Research

Psikiyatride Güncel Yaklaşımlar-Current Approaches in Psychiatry 2026; 18(Suppl 1):78-92

doi: 10.18863/pgy.1868236

Turkish Adaptation of The Multidimensional Gender Inequality Perception Scale

Çok Boyutlu Cinsiyet Eşitsizliği Algısı Ölçeğinin Türkçe Uyarlaması

Nedime Gül Doğan Özdemir[1], Gamze Yavaş[2], Özen Kulakaç[1]

[1]Ondokuz Mayıs University, Samsun
[2]Akdeniz University, Antalya

ABSTRACT

Objective: This study was conducted to adapt the Multidimensional Gender Inequality Perception Scale (MuGIPS) developed by Schwartz-Salazar et al. into Turkish and to evaluate its psychometric properties.

Method: In this methodological study, the scale underwent translation–back-translation, expert review, and pilot testing phases. The scale's construct validity was assessed using confirmatory factor analysis (CFA), and its reliability was evaluated using Cronbach's alpha and the test–retest method. A total of 322 adults were reached through an online data collection method.

Results: The CFA results confirmed that the MuGIPS-TR has a four-factor structure comprising 25 items across the dimensions of health, violence, household chores and care, and public space and power. Cronbach's alpha coefficients for the sub-dimensions ranged from 0.73 to 0.92, while the Cronbach's alpha coefficient for the total scale was found to be 0.93. The test–retest correlation was positive and significant ($r = 0.79$; $p < 0.001$).

Conclusion: The findings indicate that the MuGIPS-TR is a valid and reliable measure of multidimensional perceptions of gender inequality in this sample. It is considered suitable for research on gender inequality and for use in the nursing field.

Keywords: Multidimensional Gender Inequality Perception Scale, gender inequality, scale adaptation, validity, and reliability

ÖZ

Amaç: Bu araştırma, Schwartz-Salazar ve arkadaşları tarafından geliştirilen Çok Boyutlu Cinsiyet Eşitsizliği Algısı Ölçeğinin (ÇBCEAÖ) Türkçeye uyarlanması ve psikometrik özelliklerinin değerlendirilmesi amacıyla gerçekleştirilmiştir.

Yöntem: Metodolojik tipteki bu araştırmada ölçeğin çeviri–geri çeviri, uzman değerlendirmesi ve pilot uygulama aşamaları gerçekleştirilmiştir. Ölçeğin yapı geçerliliği doğrulayıcı faktör analizi (DFA) ile, güvenirliği ise Cronbach alfa katsayısı ve test–tekrar test yöntemi ile değerlendirilmiştir. Araştırma kapsamında çevrim içi veri toplama yöntemiyle 322 yetişkine ulaşıldı.

Bulgular: DFA sonucunda, ÇBCEAÖ-TR'nin sağlık, şiddet, ev işleri ve bakım ile kamusal alan ve güç boyutlarından oluşan 25 maddeli dört faktörlü yapısı doğrulanmıştır. Alt boyutlara ilişkin Cronbach alfa katsayıları 0,73–0,92 arasında değişirken, toplam ölçek için Cronbach alfa katsayısı 0,93 olarak bulunmuştur. Test–tekrar test korelasyonu pozitif ve anlamlıdır ($r=0,79$; $p<0,001$).

Sonuç: Bulgular, ÇBCEAÖ-TR'nin bu örneklem için çok boyutlu cinsiyet eşitsizliği algısını değerlendirmede geçerli ve güvenilir bir ölçme aracı olduğunu göstermektedir. Ölçeğin toplumsal cinsiyet eşitsizliğine yönelik araştırmalarda ve hemşirelik alanında kullanılabileceği değerlendirilmektedir.

Anahtar sözcükler: Çok Boyutlu Cinsiyet Eşitsizliği Algısı Ölçeği, toplumsal cinsiyet eşitliği, ölçek uyarlama, geçerlilik ve güvenirlik

Address for Correspondence: Nedime Gül Doğan Özdemir, Ondokuz Mayıs University, Faculty of Health Sciences, Department of Nursing, Samsun, Türkiye **e-mail:** nedime.dogan@gmail.com

Received: 22.01.2026 | **Accepted:** 11.07.2026

Introduction

The concept of multidimensional gender inequality encompasses inequalities between the sexes across various areas of life and their consequences, rather than focusing solely on inequalities in a single domain, such as income or educational attainment. Research identifies areas such as health, violence, domestic work and care, the public sphere and power, economic security, access to resources, education, employment, political participation, and social expectations as the key dimensions where gender inequalities emerge and interact with one another (Schwartz-Salazar et al. 2024, Ciaffoni et al. 2025). These areas where gender inequality manifests influence one another, and this interaction leads to more comprehensive outcomes beyond simple accumulation. In other words, inequality in one area (e.g., education) can deepen and perpetuate inequalities in another (e.g., employment). This interactive structure enables a comprehensive assessment of the inequalities faced by women and other vulnerable groups across various sectors of society (Tekgüç and Akbulut 2021).

Multidimensional gender inequality varies significantly across countries, both in terms of the extent of inequality and the areas in which it manifests. In developing Latin American countries such as Guatemala and Honduras (Stefani and Prati 2020; Begall et al. 2023), gender inequality is generally concentrated in family- and identity-based areas, including legal rights and social norms within the household. In contrast, in more developed countries, inequalities related to political participation and family life are among the primary sources. This situation demonstrates that gender inequalities have a culture-specific and multi-layered structure.

Türkiye ranks among the countries with high levels of gender inequality (EIGE 2020, Bardakçı and Oğlak 2022). In Türkiye, women face significant disadvantages not only in terms of income and employment but also in education, health, and household living conditions (Öztürk and Karataş 2020). In this context, despite having similar educational levels, women's labor force participation and political representation remain limited compared to men's (Öcal 2021, Gençtürk 2022, Arslanel 2025).

Gender-based inequalities affect not only women's positions in the labor market and political sphere but also their access to health services and their experiences in care processes (Karacan and Gökçe 2020). Therefore, understanding the multidimensional nature of gender inequality requires going beyond economic and political indicators alone. This enables a better understanding of the complex mechanisms that shape daily life and care processes (Kaya and Yerli 2024).

