



The Female Genital Self-Image Scale (FGSIS): cross-cultural adaptation and validation of psychometric properties within a Turkish population

Aski Ellibes Kaya¹ · Murat Yassa² · Ozan Dogan³ · Alper Basbug⁴ · Cigdem Pulatoglu⁵ · Eray Caliskan⁶

Received: 27 February 2018 / Accepted: 6 June 2018 / Published online: 30 June 2018
© The International Urogynecological Association 2018

Abstract

Introduction and hypothesis Women's perceived satisfaction from their own genital appearance is linked to genital image and sexual esteem. A comprehensive and easy to use scale to measure self-image was scarce in the literature. It was aimed in the present study to complement cross-culturally adapted and validated into Turkish version of the Female Genital Self-Image Scale (FGSIS) and to assess its psychometric properties.

Methods After cross-cultural adaptation, the Turkish version of the FGSI, Female Sexual Distress Scale-Revised (FSDS-R), and Female Sexual Function Index (FSFI) were administered to 461 female participants. Content/face validity, exploratory, and confirmatory factor analysis, internal consistency, and reliability were appropriately assessed. Predefined and specific hypotheses were formulated for construct validity.

Results Our findings indicated excellent content/face validity, sufficient internal consistency (Cronbach's alpha 0.818), and test-retest reliability [intra-class correlation coefficient (ICC) 0.951]. Construct validity was demonstrated by proving the hypothesis that participants who have performed at least one vaginal/clitoral masturbation for the last month reported significantly higher FGSIS scores compared with those who abstained ($Z = -6.37, p < 0.001$). Factor analyses formed one factor structure. In the proposed two-factor construct, all seven items demonstrated good to high correlations with their subdomains and lower correlations with the other domain, indicating sufficient convergent validity.

Conclusions The FGSIS was successfully validated for use in the Turkish population. The scale exhibited strong psychometric properties to assess perceived female genital image. It might be reliably used in genital cosmetic surgeries and in a variety of gynecologic conditions.

Keywords Genital perception · Female Genital Self-Image Scale · Female sexual dysfunction · Sexual well-being

✉ Aski Ellibes Kaya
askiellibes@hotmail.com

Murat Yassa
murat.yassa@gmail.com

Ozan Dogan
ozandogan02@hotmail.com

Alper Basbug
dralper23@gmail.com

Cigdem Pulatoglu
cigdempulatoglu@gmail.com

Eray Caliskan
dreraycaliskan@yahoo.com

¹ Department of Obstetrics and Gynecology, Duzce University Faculty of Medicine, Yorukler Mah., Konuralp yerleskesi, 81620 Merkez/ Duzce, Turkey

² Fatih Sultan Mehmet Training and Research Hospital, Istanbul, Turkey

³ Sisli Etfal Training and Research Hospital, Istanbul, Turkey

⁴ Duzce University Faculty of Medicine, Duzce, Turkey

⁵ Bayburt Government Hospital, Bayburt, Turkey

⁶ Bahcesehir University Faculty of Medicine, Istanbul, Turkey

Introduction

Sexual esteem is a component of psychological and behavioral aspects of sexual and perceived physical well-being [1]. Women's perceived satisfaction from their own genital appearance is linked to genital image, sexual esteem, and sexual satisfaction [1]. A higher-body appreciation reflecting a better body image represents functioning sexuality even regardless of body size [2]. Women satisfied with their own body image report more sexual activity, orgasms, and confidence with their sexual life [3]. A negative body image was found to relate with the need of physical attractiveness [3] and a positive body image was strongly correlated with sexual functioning and satisfaction [4]. High esteem for one's body can be defined as satisfaction with personal and interpersonal relations and can strongly predict one's sexual satisfaction [4].

The level of perceived body image is also known to affect very common clinical scenarios, although this relationship is mostly underestimated or goes unnoticed by clinicians. For instance, low body image at early pregnancy and postpartum periods impair sexual function [5]. On the other hand, surgical treatment of pelvic organ prolapse (POP) can recover sexual function through body image perception [6]. Women with low body image tend to delay their regular gynecological examinations, which may cause serious public health concerns. A strong relationship was found between genital self-image with attending at least one gynecological examination during the past 2 years [7].

Women find the appearance of genitalia to be very important, and increasingly more women hold concern with the appearance, probably due to the influence of media [8]. It has been suggested women with low sexual satisfaction may benefit from current treatment modalities that target the specific aspects of body image [4]. Female elective genital cosmetic surgeries have become increasingly popular; however, valid and inclusive body image scales are necessary to describe the need for or success of treatment. Most commonly used scales to measure sexual functioning, such as the Female Sexual Function Index (FSFI), seem inadequate in capturing all aspects of this issue. Women who underwent elective genital cosmetic surgeries in the study of Goodman et al. did not present significant sexual dysfunction prior to surgeries or positive postoperative response in expected subdomains when measured by the FSFI [9].

A quick and easy to use scale to measure body image is scarce in the literature. The Female Genital Self-Image Scale (FGSIS) was developed by Herbenick et al. to measure a woman's feelings toward her own genitals in a broader spectrum, including genital appearance and odor, with intrapersonal and interpersonal settings [10]. The validation of FGSIS among female college students and at a national level strengthens its generalizability and structure [7, 11]. Aims of this study were to: (1) complement the cross-culturally

adapted Turkish version of the FGSIS; (2) assess the psychometric properties of this measure in a sample of generalizable outpatient setting; and (3) further validate the structure of the FGSIS by factor analysis and testing specific independent hypotheses according to original authors' suggestions.

