ORIGINAL ARTICLE

Factorial invariance, differential item functioning, and norms of the Orgasm Rating Scale

Ana Isabel Arcos-Romero, Juan Carlos Sierra*

Centro de Investigación Mente, Cerebro y Comportamiento (CIMCYC), Universidad de Granada, Spain

Abstract *Background/Objective*: Orgasm Rating Scale (ORS) assess the subjective orgasm experience in context of sexual relationship. It is composed of four dimensions attributed to the orgasm (Affective, Sensory, Intimacy, and Rewards). The purpose is to analyse the factorial invariance of the ORS across groups, to examine the metric equivalence across sex, and to present the standard scores. *Method*: A total of 1,472 Spanish adults (715 men and 757 women) were evaluated. They were distributed across age groups (18-34, 35-49 and 50 years old and older). Factorial invariance across different groups and the differential functioning of the items across sex were analyzed, internal consistency was examined, and the standard scores were developed. *Results*: The structure of the ORS showed strict measurement invariance across sex, relationship status, sexual orientation and education level. It also reached a scalar measurement invariance across age range and duration of the relationship. Some items showed a differential functioning between sexes. *Conclusions*: The Spanish version of the ORS is invariant across different groups at a factorial level, and it shows equivalence across sex in most of its items at a metric level. The standard scores allow a more accurate assessment of the subjective orgasm experience in context of sexual relationship.

Invarianza factorial, funcionamiento diferencial de los ítems y baremos de la Orgasm Rating Scale

Resumen *Antecedentes/Objetivo*: La *Orgasm Rating Scale* (ORS) evalúa la experiencia subjetiva del orgasmo en el contexto de las relaciones sexuales en pareja. Está compuesta por cuatro dimensiones atribuidas al orgasmo (Afectiva, Sensorial, Intimidad y Recompensa). Se analiza su invarianza factorial por grupos, su equivalencia métrica por sexo y se presentan baremos en

---

* Corresponding author: Facultad de Psicología, Campus Universitario de Cartuja, s/n., 18011, Granada, Spain.
  E-mail address: jcsierra@ugr.es (J.C. Sierra).

población española. *Método*: Se evaluó a 1.472 adultos españoles (715 hombres y 757 mujeres), distribuidos en cuotas de edad (18-34, 35-49 y mayores de 50 años). Se analizó la invarianza factorial por diferentes grupos y el funcionamiento diferencial de los ítems por sexo, se examinó la consistencia interna de los factores y se establecieron sus puntuaciones baremadas. *Resultados*: La estructura de la ORS mostró un nivel de invarianza estricta por sexo, relación de pareja, orientación sexual y nivel de estudios, e invarianza fuerte por rango de edad y duración de la relación. Algunos de sus ítems mostraron indicios de funcionamiento diferencial entre sexos. *Conclusiones*: La versión española de la ORS es invariante por grupos a nivel factorial y presenta equivalencia por sexo en la mayoría de sus ítems a nivel métrico. Los baremos permiten una evaluación más precisa de la experiencia subjetiva orgásmica en el contexto de las relaciones sexuales.

Orgasm is a sensation of intense pleasure creating an altered state of consciousness accompanied by changes in the pelvic musculature and resolves the sexually-induced vasocongestion, generally with an induction of well-being and contentment (Meston, Hull, Levin, & Sipski, 2004). Subjective orgasm experience refers to the psychological perception, sensation and evaluation of orgasm (Arcos-Romero & Sierra, 2018). The Orgasm Rating Scale (ORS; Mah & Binik, 2002, 2011) is one of the few instruments that assesses the subjective orgasm experience. The Spanish version of the ORS (Arcos-Romero, Moyano, & Sierra, 2018) is comprised of 25 adjectives distributed into four dimension: Affective, which contains items that refer to the emotions experienced during orgasm (e.g., satisfying, pleasurable); Sensory, which includes items relating to the perception of physiological events of orgasm (e.g., uncontrolled, flushing); Intimacy, which contains items that reflect the intimacy component of the orgasm experience (e.g., loving, tender); and Rewards, which includes items that refer to the consequences or effects of orgasm (e.g., peaceful, relaxing). Authors informed that this version has adequate psychometric properties, it is a reliable and valid scale.

Orgasmic dysfunction refers to difficulties in sexual climax response, characterized by physiological and subjective reactions (Nobre, 2006). Its prevalence varies between 10-42% in women and between 20-30% in men (American Psychiatric Association, 2013). It has been shown that people with orgasmic difficulties report a lower level of intensity in the Affective, Sensory, and Intimacy dimension of the orgasm experience (Arcos-Romero, Moyano et al., 2018). The ORS has the ability to discriminate between those individuals with and without orgasmic difficulties (Arcos-Romero, Moyano et al., 2018). Therefore, from a clinical perspective, the ORS allows to identify which dimension or dimensions of the psychological experience of orgasm are affected in people with dysfunctions, and could be used as a guide in the treatment. Based on these findings, Arcos-Romero, Granados, and Sierra (2018) proposed a Model of the Subjective Orgasm Experience in context of sexual relationship, based on the model of Mah and Binik (2002). It includes the four components of the ORS (Affective, Sensory, Intimacy, and Rewards), and provides a conceptual framework to explain the orgasm on a psychological level in context of sex-with-partner.

