

Invariance of Spanish version of Sexual Double Standard Scale across sex, age, and education level

Ana Álvarez-Muelas¹, Carmen Gómez-Berrocal¹, Pablo Vallejo-Medina², and Juan Carlos Sierra¹

¹ Universidad de Granada and ² Fundación Universitaria Konrad Lorenz (Colombia)

Abstract

Background: The Sexual Double Standard Scale (SDSS) is one of the most widely-used scales for evaluating the Sexual Double Standard (SDS) and the Spanish version of the scale displays adequate psychometric properties in the adult Spanish population. The aim of this study is to examine the factorial invariance and differential item functioning (DIF) of the Spanish version according to sex, education level, and age, with the objective of describing and comparing SDS based on these variables. **Method:** A sample of 3,838 heterosexual participants was used (1,908 men and 1,930 women) between the ages of 18 and 90. **Results:** The results showed that the factorial structure of the SDSS is invariant in terms of sex and education level; with none of these items displaying DIF. Only items 11 and 14 display DIF in terms of age and, therefore, it is recommended that these items be excluded when making comparisons between age ranges. The association of SDSS with sex, education level, and age follows the expected results. **Conclusions:** The Spanish version of the SDSS is invariant and, as a consequence, can be used in different types of samples.

Keywords: Sexual Double Standard Scale, factorial invariance, sex, education level, age.

Resumen

Invarianza de la versión española de la Sexual Double Standard Scale por sexo, edad y nivel educativo. Antecedentes: la Sexual Double Standard Scale (SDSS) es una de las escalas más empleadas para evaluar el doble estándar sexual (DES) y la versión española presenta adecuadas propiedades psicométricas en población adulta. El objetivo de este estudio es estudiar la invarianza factorial y el funcionamiento diferencial de los ítems (DIF) de la versión en español según sexo, nivel de estudios y edad, con la finalidad de describir y comparar el DES en función de estas variables. **Método:** se empleó una muestra de 3.838 heterosexuales (1.908 hombres y 1.930 mujeres) entre 18 y 90 años de edad. **Resultados:** los resultados obtenidos muestran que la estructura factorial de la SDSS es invariante por sexo y nivel educativo, no presentando DIF ninguno de sus ítems. Solo los ítems 11 y 14 presentan DIF en función de la edad, por lo que se recomienda su eliminación en el caso de hacer comparaciones entre grupos etarios. La asociación del SDSS con el sexo, nivel de estudios y edad sigue los resultados esperados. **Conclusiones:** la versión española de la SDSS es invariante, por lo que se puede emplear en diferentes tipos de muestra.

Palabras clave: Sexual Double Standard Scale, invarianza factorial, sexo, nivel de estudios, edad.

The Sexual Double Standard (SDS) uses different criteria to evaluate the same sexual behaviour in men and women (Milhausen & Herold, 2002). More specifically, the traditional SDS prescribes that men should enjoy more sexual freedom than women. Even though, in recent decades, equality between the sexes has received increasing support for displaying or engaging in certain sexual behaviour (e.g. pre-marital sex), a traditional SDS still persists with regard to other heterosexual expressions, such as age at first sexual experience (Ortiz et al., 2011; Peixoto, Botelho, Tomada, & Tomada, 2016) or the number of sexual partners (Chi, Bongardt, & Hawk, 2015; Marks, Young, & Zaikman, 2019; Soller & Haynie, 2017). Conversely, heterosexual scripts are emerging which are aimed at sexual conservatism (Allison & Risman, 2013; Sakaluk, Todd,

Milhausen, & Lachowsky, 2014). These scripts could constitute a framework for promoting a modern SDS (as opposed to a traditional SDS) which prescribes that sexual shyness is more appropriate for women than for men (Fasula, Carry, & Miller, 2014).