Materials and methods

Cross-cultural adaptation procedure

The recommendations of the Translation and Cultural Adaptation (TCA) group and the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist were followed for the cross-cultural adaptation and validation of the Turkish version of the FGSIS [12–14]. The sequence of steps used was as follows: forward-translation, synthesis, back-translation, expert committee review, pretesting and cognitive debriefing, finalization, proofreading, and final report.

After assurance of experiential and conceptual equivalence, discrepancies between the original, forward-translation, and back-translated versions were discussed by a committee of experts composed of seven professionals (three gynecologists, one clinician with a special interest of sexual medicine, one psychiatrist working in the field of sexuality, one postgraduate nurse working in the gynecology/sexual medicine setting, and one language professional). One major and one minor revision were required; changes were carefully noted, and a final Turkish version of the FGSIS was produced.

Ten individuals were involved in the pretesting and cognitive debriefing stage to test alternative wording and understandability, interpretation, and cultural relevance of the translation. Respondents were native speakers who were believed to adequately represent the target population in terms of age and education. Pretesting stage was repeated after minor revision of one item. Less than 3 min was required for self-administration of the questionnaire.

The cognitive debriefing was assessed, and content validity was graded by the expert committee with analysis of the relevance of each item. Face validity measured researchers' and patients' comprehension and acceptance of items of the pretest sample. After finalization of the scale, proofreading of the final translation was carried out, and no errors were found. The final report, which clearly explains the reasons for all translation decisions and wording choices for cultural adaptation, was written by the head of the expert committee to inform future translations of the same instrument so they can be harmonized with the previously developed versions in other languages [15].

Study design and population

The study was conducted in a medical facility between March and October 2017. Questionnaires were administered to 461 female participants on a volunteer basis who were believed to represent the general population. Participants comprised a wide range of health workers, including doctors, medical students, nurses, secretaries, staff members, and their relatives and friends. Five patients (two cervical cancer, one uterine cancer, and one breast cancer) were excluded due to malignancies. Norm values of a reference general population and of relevant subgroups of participants who were expected to differ in scores were provided to assess interpretability. It was suggested that investigators provide mean and standard deviation (SD) scores of at least four relevant subgroups of patients with regard to obtaining information regarding what change would be considered clinically meaningful [14]. Content validity included content validity of individual items (I-CVI), content validity of overall scale (S-CVI) scores, and floor–ceiling effect. Face validity, which indicates whether the questionnaire makes sense to patients and whether all important and relevant domains were assessed at pretesting and cognitive debriefing stages. A reliability assessment of internal consistency using Cronbach's alpha and exploratory and confirmatory factor analyses were performed. Reproducibility was assessed in 32 individuals at a 2-week interval. Bland-Altman plot was analyzed to describe adequate parameter of agreement.

A total of four independent and specific hypotheses were formed. First, it was hypothesized that measure items should converge on the total score of the same construct. Second, measures of different constructs, such as the Female Sexual Distress Scale-Revised (FSDS-R) and the Female Sexual Function Index (FSFI) should not load on the same factors to prove that they are distinguishable constructs to provide evidence that items on the other constructs discriminate. Third, it was hypothesized that women who performed at least one vaginal and/or clitoral masturbation over the past month should represent higher self-image in comparison with participants not performing masturbation. The rationale for this specific hypothesis to show known-group validity was proven by Shulman and Horne [16]. Last, weak to moderate correlation strength was expected between FGSIS with FSFI and FSDS-R due to different constructs and aspects. The rationale of this hypothesis was confirmed by Goodman et al., that FSFI and body image does not necessarily need to be correlated [9].

The institutional Ethics Committee approved the study (no. 2017–124), and written informed consent was obtained from all participants in this study. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY, USA: IBM Corp., and Analysis of Moment Structures (AMOS) 23.0.0 (IBM,

SPSS Inc., USA) statistical software packages were used to conduct analyses. Specific statistical analyses regarding the psychometric assessment were comprehensively discussed within the relevant results section.

Instruments

The seven-item FGSIS assesses women's feelings and beliefs about their own genitals using a 4-point response scale (strongly agree, agree, disagree, strongly disagree). The scale has established reliability and validity in a convenience sample [10]. Respondents' scores on each item were summed for a total sum score ranging from 7 to 28, with higher scores indicating more positive genital self-image.

The FSFI is a multiple-trait scoring, self-report document used to assess female sexual function during the previous 4 weeks and consists of 19 items that encompass six separate domains: desire, arousal, lubrication, pain associated with vaginal penetration, satisfaction, and orgasm [17].

The FSDS-R assesses different aspects of sexual-activity-related distress in women. The total score, ranging from 0 to 52, can be computed by adding all 13 item scores. Higher scores indicate higher levels of sexual distress [18].

Results

Interpretability

Interpretability was defined as the degree to which one can assign qualitative meaning to quantitative scores. Interpretable demographic features of participants are shown in Table 1.

Content validity

A thorough evaluation of content validity that consists of a survey among experts in the field was performed. Experts uninvolved in checklist validation were asked to evaluate the relevance and comprehensiveness of its items by analyzing the content validity indexes (CVI) before and after changes.

For good content validity, it was determined that the all I-CVI must be >0.83 and the S-CVI average must be ≥ 0.80 based on ratings of item clarity and relevance provided by the six experts; I-CVI values ranged between 0 and 0.82 and S-CVI average was 0.52. Initial values suggest the need for substantial item improvements, and reviewers at face validity identified aspects of the construct that were not adequately covered by the item pool. Minor corrections were made to all questions and a major correction to the fifth, so the mean I-CVI and S-CVI/average values increased to 1.00 and 1.00, respectively, showing sufficient content validity and cross-cultural adaptation processes of the Turkish version of the FGSIS.

