Body Image 9 (2012) 163–171

Contents lists available at SciVerse ScienceDirect

Body Image

journal homepage: www.elsevier.com/locate/bodyimage

Analysis of the factor structure of the Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ-3) in Spanish secondary-school students through exploratory structural equation modeling

David Sánchez-Carracedo[a,*], Juan Ramón Barrada[b,1], Gemma López-Guimerà[a,2], Jordi Fauquet[c,3], Carlos A. Almenara[a,2], Esther Trepat[d,4]

[a] Dept. Clinical and Health Psychology, Universitat Autònoma de Barcelona, Barcelona, Spain
[b] Dept. Psychology and Sociology, Universidad de Zaragoza, Teruel, Spain
[c] Dept. Psychobiology and Methodology of Health Sciences, Universitat Autònoma de Barcelona, Barcelona, Spain
[d] Institute of Psychology, Private Foundation, Barcelona, Spain

ARTICLE INFO

Article history:
Received 15 March 2011
Received in revised form 6 October 2011
Accepted 13 October 2011

Keywords:
Sociocultural Attitudes Towards Appearance Questionnaire-3 (SATAQ-3)
Exploratory structural equation modeling
Media influences
Internalization of beauty ideal
Adolescent

ABSTRACT

The aims of the present study were: (1) to assess the factor structure of the SATAQ-3 in Spanish secondary-school students by means of exploratory factor analysis (EFA), confirmatory factor analysis (CFA) and exploratory structural equation modeling (ESEM) models; and (2) to study its invariance by sex and school grade. ESEM is a technique that has been proposed for the analysis of internal structure that overcomes some of the limitations of EFA and CFA. Participants were 1559 boys and girls in grades seventh to tenth. The results support the four-factor solution of the original version, and reveal that the best fit was obtained with ESEM, excluding Item 20 and with correlated uniqueness between reverse-keyed items. Our version shows invariance by sex and grade. The differences between scores of different groups are in the expected direction, and support the validity of the questionnaire. We recommend a version excluding Item 20 and without reverse-keyed items.

Introduction

Recent reviews and meta-analyses conclude that media messages and sociocultural pressures are among the principal risk factors behind body dissatisfaction, weight concerns, and disordered eating behavior (e.g., Grabe, Ward, & Hyde, 2008; Levine & Harrison, 2004; Levine & Murnen, 2009; López-Guimerà, Levine, Sánchez-Carracedo, & Fauquet, 2010). The role of the internalization of beauty ideals as mediator between sociocultural pressures and the development of disordered eating and body dissatisfaction has been shown in numerous studies, with both women and men (e.g., Blond, 2008; Cafri, Yamamiya, Brannick, & Thompson, 2005; Durkin, Paxton, & Sorbello, 2007; Thompson & Stice, 2001). Reduction in levels of internalization is thus one of the main objectives in the field of the prevention of eating disorders (e.g., Levine & Smolak, 2006; López-Guimerà et al., 2010; López-Guimerà & Sánchez-Carracedo, 2010; Stice & Shaw, 2004; Stice, Shaw, & Marti, 2007), and appears to be a good indicator for evaluating treatment efficacy in eating disorders (Heinberg, Coughlin, Pinto, Haug, Brode, & Guarda, 2008).

Development of the SATAQ and Its Different Versions

The Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ; Heinberg, Thompson, & Stormer, 1995), in its various versions, is the instrument most widely used today for assessing these sociocultural pressures and the internalization of the beauty ideal. As such, it constitutes the gold standard in this field worldwide (López-Guimerà & Sánchez-Carracedo, 2010). Originally, the SATAQ assessed awareness of the cultural ideal of beauty for women (*Awareness* subscale) and the level of acceptance and internalization of that ideal (*Internalization* subscale). The original SATAQ was slightly modified (SATAQ-R) by adding new items that took into account the recent focus on athleticism and sports in young women (Cusumano & Thompson, 1997). The third and most recent version is SATAQ-3 (Thompson, van den Berg, Roehrig, Guarda, & Heinberg, 2004), a 30-item scale with four theoretical subscales. Two of these subscales are based on different internalization factors. The first, with nine items, is *Internalization-General*,

* Corresponding author. Tel.: +34 935813855; fax: +34 935812125.
E-mail addresses: David.Sanchez@uab.es (D. Sánchez-Carracedo), Juanramon.Barrada@unizar.es (J.R. Barrada), Gemma.Lopez@uab.cat (G. López-Guimerà), Jordi.Fauquet@uab.cat (J. Fauquet), almenara.carlos@gmail.com (C.A. Almenara), etrepat@ipsicologia.cat (E. Trepat).
1 Tel.: +34 978618101; fax: +34 978618103.
2 Tel.: +34 935813855; fax: +34 935812125.
3 Tel.: +34 935811674; fax: +34 935812001.
4 Tel.: +34 932541916; fax: +34 934021362.

1740-1445/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bodyim.2011.10.002

and assesses general media influence related to TV, magazines, and movies. The second, with five items, is *Internalization-Athlete*, and assesses the internalization of athletic and sports models. The other two subscales are *Information*, with nine items, which assesses how far it is acknowledged that various media are considered important sources of information about appearance, and *Pressures*, with seven items, which assesses subjective feelings of pressure from exposure to media images and messages to modify one's appearance. The items themselves can be seen in Table 2. The response format is a Likert-type scale ranging from 1 = *completely disagree* to 5 = *completely agree*. The SATAQ-3 has shown a relatively stable internal structure, as well as boasting good indicators of reliability and validity when applied to Western women (Markland & Oliver, 2008; Thompson et al., 2004).

Considerable interest has been generated by the SATAQ and its various versions among researchers. The SATAQ has been adapted to languages such as Arabic (Madanat, Hawks, & Brown, 2006), Chinese (Jackson & Chen, 2010), French (Rousseau & Valls, 2010; Rousseau, Valls, & Chabrol, 2010), German (Knauss, Paxton, & Alsaker, 2009), and Malay (Swami, 2009). The instrument has been validated for patients with eating disorders (Calogero, Davis, & Thompson, 2004; Heinberg et al., 2008), for adolescent boys (Rousseau & Valls, 2010; Smolak, Levine, & Thompson, 2001), and for undergraduate males (Karazsia & Crowther, 2008).

The SATAQ-3: Different Structures, Versions, and Methodologies of Analysis

The original internal structure of the SATAQ-3 has not been replicated in the different versions and adaptations. A wide variety of methodological criteria have been followed to analyze the factor structure. Only the validation with Jordanian women (Madanat et al., 2006), using exploratory factor analysis (EFA) with principal axis as extraction method, replicated the four-factor structure from the original questionnaire while keeping all the items.

