**The Body Shape Questionnaire (BSQ) is Not Invariant Across Sex: Evidence from Portuguese-Speaking University Students**

Wanderson Roberto da Silva1, Viren Swami2,3, Angela Nogueira Neves4, João Marôco5, Christopher N. Ochner6, & Juliana Alvares Duarte Bonini Campos1

1Department of Food and Nutrition, School of Pharmaceutical Science of São Paulo State University (UNESP), São Paulo, Brazil. [wandersonroberto22@gmail.com](mailto:wandersonroberto22@gmail.com); [juliana.campos@unesp.br](mailto:juliana.campos@unesp.br)

2Department of Psychology, Anglia Ruskin University, Cambridge, United Kingdom [viren.swami@anglia.ac.uk](mailto:viren.swami@anglia.ac.uk)

3Centre for Psychological Medicine, Perdana University, Serdang, Malaysia

4Division of research and extension of School of Physical Education of the Army, Rio de Janeiro, Brazil. [angela.esefex@yahoo.com.br](mailto:angela.esefex@yahoo.com.br)

5William James Center for Research (WJCR), University Institute of Psychological, Social and Life Sciences (ISPA), Portugal. [jpmaroco@ispa.pt](mailto:jpmaroco@ispa.pt)

6Hospital Corporation of America, Physician Services Group, Florida, United States. [cochner@gmail.com](mailto:cochner@gmail.com)

**Abstract**

The Body Shape Questionnaire (BSQ) is a widely-used measure of body shape concerns that was originally designed for use with women but has more recently been used with boys and men. The latter use may be problematic, given that no previous study has demonstrated sex invariance for BSQ scores. To determine the extent to which BSQ scores are sex invariant, we asked Portuguese-speaking women (*n* = 1,613) and men (*n* = 871) to complete the full BSQ (34-items). Confirmatory factor analysis (CFA) indicated that a hypothesized 32-item model of BSQ scores and shorter versions had acceptable fit indices in women and men, separately. However, multi-group CFA showed that these BSQ model scores had configural but not metric, scalar, or strict sex invariance. Differential item analysis indicated significant item-functioning differences on 19 of the 32 retained BSQ items. Thus, BSQ scores are not sex invariant, making problematic the results of previous studies that have compared latent BSQ scores across sex.

*Keywords:* Latent mean comparisons; Sex differences; Measurement invariance; Confirmatory factor analysis; Portuguese students

**Introduction**

The Body Shape Questionnaire (BSQ; Cooper, Taylor, Cooper, & Fairburn, 1987) is a body image assessment tool, originally designed to assess the phenomenological experience of “feeling fat”. Its 34 items were derived from semi-structured interviews with various groups of North American women, including eating disordered patients (Cooper et al., 1987). Likewise, BSQ validity evidence for differentiating patients with eating disorders from non-patients was originally provided from samples of women (Cooper et al., 1987). From its initial development, the BSQ has been used to study both eating disordered patients (Brockmeyer et al., 2018; Laporta-Herrero, Jáuregui-Lobera, Barajas-Iglesias, & Santed-Germán, 2016) and non-clinical populations (Dondzilo, Rieger, Palermo, Byrne, & Bell, 2017; Silva, Campos, & Marôco, 2018). The factorial and construct validity of its scores have been established across a wide range of cultural and linguistic groups (e.g., Moreno, Montano, Prieto, Pérez-Acosta, 2007; Rousseau, Knotter, Barbe, Raich, & Chabrol, 2005).

Although designed to assess body shape concerns and dissatisfaction in women, the BSQ measure has occasionally been used with men and boys (Flores-Cornejo, Kamego-Tome, Zapata-Pachas & Alvarado, 2017; Laporta-Herrero et al., 2016; Najam & Ashfaq, 2012; Pokrajac-Bulian & Živčić-Bećirević, 2005; Sarwer et al., 2015), particularly among Brazilian students (Amaral & Ferreira, 2017; Branco, Hilário, & Cintra, 2006; Conti, Cordás, & de Oliveira Latorre, 2009; Di Pietro & da Silveira, 2009; Laus, Miranda, Almeida, Costa, & Ferreira, 2013). These studies have generally reported, particularly among younger age groups, that respondents were susceptible to body concerns (Silva, Costa, Pimenta, Marôco, & Campos, 2016; Silva, Dias, Marôco, & Campos, 2014). Studies examining sex differences among BSQ respondents (Amaral & Ferreira, 2017; Branco et al., 2006; Di Pietro & da Silveira, 2009; Laus et al., 2013) have typically found that men and boys obtain significantly lower BSQ scores than women and girls, leading researchers to conclude that women are more concerned with their body shape and weight than are men. However, such a conclusion may be unwarranted, given that BSQ items were originally designed to assess body concerns in women and not men (Silva et al., 2016; Silva et al., 2014). That is, while any comparison of BSQ scores across sex presupposes that the measure functions similarly across these two groups (Slof-Op t’Landt et al., 2009), no prior research has established the sex invariance of this instrument. As Swami and Barron (2018) have written, research examining sex invariance is necessary before assuming that women and men’s responses to individual items can be explained by the same latent BSQ factor.