Table 1 Interpretable features of the participants

Features	No. participants		FGSIS Item numbers							Total score	ANOVA <i>F</i>
			1	2	3	4	5	6	7		
Age (years)											.886
17–24	84	Mean ±	3.30	3.21	2.75	3.12	3.13	2.75	2.93	3.03	
		SD	0.72	0.78	1.15	0.78	1.03	1.11	0.99	0.68	
25–34	160	Mean ±	3.37	3.36	3.07	3.07	3.41	2.78	3.02	3.15	
		SD	0.74	0.76	1.05	0.94	0.82	1.08	1.12	0.62	
35–44	148	Mean ±	3.25	3.28	3.09	2.98	3.31	3.04	3.21	3.17	
		SD	0.76	0.76	0.99	0.92	0.80	0.96	0.97	0.61	
45–54	54	Mean ±	3.37	3.50	2.94	3.13	3.24	2.94	3.15	3.18	
		SD	0.76	0.64	1.05	0.95	0.87	0.96	0.94	0.65	
≥ 55	14	Mean ±	3.36	3.57	2.71	3.21	3.43	3.21	3.07	3.22	
		SD	0.75	0.65	0.99	1.12	0.85	0.89	1.00	0.52	
BMI (kg/m ²)											1.02
18.5–24.9	182	Mean ±	3.41	3.39	3.01	3.10	3.30	3.00	3.18	3.20	
		SD	0.72	0.75	1.13	0.94	0.97	1.04	1.02	0.64	
25–29.9	155	Mean ±	3.20	3.30	2.95	2.95	3.25	2.88	3.06	3.08	
		SD	0.76	0.75	1.00	0.88	0.81	1.03	1.03	0.59	
30–39.9	90	Mean ±	3.36	3.31	3.06	3.17	3.43	2.74	2.92	3.14	
		SD	0.75	0.74	1.04	0.95	0.78	1.03	1.08	0.67	
40–49.9	7	Mean ±	3.29	3.29	3.00	3.14	3.29	2.29	2.86	3.02	
		SD	0.76	0.76	0.82	0.69	0.76	1.11	1.22	0.67	
Marital status											3.902**
Single	59	Mean ±	3.05	3.02	2.24	2.83	2.88	2.63	2.88	2.79	
		SD	0.75	0.82	1.15	0.81	1.15	1.10	0.89	0.64	
Married	372	Mean ±	3.37	3.39	3.11	3.09	3.38	2.93	3.12	3.20	
		SD	0.73	0.72	0.99	0.93	0.79	1.02	1.04	0.61	
Not married, in a relationship	5	Mean ±	3.60	3.60	3.20	3.60	3.40	2.80	3.40	3.37	
		SD	0.55	0.55	1.30	0.55	1.34	1.10	1.34	0.79	
Widowed, single	4	Mean ±	3.25	3.75	3.25	3.50	3.50	3.75	3.00	3.43	
		SD	0.96	0.50	0.96	0.58	1.00	0.50	0.82	0.48	
Widowed, in a relationship	1	Mean ±	3.00	2.00	2.00	2.00	3.00	3.00	2.00	2.43	
		SD	
Divorced, single	4	Mean ±	3.00	3.00	2.00	3.00	3.00	2.25	2.75	2.71	
		SD	0.82	0.82	0.82	0.82	0.82	1.26	1.50	0.67	
Divorced, in a relationship	3	Mean ±	3.33	3.33	2.67	3.33	3.33	2.33	4.00	3.19	
		SD	0.58	0.58	1.53	0.58	1.16	0.58	0.00	0.30	
Divorced, remarried	13	Mean ±	3.15	3.23	3.08	3.00	3.31	3.08	2.92	3.11	
		SD	1.07	0.83	1.04	1.00	0.75	0.95	1.04	0.68	
Education											.794
Less than high school	241	Mean ±	3.36	3.41	3.07	3.05	3.34	2.90	3.05	3.17	
		SD	0.76	0.74	1.05	0.99	0.82	1.04	1.08	0.63	
High school	85	Mean ±	3.19	3.24	2.87	2.95	3.24	2.86	3.01	3.05	
		SD	0.81	0.72	1.02	0.75	0.85	1.09	1.02	0.62	
University	113	Mean ±	3.31	3.27	2.89	3.20	3.27	2.94	3.21	3.16	
		SD	0.67	0.76	1.10	0.79	0.98	0.99	0.94	0.65	
Bachelor's degree or higher	22	Mean ±	3.36	3.14	3.05	2.91	3.45	2.68	3.09	3.10	
		SD	0.73	0.83	1.09	1.11	0.67	0.95	0.92	0.59	

Table 1 (continued)