The French version with adolescent girls (Rousseau et al., 2010) initially found a solution of five factors. Since only Item 20 loaded on Factor 5, the four-factor solution of the original version was retained, and the authors dropped Item 20 for their version. In the French version with adolescent boys (Rousseau & Valls, 2010) an initial solution of 5 factors was again found. As no item substantially loaded in the fifth factor, the solution finally adopted was that of four factors, as in the original version. Item 20 was not cited as taking part in it and Item 18 failed to show satisfactory saturation. For both French versions, a principal components analysis (not a factor analysis) was used.

In the work with Malaysian women (Swami, 2009), an EFA was used and a solution with four factors was chosen, two of which, *Information* and *Internalization-Athlete*, mirrored the originals; the third factor was an amalgamation of the *Pressures* and *Internalization-General* factors in the original version, while the fourth factor had several items cross-loaded onto previous factors, which were eventually dropped from the analysis. Where items had relevant loadings on several factors, conceptual solutions were adopted. Items 3 and 27 were discarded from the final analysis.

The principal components analysis carried out by Calogero et al. (2004) with eating-disordered patients replicated the original four-factor solution found for a nonclinical sample, but several items cross-loaded on more than one factor. Additionally, Item 20, belonging to the *Internalization-Athlete* scale in the original version, loaded on this scale, but more strongly on *Internalization-General* (R. Calogero, personal communication, March, 15th 2011).

In the Chinese version with adolescent boys (Jackson & Chen, 2010), Item 20 was not considered in the analysis in view of its problematic performance in previous studies (Calogero et al., 2004; Markland & Oliver, 2008), and several items were modified slightly to make them gender-neutral. Two factor analyses with different samples were carried out. An initial EFA revealed a four-factor solution with dimensions reflecting *General Pressure-Internalization*, *Sources of Appearance Information*, *Pressure-Internalization of an Athletic Ideal*, and *Pressure to be Thin*. Item 1 was discarded from the analysis as it cross-loaded in two factors. This structure, without Items 1 and 20, was adopted by the authors in a subsequent confirmatory factor analysis (CFA).

Finally, the study carried out with American undergraduate males (Karazsia & Crowther, 2008) by means of a CFA replicated the original four-factor structure, but the authors did not retain the original items, rewording several items focusing on "thinness" to focus on "muscularity", so that the properties of the original version with males could not be checked. Also, for theoretical reasons, the authors decided to drop Item 15, an item referring to the internalization of the beauty ideal transmitted by music videos, from the analysis. However, Item 9, which refers to music videos as an information source for the beauty ideal, was retained.

In a footnote to the original development of the SATAQ-3, the authors recommended a new version with 8 of the 30 items reverse-keyed in order to reduce possible response bias. Most of the SATAQ-3 versions published up to now have failed to take notice of this suggestion. The only exception is the CFA carried out by Markland and Oliver (2008) with British nonclinical young women. The model with the best fit is that which excludes Item 20, the same item that performed unsatisfactorily in the versions by Calogero et al. (2004) and Rousseau et al. (2010).

Summarizing, there are four main points to highlight in the development and validation of the SATAQ-3: (a) there are different analytical approaches in the study of the internal structure; (b) versions with and without reverse-keying have been employed; (c) there are versions in which item wording has been changed for their use with certain groups (e.g., boys), so that direct comparison of scores is not possible between groups; and (d) for some versions, Item 20 has been discarded (in advance of data collection or after analysis). This may explain why the results have not been always completely congruent.

Thus, we shall examine the characteristics of a recent technique for the analysis of the internal structure of a questionnaire: exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009), an integration of CFA and EFA.

Exploratory Structural Equation Modeling (ESEM)

We shall present the characteristics of ESEM through comparison with the main limitations of other methods for the assessment of the internal structure of tests, such as EFA and CFA. EFA is usually referred to as a data-driven technique (Fabrigar, Wegener, MacCallum, & Strahan, 1999), and is commonly used with the aim of obtaining a simple and interpretable structure. Basically, and as far as this study is concerned, there are two main limitations of EFA (e.g., Brown, 2006). First, when items share any element in their wording without theoretical relevance, they may show greater covariance than can be explained merely by their relation to the measured constructs. In these cases the interpretation of the internal structure of the questionnaire becomes complex, or actually misleading. This limitation may apply in the case of the SATAQ-3, given the strong similarity in the wording of some items and the recommendation of its authors to include items with reverse wording (Thompson et al., 2004). And, second, EFA does not permit the correct evaluation of the measurement invariance across different groups (Meredith, 1993). Measurement invariance implies that the same score has the same interpretation for the different groups. The comparability of scores between groups is not something that can be assumed by default, but rather has to be supported by evidence.

CFA is considered a theory-driven technique, as the number of dimensions and the items-factors relationship with which the covariance matrix will be explained must be supported by a strong previous theory or by previous EFAs in which a simple structure has been found. In a CFA the factor loadings are usually estimated with the restriction that each item will only load in the expected factor, the other loadings being fixed to 0. Correlated uniqueness can be included in the model in such a way that the loadings are not distorted by spurious factors or redundant items. CFAs enable the testing of measurement invariance (Vandenberg & Lance, 2000). The main limitation of CFA is the restrictive assumption: The factor structure is fully simple (Asparouhov & Muthén, 2009). While in the EFA context simple structure implies no salient loadings in the secondary dimensions, in the CFA context simple structure means no loading at all. In CFA, any non-modeled loading different from 0 in the population reduces the model fit and can bias the results. This means that some theoretically consolidated assessment instruments, supported by extensive research, do not offer an acceptable fit when modeled with a CFA (Marsh, 2007). When minor cross-loadings are fixed to 0, the correlation between dimensions is spuriously inflated. This limitation could apply to the study of the factor structure of the different versions of the SATAQ-3, given the usual finding of cross-loadings in the different EFAs (e.g., Calogero et al., 2004).

ESEM (Asparouhov & Muthén, 2009) incorporates many of the advantages of CFA, but is free of its limitations. Like EFA, ESEM permits the estimation of the factor loadings of all items in all factors, so that the problem of fixing the cross-loadings to 0 disappears. When the loading matrix of the population includes cross-loadings, ESEM recovers this matrix better than CFA and is not subject to its parameter estimation bias. As such, ESEM may be the most appropriate model for the SATAQ-3 in its various versions. Like CFA, ESEM permits the testing of measurement invariance. To date, no studies have applied ESEM to the SATAQ-3. In fact, ESEM is a very recent proposal which has not been widely used up to now (e.g., Marsh et al., 2009, 2010; Rosellini & Brown, 2011).