Given that no previous study has examined the extent to which BSQ scores are sex invariant, we sought to do so in the present study. Particularly because of suspected high rates of body concerns in Brazilian and Portuguese populations (e.g., Silva et al., 2018), we recruited Brazilian and Portuguese university students to examine whether women’s and men’s scores on a Portuguese version of the BSQ would demonstrate configural similarity (i.e., a similar number of BSQ items and unifactorial model across women and men), metric similarity (i.e., similar latent BSQ variable and pattern loadings), scalar similarity (i.e., similar item loadings and item thresholds), and strict measure similarity (i.e., similar magnitude of loadings, thresholds, and covariance of the residues). This analysis is important to research advancements in the assessment of body image. Chen (2007) recommended that latent factor means between groups should not be compared in the absence of scalar invariance between the groups; that is, a failure to achieve scalar invariance in the present study would cast doubt on prior research comparisons of mean BSQ scores obtained by respondents across sex.

**Method**

**Participants**

### University students of both sexes who were enrolled in Brazilian and Portuguese higher education institution were invited to participate in this study. In Brazil, students in the Schools of Pharmaceutical Sciences, Letters and Sciences, and the Institute of Chemistry of São Paulo State University were selected for participant recruitment; while, in Portugal, students from the University Institute of Psychological, Social, and Life Sciences, the Lisbon School of Nursing, the Higher Institute of Engineering of Porto, the Higher Institute of Health Sciences Egas Moniz, and the School of Pharmacy of the University of Coimbra were recruited. Inclusion criteria included (a) enrollment in higher education courses, and (b) aged between 18- 35 years old. Exclusion criteria were (a) in receipt of treatment for an eating disorder at the time of recruitment, and (b) (for women) being in the gestational period. We restricted the age of participants because studies on body image have reported that young individuals are most likely to develop body concerns and because this age group characterizes university students relatively well. A total of 2,484 university students (64.9% women), who met the inclusion criteria and agreed to voluntarily participate in the study, were included in the sample. The mean age of participants was 20.8 years (*SD* = 2.5) for women and 21.2 years (*SD* = 3.2) for men. The mean body mass index (BMI) was 22.1 kg/m2 (*SD* = 3.4) for women and 23.7 kg/m2 (*SD* = 3.7) for men.

### Measures

### Participants were asked to complete the 34-item BSQ, designed by Cooper and colleagues (1987) to assess body concerns in women, with all items rated on a 6-point scale (1 = *never*, 6 = *always*). In the present study, we used a Portuguese version of the BSQ adapted for Brazil and Portugal by Silva and colleagues (2016). This study showed adequate estimates of factorial and convergent validity, internal consistency and composite reliability of BSQ for Portuguese and Brazilian women. All participants completed the full, 34-item version of the BSQ and psychometric properties of scores on the measure are discussed below. In addition to completing the BSQ, participants also provided certain demographic information (sex, age, height, and weight).

**Procedures**

The study was approved by the Human Research Ethics Committee of the School of Pharmaceutical Sciences (C.A.A.E. 29896214.0.0000.5426) and the Ethics Committee of the Lisbon School of Nursing (Proceeding #1413), and all participating institutions provided permission for data collection. Willing participants who fulfilled inclusion and exclusion criteria were invited to complete these paper-and-pencil questionnaires in a classroom setting. A researcher provided brief information about the project (i.e., aim and purpose of the study) and participants who agreed to take part provided written informed consent. The classroom setting allowed participants to complete the questionnaire individually. Upon completion, questionnaires were returned to the researchers. All participants took part on a voluntary basis, were not remunerated, and received no course credits for participation.

**Analytic Strategy**

Primary analyses were conducted using MPLUS v.7.2. Descriptive analysis of the items of BSQ were conducted separately for women and men, permitting us to calculate separate means, medians, modes, standard deviations, skewness, and kurtosis for these groups. We assessed univariate normality using skewness and kurtosis, with values < |3| and < |7|, respectively, considered adequate ([Marôco, 2014](#_ENREF_1)). We calculated coefficient of multivariate kurtosis (kum) and critical ratios (c.r.) to estimate multivariate normality, with values < |4| considered indicative of normality ([Marôco, 2014](#_ENREF_1)). In earlier work, Silva and colleagues (2016) reported that scores on the Portuguese version of the BSQ achieved adequate fit among university women following the removal of two items (#26 and 32) related to assessment diagnostics for eating disorders rather than body concerns, as is consistent with other work among similar cultural groups (Silva et al., 2014). However, this reduced item pool potentially creates an *ipso facto* problem *vis-à-vis* invariance, if it is assumed that a 32-item model should be fitted for men. To avoid this problem, we conducted confirmatory factor analyses (CFAs) separately for women (fitting a 32-item model) and men (initially fitting a 34-item model, since this model was not fitted previously). Moreover, shorter versions of the BSQ (16A, 16B, 8A, 8B, 8C and 8D) as studied by Evans and Dolan (1993) were also evaluated for each sex. For these CFAs, we used the Weighted Least Squares Mean and Variance Adjusted (WLSMV) method on the items’ polychoric correlation matrix. To assess model fit, we examined χ2/df, the root mean square error of approximation (RMSEA), Bentler’s Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI). Model fit was considered acceptable when χ2/df ≤ 5.00, RMSEA ≤ .10, and CFI and TLI ≥ .90 and good when χ2/df ≤ 2.00, RMSEA ≤ . 05, and CFI and TLI ≥ .95 ([Marôco, 2014](#_ENREF_1)). Factor loadings (λ) of items were considered adequate when ≥ .40. Modification indices, calculated using the Lagrange Multiplier method, were used only when necessary to improve model fit ([Marôco, 2014](#_ENREF_1)). To evaluate reliability, we used the ordinal alpha coefficient (α) and omega coefficient (ω). Values of α and ω ≥ .70 were considered adequate ([Marôco, 2014](#_ENREF_1" \o "Marôco, 2014 #99); Dunn, Baguley, & Brunsden, 2014).