Features	No. participants		FGSIS Item numbers							Total score	ANOVA <i>F</i>
			1	2	3	4	5	6	7		
Sexual orientation											2.28
Heterosexual/straight	455	Mean ± SD	3.31 0.75	3.33 0.75	2.98 1.06	3.06 0.91	3.30 0.87	2.88 1.04	3.07 1.03	3.14 0.63	
Homosexual/lesbian	0	Mean ± SD	
Bisexual	6	Mean ± SD	3.67 0.82	3.83 0.41	3.50 0.55	3.00 0.89	3.67 0.52	3.33 0.52	3.67 0.82	3.52 0.52	
Clothing style											3.251*
Chador	4	Mean ± SD	3.75 0.50	3.75 0.50	2.75 1.50	3.00 1.41	3.75 0.50	3.00 1.16	3.25 0.96	3.32 0.54	
Traditional headscarf	295	Mean ± SD	3.27 0.78	3.29 0.78	2.97 1.07	3.04 0.92	3.26 0.89	2.79 1.07	2.98 1.09	3.08 0.65	
Bareheaded	162	Mean ± SD	3.40 0.67	3.41 0.68	3.02 1.02	3.10 0.89	3.39 0.82	3.07 0.94	3.27 0.89	3.24 0.58	
Parity											.587
0	148	Mean ± SD	3.30 0.72	3.26 0.77	2.82 1.17	3.16 0.86	3.28 0.94	2.78 1.10	3.01 1.06	3.09 0.68	
1	126	Mean ± SD	3.33 0.75	3.37 0.73	3.08 0.99	3.16 0.85	3.33 0.82	2.85 1.00	3.16 1.01	3.18 0.60	
2	104	Mean ± SD	3.27 0.78	3.29 0.78	3.11 0.98	2.84 0.92	3.34 0.77	3.02 1.00	3.13 1.01	3.14 0.60	
≥3	83	Mean ± SD	3.39 0.76	3.47 0.67	3.00 1.02	3.04 1.04	3.28 0.90	2.98 1.01	3.05 1.02	3.17 0.60	
Menopausal status											.081
Postmenopause	89	Mean ± SD	3.24 0.74	3.37 0.77	2.92 1.01	3.09 0.90	3.11 0.94	3.28 0.83	3.11 0.99	3.16 0.54	
Premenopause	371	Mean ± SD	3.34 0.75	3.32 0.74	3.01 1.06	3.05 0.92	3.35 0.84	2.80 1.06	3.07 1.04	3.14 0.65	

F values were given only for total FGSIS scores

BMI body mass index, *SD* standard deviation, *FGSIS* Female Genital Self-Image Scale, *ANOVA* analysis of variance

*Correlation is significant at the 0.05 level, indicating a difference between groups on the FGSIS, **Correlation is significant at the 0.01 level

All items were found to be relevant for the study population and for the evaluative purpose of the measurement instrument. All items were found to comprehensively reflect the construct to be measured and referred to relevant aspects. The measurement aim was described by the original authors to measure female genital self-image in a reliable and valid manner to improve the understanding of any management that might affect a woman's perceptions of the way her genitals look and/or function. It was endeavored to reflect the general population as the target population after excluding known gynecologic malignancies. Completing the questionnaire did not require reading skills beyond that of a 16-year-old to avoid missing values and unreliable answers. The items were short and simple and did not contain difficult words or jargon owing

to a sufficient cross-cultural adaptation process. The time period to which the questions refer was agreed to be "in general" due to the nature of the questions.

Reliability

Internal consistency

The resulting seven-item FGSIS had a Cronbach's alpha coefficient of 0.818 and a mean score of 21.98 (SD = 4.4; *N* = 461). Floor and ceiling effects considered to be present if >15% of respondents achieved the lowest or highest possible score were not found to exist in this study.

Table 2 Item pool, retention decisions based on reliability analysis, and favor analysis of the Female Genital Self-Image Scale (FGSIS)

FGSIS Item no.	Item mean	Item standard deviation	Corrected item, total correlation	Cronbach's alpha if item deleted	Scale mean if item deleted	Factor loadings* (component = 1)
1	3.32	.746	.641	.771	18.66	.793
2	3.33	.747	.596	.778	18.65	.756
3	2.99	1.057	.639	.764	18.99	.756
4	3.06	.913	.427	.803	18.92	.570
5	3.31	.863	.582	.777	18.67	.721
6	2.89	1.034	.398	.812	19.09	.511
7	3.08	1.027	.595	.773	18.90	.715

*Principal component analysis

Factor analysis

The Kaiser–Meyer–Olkin value was high at 0.822, and the Bartlett's test of sphericity was significant ($p < 0.001$), confirming the suitability of using exploratory factor analysis. The number of participants included in a factor analysis was consisted with rules of thumb, which vary from four to ten individuals per variable, with a number of 461 participants to ensure variance–covariance matrix stability. There was no missing item for structural analysis. Exploratory principal component analysis indicated one factor that explained 48.42% (i.e., >40% cutoff) of variance. The eigenvalue reflects the amount of variance in all variables, which is accounted for by the level of a factor; loadings reflect how variables relate to each other in a factor. A large decrease was seen between the first and second eigenvalues, with small decreases thereafter (eigenvalues: 3.39, 0.96, 0.77, 0.64, 0.53, 0.42, 0.27). Factor loadings ranged from 0.51 to 0.79 (i.e., >0.40 cutoff). (Table 2) Corrected item-to-total correlations ranged from 0.39 to 0.64 (i.e., >0.30; see Table 2).

These findings provided further support for the construct validity of the FGSIS. The item “I am satisfied with the appearance of my genitals” was the most highly endorsed

(mean = 3.33, SD = 0.75). The proposed two-factor model (factor 1 = intrapersonal concerns; factor 2 = interpersonal concerns) was assessed by confirmatory factor analysis (CFA) due to the original authors' future recommendations. Results of the CFA revealed the two-factor model better fit the current data, yielding a chi-square discrepancy of 76.705, degrees of freedom (df) of 12 ($n = 457$) at a probability level of <0.001, a comparative fit index (CFI) = 0.938, and a normed fit index (NFI) = 0.928 (Fig. 1). Regression weights for the CFA model can be found in Table 3.