In order to compare a specific construct between groups it is fundamental to use equivalent instruments for avoid reporting bias in the assessment (Muñiz, Elosua, & Hambleton, 2013; Sánchez, Muñoz-Fernández, & Vega-Gea, 2017) and to obtain reliable data (Marchal-Bertrand et al., 2016). Factorial invariance analyzes the degree to which an instrument measures the same construct in different groups (Pineda, Martín-Vivar, Sandín, & Piqueras, 2018) and, as such, can make precise comparisons between them (e.g., between men and women). Therefore, an invariant measurement of the subjective orgasm experience across groups with different characteristics allows a more precise and exact assessment of this sexual response dimension. In addition, it is important that the instruments used in sexual health also use standard scores that allow professional to interpret their scores. Thus, the objective of this study is to analyze the factorial equivalence (through calculating the factorial invariance) of the ORS by different populations groups, and the metric equivalence (examining the differential item functioning) of the scale between men and women. Furthermore, the standard scores for the Spanish population will be presented.

Method

Participants

Through a non-probability quota sampling, 1,472 Spanish adults (715 men and 757 women) between 18 and 93 years of age ($M = 40.26$; $SD = 14.07$) were evaluated. The sample was randomly and uniformly distributed into three age range: 18-34 ($n = 500$; 250 men, 250 women), 35-49 ($n = 500$; 250 men, 250 women), and older than 50 years old and older ($n = 472$; 215 men, 257 women). The socio-demographic characteristics of the participants are presented divided by sex in Table 1. In regards to all of the evaluated socio-demographic variables, we only found significant differences between

Table 1 Socio-demographic characteristics of the sample.

Variables		Total (N = 1,472) M (SD)	Men (n = 715)	Women (n = 757)	t/χ^2	Cohen's d
Age		40.26 (14.07)	40.13 (14.10)	40.38 (14.04)	-0.34	
Age of the first sexual experience		17.98 (5.16)	18.07 (6.59)	17.90 (3.27)	0.64	
Number of sexual partners		10.70 (27.36)	13.43 (35.88)	8.09 (14.79)	3.32[**]	0.17
			n (%)			
Sexual activity	Yes	1,344 (91.3%)	658 (92%)	686 (90.6%)	0.92	
	No	128 (8.7%)	57 (8%)	71 (9.4%)		
Relationship (n, %)	Yes	1.156 (78.5%)	565 (79%)	591 (78.1%)	0.19	
	No	316 (21.5%)	150 (21%)	166 (21.9%)		
Living with the partner	Yes	802 (69.4%)	394 (69.7%)	408 (69%)	0,10	
	No	354 (30.6%)	171 (30.3%)	183 (31%)		
Relationship length	1-10 years	493 (46.6%)	249 (34.8%)	244 (32.2%)	2.48	
	11-20 years	211 (19.9%)	113 (15.8%)	98 (12.9%)		
	Over 20 years	354 (33.5%)	166 (23.2%)	188 (24.8%)		
Sexual orientation	Heterosexual	1,259 (85.5%)	613 (85.7%)	646 (85.3%)	0.05	
	Non-Heterosexual	213 (14.5%)	102 (14.3%)	111 (14.7%)		
Education level	University studies	1,046 (71.1%)	500 (69.9%)	546 (72.1%)	0.86	
	Non-University	426 (28.9%)	215 (30.1%)	211 (27.9%)		

Note. M = mean, SD = standard deviation, t = Student's t; χ^2 = Chi-square; Cohen's d = effect size.
[**] p < .01.

men and women in the number of sexual partners (t = 3.32; p < .01; Cohen's d = .17).

Instruments

Background questionnaire. Participants were asked about their age, sex, nationality, age of their first sexual experience (oral, vaginal, or anal intercourse), number of sexual partners, current sexual activity, partner relationship, living with their partner, length of the relationship, sexual orientation, and education level.

The Spanish version of the Orgasm Rating Scale (ORS; Mah & Binik, 2011) by Arcos-Romero, Moyano et al. (2018). It assesses the subjective orgasm experience using 25 items distributed into four factors: Affective, Sensory, Intimacy, and Rewards. In order to quantify how well the item describes their most recent orgasm experience, it uses a 6-point Likert scale, where 0 signifies ''*does not describe it at all*'' and 5 signifies ''*describes it perfectly*''. This version of the ORS has shown good psychometric properties, with its internal consistency reliability ranged between .78 (Intimacy) and .93 (Sensory). Regarding to the validity evidence, scores from the ORS have been positively and significantly correlated with sexual satisfaction and erotophilia. In addition, evidence of its discriminant validity has been showed, the scale discriminated between individuals with and without orgasmic difficulties. In the present study, examining the total sample, Cronbach's alpha coefficients for each factor were .88 (Affective), .95 (Sensory), .82 (Intimacy), and .88 (Rewards). Appendix 1.

Procedure

Two procedures were used for data collection: traditional of pen and paper, and an online questionnaire. Both strategies are typically used in human sexuality studies (e.g., Sierra, Moyano, Vallejo-Medina, & Gómez-Berrocal, 2018). With the traditional method, participants were assessed at the University, in community centers, health centers, libraries, etc. They received a copy of the instruments together with an envelope which was used to return the completed questionnaire. The online form was sent using a link created on the *LimeSurvey* platform. The link was distributed via email and through several social networks. The relevance of the study was explained and Spanish adults were invited to take part of it. Participants were informed about the voluntary and anonymous participation, and about the confidentiality of the data for the sole purpose of the study. They accepted a consent form before starting the questionnaire. The time to complete it was approximately 15 minutes. The study was previously approved by the Ethics Committee on Human Research of the University of Granada.