Further analyses were only planned in the event that comparable models (i.e., with the same items retained) were obtained for women and men separately. Specifically, further analyses consisted of examining measurement invariance across sex using multi-group CFA. Invariance was examined based on chi-square difference (Δχ2) by computing the factor loading (λ), thresholds (*T*), and the covariance of the residues (*cov*). Results were assessed stepwise, namely: (a) configural (i.e., the same structure which it was mandatory across sex and country); (b) metric (*p*Δχ2λ > .05); (c) scalar (*p*Δχ2λ and *p*Δχ2*T* > .05), and; (d) strict measure (*p*Δχ2λ, *p*Δχ2*T*, *p*Δχ2cov > .05) (Marôco, 2014). In the event that scalar invariance was not achieved across sex, we planned on conducting differential item functioning (DIF) analysis using the lordif package (Choi, 2016) in *R* (Revelle, 2018).

**Results**

Descriptive analyses of the 34 BSQ items indicated that the majority of the 34 BSQ items were univariate normal in women and men (though Items 26 and 32 were problematic; see Table 1). Multivariate kurtosis indicated that the data from women (kum = 2.32) and men (kum = 3.35) were multivariate normal.

Table 1. *Descriptive Statistics of the Body Shape Questionnaire (BSQ) for Women and Men*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Women | | | | | | Men | | | | | |
| Item | *M* | *Mdn* | *Mo* | *SD* | *Sk* | *Ku* | *M* | *Mdn* | *Mo* | *SD* | *Sk* | *Ku* |
| 1 | 2.85 | 3.00 | 3.00 | 1.40 | .46 | -.46 | 2.48 | 2.00 | 1.00 | 1.36 | .73 | -.22 |
| 2 | 3.14 | 3.00 | 3.00 | 1.61 | .24 | -1.06 | 2.48 | 2.00 | 1.00 | 1.47 | .73 | -.46 |
| 3 | 2.84 | 3.00 | 1.00 | 1.76 | .48 | -1.12 | 1.72 | 1.00 | 1.00 | 1.23 | 1.83 | 2.68 |
| 4 | 3.58 | 3.00 | 3.00 | 1.71 | -.03 | -1.24 | 2.62 | 2.00 | 1.00 | 1.61 | .70 | -.65 |
| 5 | 3.77 | 4.00 | 3.00 | 1.52 | -.10 | -.98 | 2.50 | 2.00 | 1.00 | 1.43 | .70 | -.36 |
| 6 | 3.22 | 3.00 | 3.00 | 1.66 | .22 | -1.13 | 2.26 | 2.00 | 1.00 | 1.41 | .97 | .05 |
| 7 | 1.82 | 1.00 | 1.00 | 1.31 | 1.73 | 2.28 | 1.20 | 1.00 | 1.00 | .71 | 4.26\* | 19.61\* |
| 8 | 1.57 | 1.00 | 1.00 | 1.12 | 2.26 | 4.74 | 1.27 | 1.00 | 1.00 | .81 | 3.59\* | 13.63\* |
| 9 | 2.26 | 2.00 | 1.00 | 1.46 | 1.13 | .40 | 1.49 | 1.00 | 1.00 | 1.01 | 2.44 | 5.90 |
| 10 | 2.13 | 1.00 | 1.00 | 1.51 | 1.23 | .43 | 1.30 | 1.00 | 1.00 | .83 | 3.44\* | 12.77\* |
| 11 | 1.69 | 1.00 | 1.00 | 1.12 | 1.97 | 3.81 | 1.30 | 1.00 | 1.00 | .77 | 3.24\* | 11.64\* |
| 12 | 2.95 | 3.00 | 3.00 | 1.49 | .51 | -.62 | 2.25 | 2.00 | 1.00 | 1.29 | .95 | .34 |
| 13 | 1.69 | 1.00 | 1.00 | 1.13 | 1.95 | 3.54 | 1.43 | 1.00 | 1.00 | .90 | 2.61 | 7.38\* |
| 14 | 2.33 | 2.00 | 1.00 | 1.50 | 1.02 | .06 | 1.61 | 1.00 | 1.00 | 1.05 | 1.96 | 3.65 |
| 15 | 3.07 | 3.00 | 3.00 | 1.51 | .37 | -.78 | 2.07 | 2.00 | 1.00 | 1.34 | 1.20 | .71 |
| 16 | 2.28 | 1.00 | 1.00 | 1.65 | 1.06 | -.18 | 1.49 | 1.00 | 1.00 | 1.09 | 2.53 | 6.09 |
| 17 | 2.78 | 3.00 | 1.00 | 1.62 | .60 | -.74 | 1.80 | 1.00 | 1.00 | 1.27 | 1.67 | 2.07 |
| 18 | 1.46 | 1.00 | 1.00 | 1.01 | 2.68 | 7.24 | 1.22 | 1.00 | 1.00 | .69 | 3.98\* | 17.81\* |
| 19 | 1.94 | 1.00 | 1.00 | 1.41 | 1.52 | 1.33 | 1.36 | 1.00 | 1.00 | .87 | 2.95 | 9.43\* |
| 20 | 2.46 | 2.00 | 1.00 | 1.45 | .92 | .02 | 1.80 | 1.00 | 1.00 | 1.20 | 1.69 | 2.45 |
| 21 | 2.68 | 2.00 | 1.00 | 1.66 | .62 | -.83 | 2.03 | 1.00 | 1.00 | 1.43 | 1.28 | .63 |
| 22 | 2.95 | 3.00 | 1.00 | 1.80 | .42 | -1.19 | 1.90 | 1.00 | 1.00 | 1.37 | 1.56 | 1.58 |
| 23 | 2.73 | 2.00 | 1.00 | 1.68 | .61 | -.85 | 2.10 | 1.00 | 1.00 | 1.49 | 1.23 | .45 |
| 24 | 2.97 | 3.00 | 1.00 | 1.66 | .42 | -.97 | 1.87 | 1.00 | 1.00 | 1.33 | 1.57 | 1.66 |
| 25 | 1.87 | 1.00 | 1.00 | 1.35 | 1.65 | 1.90 | 1.30 | 1.00 | 1.00 | .83 | 3.30\* | 11.44\* |
| 26 | - | - | - | - | - | - | 1.09 | 1.00 | 1.00 | .49 | 6.63\* | 49.31\* |
| 27 | 1.46 | 1.00 | 1.00 | 1.04 | 2.70 | 7.27\* | 1.26 | 1.00 | 1.00 | .77 | 3.51\* | 13.08\* |
| 28 | 2.82 | 3.00 | 1.00 | 1.61 | .56 | -.78 | 1.93 | 1.00 | 1.00 | 1.33 | 1.46 | 1.38 |
| 29 | 2.50 | 2.00 | 1.00 | 1.45 | .86 | -.14 | 1.76 | 1.00 | 1.00 | 1.16 | 1.70 | 2.48 |
| 30 | 2.91 | 3.00 | 1.00 | 1.56 | .47 | -.79 | 2.24 | 2.00 | 1.00 | 1.45 | 1.04 | .16 |
| 31 | 2.36 | 2.00 | 1.00 | 1.48 | 1.00 | .09 | 1.80 | 1.00 | 1.00 | 1.24 | 1.58 | 1.72 |
| 32 | - | - | - | - | - | - | 1.09 | 1.00 | 1.00 | .50 | 6.25\* | 42.12\* |
| 33 | 2.41 | 2.00 | 1.00 | 1.36 | .92 | .21 | 1.75 | 1.00 | 1.00 | 1.07 | 1.68 | 2.79 |
| 34 | 3.77 | 4.00 | 3.00 | 1.59 | -.19 | -1.01 | 3.19 | 3.00 | 1.00 | 1.66 | .18 | -1.11 |