In accordance with Bordini and Sperb (2013), the Sexual Double Standard Scale (SDSS; Muehlenhard & Quackenbush, 2011) is one of the most commonly-used measures for evaluating SDS and has recently been adapted for the Spanish population by Sierra, Moyano, Vallejo-Medina, and Gómez-Berrocal (2018). This adaptation has given rise to a version of 16 items grouped into two factors of eight items each: Acceptance for Sexual Freedom (ASF) and Acceptance for Sexual Shyness (ASS). In each factor the items are presented in parallel form, so that half refers to a specific sexual behaviour attributed to men, and the other half refers to the same sexual behaviour attributed to women. High ratings in the eight ASF items (e.g., item 1: "It's okay for a woman to have more than one sexual relationship at the same time"; ítem 6: "It's okay for a man to have more than one sexual relationship at the same time") indicate a positive attitude towards sexual freedom; and a positive attitude towards sexual shyness in the eight ASS items

(e.g., item 5: “A woman who initiates sex is too aggressive”; item 16: “A man who initiates sex is too aggressive”). The answers to the ASF items create the Index of Double Standard for Sexual Freedom (IDS-SF). The answers to the ASS items create the Index of Double Standard for Sexual Shyness (IDS-SS).

Compared to other scales, the SDSS has various advantages: (1) the items which are presented in parallel allow us to measure the participant’s attitude towards the same sexual behaviour in men and women (Bordini & Sperb, 2013), and by using two indices (IDS-SF and IDS-SS) it is possible to evaluate the support of SDS; (2) the factor which relates to sexual shyness allows us to detect more modern versions of SDS (Allison & Risman, 2013; Sakaluk et al., 2014; Sierra et al., 2018); and (3) both SDS indices (IDS-SF and IDS-SS) represent a bipolar measure (between -12 and +12) which allows for the evaluation of support for a traditional SDS (i.e., positive ratings) as well as an inverse SDS (i.e., negative ratings) which is characterised by the acceptance of more sexual freedom and less sexual shyness for women than for men (Fasula et al., 2014).

At present, there is a need to study differential item functioning and measurement invariance of tests (see Fonseca-Pedrero & Muñiz, 2017; Muñiz & Fonseca-Pedrero, 2019). In the SDS sphere, it is common to make comparisons by sex, age range or education level. The SDS has been found in both sexes, although it is more common in men (Gutiérrez-Quintanilla, Rojas-García, & Sierra, 2010; Sierra et al., 2018). It exists in different age ranges like adolescents (Monge, Sierra, & Salinas, 2013), young adults (Gutiérrez-Quintanilla et al., 2010) and older adults (Sierra, Monge, Santos-Iglesias, Rodríguez, & Aparicio, 2010). Moreover, it was found that the SDS decreased as the educational level increased (Sierra, Costa y Monge, 2012; Sierra et al., 2010). These comparisons were made without having previously demonstrated whether the instrument used for such purposes is invariant in terms of the group, and whether or not its items display differential functioning.

As such, it is fundamental to demonstrate that the scale for evaluating SDS allows for the comparison of ratings among the different groups, and that the measurement is equally accurate in the groups (Gómez-Benito, Sireci, Padilla, Hidalgo, & Benítez, 2018; Muñiz, Elosua, & Hambleton, 2013; Muñiz & Fonseca-Pedrero, 2019). The aim of this study is to examine the factorial equivalence and differential item functioning (DIF) of the Spanish SDSS version (Sierra et al., 2018) across sex, education level and age, and to examine the SDS based on these three variables.

Method

Participants

Through non-probability quota sampling a sample of 3,838 heterosexual Spanish adults was obtained (1,908 men y 1,930 women) between the ages of 18 and 90, with no differences between men ($M = 40.92$, $SD = 14.18$) and women ($M = 40.43$, $SD = 14.16$) ($t = 1.08$; $p = .281$). The sample was distributed incidentally into three age ranges: 18-34 years (651 men y 647 women), 35-49 years (611 men y 636 women) and over 50 years (646 men y 647 women). In regards to education level, 2.9% of men and 2.5% of women had no studies, 12.4% of men and 10.3% of women had primary school education, 30.8% of men and 24% of women had high school education, and 53.9% of men and 63.3% of women had university-level education, with significant differences between sexes in this variable ($\chi^2 = 23.47$ (3), $p < .001$).