Purposes of the Study

This study has two main purposes. First, to evaluate the factor structure of the SATAQ-3 with adolescent girls and boys in its adaptation to Spanish. The long-overdue availability of a Spanish version of the SATAQ-3 (López-Guimerà, Sánchez-Carracedo, Fauquet, Portell, & Raich, 2011) will provide research teams from Spanish-speaking countries studying psychosocial risk factors associated with eating disorders and body image, or working in eating disorder prevention, with a validated version of an instrument that currently constitutes the gold standard for assessing some of the most relevant variables in these research fields, such as media pressures towards the beauty ideal or the internalization of that ideal. In contrast to the cases of some of the previous studies, the version of the SATAQ-3 administered follows the final wording suggested by the authors, with eight reverse-keyed items. And the second main purpose is to study the measurement invariance of the instrument for boys and girls and across the four grades of Spanish Compulsory Secondary Education. In doing so, and also in contrast to previous studies, the same wording is kept for the whole sample, avoiding the use of different (and non-comparable) questionnaires for different groups. Analysis of invariance allows comparability between different groups with the same instrument, and has become increasingly common in recent years in the field of disordered eating and body image (e.g., Fonseca-Pedrero, Sierra-Baigrie, Paino, Lemos-Giráldez, & Muñiz, 2011; Rusticus, Hubley, & Zumbo, 2008; Slof-Op't Landt et al., 2009; Warren et al., 2008).

Method

Participants

The initial sample consisted of 1559 adolescents (749 girls—48.04% and 810 boys—51.94%) recruited from 5 schools (2 public and 3 grant-aided private schools) in the area of Barcelona, Spain. Participants' ages ranged from 12 to 17 years (mean = 14.3, *SD* = 1.4). The sample comprised students from the four years of Compulsory Secondary Education in the Spanish system (7th to 10th grade in the USA). Participants were roughly equally distributed across grades: 396 (25.4%), 403 (25.8%), 392 (25.1%) and 368 (23.6%) for grades from 7th to 10th. Self-reported origin of participants was as follows: 83.4% Spanish, 7.2% Latin-American, 2.3% from other European countries (Spain excluded), 1.0% African (0.9% from North Africa and 0.1% from Sub-Saharan Africa), 5.7% mixed origins, and 0.4% who did not specify their origin. After the removal of participants with missing data (see 'Results' section), the final sample numbered 1501.

Procedures

This study is part of the pilot phase of a broader research project on the prevention of disordered eating. A battery of questionnaires about body image and eating attitudes was administered in pen-and-paper format. In this study we shall consider only data from the SATAQ-3.

The research was approved by the Clinical Research Ethics Committee of the "Parc Taulí" Health Corporation in Sabadell. Adolescents were allowed to participate only if signed parental consent was obtained. Administration of the questionnaires took place during the period from February to May 2009. Those administering the questionnaires (graduate and post-graduate psychologists) received detailed verbal and written instructions on how to proceed. The translation process was in line with international criteria (Hambleton, Merenda, & Spielberger, 2005; International Test Commission, 2010).

Data Analysis

Assessment of the internal structure took place in two distinct phases. First, we studied the internal structure of the scale following three different techniques; EFA, CFA, and ESEM. The repeated inspection of the same dataset with different techniques where the next analysis is modified according to the results from the previous analysis could lead to a capitalization on chance problem (MacCallum, Roznowski, & Necowitz, 1992). We expected this to be a minor problem in our case, given: (a) the sample size (the larger it is, the lesser the problem); and (b) that all the modifications we incorporated in our analysis were not only data-driven from the initial model, but were deeply rooted in previous results and had sound theoretical bases. In other words, the reasons that lead us to prefer the ESEM approach for modeling the SATAQ-3 could be anticipated (problems with the reverse items; problems with Item 20; some minor cross-loadings of other items—see Results section); thus, our different tested models could be defined a priori, following the suggestion of MacCallum et al. (1992).

In spite of this, we preferred a conservative approach. Participants were randomly ordered and the EFA and CFA were performed with the first and second halves of the sample, respectively ($n = 751$ and 750). The ESEM analyses were carried out with the overall sample. In this way we felt we achieved a compromise between the need to cross-validate the results and the need to keep a sample size as large as possible for the final model.

Goodness-of-fit in all derived models was assessed with the common fit index (Hu & Bentler, 1999). Thus, we consider the model satisfactory if the comparative fit index (CFI) and Tucker-Lewis index (TLI) have values greater than .95, the root mean square error of approximation (RMSEA) is less than .06, and the standardized root mean square residual (SRMR) is less than .08.

After defining the internal structure of the scale, we evaluated the reliability of scores for each factor (factor scores computed as the sum of the scores to each item). As Raykov and Marcoulides (2011, pp. 155–156) indicate, the most used indicator of reliability, Cronbach's alpha, will lead to problematic estimations whenever factor loadings of the items belonging to the same scale are not homogenous or whenever there are correlated uniquenesses. Due mainly to the negative wording of some items, both problems are present within SATAQ-3 (see 'Results' section). In these conditions, a different analytical approach should be applied based on the factor loadings and the correlations between items uniqueness (e.g., Brown, 2006; Raykov, 2009).

We also carried out a factor invariance study, since an internal structure in line with what would be expected does not necessarily mean that this structure will be maintained for the different subgroups of interest in the sample, namely sex and grade. For testing the invariance, what is evaluated is the equality (or minimal difference) between the regression lines that relate the factor score (predictive variable) and the score on the item (criterion variable). In this context, the regression line is specified by the slope (factor loadings) and the criterion value when the predictor equals 0 (intercept). A series of progressive restrictions are imposed in sequentially tested models (e.g., Vandenberg & Lance, 2000). First of all, it is verified that the model fits for each group separately. Secondly, we tested equality of form. In the context of CFA, this involves fixing for the different groups the number of factors and the factors in which each item will saturate. In ESEM, given that the items load in all the factors, only the number of factors is considered. Thirdly, we tested the equality of factor loadings. For doing so, the factor loadings are fixed as equal across the different groups. We considered those restrictions as satisfactorily met if the decrease in CFI was lower than .01 (Cheung & Rensvold, 2002) and RMSEA and TLI remained constant or increased (Marsh et al., 2009, 2010). Fourthly, we analyzed the intercept invariance. The decision rule for maintaining or rejecting it is the same as in the previous step.

Measurement invariance does not necessarily mean invariance in the relationship between factors. It is possible that the interpretations of the scores of two factors are equivalent for different groups, but that those two factors correlate in a different way in those groups. Hence the distinction between measurement invariance and structure invariance. Structure invariance is tested by fixing as equal the variances and covariances between factors for the different groups.

All the analyses were performed with M*plus* 6.11 (Muthén & Muthén, 1998–2010). None of the default specifications of M*plus* were modified. For the construction of M*plus* scripts for the different analyses the examples from Dimitrov (2010) and Marsh et al. (2009) were used as guides.