Data analysis

First, the four-factor model based on the validated Spanish version of the Orgasm Rating Scale (Arcos-Romero, Moyano et al., 2018) was tested. We examined its factorial invariance (FI) using multi-group analysis with AMOS statistical software (Arbuckle, 2015). Mardia's test was used to calculate the multivariate normality of the data. The estimation

method used was the Maximum Likelihood (ML). The following indicators were taken into account: Root Mean Square Error of Approximation and its 90% confidence interval (RMSEA; Browne & Cudeck, 1993), and Comparative Fit Index (CFI; Bentler, 1990). The RMSEA is a measurement of the amount of error in the model, and evaluates how well the model fits to populations considering the observed values. Values below .05 indicate a good fit of the model, and those values below .08 indicate an adequate fit; if its 90% confidence interval includes values under .08 this would also indicate a good model fit (Browne & Cudeck, 1993). The comparative fit index (CFI) values greater than .90 indicate a good fit (Bentler, 1990). Following previous recommendations (Cheung & Rensvold, 2002; Emmerink, Van den Eijnden, Ter Bogt, & Vanwesenbeeck, 2017; Meade, Johnson, & Braddy, 2008; Putnick & Bornstein, 2016), the RMSEA fit indicators were appreciated, as well as its 90% confidence interval, but, above all, we considered changes in CFI values because it is more robust and recommended for larger sample sizes.

The different examined population groups were created based on sex, age range, relationship status and, if applicable, the length of a stable relationship (longer than one year), sexual orientation, and education level. Following Byrne (2008), FI was progressively analyzed at different levels: configural, weak, strong, and strict. In order to accept the equivalence of the factors and to compare nested models, a change in CFI equal or greater than .01 was considered to determine and adopt the less limited model and to reject the most restrictive model, being this the main fit index used to evaluate FI (Cheung & Rensvold, 2002; Milfont & Fischer, 2010; Putnick & Bornstein, 2016).

Secondly, the differential item functioning (DIF) was analyzed in order to confirm the equivalence across sex of the items of the Spanish version of the ORS. We used EASY-DIF software (González, Padilla, Hidalgo, Gómez-Benito, & Benítez, 2011). We tried to examine whether the participants' likelihood of choosing a determined value of an ORS item varied, or not, according to the group to which they belonged (i.e., reference group: men, or focal group: women). Given the fact that the items are ordinal, we observed the Mantel-Haenszel statistical test that is used as comparative criteria for the groups (Holland & Thayer, 1988; Socha, DeMars, Zilberberg, & Phan, 2015), as well as the p value and its standard deviation (SDP). The significance of p-value was interpreted in order to examine the differences between groups in the observed item functioning, taking into account the significant values ($p < .05$) as evidence that suggests the existence of DIF (González et al., 2011). In addition, the SDP indicator was interpreted in order to compare the expected mean scores of one group (reference group: men) to the other (focal group: women), in each item in relation to the total score of the scale (Gómez Benito, González, Widaman, Padilla, & Balluerka, 2017). In this case, a negative value means that mean score would be lower in women (focal group) in comparison to men (reference group), while a positive value would indicate the opposite. Furthermore, for those items that showed DIF, we compared the mean scores for two independent groups (men and women).

Then, before obtaining the standard scores, differences across sex in mean scores of each factor were analyzed, dividing the total sample into three groups based on age.

Furthermore, the internal consistency of the four factors contrasting men and women were analyzed, as well as contrasting age range. Finally, the norms of the ORS factors differentiated by sex and age groups were obtained. Histograms, skewness and kurtosis were observed, considering the normal distribution of the sub-samples of men and women. We used the direct scores of the scale.

Results

Factorial Invariance (FI)

Mardia's test (546,936) indicated the absence of multivariate normality of the data, so the Maximum Likelihood (ML) estimation robust method was used. The RMSEA and its 90% confidence interval showed that the four-factor model of the ORS predicted the data matrix in an equivalent manner across all the analyzed population groups. Furthermore, the CFI allowed the acceptance of factorial invariance across all the groups at various levels. Based on the CFI, we accepted, for each group, the model adjusted to a certain level appreciating the change in its value. In particular, as it is presented in Table 2, the factorial structure of the ORS showed strict invariance across sex [RMSEA = .058 (.056-.059); CFI = .921], partner relationship [RMSEA = .057 (.055-.059); CFI = .923], sexual orientation [RMSEA = .054 (.053-.056); CFI = .929], and education level [RMSEA = .056 (.054-.057); CFI = .926]. It also showed strong invariance across age range [RMSEA = .048 (.047-.050); CFI = .917], and relationship length [RMSEA = .048 (.046-.05); CFI = .919].

Differential item functioning (DIF) and differences by sex

Nine of the 25 items of the ORS showed DIF when comparing men and women (see Table 3). In the six items of the Sensory factor that suggested the existence of DIF, the SDP showed that all of its mean scores were higher in women group. In these items, we also found differences across sex in mean scores: Uncontrolled ($t = -4.23$; $p < .001$), Quivering ($t = -3.45$; $p < .01$), Flushing ($t = -4.30$; $p < .001$), Throbbing ($t = -3.48$; $p < .01$), Rising ($t = -3.84$, $p < .001$), and Trembling ($t = -6.11$; $p < .001$). For the Rewards factor, two items showed DIF, but in this case the SDP showed higher mean scores in men group. These two items showed significant differences between men and women: Peaceful ($t = 3,18$; $p < .01$) and Relaxing ($t = 3.10$; $p < .01$). Finally, item Tender, which belongs to the Intimacy factor, showed the least significant DIF. According to SDP, scores were slightly higher in men group, but no significant differences were found between sexes and their scores ($p > .05$). No indication of DIF was found while comparing sexes in the items of the Affective dimension.