A point estimate of the root mean square error of approximation (RMSEA) was 0.109. With ~90% confidence, the population RMSEA for the two-factor model was between 0.086 and 0.133. The standardized root mean square residual (RMR) and the goodness of fit index (GFI) were 0.051 and 0.956, respectively. An RMR of 0 and GFI value of 1 indicates a perfect fit. A total of 62.2% of variance was explained through the two-factor solution. All item-to-factor loadings were ≥ 0.60 (i.e., >0.40 cutoff) [10]; mean level of commonality was 0.724 (SD 0.116). Hoelter's critical N for a significance level of 0.01 was 156, which was the largest sample size for which one could accept at the .01 level the hypothesis that the two-factor model was correct.

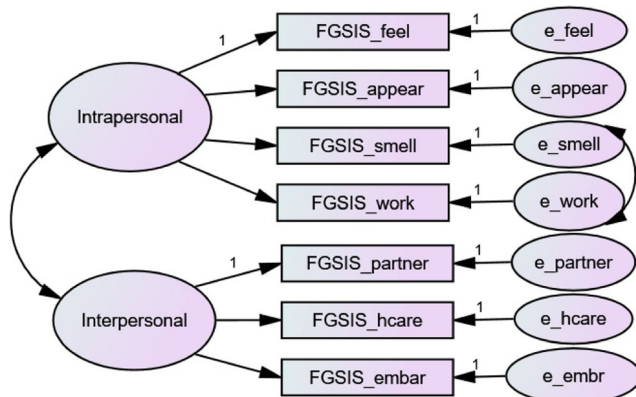


Fig. 1 Two-factor confirmatory factor analysis (CFA) of Female Genital Self-Image Scale (FGSIS)

Reproducibility

Agreement

The measurement error was adequately expressed as the standard error of measurement (SEM), which equals the square root of the error variance of an analysis of variance (ANOVA) analysis with including systematic differences ($SEM_{\text{agreement}}$). The intraclass correlation coefficient (ICC) of the total score of the FGSIS was 0.95 and the SEM 0.28. The SEM was converted into the smallest detectable change [$(SDC = 1.96 \times \sqrt{2} \times SEM)$], which reflects the smallest within-person change in score that can be interpreted as a real change above the measurement error in one individual (SDC_{ind}). SDC was

Table 3 Maximum likelihood estimates of standardized and unstandardized regression weights for proposed two-factor model based on the Female Genital Self-Image Scale (FGSIS)

FGSIS_items	Perspective	Unstandardized				Standardized
		Estimate	SE	CR	P	Estimate
Feelings	← Intrapersonal	1.000 ^a				.865
Appearance	← Intrapersonal	.944	.050	18.714	<.001	.813
Smell	← Intrapersonal	.605	.070	8.692	<.001	.429
Work related	← Intrapersonal	.785	.064	12.203	<.001	.588
Partner related	← Interpersonal	1.000 ^a				.736
Health care related	← Interpersonal	.684	.074	9.189	<.001	.514
Embarrassment	← Interpersonal	.924	.085	10.858	<.001	.700

SE standard error, CR critical ratio,

^a Fixed at 1.000; not estimated

0.78. The SDC measurable in a group of people (SDC_{group}) can be calculated by dividing the SDC_{ind} by \sqrt{n} . SDC_{group} was 0.052. Bland–Altman plot analysis of agreement is shown in Fig. 2. The limits of agreement equal the mean change in scores of repeated measurements (mean change) \pm 1.96 \times SD of those changes.

Reliability

Three participants were excluded, 32 participants were included in test–retest analysis. $ICC_{agreement}$ (two-way random effects model) was preferred to test reliability. Test–retest reliability was excellent, with all single items and total score

being consistent between the two measurement points and being significantly correlated ($P < 0.05$) (Table 4). Test–retest correlation coefficient of total score was 0.951; ICCs of each items were between 0.800 and 0.945, with a mean >0.890 ($SD = .06$).

Construct validity

We hypothesized that items within the construct should converge and items across constructs should discriminate. In the proposed two-factor construct, all seven items demonstrated good to high correlations with their subdomains and lower correlations with the other domain, indicating sufficient

Fig. 2 Bland–Altman plot visualizing agreement for test–retest with the limits marked as mean difference \pm standard deviation (SD) on a four-point scale

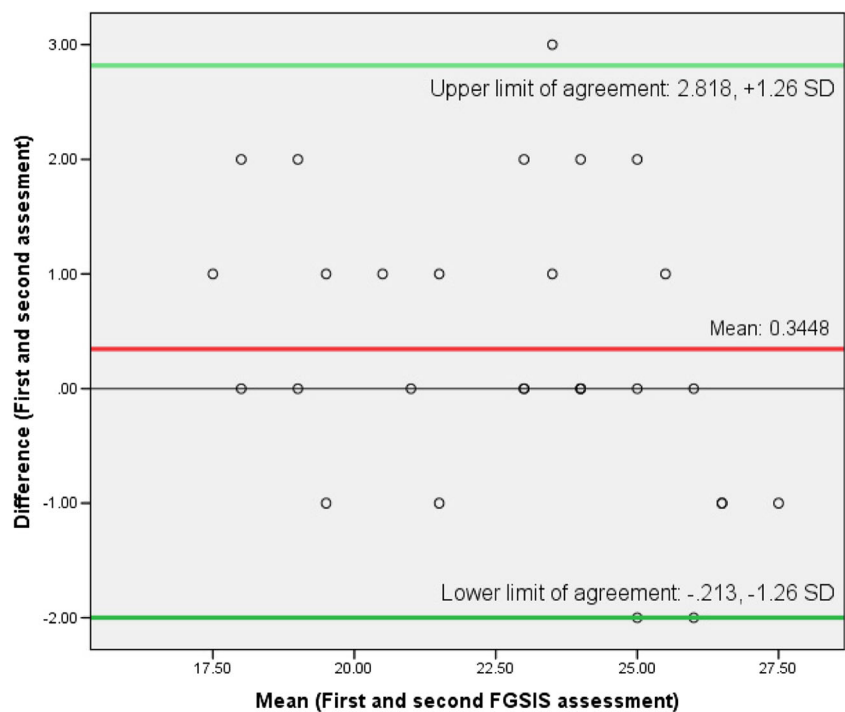


Table 4 Reproducibility of Female Genital Self-Image Scale (FGSIS)