Results

Skewness and kurtosis of item scores were assessed in order to determine the estimator to apply in the factor analysis. The absolute value of skewness of all the items was lower than 2 (mean = .68; minimum = .06; maximum = 1.98), while the absolute value of kurtosis was always lower than 7 (mean = 1.00; minimum = .04; maximum = 2.89). These modest deviations from normality allowed us to employ maximum-likelihood (ML) estimation in the factor models (Russell, 2002; West, Finch, & Curran, 1995).

Participants whose percentage of missing data was over 10% in any of the questionnaires from the battery were excluded from the analysis, so that the effective sample size was of 1501 participants (96.28% of the original sample). Currently, when ML can be applied, the full information maximum likelihood (FIML) algorithm is considered to be among the most acceptable methods for accommodating missing data (Enders & Bandalos, 2001). We used FIML to deal with missing data.

To confirm that the results did not depend on the estimator used or the method used with the missing data, all the analyses were replicated with robust maximum-likelihood and multiple imputation, and an identical pattern of results found.

Internal Structure of the SATAQ-3 and Reliability

In a first EFA, with the first half of the sample, we fixed at four the number of factors to extract, following the theoretical model used in the construction of the questionnaire. The model fit was poor (see M1 in Table 1). We checked a solution with five factors (see M2 in Table 1), which brought the fit up to satisfactory levels. Factor loadings are shown in Table 2. There are two main aspects to discuss regarding the results from M2: (a) the emergence of a method factor (a statistical artifact reflecting the excess of covariance between items sharing a characteristic in their wording) that groups the reverse-keyed items; and (b) the presence of some cross-loadings. Item 20 presents important cross-loadings: Although originally designed to assess the *Internalization-Athlete* dimension, it presents significant loadings in the four substantive factors. Considering the loadings from the statistical significance perspective, five items had significant loadings in more than one substantive factor. The *Internalization-General* and *Pressures* factors are those with the greatest presence of significant cross-loadings, with three items. Not considering Item 20 and the method factor, the mean value of loadings that are significant in a factor other than the principal factor is .22, with a maximum of .26 for Items 4 and 16 in *Pressures*.

The results clearly indicate that EFA is not the appropriate technique for analyzing this data. The emergence of a method factor, irrelevant from a theoretical point of view, indicates that part of the covariance between reverse items cannot be described by means of the theoretical latent factors alone. In view of these results we decided to remove Item 20. It is questionable for an item to remain in a test when that item presents similar loadings in the factor to which it was theoretically linked at the time of its construction and also in other factors. Furthermore, from the applied perspective, this way of working would involve problems when it comes to scoring the test. The majority of test uses are based not on latent scores but on the sum of observed scores. In such cases, the treatment of an item that belongs simultaneously and markedly to several factors would be problematic.

The previous analyses clearly indicate the need to correlate the uniqueness between reverse items and the appropriateness of removing Item 20 from the instrument. Given the cross-loadings found, we expected that the CFA model would not fit. In spite of this, we tested a CFA to gain comparability with previous studies where this technique has been applied with SATAQ-3 (Jackson & Chen, study 2, 2010; Karazsia & Crowther, 2008; Markland & Oliver, 2008). The fit of this model (M3), with the second half of the sample, did not reach the established thresholds.

Next, we applied ESEM with the overall sample, with correlated uniqueness and discarding Item 20. The model fit indices support the modeling of the SATAQ-3 internal structure using this technique (M4). Even though the TLI is slightly below the proposed cut-off value (.94 versus .95), the values obtained in the remaining indices lead us to consider the model's fit as satisfactory.

Table 1
Goodness of fit indices for the different models.

Models		χ^{2a}	*df*	RMSEA	SRMR	TLI	CFI	ΔRMSEA	ΔTLI	ΔCFI
EFA										
M1	1st half of the sample	1535.959	321	.071	.037	.858	.895			
M2	1st half of the sample – 5F	897.777	295	.052	.023	.924	.948			
CFA										
M3	2nd half of the sample – CU – NO20	1043.097	316	.055	.048	.917	.931			
ESEM										
M4	Overall sample – CU – NO20	1141.524	268	.047	.020	.940	.960			
By sex – CU – NO20										
M5	Girls	742.255	268	.049	.022	.935	.957			
M6	Boys	750.671	268	.048	.026	.924	.950			
M7	Equal form	1492.926	536	.049	.024	.930	.954			
M8	Equal form and equal factor loadings	1731.088	636	.048	.032	.932	.947	−.001	.002	−.007
M9	Equal form, equal factor loadings and equal intercepts	1794.560	661	.048	.034	.933	.945	0	.001	−.002
M10	Equal form, equal factor loadings, equal intercepts and equal covariances	2102.461	671	.053	.088	.916	.931	.005	−.017	−.014
By grade – CU – NO20										
M11	Seventh	502.033	268	.048	.029	.925	.951			
M12	Eighth	598.201	268	.057	.031	.906	.938			
M13	Ninth	556.471	268	.053	.026	.928	.952			
M14	Tenth	594.624	268	.058	.027	.920	.947			
M15	Equal form	2251.330	1072	.054	.028	.920	.947			
M16	Equal form and equal factor loadings	2704.686	1372	.051	.042	.929	.940	−.003	.009	−.007
M17	Equal form, equal factor loadings and equal intercepts	2814.370	1447	.050	.043	.931	.939	−.001	.001	−.001
M18	Equal form, equal factor loadings, equal intercepts and equal covariances	2932.191	1477	.051	.061	.928	.935	.001	−.003	−.004

Notes: EFA = exploratory factor analysis; CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; 5F = model with 5 factors extracted (all the other models were with 4 factors); CU = correlated uniqueness; NO20 = deletion of Item 20; χ^2 = chi-square test; *df* = degrees of freedom; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; TLI = Tucker–Lewis index; CFI = comparative fit index; Δ = increment in fit index with respect to previous model.

[a] χ^2 values of EFA, CFA, and ESEM cannot be directly compared, as the analysis are based on different samples sizes.

The factor loadings for the M4 are shown in Table 2. All the items loaded significantly and principally in the factor with which they were associated. The average loading of the direct items in their principal factors was clearly greater than that of the reverse items (.72 versus .37). Of the 87 secondary loadings (29 items × 3 secondary factors), 15 were statistically significant, with an average, taking their absolute values, of .15 (maximum = .30; minimum = .07). The different number of secondary loadings with statistical significance in EFA and ESEM can be explained mainly by the different sample size, and hence, the different statistical power. As we found with EFA, the highest number of cross-loadings is found between the factors *Pressures* and *Internalization-General*, in this case with five items. These results show that the CFA requirement of fixing the secondary loadings to 0 is excessively restrictive for the case of the SATAQ-3.