Standard scores

Significant differences between sexes were found in the scores for the youngest age group (18-34 years) in the Rewards dimension ($t = -4.54$; $p < .001$; $d = .20$), where men ($M = 11.40$; $SD = 3.79$) indicated higher scores than women

Table 2 Test of factorial invariance across sex, age range, partner relationship, relationship length, sexual orientation, and education level.

IF test	RMSEA	90% CI RMSEA	CFI	Decision
Sex (men, women)				
1. Configural	.058	.056-.060	.928	Invariance accepted
2. Weak	.057	.055-.059	.927	Invariance accepted
3. Strong	.057	.055-.059	.926	Invariance accepted
4. Strict	.058	.056-.059	.921	Invariance accepted
Age range (18-34, 35-49, > 50 years old)				
1. Configural	.048	.046-.050	.923	Invariance accepted
2. Weak	.048	.046-.049	.920	Invariance accepted
3. Strong	.048	.047-.050	.917	Invariance accepted
4. Strict	.052	.050-.053	.898	Invariance rejected
Relationship (yes, no)				
1. Configural	.057	.055-.059	.929	Invariance accepted
2. Weak	.056	.054-.058	.928	Invariance accepted
3. Strong	.056	.054-.058	.928	Invariance accepted
4. Strict	.057	.055-.059	.923	Invariance accepted
Relationship length (1-10, 11-20, > 20 years)				
1. Configural	.050	.048-.052	.920	Invariance accepted
2. Weak	.049	.047-.051	.919	Invariance accepted
3. Strong	.048	.046-.050	.919	Invariance accepted
4. Strict	.050	.048-.052	.908	Invariance rejected
Sexual orientation (heterosexual, non-heterosexual)				
1. Configural	.057	.055-.059	.929	Invariance accepted
2. Weak	.056	.054-.058	.929	Invariance accepted
3. Strong	.056	.054-.057	.929	Invariance accepted
4. Strict	.054	.053-.056	.929	Invariance accepted
Education level (university studies, non-university)				
1. Configural	.058	.056-.060	.928	Invariance accepted
2. Weak	.057	.055-.059	.927	Invariance accepted
3. Strong	.056	.054-.058	.927	Invariance accepted
4. Strict	.056	.054-.057	.926	Invariance accepted

Note. Fit indexes included: *Root mean Square Error of Approximation* (RMSEA), 90% confidence interval for RMSEA (90% CI RMSEA), and *Comparative Fit Index* (CFI).

($M = 10.59$; $SD = 4.16$), and in the Sensory dimension ($t = 2.26$; $p < .05$; $d = .41$), where women ($M = 47.09$; $SD = 13.78$) indicated higher scores than men ($M = 41.12$; $SD = 15.53$). In the 35-49 age group, the most significant differences were noted in the Affective ($t = -3.13$; $p < .01$; $d = .28$) and Sensory ($t = -2.39$; $p < .05$; $d = .21$) dimensions. In both cases women indicating higher scores (Affective: $M = 25.38$, $SD = 5.29$; Sensory: $M = 39.08$, $SD = 17.06$) than men (Affective: $M = 23.80$, $SD = 5.98$; Sensory: $M = 35.44$, $SD = 16.99$). No significant differences between men and women in the 50+ age group were found for the ORS factors. Tables 4 and 5 present the standard scores of the four factors of the ORS, differentiated by sex and age groups.

Discussion

The main objective of this study was to analyze the factorial equivalence of the Spanish version of the ORS across various population groups, as well as the metric equivalence of the scale across sex. As a result, the invariance of the four factors was progressively evaluated across different groups, and the differential functioning of the 25 items across sex.

Once the factorial invariance was tested at different levels, and its metric equivalence, the standard scores for men and women were presented, distributed in three age range.

On the one hand, the equivalence of the factorial structure of the ORS has been confirmed by comparing groups created according to different socio-demographic characteristics. The four-factor model that describes the subjective orgasm experience in context of sexual relationship appears to be strictly invariant across sex, partner relationship, sexual orientation and education level. It reaches a strong level of equivalence across age range and relationship length. This means that being involved, or not, in a relationship with a partner, being heterosexual or having a different sexual orientation, as well as having a certain education level, does not change the perception of the subjective orgasm experience measured through the ORS. Having invariant measures in terms of sexual orientation is fundamental for the study of sexual diversity (Calvillo, Sánchez-Fuentes, & Sierra, 2018). Furthermore, being part of a certain age range, and, if applicable, the length of a relationship, do not impact the assessment of the four dimensions that constitute the scale (Affective, Sensory, Intimacy, and Rewards). Just as with other measurements

Table 3 Differential functioning of the items of the ORS across sex.