*Note.* *M* = mean, *Mdn* = median, *Mo* = mode, *SD* = standard deviation, *Sk* = skewness, *Ku* = kurtosis.

\* Discrepant values of skewness and/or kurtosis.

Table 2 presents the factor loadings and the residual variances of the items estimated with the CFA for women (32-item) and men (34-item). All factor loadings were > .50 and the residual variances were relatively stable in both subsamples.

Table 2. *Factor Loadings of the Items of the Body Shape Questionnaire (BSQ) in Women and Men obtained in Confirmatory Factorial Analysis and the Results of Differential Item Functioning (DIF) Analysis*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Women | | Men | | DIF (Men vs. Women) |
| Item | λ | e | λ | e | *p*-value for χ2 |
| 1 | .56 | .69 | .55 | .70 | .049 |
| 2 | .84 | .29 | .79 | .38 | .001 |
| 3 | .67 | .55 | .65 | .57 | <.001 |
| 4 | .84 | .30 | .79 | .37 | .021 |
| 5 | .71 | .49 | .70 | .51 | <.001 |
| 6 | .77 | .41 | .74 | .45 | <.001 |
| 7 | .77 | .41 | .84 | .30 | <.001 |
| 8 | .68 | .54 | .78 | .38 | .001 |
| 9 | .85 | .27 | .86 | .26 | <.001 |
| 10 | .76 | .42 | .82 | .32 | <.001 |
| 11 | .78 | .39 | .83 | .32 | .076 |
| 12 | .80 | .36 | .71 | .50 | <.001 |
| 13 | .74 | .46 | .76 | .41 | <.001 |
| 14 | .86 | .26 | .86 | .27 | .013 |
| 15 | .79 | .38 | .78 | .39 | <.001 |
| 16 | .77 | .41 | .82 | .32 | <.001 |
| 17 | .84 | .30 | .84 | .29 | <.001 |
| 18 | .75 | .44 | .82 | .33 | .029 |
| 19 | .87 | .24 | .86 | .25 | .195 |
| 20 | .85 | .28 | .81 | .34 | .168 |
| 21 | .79 | .38 | .75 | .44 | .423 |
| 22 | .66 | .57 | .70 | .51 | <.001 |
| 23 | .77 | .41 | .74 | .45 | .053 |
| 24 | .85 | .27 | .85 | .27 | <.001 |
| 25 | .74 | .44 | .81 | .35 | <.001 |
| 26 | - | - | .86 | .27 | - |
| 27 | .72 | .48 | .77 | .40 | .043 |
| 28 | .85 | .28 | .86 | .25 | <.001 |
| 29 | .86 | .25 | .84 | .29 | <.001 |
| 30 | .70 | .51 | .74 | .45 | .015 |
| 31 | .78 | .39 | .80 | .36 | .042 |
| 32 | - | - | .78 | .39 | - |
| 33 | .88 | .23 | .83 | .31 | .129 |
| 34 | .73 | .46 | .67 | .55 | <.001 |

Note. λ = Factor Loading, e = Residual Variance (1- λ2). χ2 = Chi-square test.

p < .01.