FGSIS item no.	Intraclass correlation coefficient (ICC)		
	ICC ^a	95% confidence interval	
		Lower bound	Upper bound
1	.945	.883	.974
2	.976	.950	.989
3	.843	.664	.926
4	.901	.791	.953
5	.800	.555	.908
6	.879	.744	.943
7	.889	.709	.952
Total score	.951	.896	.977

Two-way random effects model: people effects and measures effects are random

^aType A ICC using an absolute agreement definition

convergent validity. Variance extracted between the two factors and the correlation square were 0.536 and 0.251, respectively, hence establishing discriminant validity (Table 5).

Hypothesis testing found a high positive correlation ($r_{\text{mean}} = 0.69 \pm 0.07$; range = 0.59–0.77; $p < 0.001$) with the total score for their own scale, which confirmed the correlation hypothesis (Table 6). Results of known-group validity are summarized in Table 7. As expected, participants who performed at least one vaginal/clitoral masturbation for the prior month reported significantly higher FGSIS scores compared with women who abstained.

As hypothesized, weak to moderate correlation was found between FGSIS with FSFI ($r = 0.597$) and FSDS-R ($r = -4.51$) (Table 8).

Discussion

This study successfully cross-culturally adapted and validated the reliability and construct of FGSIS in a large sample for use to measure female genital self-image. The seven-item FGSIS had a Cronbach's alpha coefficient of 0.88 in the original study, a range of 0.82 and 0.89 in the two-factor model, and 0.86 in the four-item model in subsequent validation studies. Our result (0.82) was comparable with another external validation study (0.86) [19]. The major strength of our study was the vigorous cross-cultural adaptation and the validation process, which were fully compatible with current guidelines. The Turkish version of the FGSIS exhibited a one-factor model in the current study. However, two-factor (interpersonal and intrapersonal) structure as suggested by the original authors was also sufficiently fit to the model [7]. This flexibility should be further analyzed in future studies.

Table 5 Pattern matrix for convergent and discriminant validity in the proposed two-factor model

	Component				
	1	2	3	4	5
Convergent validity ^a					
FGSIS 1	.870				
FGSIS 2	.851				
FGSIS 3	.380	.526			
FGSIS 4	.673				
FGSIS 5	.662	.126			
FGSIS 6	-.259	1			
FGSIS 7	.303	.568			
Discriminant validity ^b					
FSDS-R1	.770	-.128	.133	.123	-.059
FSDS-R2	.882	.110	.022	.194	.097
FSDS-R3	.763	.182	-.008	.155	.269
FSDS-R4	.782	-.049	.068	.049	.283
FSDS-R5	.804	.002	.076	-.136	.018
FSDS-R6	.816	-.063	.058	.047	.115
FSDS-R7	.714	-.032	-.073	-.024	.187
FSDS-R8	.675	-.063	.062	-.048	.269
FSDS-R9	.820	-.031	-.041	.093	-.231
FSDS-R10	.857	.078	-.109	-.079	-.277
FSDS-R11	.737	-.029	-.041	-.006	-.287
FSDS-R12	.784	.020	-.093	-.078	-.048
FSDS-R13	.636	.025	.032	-.227	.401
FSFI Desire	-.022	.636	-.167	.449	-.071
FSFI Arousal	.006	.943	-.150	.085	.054
FSFI Lubrication	.009	.952	-.028	-.147	-.049
FSFI Orgasm	.079	.880	.118	-.251	-.089
FSFI Satisfaction	-.101	.693	.202	-.324	-.061
FSFI Pain	.014	.894	-.093	.109	-.003
FGSIS 1	-.009	-.163	.954	.078	-.191
FGSIS 2	.099	-.036	.865	.148	-.258
FGSIS 3	.070	.462	.389	.255	-.110
FGSIS 4	-.039	.170	.426	-.112	.047
FGSIS 5	-.107	.119	.345	-.141	-.664
FGSIS 6	.106	-.123	.151	.820	.112
FGSIS 7	-.303	.247	.211	.326	.283

Extraction method, principal component analysis; rotation method, Promax with Kaiser normalization

Kaiser-Meyer-Olkin measure of sampling adequacy: .865

Bartlett's test of sphericity; $p < 0.001$, $df = 325$, $X^2 = 2211.2$

FGSIS Female Genital Self Image Scale, FSDS-R Female Sexual Distress Scale, FSFI Female Sexual Function Index

^aRotation converged in 3 iterations

^bRotation converged in 6 iterations

There was a positive relationship between women's sexual self-pleasuring and positive body image among European

Table 6 Correlation hypothesis testing ($n = 461$). One-trait scaling analysis was used. Correlations between each questionnaire item and total questionnaire score were generated. Validity was assumed if all questionnaire items showed correlation at $r > 0.40$ with the total score for their own scale

FGSIS item no.	1	2	3	4	5	6	7	FGSIS total score
1	–	.723**	.471**	.357**	.480**	.243**	.443**	.735**
2		–	.416**	.330**	.424**	.237**	.426**	.700**
3			–	.323**	.531**	.405**	.482**	.771**
4				–	.384**	.162**	.315**	.589**
5					–	.229**	.398**	.704**
6						–	.402**	.587**
7							–	.736**