Items		Mantel-Haenszel (χ^2)	p	SDP
1	Elated	.29	.59	-.00
2	Flooding	1.16	.28	-.08
3	Pulsating	3.36	.06	.09
4	Satisfying	.41	.52	.01
5	**Uncontrolled**	**13.40**	**.00*****	**.20**
6	Blissful	.71	.40	.05
7	Loving	1.13	.29	-.08
8	**Quivering**	**9.22**	**.00*****	**.09**
9	Shooting	3.24	.07	-.11
10	Euphoric	1.90	.17	.04
11	**Flushing**	**10.40**	**.00*****	**.20**
12	**Tender**	**4.80**	**.03***	**-.17**
13	Close	2.03	.15	-.09
14	Exciting	.01	.92	-.05
15	Fulfilling	.09	.76	-.00
16	**Peaceful**	**32.98**	**.00*****	**-.41**
17	**Relaxing**	**26.50**	**.00*****	**-.35**
18	Soothing	3.49	.06	-.13
19	**Throbbing**	**10.63**	**.00*****	**.09**
20	Exploding	1.15	.29	-.07
21	Pleasurable	3.77	.05	.03
22	**Rising**	**9.28**	**.00*****	**.21**
23	Spreading	2.98	.08	.09
24	**Trembling**	**32.49**	**.00*****	**.44**
25	Wild	.48	.49	-.01

Note. Items with DIF in **bold**. SDP = p standardized difference.
* p < .05;
*** p < .001.

of sexual health dimensions, such as sexual assertiveness (Sierra, Santos-Iglesias, & Vallejo-Medina, 2012), or sexual double standard (Álvarez-Muelas, Vallejo-Medina, Gómez-Berrocal, & Sierra, 2018; Monge, Sierra, & Salinas, 2013), the ORS has shown to be invariant between men and women, which allows us to make valid comparisons between sexes.

On the other hand, after analyzing the differential functioning of the items of the ORS, it has been shown that all the items from Affective factor are equivalent for both men and women. This means that items related to the emotions experienced during orgasm in context of sexual relationship (e.g., satisfaction) do not differ between sexes. Items from Intimacy factor (e.g., loving) also do not present a different functioning across sex. Only one of these items showed weak signs of DIF, but does not reflect significant differences in the scores between men and women. Nevertheless, in accordance with previous studies (Arcos-Romero, Granados et al., 2018; Dubrai, Gérard, Beaulieu-Prévost, & Courtois, 2017; Mah & Binik, 2002), it has been shown the differences between sexes with regards to the perception of physiological feelings of orgasm. Some items from Sensory factor showed DIF and also significant differences between men and women. In this study, most of the items that show DIF comparing men and women belong to the Sensory factor (e.g., uncontrolled). In all of these items scores tend to be higher for women, which maybe supports the fact that men subjectively evaluate the physiological aspects of orgasm differently than women.

Finally, it should be noted that two of the three items from Rewards factor (i.e., peaceful and relaxing) have also showed a different functioning and mean scores across sex; in this case they tend to be higher in men. This would also suggest that the consequences caused by the subjective orgasm experience are evaluated differently by men and women. Similarly, Paterson, Jin, Amsel, and Binik (2014) also found significant differences across sex in the effects or physical and psychological changes experienced after orgasm.

If it is true that there is a concordance between the presence of DIF and the significant different mean scores of a specific item across sex, but they are completely independent analysis and we cannot conclude that these differences between men and women are necessarily due to the DIF. Differences found between sexes could be attributed to: (a) groups really differ in the evaluated variable; (b) items function differently between groups; and (c) a combination of both, that is to say, there are real differences between groups, and they also differ because the items do not function in the same way for both groups. Thus, we have to be cautious when interpreting the results of the DIF; attributing all the found differences in an assessment between groups to a real difference in the evaluated variable, can be as wrong as attributing the differences exclusively to the differential functioning of the items.

Based on the standard scores of the factors of the ORS, it is observed that comparisons by sex show results that

Table 4 Standard scores of the ORS in men of different age groups.

	Men (N = 715)											
Age	18-34 years	(n = 250)			35-49 years	(n = 250)			50-93 years	(n = 215)		
	A	S	I	R	A	S	I	R	A	S	I	R
α	.85	.95	.84	.88	.88	.95	.85	.89	.89	.96	.87	.93
M	26.24	41.12	11.0	11.4	23.8	35.44	10.6	10.36	22.83	33.06	10.26	10.33
SD	4.34	15.52	3.73	3.8	5.98	16.98	3.87	4.08	6.91	17.5	4.14	4.28
Min.	0	0	0	0	0	0	0	0	0	0	0	0
Max.	30	65	15	15	30	65	15	15	30	65	15	15
Percentile												
1	14	0	0	0	2	0	0	0	0	0	0	0
5	17	12	4	3	11	5	3	3	8	1	0	0
10	20	18	5	6	16	11	5	4	12	6	4	3
15	21	23	7	8	19	15	6	6	15	13	6	6
20	23	27	8	9	20	18	8	6	17	15	7	7
25	24	29	9	9	21	23	8	7	19	18	8	8
30	25	35	9	10	22	26	9	9	21	23	9	9
35	25	38	10	10	23	28	10	9	22	26	9	9
40	26	39	11	11	24	31	10	10	23	29	10	10
45	27	41	12	12	24	35	11	11	24	34	11	11
50	28	43	12	12	25	37	11	11	24	36	11	12
55	28	47	12	13	26	40	12	12	25	38	12	12
60	29	48	13	13	26	42	12	12	2	39	12	12
65	29	50	13	14	27	44	13	12	27	41	13	13
70	30	51	14	15	28	47	13	13	28	44	13	13
75	30	52	14	15	29	49	14	14	29	47	13	14
80	30	54	14	15	29	50	14	15	29	49	14	15
85	30	57	15	15	30	53	15	15	30	52	15	15
90	30	61	15	15	30	56	15	15	30	55	15	15
95	30	65	15	15	30	63	15	15	30	61	15	15
99	30	65	15	15	30	65	15	15	30	65	15	15