Table 3 presents the results of the CFA and the reliability of BSQ scores to the complete and to the shorter BSQ versions in women and men. For women, we found acceptable fit indices in the 32-item and 8B-item models. For men, we found acceptable fit indices in the 34-item, 32-item, 16A-item, 16B-item, and 8C-item models. It should be highlighted that χ2/df values were above the cut-off because this index is usually sensitive to sample size. Regarding reliability, in all models the estimates of ordinal alpha coefficient (α = .90-.98) and of omega coefficient (ω = .86-.97) were adequate. To allow for further analyses, we performed tests for sex invariance using the 32-item model of BSQ scores (i.e., omitting Items 26 and 32, since they showed discrepant values of skewness and kurtosis in the male sample) and all shorter versions. The results indicated that the structure of models was equal among sex and country with adequate configural invariance. This result allowed us to continue to analyze the next levels of invariance across sex.

Table 3. *Fit Indices for the Models of the Body Shape Questionnaire (BSQ) in Women and Men*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample | Model | λ | χ2/*df* | CFI | TLI | RMSEA [90% CI] | α | ω |
| Women | 32-item | .56-.88 | 16.79 | .93 | .93 | .09 [.09, .10] | .98 | .97 |
|  | 16A-item | .32-.74 | 17.88 | .95 | .94 | .10 [.10, .11] | .95 | .94 |
|  | 16B-item | .47-.77 | 24.32 | .96 | .95 | .12 [.12, .12] | .96 | .95 |
|  | 8A-item | .30-.73 | 32.52 | .95 | .93 | .14 [.13, .15] | .90 | .87 |
|  | 8B-item | .45-.69 | 11.86 | .98 | .98 | .08 [.07, .09] | .91 | .89 |
|  | 8C-item | .49-.77 | 31.70 | .97 | .96 | .14 [.13, .15] | .93 | .92 |
|  | 8D-item | .50-.76 | 17.29 | .98 | .97 | .10 [.09, .11] | .92 | .90 |
| Men | 34-item | .55-.87 | 6.25 | .93 | .93 | .08 [.07, .08] | .98 | .97 |
|  | 32-item | .55-.87 | 6.65 | .93 | .93 | .08 [.07, .08] | .98 | .97 |
|  | 16A-item | .31-.74 | 6.44 | .96 | .95 | .08 [.07, .08] | .95 | .93 |
|  | 16B-item | .51-.74 | 9.46 | .95 | .94 | .10 [.09, .10] | .96 | .94 |
|  | 8A-item | .31-.76 | 11.17 | .96 | .94 | .11 [.10, .12] | .90 | .86 |
|  | 8B-item | .49-.76 | 8.74 | .97 | .96 | .09 [.08, .11] | .92 | .87 |
|  | 8C-item | .49-.74 | 4.64 | .99 | .98 | .06 [.05, .08] | .93 | .90 |
|  | 8D-item | .52-.73 | 15.19 | .96 | .95 | .13 [.11, .14] | .92 | .88 |

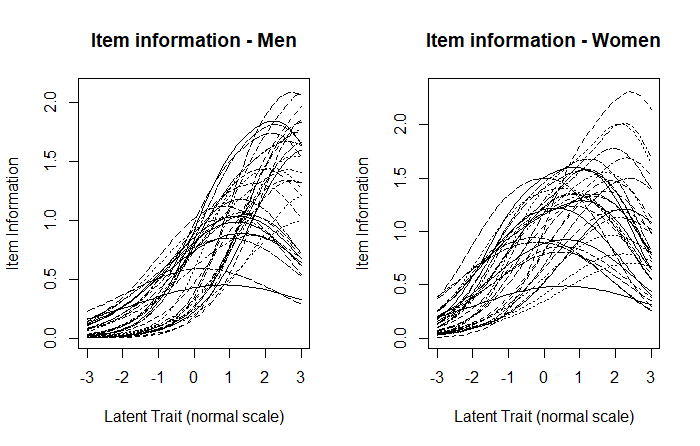
*Note*. λ = Factor Loading, χ2/*df* = Value of Chi-square divided by Degrees of Freedom, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation [90% CI = confidence interval of 90%], α = ordinal alpha coefficient, ω = omega coefficient.

Table 4 presents the remaining invariance tests across sex. We were unable to show metric, scalar or strict invariance in any model across sex. For this reason, we followed-up with DIF analysis, which indicated that the responses for 19 of the 32 BSQ items were significantly different between women and men (see Table 2 for *p*-values for the χ2 test for the DIF analysis). For women, responses ranged around the standardized mean (*M* = 2.53, *SD* = 1.07) regardless of whether they chose low or high values; that is, BSQ items were informative for the women. For men, most responses were below the standardized mean (*M* = 1.83, *SD* = .81) and were only informative for those who chose high values on BSQ items (see Figure 1).