The reliability estimates were .840, .830, .859 and .775 for the scales *Internalization-General*, *Information*, *Pressures* and *Internalization-Athlete*, respectively. All these values can be considered as satisfactory.

Inter-factor correlations for this model can be seen in Table 3. As in the original study (Thompson et al., 2004), the highest correlation (.64) was between *Pressures* and *Internalization-General* factors, indicating a moderate degree of overlap.

Having determined that the model which best permits the description of covariances between items in the SATAQ-3 is ESEM, we proceeded to analyze the measurement invariance by both sex and grade.

Measurement Invariance

As previously described, the usual approach involves following a series of steps. First, we checked the model for each group separately. In our case, and as can be seen in Table 1, the fit was satisfactory for both males and females (M5 and M6), and for the sample divided into the different grades (M11–M14). The lower degree of fit for the sample from grade 8 should be noted, but even so we consider that this step can be deemed satisfactorily completed. Secondly, we tested equality of form. The fit was satisfactory for this step, both in the case of the sample segmented by sex (M7) and in that segmented by grade (M15). Thirdly, we tested the equality of factor loadings. Given the minimal change in model fit, we can maintain the invariance of factor loadings both by sex (M8) and by grade (M16). Fourthly, we analyzed the intercept invariance. We added to the invariance in factor loadings the invariance in the intercepts (M9 and M17) and found no relevant change in the fit indices.

After this, we analyzed whether the inter-factor variance–covariance matrix was equal across groups. As can be seen in Table 1, we cannot sustain structure invariance for the SATAQ-3 by sex, given that the TLI and CFI fall by more than .01 (M10). However, this kind of invariance holds for the different grades (M18). Thus, the correlations between factors for the overall sample presented in Table 3 will have to be qualified taking sex into account.

Comparison by Sex and Grade

Invariance of loadings and intercepts permits identical interpretation for the same factor score in individuals from different groups, so that it is possible to compare means between groups. These results are those shown in Table 4.

The metric of the factor scores is determined by fixing the mean and standard deviation of the reference group (females or seventh grade in these cases) to 0 and 1. The differences in means can be approximately interpreted as effect sizes according to Cohen's (1988) criteria. The pattern of results with regard to sex shows what would be expected theoretically. For all the SATAQ-3 dimensions except *Internalization-Athlete* males present lower mean values than females. On analyzing the differences in scores by grade, we see that both *Internalization-General* and *Information* increase with increasing grade. For the *Pressures* factor the pattern is somewhat more confused, since the mean for grade 3 lies between grades 2 and 4, but at the end of compulsory secondary education the mean is higher than at the beginning of it. Only the *Internalization-Athlete* dimension seems not to vary across the different grades.

Table 2
Factor loadings for EFA (M2) and ESEM (M4).

		EFA Factor loadings					ESEM Factor loadings			
		I-GEN	INFO	PRES	I-ATL	NEG	I-GEN	INFO	PRES	I-ATL
01	TV programs are an important source of information about fashion and "being attractive."	.004	**.671**	−.047	.041	.083	.004	**.696**	−.036	.021
02	I've felt pressure from TV or magazines to lose weight	.028	−.024	**.791**	−.101	−.049	.006	−.019	**.758**	−**.092**
03*	I wouldn't like my body to look like the people who are on TV	**.275**	−.077	.131	−.005	**.499**	**.402**	−.030	.124	−.036
04	I compare my body to the bodies of TV and movie stars	**.604**	.014	**.257**	−.067	−.013	**.582**	.030	**.210**	−.050
05	TV commercials are an important source of information about fashion and "being attractive."	−.014	**.752**	.006	.026	.024	−.032	**.774**	−.033	.009
06*	I haven't felt pressure from TV or magazines to look pretty	−.054	.028	**.310**	.048	**.520**	.069	.003	**.353**	−.012
07	I would like my body to look like the models who appear in magazines	**.805**	−.031	−.107	.051	.032	**.819**	−.019	−**.123**	.071
08	I compare my appearance to the appearance of TV and movie stars	**.657**	.010	.178	.012	−.005	**.656**	.019	**.164**	−.008
09*	Music videos on TV aren't an important source of information about fashion and "being attractive"	−.118	**.269**	−.012	.002	**.492**	.010	**.286**	.022	−.035
10	I've felt pressure from TV and magazines to be thin	.086	−.021	**.730**	−.055	−.004	.036	−.025	**.773**	−.033
11	I would like my body to look like the people who are in the movies	**.738**	.021	.028	.053	.010	**.767**	−.004	.024	**.087**
12*	I don't compare my body to the bodies of people who appear in magazines	**.222**	.023	.098	−.022	**.554**	**.330**	.003	**.175**	−.069
13*	Magazine articles aren't an important source of information about fashion and "being attractive"	−.011	**.431**	−.024	−.053	**.423**	.051	**.457**	.002	−.068
14	I've felt pressure from TV or magazines to have a perfect body	**.177**	.013	**.661**	.030	.038	**.152**	.020	**.670**	.049
15	I wish I looked like the models in music videos	**.718**	.074	−.039	.049	.000	**.730**	.055	−.009	**.072**
16	I compare my appearance to the appearance of people in magazines	**.617**	.051	**.260**	−.034	.066	**.584**	.045	**.303**	−.034
17	Magazine advertisements are an important source of information about fashion and "being attractive"	−.003	**.832**	.037	.002	−.014	−.049	**.828**	.021	.019
18	I've felt pressure from TV or magazines to diet	−.055	−.025	**.756**	.079	.030	−.063	.003	**.786**	.050
19*	I don't wish I looked as athletic as the people in magazines	.113	−.071	−.078	**.353**	**.443**	**.219**	−.034	−.024	**.341**
20	I compare my body to that of people in "good shape"	**.286**	**.136**	**.213**	**.290**	−.071	NA	NA	NA	NA
21	Pictures in magazines are an important source of information about fashion and "being attractive"	.072	**.763**	.055	−.005	−.027	.046	.777	.021	−.014
22	I've felt pressure from TV or magazines to exercise	−.002	.075	**.603**	**.197**	.032	−.005	.032	**.621**	**.216**
23	I wish I looked as athletic as sports stars	**.005**	.001	−.082	**.781**	−.015	.028	.008	−.085	**.809**
24	I compare my body to that of people who are athletic	.084	.060	**.143**	**.623**	−.025	.113	.031	**.154**	**.595**
25	Movies are an important source of information about fashion and "being attractive"	.059	**.722**	.053	.035	−.016	−.002	**.727**	.081	.040
26	I've felt pressure from TV or magazines to change my appearance	.051	.078	**.655**	.041	.032	.063	.048	**.691**	**.082**
27*	I don't try to look like the people on TV	**.219**	.040	.065	.017	**.524**	**.346**	.022	.136	.008
28*	Movies aren't an important source of information about fashion and "being attractive"	.000	**.468**	−.062	−.050	**.410**	.103	**.453**	−.025	−.042
29	Famous people are an important source of information about fashion and "being attractive"	.141	**.618**	.050	−.004	−.002	**.123**	**.640**	.026	−.006
30	I try to look like sports athletes	−.080	−.037	.040	**.836**	.028	−.066	−.019	.055	**.806**

Notes: I-GEN: *Internalization-General*; INFO: *Information*; PRES: *Pressures*; I-ATL: *Internalization-Athlete*. Items with asterisk correspond to reverse-keyed items. Shaded cells indicate the factor where the item theoretically belongs. Bold loadings indicate statistically significant loadings, applying correction for multiple comparisons.