Pearson Correlation

FGSIS Female Genital Self Image Scale

** Correlation is significant at the 0.01 level

American women [16]. As expected, women who had at least one masturbation history for the prior month exhibited higher self-image. This study points to another important issue with hypothesis testing: As hypothesized, FGSIS showed weak to moderate correlation with FSFI and FSDDS-R scores, contrary to other studies [11]. We speculate it might be related to cultural differences. FSFI captures sexual “functionality,” which depends on partner-related variables rather than self-image. In the same manner, FSDDS-R is also related more to partner-related distress and may be affected by the general happiness of the woman about her life or relationship. Therefore, we believe that the FGSIS connotes a different concept to other questionnaire tools. Similarly, in their prospective long-term follow-up cohort study, Goodman et al. found that women did not have symptoms of sexual dysfunction prior to vulvovaginal esthetic surgery. Besides, FSFI scores did not

alter after surgery, except for the satisfaction subdomain [9]. We believe results of that study support our theory and therefore weak to moderate correlation should be expected with women’s bodily self-image.

Findings from this study have several implications. The proportion of women seeking elective cosmetic surgery increases and seems to be strongly influenced by the media [8, 20]. FGSIS might capture the progress of the individual after cosmetic surgery. The minimal amount of change we found in the scale is considered important, and therefore, responsiveness ability of FGSIS will be assessed in a future study.

Schick et al. observed a shift in genital appearance ideals across five decades, and current perception—primarily created by the media—fosters significant body-image disturbance among women [21]. Laan et al. showed pictures of natural vulvas to college-educated women to assess their influence and women’s self-awareness [22]. They found that exposure to pictures of natural vulvas positively affected genital self-image. Hummel et al. found that Internet-based cognitive behavioral therapy was every effective for treating sexual dysfunction and body image in breast cancer survivors [23]. FGSIS has the potential to be used in such approaches prior to female genital elective cosmetic surgeries.

Body image and sexual function might be influenced by conditions commonly seen at outpatient gynecology settings, such as dyspareunia, endometriosis, pregnancy, gestational diabetes, and infertility [5, 6, 24–26]. It would be interesting to see the alteration between self-image and treatment in the obstetrics and gynecology setting.

A limitation of this study was the absence of testing gynecological examination behavior, as suggested by the original authors. That hypothesis was not consciously tested because the sample was drawn around the faculty, and it might have been a cause of bias. The sample did not contain patients but medical staff, relatives and friends, thereby ensuring generalizability. Participants with an education level of less than high school represents 52.5% (242/461) of the cohort; therefore, we believe this instrument is also valid among more- and

Table 7 Discrimination hypothesis regarding masturbation ($n = 461$). Known-groups validation was used. Differences between item scores were examined in terms of the presence or absence of vaginal and/or clitoral masturbation. We hypothesized that women performed at least one masturbation during the prior month would report higher scores across FGSIS items in comparison with participants not performing masturbation

FGSIS item no.	Absent ($n = 360$) Mean rank	≥ 1 / month ($n = 101$) Mean rank	Z	P value
1	219.6	271.8	–3.81	<0.001
2	219.5	272	–3.85	<0.001
3	218	277.2	–4.17	<0.001
4	218.5	275.5	–4.04	<0.001
5	219.1	273.5	–3.99	<0.001
6	212	298.8	–6.06	<0.001
7	215.5	286.2	–5.04	<0.001
FGSIS Total score	210.1	305.3	–6.37	<0.001

Mann-Whitney U test

FGSIS Female Genital Self Image Scale

Table 8 Relationship between FGSIS score and FSFI and FSDS-R test scores alone and after controlling for stratified ages of participants

FSFI and FSDS-R test scores	FGSIS item correlations							FGSIS total score
	1 Feelings	2 Appearance	3 Partner related	4 Smell	5 Work related	6 Healthcare	7 Embarrassment	
FSFI total	.319**	.353**	.595**	.324**	.468**	-.100	.447**	.597**
Desire	.117	.194	.477**	.155	.281**	.033	.345**	.409**
Arousal	.239*	.241*	.465**	.288**	.319**	-.035	.334**	.461**
Lubrication	.284**	.337**	.524**	.316**	.415**	-.161	.385**	.519**
Orgasm	.361**	.357**	.533**	.290**	.479**	-.172	.415**	.556**
Satisfaction	.384**	.314**	.582**	.361**	.526**	-.236*	.412**	.576**
Pain	.244*	.354**	.444**	.256**	.346**	.047	.380**	.512**
FSDS-R	-.213*	-.182	-.426**	-.339**	-.316**	.116	-.427**	-.451**
Variables controlled for age								
FSFI total	.315**	.361**	.595**	.340**	.485**	-.120	.458**	.597**
Desire	.094	.212*	.458**	.190	.317**	-.006	.367**	.412**
Arousal	.224*	.256*	.448**	.321**	.351**	-.073	.352**	.463**
Lubrication	.288**	.341**	.535**	.317**	.416**	-.161	.394**	.522**
Orgasm	.363**	.361**	.534**	.302**	.493**	-.192	.421**	.555**
Satisfaction	.370**	.336**	.577**	.376**	.541**	-.246*	.444**	.586**
Pain	.274**	.342**	.473**	.260**	.356**	.037	.364**	.514**
FSDS-R	-.234*	-.177	-.457**	-.330**	-.309*	.100	-.431**	-.459**

Pearson correlation; partial correlation

FGSIS Female Genital Self Image Scale, FSFI Female Sexual Function Index, FSDS-R Female Sexual Distress Scale-Revised

**Correlation is significant at the 0.01 level

*Correlation is significant at the 0.05 level

less-educated women. However, future work should validate among less-skilled women, since most participants were skilled employees.