Note. A: Affective; S: Sensory; I: Intimacy; R: Rewards; M: mean; SD: standard deviation; Min.: minimum; Max.: maximum.
Skewness = -.59.
Kurtosis = .03.

are largely in accordance with those of the DIF and the differences between the measurements of the individual items. Despite the small effect size of the differences found, men and women under 50 differ in the Sensory dimension. Women show a greater intensity in the physiological perception of the orgasm experience. In accordance with the study of Mah and Binik (2005), and considering that this version of the ORS assesses a recent, unique and with-a-partner orgasm experience, it could be interpreted as woman have a higher perception of the anatomical localization of the orgasm in different parts of the body (e.g., heart rate, muscular spasms). Total scores of the Affective dimension are also higher in women, but only in the 35-49 age group. This could be associated to sexual scripts and gender roles, which are understood to be sociocultural constructs (Wiederman, 2015). This assign to women a more passive and submissive role during their sexual experiences giving them an emotional and sensitive character, in comparison to the more traditional dominant and controlling role of men (Bringas-Molleda et al., 2017; Garrido-Macías, Valor-Segura, & Expósito, 2017; Moyano, Monge, & Sierra, 2017; Ramiro-Sánchez, Ramiro, Bermúdez, & Buela-Casal, 2018). This fact would not be happening in youngest people.

However, in this age group, scores from the Rewards dimension are higher in men, which could mean that the gratification that the orgasm experience implies is more valued by men. It is possible that the lack of sexual experience in younger people leads men to exaggerate this evaluation of rewards, while women tend to undervalue it. Furthermore, it is possible that the female orgasm is less established than the male orgasm at this age. Sexual differences disappear in people over fifty, perhaps, as it is indicated below, due to a reduction in the intensity of the subjective orgasm experience as one gets older.

The standard scores also indicate that, as people get older, both men and women tend to have lower scores in all the four dimensions. Generally, the subjective assessment of the orgasm experience decreases with the passing of the years. As happens with other dimensions of sexual functioning, for example sexual desire (Moyano, Vallejo-Medina, & Sierra, 2017), these findings are not surprising considering that with age sexual inhibition (Granados, Salinas, & Sierra, 2017) and sexual dysfunctions (Sierra, Vallejo-Medina, Santos-Iglesias, & Lameiras-Fernández, 2012) tend to increase. It could lead to a lower perception and, as a result, a lower assessment of the orgasm experience. This

Table 5　Standard scores of the ORS in women of different age groups.

Age	Women (N = 757)											
	18-34 years	(n = 250)			35-49 years	(n = 250)			50-85 years	(n = 257)		
	A	S	I	R	A	S	I	R	A	S	I	R
α	.88	.93	.76	.87	.88	.95	.83	.87	.85	.96	.76	.89
M	26.92	47.1	10.87	10.6	25.38	39.08	10.0	10.0	22.76	34.1	10.9	9.95
SD	4.04	13.78	3.68	4.16	5.29	17.06	4.0	4.20	6.48	17.3	3.54	4.38
Min.	8	3	0	0	0	0	0	0	0	0	0	0
Max.	30	65	15	15	30	65	15	15	30	65	15	15
Percentile												
1	12	10	0	0	2	1	0	0	4	0	0	0
5	18	21	4	2	14	5	2	3	10	2	4	0
10	22	27	5	4	19	14	4	4	12	9	6	3
15	23	32	6	6	21	18	6	5	16	13	7	5
20	24	38	7	7	22	23	6	6	18	17	8	6
25	25	40	8	8	23	25	7	7	19	22	9	7
30	26	41	9	9	24	29	8	8	20	24	10	8
35	27	42	10	10	25	35	9	9	22	26	10	9
40	27	44	10	10	26	37	9	9	23	29	11	9
45	28	48	11	11	26	39	10	10	24	31	12	10
50	28	50	12	11	27	42	10	11	24	35	12	11
55	29	51	12	12	28	43	11	12	25	38	12	11
60	29	52	13	12	28	47	12	12	26	39	12	12
65	30	54	13	13	29	49	12	12	27	42	13	12
70	30	55	14	14	29	51	13	13	27	47	13	13
75	30	57	14	15	29	52	13	14	28	48	14	14
80	30	60	15	15	30	54	14	15	28	50	14	14
85	30	62	15	15	30	56	15	15	29	52	15	15
90	30	64	15	15	30	60	15	15	30	56	15	15
95	30	65	15	15	30	65	15	15	30	63	15	15
99	30	65	15	15	30	65	15	15	30	65	15	15

Note. A: Affective; S: Sensory; I: Intimacy; R: Rewards; *M*: mean; *SD*: standard deviation; Min.: minimum; Max.: maximum.
Skewness = -.69.
Kurtosis = .12.

also could be interpreted as a lower intensity of the subjective orgasm experience as one gets older associated with decreased strength in physiological changes that characterize the sexual response (López Sánchez, 2012). Notably, scores from the Affective and Sensory dimensions of orgasm are the ones that show a drastic decrease as people get older, while scores from the Intimacy and Rewards dimensions decrease more slowly in both men and women.