Table 3
Inter-factor correlations for the overall sample (M4) and by sex (M10).

	I-GEN	INFO	PRES	I-ATL
Overall				
I-GEN				
INFO	.547			
PRES	.637	.485		
I-ATL	.304	.159	.190	
Girls				
I-GEN				
INFO	.643			
PRES	.592	.525		
I-ATL	.248	.252	.136	
Boys				
I-GEN				
INFO	.532			
PRES	.439	.384		
I-ATL	.692	.447	.338	

Notes: I-GEN: *Internalization-General*; INFO: *Information*; PRES: *Pressures*; I-ATL: *Internalization-Athlete*.

Table 4
Factor means (and standard deviations) by sex (M9) and by grade (M18).

	By sex[a]		By grade[b]			
	Girls	Boys	Grade 7	Grade 8	Grade 9	Grade 10
INTER-G	0 (1)	−0.552 (0.617)	0 (1)	0.284 (1)	0.343 (1)	0.437 (1)
INFO	0 (1)	−0.623 (0.320)	0 (1)	0.291 (1)	0.334 (1)	0.563 (1)
PRESS	0 (1)	−0.445 (0.965)	0 (1)	0.283 (1)	0.154 (1)	0.277 (1)
INTER-A	0 (1)	0.766 (1.184)	0 (1)	0.066 (1)	−0.058 (1)	0.051 (1)

Notes: I-GEN: *Internalization-General*; INFO: *Information*; PRES: *Pressures*; I-ATL: *Internalization-Athlete*.
The metric of the model is determined by fixing the mean and standard deviation of the reference group (girls or grade 7) to 0 and 1.
[a] Descriptives for M9.
[b] Descriptives for M18.

Table 3 shows the matrix of correlations between factors by sex. These results qualify the correlations between factors presented previously. It is observed how, for the boys, *Internalization-Athlete* is much more of a core component within the SATAQ-3 scale than it is for the girls, since for males this scale is much more closely correlated with the others. Thus, for example, the correlation between *Internalization-General* and *Internalization-Athlete* is .25 for females and .69 for males.

Discussion

The present study had two basic objectives. The first was to study the factor structure of the Spanish version of the SATAQ-3 with adolescent boys and girls maintaining the original text of the items for both sexes, with the reverse-keyed items suggested by the authors of the original version. Our data replicate the original four-factor structure and support the appropriateness of our SATAQ-3 version (see Appendix in supplementary content, available online) for use with Spanish adolescent population. This is the first version of this questionnaire that shows satisfactory fit for adolescents by means of an ESEM model. To date, only the French versions (Rousseau et al., 2010; Rousseau & Valls, 2010) had been applied to adolescent population, but using principal components analysis and without reverse-keyed items.

Our results show an anomalous functioning of the negative items. Thus, in the EFA models the negative items present relevant cross-loadings between the substantive factor to which each item corresponds and a fifth, method factor, the loading in the method factor being greater than in the substantive factor. Moreover, in the ESEM model, the factor loadings of the reverse-keyed items are always lower than those of the direct items. Therefore, the correlation between negative items could be better explained by an irrelevant aspect (wording) than by their theoretical content. Problems with the use of negative items have already been mentioned in other studies (e.g., Barnette, 2000; Schriesheim, Eisenbach, & Hill, 1991). We therefore recommend the use of a version without reverse-keyed items.

The factors *Internalization-General* and *Pressures* are those with the most cross-loadings in the different models, in line with the findings of other studies in which the two factors have been grouped in a single factor (Jackson & Chen, 2010; Swami, 2009). These results suggest the difficulty of developing items assessing purely a single dimension, probably due to the close conceptual link between the two factors. Subsequent studies should address the conceptual links and/or differences between these two factors.

The factor *Internalization-Athlete* presents the lowest correlations with the rest of the factors in the different models used, as already observed in the original study (Thompson et al., 2004) and in the CFA carried out by Markland and Oliver (2008), even though this result is modulated by the sex variable.

Item 20, designed to assess the *Internalization-Athlete* dimension, is the only item that presents significant cross-loadings in all the substantive dimensions in the EFA. This item already showed similar problems in previous work (Calogero et al., 2004; Markland & Oliver, 2008; Rousseau et al., 2010), leading other validation studies to remove it from the outset (e.g., Jackson & Chen, 2010). The explanation may lie in the semantic differences between this item and the rest of the *Internalization-Athlete* factor items (Markland & Oliver, 2008). Whereas the other *Internalization-Athlete* items refer to sporty or athletic individuals as a source of comparison, Item 20 refers to comparisons with others who are "in good shape". For all of these reasons, we recommend the removal of Item 20 from the SATAQ-3.

The data show how the limitations of EFA (incapacity to establish correlations between uniqueness; impossibility of assessing invariance) and CFA (excessively restrictive model on defining all the secondary loadings as equal to 0) are overcome by ESEM on analyzing the internal structure of the SATAQ-3. With EFA we obtained a factor solution that was difficult to interpret, due to the appearance of a method factor that grouped the negative items. With CFA model fit indices did not meet the specified thresholds. ESEM reduces the restrictions to impose and provides the best fit of all the models analyzed.

The second objective of the present work was to study the invariance of the instrument. The results support the invariance both by sex and by grade, providing substantial endorsement for the possibility of using the same version of the questionnaire (without Item 20) with both boys and girls and throughout a large portion of adolescence, without adaptations or changes from previous versions, as has been the case in other studies (e.g., Karazsia & Crowther, 2008). Once the invariance of scores by sex and grade had been guaranteed, we proceeded to compare the different scores of the groups.