In Türkiye, gender inequality is primarily assessed through dimensions such as gender roles, educational opportunities, economic status, or discrimination (Karacan and Gökçe 2020, Şahin and Bayhan 2020, Güneş and Binboğa 2021, Berberoğlu 2024). It is observed that the measurement tools used in Türkiye primarily focus on gender equality (Şanlı and Karakuş 2025), gender perceptions (Altınova and Duyan 2013), and gender stereotypes (Tarhan et al. 2024). This situation indicates that the multidimensional nature of gender inequality is not adequately reflected and highlights the need for a multidimensional gender inequality perception scale.

A review of the international literature reveals that the Multidimensional Gender Inequality Perception Scale (MuGIPS), developed by Schwartz-Salazar and colleagues (2024), enables the assessment of the multidimensional structure of gender inequality. This study aims to adapt the Turkish version of the scale (MuGIPS-TR) to a Turkish sample and to evaluate its validity and reliability. Accordingly, the study aims to adapt the Turkish version of the MuGIPS-TR and examine its psychometric properties. The study predicts that the scale's original four-factor structure will be confirmed in the Turkish sample and that it will demonstrate acceptable levels of validity and reliability.

The adaptation of the MuGIPS into Turkish will help address the lack of measurement tools to assess perceptions of multidimensional gender inequality in Türkiye. Making the scale available as a valid and reliable measurement tool could help assess gender inequality across different dimensions and support research in this field.

Method

A cross-sectional research design was used in this study. The processes of adapting the scale to the Turkish sample were conducted in accordance with the guidelines provided by relevant sources that guide researchers in scale development and adaptation processes (WHO 2017, Boateng et al. 2018, Carpenter 2018). The COSMIN checklist was considered in reporting the results. Within the scope of the COSMIN checklist, content validity, construct validity, internal consistency, and test-retest reliability were specifically evaluated. The scale's translation, expert review, pilot testing, DFA, Cronbach's alpha, and test-retest processes were reported in accordance with COSMIN recommendations (Table 1).

Table 1. Characteristics evaluated within the COSMIN framework

Characteristic evaluated	Evaluation method
Content validity	Expert opinion and content validity index
Construct validity	Confirmatory factor analysis
Internal consistency	Cronbach's alpha coefficient
Test-retest reliability	Test-retest method

Sample

The study included individuals aged 18–65 who were literate, capable of using a smartphone or tablet, and who provided online consent to participate. Since all questions on the data collection form were marked as mandatory, no records containing missing data were generated. In scale adaptation studies, it is recommended that the sample size be 5–10 times the number of items (Boateng et al. 2018, Carpenter 2018) or that at least 300 participants be included (White 2022). In this study, 322 participants were recruited, exceeding the recommended sample size for a 25-item scale. This sample size was deemed sufficient for confirmatory factor analysis and psychometric evaluations.

Procedure

The study was conducted in accordance with the principles of the Declaration of Helsinki and following ethical approval from the Akdeniz University Medical Scientific Research Ethics Committee, as per its decision dated July 25, 2024, and numbered TBAEK-508. Prior to initiating the online data collection process, participants were provided with an informational text explaining the study's purpose, scope, the population involved, how data would be stored and used, and that participation was voluntary. Participants who checked the consent box were included in the study. Additionally, permission was obtained via email from the scale's original author, Sofía Schwartz-Salazar, to begin the Turkish adaptation of the scale.

This study was conducted between August and November 2024 using an online survey method and non-probability sampling. Since achieving an adequate sample size for evaluating psychometric properties is a primary objective in scale adaptation studies, non-probability sampling is recommended for such research (Boateng et al. 2018, Carpenter 2018).

The link to the Google Forms-based online survey created by the researchers was shared via social media platforms and the WhatsApp app. The survey used in the study consisted of a sociodemographic data collection form and the latest Turkish adaptation of the MuGIPS. An IP restriction was implemented to prevent participants from filling out the same form more than once. The research data was stored in a password-protected electronic environment accessible only to the researchers and was safeguarded in accordance with privacy principles. The data collection process, including participant informed consent and data management, was conducted by the researchers.

Measures

Socio-Demographic Data Collection Form

The socio-demographic data collection form includes four questions regarding participants' age, gender, marital status, and educational background.

Multidimensional Gender Inequality Perception Scale

In this study, data were collected using the Turkish adaptation of the Multidimensional Gender Inequality Perception Scale (MuGIPS) developed by Schwartz-Salazar et al. (2024). The purpose of the scale is to assess individuals' perceptions of multidimensional gender inequality. The authors initially developed the scale as a six-dimensional construct encompassing health, violence, household chores and care, education, economy/work, and public sphere/representation; however, confirmatory factor analysis indicated a four-dimensional structure. The scale consists of 4 subscales and 25 items. These sub-dimensions are as follows:

1. Health (2 items: 1–2),
2. Violence (6 items: 3–8),
3. Household work and caregiving (3 items: 9–11),
4. Public sphere and power (14 items: items 12–25).

The scale uses a 7-point Likert scale (1 = never, 7 = always), with a possible score range of 25 to 175. Higher scores on the scale indicate that individuals perceive gender inequalities more acutely across various areas of life.

Scale Adaptation and Determination of Psychometric Properties

The process of adapting the MuGIPS-TR into Turkish was conducted in accordance with the scale-adaptation guidelines recommended by the World Health Organization (WHO, 2017). Each stage outlined in the guidelines was followed:

Language Validity

To ensure that valid measurement tools can be adapted to different cultural contexts, language validity must first be established (Karaçam 2019). Language validity encompasses not only the process of direct translation but also the restructuring of each scale item to preserve its meaning in the target culture. Accordingly, a multi-stage process involving translation, back-translation, and expert reviews was conducted to ensure the scale's language validity. At each stage, the meaning, conceptual equivalence, and cultural appropriateness of the items were considered (WHO 2017). Details regarding the adaptation process are provided below.