Table 4. *Invariance Tests for the Models of the Body Shape Questionnaire (BSQ) Across Sex*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | χ2/*df* | CFI | TLI | RMSEA [90% CI] | Δχ2λ *(df)* | Δχ2*T (df)* | Δχ2cov *(df)* | *p* |
| 32-item | 11.12 | .94 | .93 | .09 [.09, .09] | 374.86 (31) | 618.15 (158) | 411.51 (127) | < .001 |
| 16A-item | 11.63 | .96 | .95 | .09 [.09, .10] | 191.17 (15) | 372.91 (78) | 253.19 (63) | < .001 |
| 16B-item | 16.45 | .96 | .95 | .11 [.11, .11] | 104.70 (15) | 220.53 (78) | 150.90 (63) | < .001 |
| 8A-item | 21.28 | .95 | .93 | .13 [.12, .13] | 89.94 (7) | 232.88 (38) | 166.12 (31) | < .001 |
| 8B-item | 8.12 | .99 | .98 | .08 [.07, .08] | 50.88 (7) | 133.85 (38) | 96.35 (31) | < .001 |
| 8C-item | 23.06 | .97 | .96 | .13 [.13, .14] | 29.16 (7) | 70.25 (38) | 48.84 (31) | < .05 |
| 8D-item | 12.78 | .98 | .97 | .10 [.09, .10] | 56.44 (7) | 128.36 (38) | 85.70 (31) | < .001 |

*Note.* λ = Factor Loading, χ2/*df* = Value of Chi-square divided by Degrees of Freedom, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation [90% CI = confidence interval of 90%],*T* = Threshold, Cov = Covariance, Δχ2λ = Metric Invariance, Δχ2*T* = Scalar Invariance, Δχ2cov = Strict Measure Invariance.



*Figure 1. Differential item Functioning (DIF) of the Items of the Body Shape Questionnaire (BSQ) in Men and Women*

**Discussion**

The BSQ is a widely-used measure of body shape concerns designed for women respondents, and it has been occasionally applied without appropriate care for use with boys and men; that is, previous studies have neglected to examine the extent to which BSQ scores are invariant across sex (e.g., Amaral & Ferreira, 2017; Branco et al., 2006; Conti et al., 2009; Laus et al., 2013; Di Pietro & da Silveira, 2009). The results of the present study, conducted with a Portuguese-speaking sample of Portuguese and Brazilian university students, suggests that BSQ scores in all models (32-item and shorter versions) should not be considered invariant across sex in this population. The results of our multi-group CFA indicated that invariance could not be supported at the metric, scalar, or strict levels of invariance, posing concerns for conclusions drawn by earlier studies comparing women’s and men’s BSQ scores. Given that BSQ items may function differently for women and men respondents, any sex comparisons and interpretations of the latent meaning of BSQ scores is questionable (cf. Slof-Op t’Landt et al., 2009).

Indeed, the results of the present study suggest that there are differences in item functioning across sex for the majority of BSQ items. The most parsimonious explanation of these findings is that the item content of the BSQ may not adequately represent the body concerns of men, which are likely to be heavily focused on upper-body muscularity (e.g., Swami & Tovée, 2005; Dakanalis et al., 2015). This is unsurprising since the BSQ was originally designed to assess body shape concerns in women with emphasis on their frequent dissatisfaction with lower body shape (e.g., thighs and hips). As a result, it is quite likely that previous studies using the BSQ in male and female comparisons have grossly underestimated body image concerns in men and/or over-stated the case for sex differences in body image dissatisfaction. While our findings might be unsurprising to some, they are important in the context of understanding prior research using the BSQ with boys and men under the assumption that items will function similarly with them as when the instrument is completed by girls and women. From a practical point-of-view, we strongly recommend that scholars desist from further tests of sex differences on BSQ scores in the absence of demonstrated sex invariance of the BSQ when used in the specific populations of interest to the scholars. Unless future research demonstrates the appropriate use of the BSQ with men of certain populations, our results suggest that the BSQ is not appropriate for use with men.

The present study is not without its limitations. First, the use of a non-probability recruitment design, targeting university students, means that caution should be exercised when generalizing the present findings to a wider array of Portuguese-speaking groups or, indeed, to other national or linguistic populations. It should be noted, however, that the BSQ has been used with boys and men outside Portuguese-speaking samples (Flores-Cornejo et al., 2017; Laporta-Herrero et al., 2016; Najam & Ashfaq, 2012; Pokrajac-Bulian & Živčić-Bećirević, 2005; Sarwer et al., 2015), and we suspect our broad conclusion would be applicable in other cultural groups given the item content of the BSQ. Second, the present study was focused on sex invariance, and we did not conduct an examination of additional indices of validity across women and men. For example, it might be instructive to examine the extent to which BSQ scores in men demonstrate adequate convergent validity, which could be examined through associations with measures designed specifically to tap body image issues of particular interest to men (e.g., drive for muscularity). Third, we included individuals from two countries in the samples of women and men and confirmed the BSQ’s configural invariance between respondents from Brazil and Portugal before we could continue analyzing other levels of invariance. Cultural differences between these countries can exist, but a previous study (see Silva et al., 2016) reported the transnational invariance of the BSQ in separate samples of women, leading us to focus this study on sex invariance. Transnational invariance, even among women alone may be more difficult to establish when comparing other nationalities that may be less similar in their ancestral heritage and cultural customs.