As regards sex, the pattern of results reflects the theoretical expectations, so that for all the SATAQ-3 dimensions except *Internalization-Athlete*, the boys present lower mean values than the girls. These results are in accordance with those of previous studies. In this regard, the only meta-analysis published to date for determining the extent to which pressure from the mass media to conform to the muscular "ideal" male body affects men's self-images (Barlett, Vowels, & Saucier, 2008) found effect sizes smaller than those found by other meta-analyses in relation to the effects of media exposure and media influence on body image and disordered eating in females (Grabe et al., 2008; Groesz, Levine, & Murnen, 2002). In the case of *Internalization-Athlete* it is boys who score higher. This finding, together with the pattern of higher correlations of this factor with the remaining factors in the case of boys, may indicate that this subscale would be much more of a core component within the SATAQ-3 in boys than it is for girls. Although the items making up this subscale were added in view of the recent focus on athleticism and sports in young women (Cusumano & Thompson, 1997), the data indicate that in the case of males the subscale might also assess the internalization of the muscular ideal, which is encouraged more among boys (Blond, 2008; Calogero & Thompson, 2010; Warren, 2008).

As far as grade is concerned, except in the case of *Internalization-Athlete*, the scores in the factors increase with age, and especially so for the *Information* and *Internalization-General* factors. Prospective studies suggest that eating pathology is most likely to emerge between the ages of 15 and 19 in adolescent girls (Lewinsohn, Striegel-Moore, & Seeley, 2000; Stice, Killen, Hayward, & Taylor, 1998). In the field of eating disorders prevention, selected prevention programmes produced significantly larger decreases in thin-ideal internalization than universal programmes, and significantly larger effects were observed for trials focusing on participants over age 15 than for trials focusing on younger participants (Stice et al., 2007; Stice & Shaw, 2004). Interventions may be more effective for the former because they were delivered during the period of greatest risk for emergence of eating disturbances, because younger adolescents may have limited insight, or because of a floor effect caused by the low levels of eating pathology during early adolescence (Stice et al., 2007; Stice & Shaw, 2004). In line with these findings is the observation of the general increase with age that occurs in scores for *Internalization-General* and *Information*.

These results, with regard to both sex and grade, which are in accordance with what was expected, lend support to the validity of our version of the SATAQ-3. Likewise, the invariance found suggests that the construction of different versions of a questionnaire for its applications to different populations should be confined to those cases in which there is evidence of such a need. The fact of having administered different tests to boys and girls could have led to their results being impossible to compare.

This study has several limitations, while also pointing the way to future lines of research. On the one hand, we did not find satisfactory explanations for the relative stability of scores on the *Internalization-Athlete* scale across the different grades. This result should be checked in future studies, exploring the way scores on the different SATAQ-3 subscales develop with increasing age. On the other hand, our findings clearly suggest the pertinence of using a version without reverse-keyed items, but we cannot guarantee that the results obtained with the tested version (fit of the model and invariance) are generalizable for a version with all direct items. Therefore, it would be necessary to replicate the study using a version without reverse-keyed items.

We should point out several limitations of ESEM. First, in terms of usability, to date ESEM is implemented exclusively in M*plus*. Second, in terms of historical background, while research on EFA and CFA (and with EFA and CFA as statistical tools) dates back over decades, ESEM is a much newer approach, and this has certain consequences. For instance, we have used as cut-points for evaluating model fit and measurement invariance those developed in the CFA context, as no specific guidelines have been proposed for ESEM. Third, ESEM also has its own statistical limitations, some of which could be relevant for the analysis of SATAQ-3. In the CFA context the excess of covariance between negative items can be modeled with correlated uniqueness or with a method factor. In ESEM this second option is not available. In EFA and CFA, a common recommendation when correlated factors are obtained is to perform a second-order factor analysis, but with ESEM this is not possible.

To summarize, the results of the present study suggest that the SATAQ-3 can be used with Spanish adolescent populations. The invariance demonstrated, estimated for the first time with the SATAQ-3, will also permit the instrument as originally constructed to be administered to both sexes and throughout a large portion of the adolescent period.

Funding source

Preparation of this article was supported by the Research Grants from Ministry of Science & Innovation (PSI2009-08956) of the Spanish Government.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bodyim.2011.10.002.

References

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.

Barlett, C. P., Vowels, C. L. & Saucier, D. A. (2008). Meta-analyses of the effects of media images on men's body-image concerns. *Journal of Social & Clinical Psychology*, *27*, 279–310.

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*, 361–370.

Blond, A. (2008). Impacts of exposure to images of ideal bodies on male body dissatisfaction: A review. *Body Image*, *5*, 244–250.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Cafri, G., Yamamiya, Y., Brannick, M. & Thompson, J. K. (2005). The influence of sociocultural factors on body image: A meta-analysis. *Clinical Psychology: Science and Practice*, *12*, 421–433.

Calogero, R. M., Davis, W. N. & Thompson, J. K. (2004). The Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ-3): Reliability and normative comparisons of eating disordered patients. *Body Image*, *1*, 193–198.

Calogero, R. M. & Thompson, J. K. (2010). Gender and body image. In J. C. Chrisler & D. R. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 153–184). New York: Springer.

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cusumano, D. & Thompson, J. (1997). Body image and body shape ideals in magazines: Exposure, awareness, and internalization. *Sex Roles*, *37*, 701–721.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*, 121–149.

Durkin, S. J, Paxton, S. J. & Sorbello, M. (2007). An integrative model of the impact of exposure to idealized female images on adolescent girls' body satisfaction. *Journal of Applied Social Psychology*, *37*, 1092–1117.

Enders, C. K. & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*, 430–457.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299.

Fonseca-Pedrero, E., Sierra-Baigrie, S., Paino, M., Lemos-Giráldez, S. & Muñiz, J. (2011). Factorial structure and measurement invariance of the Bulimic Investigatory Test Edinburgh across gender and age. *International Journal of Clinical and Health Psychology*, *11*, 109–123.

Grabe, S., Ward, L. M. & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin*, *134*, 460–476.

Groesz, L. M., Levine, M. P. & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders*, *31*, 1–16.

Hambleton, R. K., Merenda, P. F. & Spielberger, C. D. (Eds.). (2005). *Adapting educational and psychological tests for cross-cultural assessment*. London: Erlbaum.

Heinberg, L. J., Coughlin, J. W., Pinto, A. M., Haug, N. A., Brode, C. & Guarda, A. S. (2008). Validation and predictive utility of the Sociocultural Attitudes Towards Appearance Questionnaire for Eating Disorders (SATAQ-ED): Internalization of sociocultural ideals predicts weight gain. *Body Image*, *5*, 279–290.

Heinberg, L., Thompson, J. K. & Stormer, S. (1995). Development and validation of the Sociocultural Attitudes Towards Appearance Questionnaire. *International Journal of Eating Disorders*, *17*, 81–89.

Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

International Test Commission. (2010). *International Test Commission guidelines for translating and adapting tests*. http://www.intestcom.org/upload/sitefiles/40.pdf

Jackson, T. & Chen, H. (2010). Factor structure of the Sociocultural Attitudes Towards Appearance Questionnaire-3 (SATAQ-3) among adolescent boys in China. *Body Image*, *7*, 349–355.