Translation–Back-Translation

A structured translation-back-translation method was applied in accordance with the translation guidelines recommended in the international literature (WHO 2017). First, the items in the original scale were translated into Turkish by two independent experts knowledgeable about the subject and fluent in Turkish. Subsequently, these translations were combined to create a unified text in terms of meaning and language. In the second stage, the resulting Turkish version was translated back into English by two experts proficient in Turkish but native English speakers. The text produced through back-translation was compared with the original scale to assess consistency in terms of meaning, conceptual overlap, and cultural appropriateness. As a result of these comparisons, revisions were made to items translated into the target language when meaning shifted or the text was out of context.

This multi-step approach is considered a fundamental safeguard for ensuring validity across different cultural contexts (Sousa and Rojjanasrirat 2011). In this study, both the grammatical accuracy of the MuGIPS-TR items and their conceptual integrity with the original scale have been ensured.

Expert Review

To ensure the content and cultural validity of the scale items, 9 experts conducted a review—each holding at least a doctoral degree, fluent in both Turkish and English, and with experience in gender studies and/or scale development or adaptation—following the translation and back-translation process. Eight of the

experts work in obstetrics and gynecology, while one works in nursing administration. Necessary revisions were made to scale items that did not fully align with the cultural context or were ambiguous in meaning, based on the experts' opinions. The Davis technique was used to assess the content validity of the MuGIPS-TR quantitatively. Accordingly, the content validity index (CVI) was calculated as 0.89.

Pilot Study

To evaluate the comprehensibility and feasibility of the MuGIPS-TR in the target language, a pilot study was conducted using a face-to-face survey method with 30 adults. The individuals participating in the pilot study were not included in the research sample. As a result of the pilot study, it was determined that the scale items were understandable to the participants and that the scale was ready for implementation. Accordingly, it was decided to use the MuGIPS-TR in its current form during data collection.

Statistical Analysis

SPSS 25.0 and AMOS 24.0 were used to analyze the research data. The sociodemographic characteristics of the participants were presented as frequencies, percentages, means, and standard deviations. To assess the construct validity of the scale, a confirmatory factor analysis was conducted; Cronbach's alpha was used to assess internal consistency, and adjusted item-total correlations were calculated as part of the item analysis. To evaluate test-retest reliability, the test-retest method was employed, and the relationship between the two administrations was examined using Pearson's correlation.

Since the factor structure of the scale's original form is clearly defined both theoretically and empirically, it is recommended that exploratory factor analysis (EFA) not be conducted in scale adaptation studies. It is noted that conducting confirmatory factor analysis (CFA) to validate the existing factor structure across different samples is directly appropriate (Kline, 2023). Although some studies recommend using AFA and DFA together (Flora and Flake 2017), this approach is typically preferred in new scale development studies or when the factor structure is unclear. In this study, however, the scale's original four-factor structure was evaluated using DFA to test its validity in a Turkish sample. Since the factor structure of the scale's original form is explained both theoretically and empirically, DFA was directly preferred, as recommended in cross-cultural adaptation studies. This approach is considered important for preserving the scale's original structure and maintaining cross-cultural comparability.

The following fit indices were used: Chi-square/degrees of freedom (CMIN/DF), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Normed Fit Index (NFI), and Tucker-Lewis Index (TLI) (Tabachnick and Fidell 2013). These fit indices were selected to evaluate model fit from different perspectives. While RMSEA evaluates the absolute fit of the model, CFI, NFI, and TLI provide information on comparative fit. GFI and AGFI are among the commonly used fit indices for assessing the extent to which the model fits the data (Tabachnick and Fidell 2013, Kline 2023).

To assess the scale's reliability, Cronbach's alpha was used to measure internal consistency, and the test–retest method was used to assess test–retest reliability. Test–retest reliability was evaluated by calculating the correlation coefficient between the two administrations.

Results

The 322 participants included in the study ranged in age from 18 to 64, with a mean age of 34.48 ± 9.48. Of the participants, 48.1% were married, and 78.6% were female. Regarding educational level, 50.3% of participants held a bachelor's degree, while 44.1% held a graduate degree.

The mean total score on the MuGIPS-TR was relatively high among the participants. Upon examining the sub-dimensions, it was determined that the highest average is in the public sphere and power dimension, while the other dimensions remain at relatively lower levels. An increase in the scale score indicates that individuals' perceptions and awareness of multidimensional gender inequality have increased (Table 2).

Table 2. Subdimensions, number of items, score ranges, and mean scores of the MuGIPS-TR				
Subdimensions	**Number of Items**	**Min-Max.**	**Mean**	**SD**
Health	2	2-14	6.67	3.32
Violence	6	6-42	36.52	5.37
Household work and caregiving	3	3-21	18.11	2.93
Public sphere and power	14	14-98	74.14	14.98
MuGIPS-TR	25	25-175	135.45	22.64

Mean: Mean, SD: Standard deviation, Min.: Minimum, Max.: Maximum. The number of items in the subscales is consistent with the scale's original factor structure (Schwartz-Salazar et al. 2024)

Item Analyses

Item analyses were conducted to assess the quality and discriminant validity of the MuGIPS-TR items. Item-total score correlation coefficients above 0.30 were considered acceptable in terms of item discriminant validity. The analysis revealed that the corrected item-total correlation coefficients for the scale items ranged from 0.38 (Item 1) to 0.78 (Item 23); no difference in Cronbach's alpha was observed when an item was removed (Table 3).