Instruments

Sociodemographic Questionnaire. The questionnaire contains questions about sex, age, nationality, education level, and sexual orientation.

Spanish version of the Sexual Double Standard Scale (SDSS; Sierra et al., 2018). The Spanish version consists of 16 items which are answered on a Likert-type scale of four points, from 0 (*strongly disagree*) to 3 (*strongly agree*) and are grouped into two sub-scales: Acceptance for sexual freedom and Acceptance for sexual shyness. The indices of internal structure oscillated between .84 and .87 respectively, and the test-retest coefficients are placed above .70 at four and eight weeks. When correlated, the ratings from this version display adequate evidence of internal structure in the expected way for related constructs (Sierra et al., 2018). The sum of the ratings in the eight items, which make up each sub-scale, allows us to obtain two SDS indexes: one for sexual freedom (IDS-SF) and another for sexual shyness (IDS-SS). For the IDS-SF the items for women are inverted, and for the IDS-SS the items for men are inverted.

Procedure

The participants were from the general Spanish population and their answers were collected in two formats (pencil and paper, and online) which is common for questionnaires on sexual behaviour. There were no differences between the two versions (Sierra et al., 2018; Velten, Scholten, & Margraf, 2018). The pencil and paper format were managed by a trained evaluator in classrooms, foundations, and community centres; in small groups or individually. Once the participants had completed the questionnaire, it was returned in a sealed envelope. For the online questionnaire, a URL was sent via social network platforms and the IP address for each questionnaire was controlled; in order to access the questionnaire, the participant had to validate their access by answering a security question consisting of a random addition question. Both formats included an informed consent form which described the purpose of the study and included an explanation of what their participation entailed. Anonymity and confidentiality of answers were guaranteed. The study received prior approval from the Human Research Ethics Committee of the University of Granada.

Data analysis

The results were calculated using the R programme (Version 3.6.0; R Core Team, 2017), and the RStudio interface (Version 1.1.463; RStudio Team, 2016). The progressive invariance of the bifactorial model of the Spanish version of the SDSS was tested across sex, age range, and education level, which has six covariances between errors: configural invariance, weak, strong and strict, of the model on the polychoric matrix. We decided to follow these classical invariance method due to low cross-weighted loadings observed previously between both factors (Sierra et al., 2018). Thus, model should not be compromise for the non-target factor loadings constrained to zero across groups. Therefore, more flexible models as ESEM (Asparouhov & Muthén, 2009) or the maximum likelihood alignment approach (Byrne & van de Vijver, 2017) were not considered for the current paper; especially, considering that the maximum number of groups for

comparison in this study is not large ($n = 4$). The estimation model was the Diagonally Weighted Least Squares (WLSMV); a robust estimator for non-compliance of multivariate normality, designed for ordinal data (Li, 2016). It was considered for a good model fit an RMSEA $\leq .08$ and a CFI $\geq .95$ (Hu & Bentler, 1999). In order to obtain the invariance, a $-.01$ change in CFI was analysed, paired with changes in RMSEA of $.015$ with regard to the least restrictive model (Chen, 2007). The semTools package (R package, Version 0.4-14; semTools Contributors, 2016) and lavaan (Rosseel, 2012) were used for these analyses. The Differential Item Analysis (DIF) was calculated with the Lordif package (Choi, Gibbons, & Crane, 2011). The Lordif package is based on the DIF with par framework (Crane, Gibbons, Jolley, & van Belle, 2006) and allows for DIF detection with a flexible iterative hybrid /IRT framework. Uniform (model1- model2) and non-uniform (model 2- model3) DIF were tested; a single omnibus test of both uniform and non-uniform DIF (model 1- model 3) was also tested. DIF presence was considered based on the likelihood ratio χ^2 test, but due to the big sample size some DIF magnitude estimator was also considered. DIF magnitude will be evaluated with 3 pseudo R^2 measures Cox and Snell (1989), Nagelkerke (1991), and McFadden (Menard, 2000). $R^2 \geq .035$ will highlight DIF presence (Jodoin & Gierl, 2001). In addition, the β_1 was considered to be an estimator of the presence of DIF; an increase of 5% between model 1 and 2 would indicate the presence of DIF (Crane et al., 2007). Then, where there was differential item functioning, a Monte Carlo (1000 replications) procedure was also run to identify empirically-based thresholds for DIF detection (Choi et al., 2011). The use of the packages ggplot2 (R package, Version 3.1.1; Wickham, 2009) and gridExtra (R package, Version 2.3; Auguie, 2017) must also be highlighted. The corrected item-total correlations are polychoric and the presented alpha is the ordinal. Finally, ANOVA data analysis was performed in order to examine the differences of SDS across sex, education level, and age.