Despite these limitations, the results of the present study raise important concerns about the use of the BSQ with men and boys; and these findings raise important concerns about prior comparisons of the latent meaning of BSQ scores obtained by women and men. When BSQ differences across sex have been found, far from indicating manifest sex differences in body shape concerns, our results suggest that earlier conclusions are very likely to be an artifact of differential BSQ item coverage of body shape concerns of respondents of a different sex. More worryingly, our results suggest that scholars should avoid drawing comparisons between women and men on their BSQ scores, given the failure in the present study to obtain satisfactory indices of sex invariance. We recommend avoiding the use of BSQ in men, as it may not capture the primary body concerns of this population segment. In addition, we conclude that scholars in body image research should avoid making between-group sex comparisons with the BSQ in the absence of its measurement invariance for these groups.

**Declaration of conflicting interests**

The authors declare that there is no conflict of interest.

**Funding and Acknowledgement**

The authors thank the São Paulo Research Foundation (FAPESP; grant #2014/03093-2, #2015/00228-7), the Coordination for the Improvement of Higher Education Personnel (CAPES), and the National Counsel of Technological and Scientific Development (CNPq; 142315/2014-1) for financing the study. We would also like to thank Professors Afonso Cavaco, Filipa Costa, José Santos, Maria Caramona, Célia Oliveira, Maria Costa, and Maria da Graça Vinagre, as well as Bianca Martins, Andreia Caldeira, Moema Santana, and Fernanda Cristina Maurício for their contributions to the study.

**References**

Amaral, A. C. S., & Ferreira, M. E. C. (2017). Body dissatisfaction and associated factors among Brazilian adolescents: A longitudinal study. *Body Image, 22,* 32-38. doi:10.1016/j.bodyim.2017.04.006

Branco, L. M., Hilário, M. O. E., & Cintra, I. P. (2006). Percepção e satisfação corporal em adolescents relação com seu estado nutricional [Perception and satisfaction with body image in adolescents in relation to nutritional status]. *Revista de Psiquiatria Clinica*, *33*, 292-296. doi:10.1590/S0101-60832006000600001

Brockmeyer, T., Anderle, A., Schmidt, H., Febry, S., Wunsch-Leiteritz, W., Leiteritz, A., & Friederich, H. C. (2018). Body image related negative interpretation bias in anorexia nervosa. *Behavioral Research Therapy*, *104*, 69-73. doi:10.1016/j.brat.2018.03.003

Chen, F. F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464-504. doi:10.1207/s15328007sem1203\_7

Choi, S. W. (2016) lordif: Analysis of Differential Item Functioning (DIF) for dichotomous and polytomous items using an iterative hybrid of ordinal logistic regression and item response theory (IRT). Available online at: [https://CRAN.R-project.org/package=lordif](https://cran.r-project.org/package=lordif)

Conti, M. A., Cordás, T. A., & de Oliveira Latorre, M. R. D. (2009). Estudo de validade e confibilidade da versão brasileira do Body Shape Questionnaire (BSQ) para adolescentes [A study of the validity and reliability of the Brazilian version of the Body Shape Questionnaire (BSQ) among adolescents]. *Revista Brasileira de Saúde Materno Infantil*, *9*, 331-338. doi:10.1590/S1519/28292009000300012

Cooper, P. J., Taylor, M. J., Cooper, Z., & Fairburn, C. G. (1987). The development and validation of the Body Shape Questionnaire. *International Journal of Eating Disorders*, *6*, 485-494. doi:10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O

Dakanalis, A., Favagrossa, L., Clerici, M., Prunas, A., Colmegna, F., Zanetti, M. A., & Riva, G. (2015). Body dissatisfaction and eating disorder symptomatology: A latent structural equation modeling analysis of moderating variables in 18-to-28-year-old males. *Journal of Psychology*, *149*, 85-112. doi:10.1080/00223980.2013.842141

Di Pietro, M., & da Silveira, D. X. (2009). Internal validity, dimensionality, and performance of the Body Shape Questionnaire in a group of Brazilian college students. *Revista Brasileira de Psiquiatria*, *31*, 21-24. doi:10.590/S1516-44462008005000017

Dondzilo, L., Rieger, E., Palermo, R., Byrne, S., & Bell, J. (2017). The mediating role of rumination in the relation between attentional bias towards thin female bodies and eating disorder symptomatology. *PLoS One, 12*, e0177870. doi:10.1371/journal.pone.0177870

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practicalsolution to the pervasive problem of internal consistency estimation. British Journal of Psychology, 105, 399–412. doi:10.1111/bjop.12046

Evans, C., & Dolan, B. (1993). Body Shape Questionnaire: Derivation of shortened “alternate forms”. *International Journal of Eating Disorders*, *13*, 315-321. doi:10.1002/1098-108X(199304)13:3<315::AID-EAT2260130310>3.0.CO;2-3