Table 3. Item analysis results for the MuGIPS-TR			
Items	**Mean (SD)**	**Adjusted item-total correlation**	**Cronbach's alpha value when the item is removed**
Item 1	3.37 (1.75)	0.38	0.94
Item 2	3.31 (1.81)	0.45	0.94
Item 3	6.31 (1.07)	0.60	0.94
Item 4	5.94 (1.25)	0.65	0.94
Item 5	6.29 (0.94)	0.59	0.94
Item 6	5.87 (1.31)	0.59	0.94
Item 7	6.18 (1.13)	0.61	0.94
Item 8	5.93 (1.07)	0.68	0.94
Item 9	5.88 (1.46)	0.49	0.94
Item 10	6.20 (1.01)	0.52	0.94
Item 11	6.05 (1.12)	0.62	0.94
Item 12	5.24 (1.75)	0.35	0.94
Item 13	5.93 (1.08)	0.62	0.94
Item 14	5.66 (1.32)	0.66	0.94
Item 15	5.33 (1.51)	0.74	0.94
Item 16	5.56 (1.44)	0.73	0.94
Item 17	5.11 (1.55)	0.76	0.93
Item 18	5.78 (1.23)	0.62	0.94
Item 19	5.59 (1.32)	0.50	0.94
Item 20	4.26 (1.86)	0.68	0.94
Item 21	5.11 (1.64)	0.72	0.94
Item 22	5.08 (1.60)	0.77	0.93
Item 23	5.10 (1.58)	0.78	0.93
Item 24	4.58 (1.77)	0.65	0.94
Item 25	5.82 (1.32)	0.57	0.94

SD: Standard Deviation

Validity Analyses

In this study, DFA was applied to evaluate the validity of the current MuGIPS-TR structure in the Turkish sample.

Confirmatory Factor Analysis

In the confirmatory factor analysis, the original four-factor structure of the MuGIPS-TR served as the basis for evaluation in AMOS. The fit indices from the CFA indicated that the overall model fit was marginally

acceptable (Table 4). The fit indices revealed that the model demonstrated overall fit at a marginally acceptable level.

Table 4. Fit indices for the Confirmatory Factor Analysis of the MuGIPS-TR

Scale fit indices	Excellent fit indices	Acceptable fit indices	Scale fit indices	Results
CMIN/DF	0≤χ2/df≤3	3≤χ2/df≤5	4.96	Acceptable fit
RMSEA	0.0≤RMSEA≤0.05	0.05 < RMSEA ≤ 0.08	0. 08	Acceptable fit
GFI	0.90≤GFI	0.80≤GFI	0. 80	Borderline acceptable fit
NFI	0.95≤NFI	0.80≤NFI	0. 86	Borderline acceptable fit
TLI	0.90≤TLI	0.80≤TLI	0. 87	Borderline acceptable fit
AGFI	0.90≤AGFI	0.80≤AGFI	0. 80	Borderline acceptable fit
CFI	0.95≤CFI	0.85≤CFI	0. 86	Borderline acceptable fit

CMIN/DF: Chi-square/Degrees of Freedom; RMSEA: Root Mean Square Error of Approximation; GFI: Goodness-of-Fit Index; AGFI: Adjusted Goodness-of-Fit Index; NFI: Normed Fit Index; TLI: Tucker-Lewis Index; CFI: Comparative Fit Index

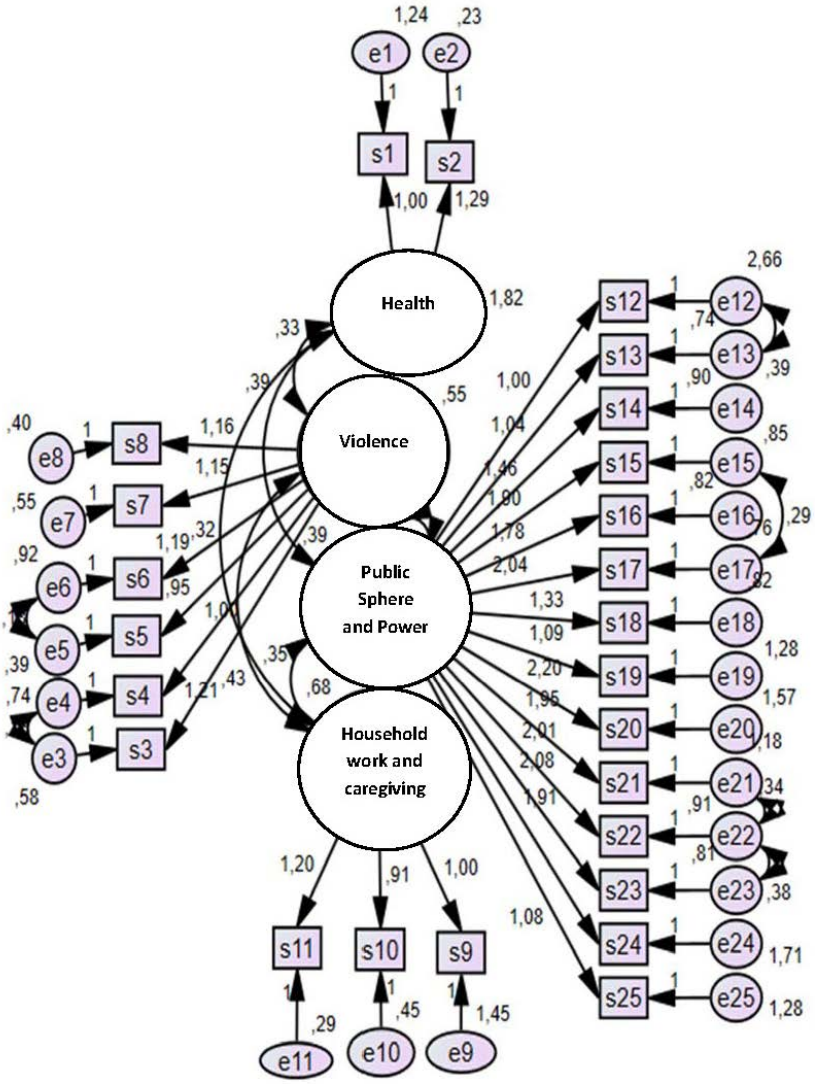


Figure 1. AMOS regression-weighted model for MuGIPS-TR (n=322)

It is recommended to use multiple fit indices together when evaluating model fit (Tabachnick and Fidell 2013, Kline 2023). Accordingly, this study utilized both absolute fit indices (RMSEA, GFI, AGFI) and comparative fit indices (CFI, NFI, TLI) to evaluate model fit from different perspectives. In particular, CFI and TLI values below the ideal cutoffs warrant cautious interpretation of model fit. This approach enables a more comprehensive and reliable assessment of the model's fit with the data (Kline 2023).

