlu G, Büyüköztürk Ş. Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları. Ankara: Pegem Akademi. 2010.
21. Kline RB. Principles and practice of structural equation modeling. 3rd ed. New York, NY: Guilford. 2011;7–342.
22. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1(2):130–149. [CrossRef]
23. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health. 2007;30(4):459–467. [CrossRef]