Flores-Cornejo, F., Kamego-Tome, M., Zapata-Pachas, M. A., & Alvarado, G. F. (2017). Association between body image dissatisfaction and depressive symptoms in adolescents. *Brazilian Journal of Psychiatry, 39*, 316-322. doi:10.1590/1516-4446-2016-1947

[Laporta-Herrero, I](https://www.ncbi.nlm.nih.gov/pubmed/?term=Laporta-Herrero%20I%5BAuthor%5D&cauthor=true&cauthor_uid=28039667)., [Jáuregui-Lobera, I](https://www.ncbi.nlm.nih.gov/pubmed/?term=J%C3%A1uregui-Lobera%20I%5BAuthor%5D&cauthor=true&cauthor_uid=28039667)., [Barajas-Iglesias, B](https://www.ncbi.nlm.nih.gov/pubmed/?term=Barajas-Iglesias%20B%5BAuthor%5D&cauthor=true&cauthor_uid=28039667)., & [Santed-Germán, M. Á](https://www.ncbi.nlm.nih.gov/pubmed/?term=Santed-Germ%C3%A1n%20M%C3%81%5BAuthor%5D&cauthor=true&cauthor_uid=28039667). (2016). Body dissatisfaction in adolescents with eating disorders. [*Eating and Weight Disorders*.](https://www.ncbi.nlm.nih.gov/pubmed/28039667) *23*, 339-347. doi:10.1007/s40519-016-0353-x

Laus, M. F., Miranda, V. P. N., Almeida, S. S., Costa, T. M. B., & Ferreira, M. E. C. (2013). Geographic location, sex, and nutritional status play an important role in body image concerns among Brazilian adolescents. *Journal of Health Psychology*, *18*, 332-338. doi:10.1177/1359105311434755

Marôco, J. (2014). *Análise de equações estruturais* [*Structural equation analysis*]. Lisbon: ReportNumber.

Moreno, D. C., Montaño, I. L., Prieto, G. A., & Pérez-Acosta, A. M. (2007). Validación del Body Shape Questionnaire (Cuestionario de la figura corporal) BSQ para la población Colombiana [Validation of the Body Shape Questionnaire (body image questionnaire) in a Colombian population]. *Acta Colombiana de Psicologia*, *10*, 15-23.

Najam, N., & Ashfaq, H. (2012). Gender Differences in Physical Fitness, Body Shape Satisfaction, and Body Figure Preferences. *Pakistan Journal of Psychological Research, 27*, 187-200.

Pokrajac-Bulian, A., & Živčić-Bećirević, I. (2005). Locus of control and self-esteem as correlates of body dissatisfaction in Croatian university students. *European Eating Disorders Review, 13*,54-60. doi:10.1002/erv.606

Revelle, W. (2018) psych: Procedures for personality and psychological research, Northwestern University, Evanston, IL. Available online at: [https://CRAN.R-project.org/package=psych](https://cran.r-project.org/package=psych) Version = 1.8.4.

Rousseau, A., Knotter, A., Barbe, P., Raich, R. M., & Chabrol, H. (2005). Étude de validation de la version française du Body Shape Questionnaire [Validation of the French version of the Body Shape Questionnaire]. *L’Encéphale*, *31*, 162-173.

Sarwer, D. B., Spitzer, J. C., Wadden, T. A., Rosen, R. C., Mitchell, J. E., Lancaster, K., … & Christian, N. J. (2015). *Surgery for Obesity and Related Diseases, 11*, 643-51. doi:10.1016/j.soard.2014.12.014

Silva, Dias, J. C. R., Marôco, J., & Campos, J. A. D. B. (2014). Confirmatory factor analysis of different versions of the Body Shape Questionnaire applied to Brazilian university students. *Body Image*, *11*, 384-390. doi:10.1016/j.bodyim.2014.06.001

Silva, W. R., Costa, D., Pimenta, F., Marôco, J., & Campos, J. A. D. B. (2016). Psychometric evaluation of a unified Portuguese-language version of the Body Shape Questionnaire in female university students. *Cadernos de Saúde Pública*, *32*, 21. doi:10.1590/0102-311X00133715

Silva, W. R., Dias, J. C. R., Pimenta, F., & Campos, J. A. D. B. (2014). Assessment of the body shape concern: A challenge. *Psychology, Community, and Health*, *3*, 2. doi:10.5964/pch.v3i2.89

Silva, W. R. D., Campos, J., & Maroco, J. (2018). Impact of inherent aspects of body image, eating behavior and perceived health competence on quality of life of university students. *PLOS One*, *13*, e0199480. doi:10.1371/journal.pone.0199480

Slof-Op t’Landt, M. C. T., van Furth, E. F., Rebollo-Mesa, I., Bartels, M., van Beijsterveldt, C. E. M., Slagboom, P. E., … & Dolan, C. V. (2009). Sex differences in sum scores may be hard to interpret: The importance of measurement invariance. *Assessment*, *16*, 415-423. doi:10.1177/1073191109344827

Swami, V., & Barron, D. (2018). Translation and validation of body image instruments: Challenges, good practice guidelines, and reporting recommendations for test adaptation. *Body Image*, advanced online publication. doi:10.1016/j.bodyim.2018.08.014

Swami, V., & Tovée, M. J. (2005). Male physical attractiveness in Britain and Malaysia: A cross-cultural study. *Body Image*, *2*, 383-393. doi:10.1016/j.bodyim.2005.08.001