In accordance with the modification indices examined during the DFA process, only those between items within the same subscale and showing content similarity were evaluated. Accordingly, modifications were made between two items in the violence subscale and between four item pairs in the public space and power subscale. As a result of these limited modifications, model fit indices improved, and the model reached a marginal level of fit. However, interventions aimed at excessively improving model fit were avoided, and care was taken to preserve the scale's original factor structure. Modifications were made only among items within the same subscale. Specifically, error covariances were defined between similar items expressing experiences of violence within the violence subscale, and between items with similar content regarding social representation, decision-making processes, and power relations within the public sphere and power subscale.

According to the criteria proposed in the literature, an RMSEA value below 0.08 and CFI and TLI values of 0.90 or higher indicate acceptable fit (Carpenter 2018). When the MuGIPS-TR data were evaluated against these criteria, the model generally exhibited a marginal fit. The regression weights for the scale in the DFA model are shown in Figure 1.

Upon examining the standardized factor loadings of the MuGIPS-TR, all items loaded onto their respective subscales with loadings above 0.30. It was determined that the standardized factor loadings of the items in the public sphere and power subscales ranged from 0.35 to 0.74. In contrast, in the other subscales, some items' factor loadings were relatively lower but remained within acceptable limits.

Reliability Analyses

Internal consistency and test-retest reliability analyses were conducted to assess the scale's reliability. In this context, Cronbach's alpha and the test-retest method were used.

Internal Consistency Analysis

The Cronbach's alpha coefficient was calculated to assess the internal consistency of the MuGIPS-TR. In this study, the Cronbach's alpha coefficient for the total scale was 0.93. It was determined that the Cronbach's alpha coefficients for the scale's subscales ranged from 0.73 to 0.92. In the original study, the Cronbach's alpha coefficients for the total scale were 0.97, for the violence subscale 0.90, for the household chores and care subscale 0.82, and for the public space and power subscale 0.95. The Cronbach's alpha reliability coefficients for the total and subscales of the MuGIPS-TR are presented in Table 5.

Table 5. Reliability coefficients for the total and subscales of the MuGIPS-TR

Scale and sub-dimensions	α
MuGIPS-TR	0.93
Health	0.85
Violence	0.88
Household work and caregiving	0.73
Public sphere and power	0.92

α: Cronbach's alpha coefficient

Table 6. Information on the pre-test and post-test results of the MuGIPS-TR (n=30)

Scale and sub-dimensions	MuGIPS-TR (n=30)		Statistics	
	First application $\bar{x}$±SD	Second application $\bar{x}$±SD	r	p
MuGIPS-TR	133.66±24.35	127.80±24.06	0.791	0.000
Health	7.26±3.30	5.77±3.29	0.758	0.000
Violence	36.26±4.58	35.51±6.56	0.853	0.000
Household work and caregiving	16.88±3.05	16.51±3.44	0.841	0.000
Public sphere and power	73.26±16.44	70.00±15.23	0.794	0.000

$\bar{x}$: Mean; SD: Standard Deviation; r: Pearson correlation coefficient; p: Significance level

Test-Retest Method

The test-retest method was used to assess the MuGIPS-TR's test-retest reliability. The same scale link was resent 4 weeks after the first administration to 30 volunteer participants who had participated in the first administration and had provided their contact information, indicating their intention to participate in the second administration. The evaluation revealed that the Pearson correlation coefficients between the two administrations of the scale were positive and statistically significant. The correlation coefficients ranged from 0.75 to 0.85. Findings regarding the test-retest results are presented in Table 6.

Discussion

The literature emphasizes that gender inequality manifests across various dimensions—such as health, violence, domestic work and care, education, employment, and the public sphere—and that these dimensions affect individuals' living conditions and well-being in different ways (Schwartz-Salazar et al. 2024, Hua 2024). Therefore, gender inequality should be addressed not through a single indicator but within a multidimensional framework. This multidimensional approach allows for a more comprehensive understanding of the different manifestations of inequality reflected in individuals' lives (Ciaffoni et al. 2025). Furthermore, it is important not only to identify the existence of inequality but also to determine in which areas and to what extent it is experienced. In this context, it is necessary to examine the psychometric validity and reliability of measurement tools capable of assessing perceptions of multidimensional gender inequality across diverse cultural contexts (Hua 2024). Accordingly, this study evaluates the psychometric properties of the MuGIPS-TR.

The Davis technique demonstrates expert consensus that the items are related to the construct being measured, with values of 0.80 or higher considered acceptable (Almanasreh et al. 2019). In this study, it was determined that the content validity coefficients for all items of the scale were above the acceptable level. The findings indicate that the MuGIPS-TR form adequately represents the construct of interest and is suitable for the Turkish sample.

In scale adaptation studies, a pilot study is conducted to assess the scale's applicability to the target population and the clarity of the items. While the recommended number of participants in pilot studies varies depending on the context and the study's purpose, it is generally recommended to include at least 30 participants (Boateng et al. 2018). The findings from the pilot study in this research indicate that the MuGIPS-TR items are understandable and feasible.

The literature recommends using adjusted item-total score correlations in item analysis studies (Karaçam 2019). This is because unadjusted item-total correlations can yield higher values than they actually are. After all, the item score is added to the total score. Adjusted item–total score correlation coefficients above 0.30 are considered an acceptable criterion for retaining an item on the scale (Streiner et al. 2015). It was determined that the adjusted item–total correlation coefficients obtained in this study were above 0.30 for all items. This finding is important as it indicates that the scale items are sufficiently related to the construct being measured and contribute to the scale's overall structure.

Since the four-factor structure of the original MuGIPS (Schwartz-Salazar et al., 2024) is clearly defined both theoretically and empirically, Direct Factor Analysis (DFA) was applied to test its validity in the Turkish sample. In scale adaptation studies, it is crucial that the measurement model maintains the same structure across cultures and supports cross-cultural comparisons. The DFA findings from this study indicate that the scale's original structure is preserved in the Turkish sample and that the model demonstrates acceptable fit. DFA allows for testing the extent to which the observed variables represent the theoretically defined measurement model (Flora and Flake 2017).