In conclusion, the Spanish version of the Orgasm Rating Scale is an instrument which is invariant across group on a factorial level, and shows equivalence across sex in the majority of its items on a metric level. Despite the fact that the scale factors as a whole are invariant across sex, some differential functioning is found between men and women when the items that make up each dimension are examined one by one. Definitely, the creation of the standard scores allows for a more precise multidimensional assessment of the subjective orgasm experience in men and women. Other generic scales which assess sexual functioning, as the Massachusetts General Hospital-Sexual Functioning Questionnaire (MGH-SFQ; Sierra et al., 2012) or the Arizona Sexual Experience Scale (ASEX; Sánchez-Fuentes, Moyano, Granados, & Sierra, 2019), only just interpret the presence

or absence, or the level of affectation of the orgasm. In general, the analysis carried out supports the tetra-factorial structure of the ORS. The adequate fit indicators, reliability, and validity endorse the model of the subjective orgasm experience that sustains the scale (Arcos-Romero, Granados et al., 2018). As a study limitation, the degree of accuracy of the results and their generalization would be modest because using the non-probability sampling we fix exclusively the size of the quotas by sex and age.

Note

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijchp. 2018.11.001.

References

Álvarez-Muelas, A., Vallejo-Medina, P., Gómez-Berrocal, C., & Sierra, J. C. (2018). *Factorial Invariance and Differential Item Functioning of Spanish version of Sexual Double Standard Scale*. Manuscript submitted for publication

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Arbuckle, J. R. (2015). *IBM SPSS Amos 24.0 User's guide*. Chicago, IL: Marketing Department, SPSS Incorporated.

Arcos-Romero, A. I., Granados, R., & Sierra, J. C. (2018). Relationship between orgasm experience and sexual excitation: Validation of the Model of the Subjective Orgasm Experience. *International Journal of Impotence Research*, http://dx.doi.org/10.1038/s41443-018-0095-6

Arcos-Romero, A. I., Moyano, N., & Sierra, J. C. (2018). Psychometric properties of the Orgasm Rating Scale (ORS) in context of sexual relationship in a Spanish sample. *Journal of Sexual Medicine*, *15*, 741–749. http://dx.doi.org/10.1016/j.jsxm.2018.03.005

Arcos-Romero, A. I., & Sierra, J. C. (2018). Revisión sistemática sobre la experiencia subjetiva del orgasmo. *Revista Internacional de Andrología*, *16*, 75–81. http://dx.doi.org/10.1016/j.androl.2017.09.003

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bringas-Molleda, C., Estrada-Pineda, C., Suárez-Álvarez, J., Torres, A., Rodríguez-Díaz, F. J., García-Cueto, E., & Rodríguez-Franco, L. (2017). Actitud sexista y trascedente durante el noviazgo entre universitarios lationoamericanos. *Revista Iberoamericana de Psicología y Salud*, *8*, 44–55. http://dx.doi.org/10.23923/j.rips.2017.08.005

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136–162). Thousand Oaks: Sage.

Byrne, B. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, *20*, 872–882.

Calvillo, C., Sánchez-Fuentes, M. M., & Sierra, J. C. (2018). Revisión sistemática sobre la satisfacción sexual en parejas del mismo sexo. *Revista Iberoamericana de Psicología y Salud*, *9*, 115–136. http://dx.doi.org/10.23923/j.rips.2018.02.018

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5

Dubrai, S., Gérard, M., Beaulieu-Prévost, D., & Courtois, F. (2017). Validation of a self-reported questionnaire assessing the bodily and physiological sensations of orgasm. *Journal of Sexual Medicine*, *14*, 255–263. http://dx.doi.org/10.1016/j.jsxm.2016.12.006

Emmerink, P. M., Van den Eijnden, R. J., Ter Bogt, T. F., & Vanwesenbeeck, I. (2017). A scale for the assessment of sexual standards among youth: Psychometric properties. *Archives of Sexual Behavior*, *46*, 1699–1709. http://dx.doi.org/10.1007/s10508-017-1001-x

Garrido-Macías, M., Valor-Segura, I., & Expósito, F. (2017). ¿Dejaría a mi pareja? Influencia de la gravedad de la transgresión, la satisfacción y el compromiso en la toma de decisión. *Psychosocial Intervention*, *26*, 111–116. http://dx.doi.org/10.1016/j.psi.2016.12.001

Gómez Benito, J., González, A., Widaman, K. F., Padilla, J. L., & Balluerka, N. (2017). Detecting differential item functioning in behavioral indicators across parallel forms. *Psicothema*, *29*, 91–95. http://dx.doi.org/10.7334/psicothema2015.112

González, A., Padilla, J. L., Hidalgo, M. D., Gómez-Benito, J., & Benítez, I. (2011). EASY-DIF: Software for analyzing differential item functioning using the Mantel-Haenszel and standardization procedures. *Applied Psychological Measurement*, *35*, 483–484. http://dx.doi.org/10.1177/0146621610381489

Granados, M. R., Salinas, J. M., & Sierra, J. C. (2017). Spanish version of the Sexual Excitation/Sexual Inhibition Inventory for Women: Factorial structure, reliability and validity evidence. *International Journal of Clinical and Health Psychology*, *17*, 65–76. http://dx.doi.org/10.1016/j.ijvhp.2016.09.003

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel Procedure. In H. Wainer, & H. I. Braun (Eds.), *Test Validity* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum Associates.

López Sánchez, F. (2012). *Sexualidad y afectos en la vejez*. Madrid: Pirámide.

Mah, K., & Binik, Y. M. (2002). Do all orgasms feel alike? Evaluating a two-dimensional model of the orgasm experience across gender and sexual context. *Journal of Sex Research*, *39*, 104–113. http://dx.doi.org/10.1080/00224490209552129

Mah, K., & Binik, Y. M. (2005). Are orgasms in the mind or the body? Psychosocial versus physiological correlates of orgasmic pleasure and satisfaction. *Journal of Sex & Marital Therapy*, *31*, 187–200. http://dx.doi.org/10.1080/00926230590513401

Mah, K., & Binik, Y. M. (2011). Orgasm Rating Scale. In T. D. Fisher, C. M. Davis, W. L. Yarber, & S. L. Davis (Eds.), *Handbook of Sexuality-Related Measures* (3rd ed., pp. 500–502). New York, NY: Routledge.