Karazsia, B. T. & Crowther, J. H. (2008). Psychological and behavioral correlates of the SATAQ-3 with males. *Body Image*, *5*, 109–115.

Knauss, C., Paxton, S. J. & Alsaker, F. D. (2009). Validation of the German version of the Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ-G). *Body Image*, *6*, 113–120.

Levine, M. P. & Harrison, K. (2004). Media's role in the perpetuation and prevention of negative body image and disordered eating. In J. K. Thompson (Ed.), *Handbook of eating disorders and obesity* (pp. 695–717). New York: Wiley.

Levine, M. P. & Murnen, S. K. (2009). "Everybody knows that mass media are/are not [pick one] a cause of eating disorders": A critical review of evidence for a causal link between media, negative body image, and disordered eating in females. *Journal of Social & Clinical Psychology*, *28*, 9–42.

Levine, M. P. & Smolak, L. (2006). *The prevention of eating problems and eating disorders: Theory, research and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lewinsohn, P. M., Striegel-Moore, R. H. & Seeley, J. R. (2000). Epidemiology and natural course of eating disorders in young women from adolescence to young adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 1284–1292.

López-Guimerà, G., Levine, M. P., Sánchez-Carracedo, D. & Fauquet, J. (2010). Influence of mass media on body image and eating disordered attitudes and behaviors in females: A review of effects and processes. *Media Psychology*, *13*, 387–416.

López-Guimerà, G. & Sánchez-Carracedo, D. (2010). *Prevención de las alteraciones alimentarias: Fundamentos teóricos y recursos prácticos[nl]Disordered eating prevention: Theoretical basis and practical resources*. Madrid: Pirámide.

López-Guimerà, G., Sánchez-Carracedo, D., Fauquet, J., Portell, M. & Raich, R. M. (2011). Impact of a school-based disordered eating prevention program in adolescent girls: General and specific effects depending on adherence to the interactive activities. *Spanish Journal of Psychology*, *14*, 293–303.

MacCallum, R. C., Roznowski, M. & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*, 490–504.

Madanat, H. N., Hawks, S. R. & Brown, R. B. (2006). Validation of the Sociocultural Attitudes Towards Appearance Questionnaire-3 among a random sample of Jordanian women. *Body Image*, *3*, 421–425.

Markland, D. & Oliver, E. J. (2008). The Sociocultural Attitudes Towards Appearance Questionnaire-3: A confirmatory factor analysis. *Body Image*, *5*, 116–121.

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport and exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., Vol. 2, pp. 774–798). Hoboken, NJ, USA: John Wiley & Sons Inc.

Marsh, H. W., Ludtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., et al. (2010). A new look at the Big Five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471–491.

Marsh, H. W., Muthén, B., Asparouhov, T., Ludtke, O., Robitzsch, A., Morin, A. J. S., et al. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*, 439–476.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.

Muthén, L. K. & Muthén, B. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Raykov, T. (2009). Evaluation of scale reliability for unidimensional measures using latent variable modeling. *Measurement and Evaluation in Counseling and Development*, *42*, 223–232.

Raykov, T. & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. Taylor & Francis: New York.

Rosellini, A. J. & Brown, T. A. (2011). The NEO five-factor inventory: Latent structure and relationships with dimensions of anxiety and depressive disorders in a large clinical sample. *Assessment*, *18*, 27–38.

Rousseau, A. & Valls, M. (2010). Étude de validation de la version française du Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ-3) chez les adolescents garçons [Validation of the Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ-3) in a sample of French adolescent boys]. *Journal de Thérapie Comportementale et Cognitive*, *20*, 56–60.

Rousseau, A., Valls, M. & Chabrol, H. (2010). The Sociocultural Attitudes Towards Appearance Questionnaire (SATAQ-3): Étude de validation de la version française [Validation of the French version of the Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3)]. *L'Écéphale*, *36*, 270–276.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis. *Personality and Social Psychology Bulletin*, *28*, 1629–1646.

Rusticus, S. A., Hubley, A. M. & Zumbo, B. D. (2008). Measurement invariance of the Appearance Schemas Inventory—Revised and the Body Image Quality of Life Inventory across age and gender. *Assessment*, *15*, 60–71.

Schriesheim, C. A., Eisenbach, R. J. & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement*, *51*, 67–78.

Slof-Op't Landt, M. C., van Furth, E. F., Rebollo-Mesa, I., Bartels, M., van Beijsterveldt, C. E., Slagboom, P. E., et al. (2009). Sex differences in sum scores may be hard to interpret: The importance of measurement invariance. *Assessment*, *16*, 415–423.

Smolak, L., Levine, M. P. & Thompson, J. K. (2001). The use of the Sociocultural Attitudes Towardss Appearance Questionnaire with middle school boys and girls. *International Journal of Eating Disorders*, *29*, 216–223.

Stice, E., Killen, J. D., Hayward, C. & Taylor, C. B. (1998). Age of onset for binge eating and purging during adolescence: A 4-year survival analysis. *Journal of Abnormal Psychology*, *107*, 671–675.

Stice, E. & Shaw, H. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin*, *130*, 206–227.

Stice, E., Shaw, H. & Marti, C. N. (2007). A meta-analytic review of eating disorder prevention programs: Encouraging findings. *Annual Review of Clinical Psychology*, *3*, 207–231.

Swami, V. (2009). An examination of the factor structure of the Sociocultural Attitudes Towards Appearance Questionnaire-3 in Malaysia. *Body Image*, *6*, 129–132.

Thompson, J. K. & Stice, E. (2001). Internalization of the thin-ideal: Mounting evidence for a new risk factor for body image disturbance and eating pathology. *Current Directions in Psychological Science*, *10*, 181–183.

Thompson, J. K., van den Berg, P., Roehrig, M., Guarda, A. S. & Heinberg, L. J. (2004). The Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3): Development and validation. *International Journal of Eating Disorders*, *35*, 293–304.

Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4–70.

Warren, C. S. (2008). The influence of awareness and internalization of western appearance ideals on body dissatisfaction in Euro-American and Hispanic males. *Psychology of Men & Masculinity*, *9*, 257–266.

Warren, C. S., Cepeda-Benito, A., Gleaves, D. H., Moreno, S., Rodríguez, S., Fernández, M. C., et al. (2008). English and Spanish versions of the Body Shape Questionnaire: Measurement equivalence across ethnicity and clinical status. *International Journal of Eating Disorders*, *41*, 265–272.

West, S. G., Finch, J. F. & Curran, P. J. (1995). Structural equation models with nonnormal variables. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56–75). Thousand Oaks, CA: Sage.