Within the framework of DFA, it is recommended that multiple fit indices be reported together to allow for a more comprehensive evaluation of model fit (Goretzko et al. 2023). The DFA findings from this study indicate that the MuGIPS-TR possesses a factor structure consistent with its original scale structure. It has been noted that the fit indices reported in the original study were at good and acceptable levels (Schwartz-Salazar et al. 2024), and the findings of this study also indicate similar model fit. Additionally, it

was determined that the standardized factor loadings of the scale's items exceeded the lower threshold of 0.30 recommended in the literature (Boateng et al. 2018; Carpenter 2018). These results indicate that the four-factor structure of the MuGIPS -TR is supported in the Turkish sample as well. However, it is known that DFA results may be influenced by sample size. The sample size used in this study meets the lower limits recommended in the literature; however, it is believed that studies with larger, more heterogeneous samples could yield a more robust evaluation of model fit indices.

The factor loadings obtained through DFA reflect the strength of the relationship between the scale items and the latent structures they are intended to measure. In the literature, an absolute value of factor loadings above a certain threshold indicates that the observed variable adequately represents the relevant latent factor (Tang et al. 2024). It is generally noted that factor loadings above the 0.30–0.40 range are considered acceptable (Boateng et al. 2018, Carpenter 2018). In this study, the standardized factor loadings for the MuGIPS-TR items ranged from 0.35 to 0.74. These values indicate that the scale items adequately represent the relevant subscales and support the scale's factor structure. In particular, items with low factor loadings (around 0.35) were re-evaluated. Since these items were deemed important to the scale's conceptual integrity and their removal did not result in a significant improvement in Cronbach's alpha or model fit, it was decided to retain them.

These findings are also consistent with the factor loadings reported by Schwartz-Salazar et al. (2024). It was found that some items on the scale had relatively low factor loadings (e.g., around 0.35). However, the literature indicates that factor loadings of 0.30 or higher are considered acceptable (Boateng 2018, Carpenter 2018). Additionally, a sensitivity analysis was conducted to assess the impact of items with low factor loadings on the model. However, removing these items could weaken the scale's theoretical coherence and yield only limited improvements in model fit. Therefore, the original four-factor structure was retained.

Similarly, it has been reported that the DFA fit indices for gender scales developed or adapted in Türkiye are mostly at acceptable or borderline acceptable levels (Altınova and Duyan 2013, Tarhan et al. 2024, Şanlı and Karakuş 2025). The factor loadings and fit indices obtained in this study generally align with those in the aforementioned studies. However, the relatively low factor loadings observed for some items may be due to differences in cultural context and sample characteristics.

In this study, Cronbach's alpha was calculated to assess the MuGIPS-TR's reliability. Cronbach's alpha allows for the evaluation of the internal consistency of scale items and the extent to which the scale reliably measures the intended construct. The literature indicates that a Cronbach's alpha coefficient of 0.70 or higher is considered an acceptable level of reliability (Yun et al. 2023). It was determined that all Cronbach's alpha coefficients for the total scale and subscales obtained in this study were above 0.70. These findings are consistent with the values reported by Schwartz-Salazar et al. (2024). The results support the conclusion that the scale's items demonstrate sufficient internal consistency for the construct and that the scale is a reliable measurement tool in the Turkish sample. Additionally, the high Cronbach's alpha coefficient for the total scale (0.93) is comparable to the reliability coefficient reported in the original study (0.97). The attainment of high internal consistency coefficients in similar scale adaptation studies supports the notion that scales can provide reliable measurements across different cultural contexts (Altınova and Duyan 2013, Tarhan et al. 2024, Şanlı and Karakuş 2025).

Upon examining the MuGIPS-TR subscales, an imbalance in the number of items across the factors is evident. In particular, it is noteworthy that while the "public sphere and power" subscale consists of 14 items, the "Health" subscale consists of only 2 items. However, this structure is consistent with the scale's original form, and it is recommended that the original factor structure be preserved in scale adaptation studies. The literature indicates that two-item factors are acceptable, particularly when they are theoretically meaningful and have sufficient internal consistency (Tabachnick and Fidell 2013). In this study, the Cronbach's alpha coefficient for the "Health" subscale was found to be 0.85. However, it is known that two-item factors may be more susceptible to measurement error and may represent the structure only to a limited extent (Eisinga et al. 2013). Therefore, it is recommended that the "Health" subscale, in particular, be re-evaluated in future studies with different samples and supported by new items.

The test-retest method involves measuring the same individuals with the same assessment tool at two different time points. Test-retest reliability is typically assessed using correlation coefficients; values above 0.70 are considered acceptable (Polit and Beck 2017; Min et al. 2022). In this study, the scale was re-administered to 30 participants four weeks apart to assess its test-retest reliability. The correlation coefficient was r = 0.79 and was statistically significant (p <0.001). Correlation coefficients for the subscales were similarly high and significant. These findings indicate that the scale maintains the ranking among individuals across repeated measurements and provides consistent measurements over time. However, the differences observed in average scores between the two administrations suggest that environmental and social factors may influence individuals' perceptions of gender inequality over time.

This study has certain limitations. First, collecting data via social media and online surveys increases the likelihood of selection bias in the sample and limits the generalizability of the findings. The fact that the vast majority of participants are women and highly educated individuals weakens the sample's ability to represent Türkiye's general population. In particular, the insufficient representation of individuals living in rural areas, those with low educational levels, and those with limited digital access in the sample may limit the generalizability of the findings to different sociodemographic groups. This situation limits the ability to evaluate whether the scale's psychometric properties are equally valid across different groups.

Furthermore, the online data collection method carries certain limitations regarding data quality. Although IP restrictions were applied during data collection to prevent duplicate participation, the inability to verify participants' identities and fully validate response accuracy creates a potential limitation to data reliability. Additionally, this method may limit researchers' direct control over the sample and data collection processes. To improve data quality, repeated participation from the same IP address was blocked, incomplete data entry was not permitted, and answering all questions was mandatory. However, the lack of control questions was identified as a major limitation of the online data collection process.