Marchal-Bertrand, L., Espada, J. P., Morales, A., Gómez-Lugo, M., Soler, F., & Vallejo-Medina, P. (2016). Adaptation, validation and reliability of the Massachusetts General Hospital-Sexual Functioning Questionnaire in a Colombian sample and factorial equivalence with the Spanish version. *Revista Latinoamericana de Psicología*, *48*, 88–97. http://dx.doi.org/10.1016/j.rlp.2016.01.001

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and Sensitivity of Alternative Fit Indices in Tests of Measurement Invariance. *Journal of Applied Psychology*, *93*, 568–592. http://dx.doi.org/10.1037/0021-9010.93.3.568

Meston, C. M., Hull, L., Levin, R. J., & Sipski, M. (2004). Disorders of orgasm in women. *Journal of Sexual Medicine*, *1*, 66–68. http://dx.doi.org/10.1111/j.1743-6109.2004.10110.x

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in crosscultural research. *International Journal of Psychological Research*, *3*, 111–121.

Monge, F. S., Sierra, J. C., & Salinas, J. M. (2013). Factorial and metric equivalence of Double Standard Scale by gender and age. *Suma Psicológica*, *20*, 7–14.

Moyano, N., Monge, F. S., & Sierra, J. C. (2017). Predictors of sexual aggression in adolescents: Gender dominance vs. rape supportive attitudes. *The European Journal of Psychology Applied to Legal Context*, *9*, 25–31. http://dx.doi.org/10.1016/j.ejpal.2016.06.001

Moyano, N., Vallejo-Medina, P., & Sierra, J. C. (2017). Sexual Desire Inventory: Two or three dimensions. *Journal of Sex Research*, *54*, 105–116. http://dx.doi.org/10.1080/00224499.2015.1109581

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, *25*, 151–157.

Nobre, P. J. (2006). *Disfunções sexuais – Teoria, Investigação e Tratamento*. Lisboa: Climepsi Editores.

Paterson, L., Jin, E. S., Amsel, R., & Binik, Y. M. (2014). Gender similarities and differences in sexual arousal, desire, and orgasmic pleasure in the laboratory. *Journal of Sex Research*, *51*, 801–813. http://dx.doi.org/10.1080/00224499.2013.867922

Pineda, D., Martín-Vivar, M., Sandín, B., & Piqueras, J. A. (2018). Factorial invariance and norms of the 30-item shortened-version of the Revised Child Anxiety and

Depression Scale (RCADS-30). *Psicothema*, *30*, 232–237. http://dx.doi.org/10.7334/psicothema2017.276

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. http://dx.doi.org/10.1016/j.dr.2016.06.004

Ramiro-Sánchez, T., Ramiro, M. T., Bermúdez, M. P., & Buela-Casal, G. (2018). Sexism and sexual behavior in adolescents: Gender differences. *International Journal of Clinical and Health Psychology*, *18*, 245–253. http://dx.doi.org/10.1016/j.ijchp.2018.04.002

Sánchez, V., Muñoz-Fernández, N., & Vega-Gea, E. (2017). Peer sexual cibervictimization in adolescents: Development and validation of a scale. *International Journal of Clinical and Health Psychology*, *17*, 171–179. http://dx.doi.org/10.1016/j.ijchp.2017.01.001

Sánchez-Fuentes, M. M., Moyano, N., Granados, M. R., & Sierra, J. C. (2019). Validation of the Spanish version of the Arizona Sexual Experience Scale (ASEX) using self-reported and physiological measures. *Revista Iberoamericana de Psicología y Salud*, *10*, 1–14. http://dx.doi.org/10.23923/j.rips.2019.01.021

Sierra, J. C., Moyano, N., Vallejo-Medina, P., & Gómez-Berrocal, C. (2018). An abridged Spanish version of Sexual Double Standard Scale: Factorial structure, reliability and validity evidence. *International Journal of Clinical and Health Psychology*, *18*, 69–80. http://dx.doi.org/10.1016/j.ijchp.2017.05.003

Sierra, J. C., Santos-Iglesias, P., & Vallejo-Medina, P. (2012). Evaluación de la equivalencia factorial y métrica de la Sexual Assertiveness Scale (SAS) por sexo. *Psicothema*, *24*, 316–322.

Sierra, J. C., Vallejo-Medina, P., Santos-Iglesias, P., & Lameiras-Fernández, M. (2012). Validación del Massachusetts General Hospital-Sexual Functioning Questionnaire (MGH-SFQ) en población española. *Atención Primaria*, *44*, 516–526. http://dx.doi.org/10.1016/j.aprim.2012.02.004

Socha, A., DeMars, C. E., Zilberberg, A., & Phan, H. (2015). Differential item functioning detection with the Mantel-Haenszel procedure: The effects of matching types and other factors. *International Journal of Testing*, *15*, 193–215. http://dx.doi.org/10.1080/15305058.2014.984066

Wiederman, M. W. (2015). Sexual script theory: Past, present, and future. In J. DeLamater, & R. F. Plante (Eds.), *Handbook of the Sociology of Sexualities* (pp. 7–22). New York, NY: Springer.