Conclusion

The FGSIS was successfully validated for use in the Turkish population. The scale exhibited strong psychometric properties in a diverse population. The quick, simple, and comprehensive structure of the scale might aid clinicians in sexual medicine, psychiatry, obstetrics, and gynecology settings to better understand women's issues around genital self-image and areas in which self-image may play a positive or negative role.

Compliance with ethical standards

We have read and understood the journal's policies on copyright, ethics, etc., and believe that neither the manuscript nor the study violates any of these.

Conflicts of interest None.

Financial disclaimer None.

References

- Schick VR, Calabrese SK, Rima BN, Zucker AN. Genital appearance dissatisfaction: implications for women's genital image self-consciousness, sexual esteem, sexual satisfaction, and sexual risk. *Psychol Women Q.* 2010;34(3):394–404.
- Satinsky S, Reece M, Dennis B, Sanders S, Bardzell S. An assessment of body appreciation and its relationship to sexual function in women. *Body Image.* 2012;9(1):137–44.
- Ackard DM, Kearney-Cooke A, Peterson CB. Effect of body image and self-image on women's sexual behaviors. *Int J Eat Disord.* 2000;28(4):422–9.
- Pujols Y, Meston CM, Seal BN. The association between sexual satisfaction and body image in women. *J Sex Med.* 2010;7(2pt2):905–16.
- Pauls RN, Occhino JA, Dryfhout VL. Effects of pregnancy on female sexual function and body image: a prospective study. *J Sex Med.* 2008;5(8):1915–22.
- Lowenstein L, Gamble T, Sanses TVD, Van Raalte H, Carberry C, Jakus S, et al. Changes in sexual function after treatment for prolapse are related to the improvement in body image perception. *J Sex Med.* 2010;7(2pt2):1023–8.
- DeMaria AL, Hollub AV, Herbenick D. The female genital self-image scale (FGSIS): validation among a sample of female college students. *J Sex Med.* 2012;9(3):708–18.
- Koning M, Zeijlmans IA, Bouman TK, van der Lei B. Female attitudes regarding labia minora appearance and reduction with consideration of media influence. *Aesthet Surg J.* 2009;29(1):65–71.

9. Goodman M, Fashler S, Miklos JR, Moore RD, Brotto LA. The sexual, psychological, and body image health of women undergoing elective vulvovaginal plastic/cosmetic procedures: a pilot study. *Am J Cosmet Surg*. 2011;28(4):219–26.
10. Herbenick D, Reece M. OUTCOMES ASSESSMENT: development and validation of the female genital self-image scale. *J Sex Med*. 2010;7(5):1822–30.
11. Herbenick D, Schick V, Reece M, Sanders S, Dodge B, Fortenberry JD. The female genital self-image scale (FGSIS): results from a nationally representative probability sample of women in the United States. *J Sex Med*. 2011;8(1):158–66.
12. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186–91.
13. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539–49.
14. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.
15. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health*. 2005;8(2):94–104.
16. Shulman JL, Home SG. The use of self-pleasure: masturbation and body image among African American and European American women. *Psychol Women Q*. 2003;27(3):262–9.
17. Verit FF, Verit A. Validation of the female sexual function index in women with chronic pelvic pain. *J Sex Med*. 2007;4(6):1635–41.
18. Aydın S, Onaran ÖI, Topalan K, Aydın ÇA, Dansuk R. Development and validation of Turkish version of the female sexual distress scale-revised. *J Sex Med*. 2016;4(1):e43–50.
19. Pakpour AH, Zeidi IM, Ziaeiha M, Burri A. Cross-cultural adaptation of the female genital self-image scale (FGSIS) in Iranian female college students. *J Sex Res*. 2014;51(6):646–53.
20. Müllerová J, Weiss P. Motivational factors of women seeking Labiaplasty and body dysmorphic disorder risk: a review. *J Womens Health Issues Care*. 2017;6:1.
21. Schick VR, Rima BN, Calabrese SK. E vulva lution: the portrayal of women's external genitalia and physique across time and the current Barbie doll ideals. *J Sex Res*. 2011;48(1):74–81.
22. Laan E, Martoredjo DK, Hesselink S, Snijders N, van Lunsen RH. Young women's genital self-image and effects of exposure to pictures of natural vulvas. *J Psychosom Obstet Gynecol*. 2017;38(4):249–55.
23. Hummel SB, van Lankveld JJ, Oldenburg HS, Hahn DE, Kieffer JM, Gerritsma MA, Kuenen MA, Bijker N, Borgstein PJ, Heuff G. Internet-based cognitive behavioral therapy realizes long-term improvement in the sexual functioning and body image of breast Cancer survivors: internet-based CBT for sexual dysfunctions. *J Sex Marit Ther*. 2018;0(0):1–12.
24. Aydın S, Kurt N, Mandel S, Kaplan MA, Karaca N, Dansuk R. Female sexual distress in infertile Turkish women. *Turk J Obstet Gynecol*. 2015;12(4):205.
25. Melis I, Litta P, Nappi L, Agus M, Melis GB, Angioni S. Sexual function in women with deep endometriosis: correlation with quality of life, intensity of pain, depression, anxiety, and body image. *Int J Sex Health*. 2015;27(2):175–85.
26. Sargin MA, Yassa M, Taymur BD, Taymur B, Akca G, Tug N. Female sexual dysfunction in the late postpartum period among women with previous gestational diabetes mellitus. *J Coll Physicians Surg Pak*. 2017;27(4):203–8.