Furthermore, the study's cross-sectional design limits the interpretation of the findings in terms of causal relationships. In addition, because the scale has been tested only in adults, its psychometric properties need to be evaluated across different age groups. Therefore, it is recommended that future studies be replicated with more representative and heterogeneous samples, employ different data collection methods, and, if possible, use longitudinal designs.

Conclusion

This study determined that the Turkish version of the Multidimensional Gender Inequality Perception Scale is a valid and reliable measurement tool. The findings revealed that the scale's original four-factor structure was confirmed in the Turkish sample and that it demonstrated adequate internal consistency and test-retest reliability. It is believed that the MuGIPS-TR is a useful measurement tool for research aimed at assessing perceptions of gender inequality in a multidimensional manner, as well as for use in the nursing field.

In this regard, the MuGIPS-TR is evaluated as a valid and reliable measurement tool for this sample, suitable for use in assessing the various dimensions of gender inequality in research conducted in the health sector—particularly in nursing education and practice—as well as in social work and social policy development processes. In this regard, the study fills an important gap in the literature by assessing perceptions of multidimensional gender inequality within the Turkish context, thereby offering a unique contribution to the field.

References

Almanasreh E, Moles R, Chen T (2019) Evaluation of methods used for estimating content validity. Res Social Adm Pharm, 15:214-221.

Altınova HH, Duyan V (2013) Toplumsal cinsiyet algısı ölçeğinin geçerlik-güvenirlik çalışması. Toplum ve Sosyal Hizmet, 24:9-22.

Arslanel MN (2025) Ataerkil düzende kadınların siyasal katılımı meselesi. J Int Manag Educ Econ Perspect, 13:76-84.

Bardakçı Ş, Oğlak S (2022) Toplumsal cinsiyet eşitsizliği endeksi ve Türkiye. Toplumsal Politika Dergisi, 3:71-90.

Begall K, Grunow D, Buchler S (2023) Multidimensional gender ideologies across Europe: evidence from 36 countries. Gender Soc, 37:177-207.

Berberoğlu B (2024) Türkiye'de eğitimde istihdam açısından toplumsal cinsiyet eşitliği kavramının incelenmesi. J Soc Humanit Adm Sci, 6:2271-2289.

Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL (2018) Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health, 6:149.

Carpenter S (2018) Ten steps in scale development and reporting: a guide for researchers. Commun Methods Meas, 12:25-44.

Ciaffoni S, Rubini M, Moscatelli S (2025) The unbearable weight of gender inequalities: development and validation of the Social Treatment and Experiences of Women (STEW) scale. Sex Roles, 91:3.

WHO (2017) Process of translation and adaptation of instruments. https://www.who.int/substance_abuse/research_tools/translation/en/ (Accessed 12.01.2026).

Eisinga R, Grotenhuis M, Pelzer B (2013) The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? Int J Public Health, 58:637-642.

EIGE (European Institute for Gender Equality) (2020) Gender equality indices in the Western Balkans and Turkey. https://eige.europa.eu/gender-equality-indices-western-balkans-and-turkey (Accessed 12.01.2026).

Flora D, Flake J (2017) The purpose and practice of exploratory and confirmatory factor analysis in psychological research: decisions for scale development and validation. Can J Behav Sci, 49:78-88.

Gençtürk Z (2022) Türkiye'de kadının işgücüne katılımını etkileyen faktörler. Süleyman Demirel Üniv Sosyal Bilimler Enstitüsü Dergisi, 43:361-394.

Goretzko D, Siemund K, Sterner P (2023) Evaluating model fit of measurement models in confirmatory factor analysis. Educ Psychol Meas, 84:123-144.

Güneş ST, Binboğa G (2021) Toplumsal cinsiyet eşitsizliği bağlamında kadın yoksulluğu ve sürdürülebilir kalkınma: Türkiye ölçeğinde bir inceleme. Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 23:1589-1615.

Hua W (2024) Gender inequality: a multi-perspective analysis. Lecture Notes in Education Psychology and Public Media, 42:200-204.

Karacan E, Gökçe S (2020) Toplumsal cinsiyet eşitsizliği ve kadın sağlığı. Sosyal Politika ve Sosyal Hizmet Çalışmaları Dergisi, 1:39-59.

Karaçam Z (2019) Ölçme araçlarının Türkçeye uyarlanması. J Midwifery Health Sci, 2:28-37.

Kaya SP, Yerli K (2024) Factors affecting nursing students' attitudes toward child brides. Journal of Education and Research in Nursing, 21:325-333.

Kline RB (2023) Principles and Practice of Structural Equation Modeling. New York, Guilford Press.

Min S, Duangthip D, Gao S, Detsomboonrat P (2022) Quality of the adaptation procedures and psychometric properties of the SOHO-5 scale: a systematic review. Qual Life Res, 32:1537-1547.

Öcal AT (2021) Türkiye'de işgücü piyasasında toplumsal cinsiyet eşitsizliği. World Women Conference-II, 11-12 Şubat 2021, Bakü, Azerbaycan. Kongre bildiri kitabı, 2:813-825.

Öztürk S, Karataş M (2020) Çalışma yaşamında kadınların karşılaştığı sorunlar. Hak İş Uluslararası Emek ve Toplum Dergisi, 9(23):103-120.

Polit DF, Beck CT (2017) Nursing Research: Generating and Assessing Evidence for Nursing Practice. Philadelphia, Wolters Kluwer.

Schwartz-Salazar S, García-Sánchez E, Martínez R, Rodríguez-Bailón R (2024) Development and validation of the Multidimensional Gender Inequality Perception Scale (MuGIPS). PLoS One, 19:e0301755.

Sousa VD, Rojjanasrirat W (2011) Translation, adaptation, and validation of instruments for cross-cultural research. J Eval Clin Pract, 17:268-274.

Stefani S, Prati G (2020) Are dimensions of gender inequality uniformly associated with human values? Eur J Psychol, 17:92-102.

Streiner DL, Norman GR, Cairney J (2015) Health Measurement Scales: A Practical Guide to Their Development and Use. 5th ed. New York, Oxford University Press.

Şahin S, Bayhan AC (2020) Türkiye'de toplumsal cinsiyet eşitsizliğinin ekonomik boyutu. Lectio Socialis, 4:59-74.

Şanlı B, Karakuş F (2025) Toplumsal Cinsiyet Eşitliği Farkındalık Ölçeği. Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 69:85-99.

Tabachnick BG, Fidell LS (2013) Using Multivariate Statistics. 6th ed. Boston, Pearson.

Tang D, Boker S, Tong X (2024) Are the signs of factor loadings arbitrary in CFA? Struct Equ Modeling, 32:142-154.

Tarhan S, Ünal F, Köksal EÇ (2024) Toplumsal cinsiyet kalıp yargıları ölçeği geliştirme çalışması. Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 62:353-367.

Tekgüç H, Akbulut B (2021) A multidimensional approach to the gender gap in poverty: Turkey application. Fem Econ, 28:119-151.

White M (2022) Sample size in quantitative instrument validation studies. Heliyon, 8:e12223.

Yun VWS, Ulang NM, Husain SH (2023) Measuring the internal consistency and reliability of the hierarchy of controls in preventing infectious diseases on construction sites: the Kuder-Richardson (KR-20) and Cronbach's alpha. J Adv Res Appl Sci Eng Technol, 33:392-405.

Authors Contributions: The author(s) have declared that they have made a significant scientific contribution to the study and have assisted in the preparation or revision of the manuscript
Peer-review: Externally peer-reviewed.
Ethical Approval: Ethical approval for this study was obtained from the Akdeniz University Medical Research Ethics Committee pursuant to Decision No. TBAEK-508 dated July 25, 2024. Participants provided informed consent prior to participating in the study.
Conflict of Interest: No conflict of interest was declared.
Financial Disclosure: No financial support was declared for this study.
Acknowledgment: Permission was obtained via email from Sofía Schwartz-Salazar, the author of the original scale, to begin work on the Turkish adaptation of the scale. Authorsw would like to thank Sofía Schwartz-Salazar, the developer of the MuGIPS, for granting permission to adapt the scale into Turkish for use in this study.

Addendum 1. Multidimensional Gender Inequality Perception Scale (MuGIPS-TR)

Instructions

Below are various statements regarding gender inequalities observed between women and men in our country. Please read each statement carefully and, considering how often these situations occur in our country, select the option that best reflects your opinion.

The scale is scored as follows: 1=Never, 2=Rarely, 3=Sometimes, 4=Occasionally, 5=Often, 6=Very often, 7=Always.

Items	1	2	3	4	5	6	7
1. Kadınlarda aynı sağlık sorunu, erkeklere göre daha az dikkate alınmaktadır.							
2. Kadınlara özgü sağlık sorunları, erkeklere özgü sağlık sorunlarına kıyasla daha az önemsenmektedir.							
3. Kadınlar, gündelik yaşamlarında (sokakta, işte...) saldırıya uğrama olasılıkları nedeniyle erkeklere göre daha fazla güvensizlik duygusu yaşamaktadır.							
4. Kadınlar, kadın oldukları için erkeklerden daha fazla şiddete maruz kalmaktadır.							
5.Kadınlar erkeklere göre daha fazla cinsel saldırıya uğramaktadır.							
6. Kadınlar iş yerinde erkeklere göre daha fazla tacize uğramaktadır.							
7. Kadınlar erkeklere göre daha fazla obje olarak görülmektedir.							
8. Kadınlar, yakın partner ilişkilerinde erkeklere göre daha fazla şiddete maruz kalmaktadır							
9. Kadınlar, çocuklarına bakım verme konusunda, erkeklerden daha fazla sorumluluk üstlenmektedir.							
10. Kadınlar erkeklerden daha fazla ev işi yapmaktadır.							
11. Kadınların aile üyelerine ve yakınlarına bakım verme olasılığı, erkeklere göre daha yüksektir.							
12. Kadınların iş ve yaşam arasında bir denge kurması erkeklere göre daha zordur.							
13. Kadınlar, aileleriyle ilgilenmek için kariyerlerinden vazgeçme konusunda erkeklere göre daha fazla baskı altındadır.							
14. Erkekler, özel ve kamu kurumlarında kadınlara göre daha fazla temsil ve güce sahiptir.							
15. Aynı haklara sahip olmalarına rağmen erkekler, toplumsal olarak kadınlardan daha fazla saygı görmektedir.							
16. Kadınlar kendilerinden bekleneni yerine getirmediklerinde erkeklere göre daha fazla sorgulanmaktadır.							
17. Erkeklerin görüş ve fikirlerine kadınlarınkinden daha fazla değer verilmektedir.							
18. Genel olarak toplumumuzda erkekler, kadınlara göre daha fazla güce sahiptir.							
19. Kız ve erkek çocukları, toplumdaki rolleri konusunda farklı şekilde eğitilmektedir.							
20. Erkekler tarafından alınan üniversite eğitimi, kadınların aldığı üniversite eğitimine göre genellikle daha fazla değer görmektedir.							
21. Kadınlar öğrenimlerini sürdürürken erkeklere göre daha fazla engelle karşılaşmaktadır.							
22. Kadınlar iş bulma konusunda erkeklere göre daha fazla engelle karşılaşmaktadır.							
23. Kadınlar, toplum tarafından değerli görülen işlere erişim konusunda erkeklere göre daha fazla engelle karşılaşmaktadır.							
24. Toplumumuzda cinsiyete dayalı bir ücret farkı vardır, yani aynı işi yapmalarına rağmen erkekler kadınlardan daha fazla ücret almaktadır.							
25.Erkeklerin daha iyi koşullara sahip bir işte çalışma olasılığı kadınlara göre daha yüksektir.							

Scoring

The MuGIPS-TR consists of a total of 25 items and four subscales:

1. Health (Items 1–2)
2. Violence (Items 3–8)
3. Household work and caregiving (Items 9–11)
4. Public sphere and power (Items 12–25)

There are no reverse-scored items on the scale. Each item is scored on a scale of 1–7. Subscale scores are calculated by summing the relevant items. The total score range for the scale is 25 to 175. Higher scores indicate that the individual perceives gender inequalities more acutely across various areas of life.

Permission must be obtained from the scale developers to use the scale.