Results

Initially, the factorial invariance was evaluated for the previously observed bifactorial model of the SDSS (Sierra et al., 2018). As this is a variable which could be affected – or which is usually compared – by sex, age or education level, we wanted to

ensure that its dimensionality was stable by using these categories. As can be seen in Table 1, it appeared that these two factors were strictly comparable with regards to sex. In terms of education level, the observed data indicated an almost total equivalence between the four groups (no studies, primary, secondary, and university). Albeit, for the last education level – the strict level – the CFI reduced by $.002$ more than the allowable threshold; something which did not happen for RMSEA Δ , which remained within the specified range. This could indicate a distribution of the non-invariant errors by education level; something that would not affect the direct comparisons of the total of the sub-scales between the sexes. With regard to age, the results initially cast doubt on the configural model with an RMSEA above the required minimum threshold. In addition, there was no equivalence between the errors of the three evaluated groups, or at least that is the result of the values - much higher than allowed - for the CFI Δ , and RMSEA Δ . This compelled us to consider that there may be problems with some of the items; most likely the existence of DIF (Dimitrov, 2010). As such, DIF was analysed in order to find more concrete problems for the scores of this variable, and other variable that were compared.

After analysing the factorial invariance, the presence of DIF was tested across sex, age range, and educational level. The main criteria that we selected for indicating the presence of DIF was the Likelihood Ratio χ^2 test; however, it is normal to observe false DIF detection in large samples and, as a consequence, the results were supplemented with other estimators. With regards to sex, five out of eight items displayed DIF using the Likelihood Ratio χ^2 test for the ASF sub-scale. Nevertheless, as can be seen in Figure 1, the characteristic curves of the items with DIF were identical to those items without DIF. In fact, the highest observed R^2 – the Nagelkerke – is $.020$ from the item in sixth position (item 14), which would be the same as what is observed in Figure 1D, for 1,000 replicas simulated without DIF. For the same dimension, Figures 1B and 1C would confirm the true absence of DIF. Figure 1B showed a trait distribution which was practically identical for men and women, and in Figure 1C the lines for purified and non-purified DIF overlap (only one line can be seen in Figure 1C). All of the above would indicate that the presence of DIF is false and that there was no bias in the ASF dimension with regard to the sex of the participants. Also, with regard to sex, a total of four out of

Table 1
Fit indices and invariance indicators for the bifactorial model

		df	R χ^2	$\Delta R\chi^2$	p	CFI	ΔCFI	RMSEA	$\Delta RMSEA$
Sex	Configural	194	2098	NA	<.01	.971	NA	.080	NA
	Weak	208	2277	170.70	<.01	.968	-.002	.080	.000
	Strong	238	2586	258.91	<.01	.966	-.003	.078	-.002
	Strict	240	2760	18.02	<.01	.966	.000	.077	-.001
Studies	Configural	388	2457	NA	<.01	.969	NA	.079	NA
	Weak	430	2533	56.47	<.01	.969	.000	.074	-.004
	Strong	520	2790	144.70	<.01	.971	.001	.066	-.008
	Strict	526	4123	114.18	<.01	.959	-.012	.078	.012
Age	Configural	291	2473	NA	<.01	.964	NA	.081	NA
	Weak	319	2546	100.40	<.01	.962	-.002	.080	-.002
	Strong	379	3127	408.54	<.01	.957	-.005	.078	-.002
	Strict	383	5530	253.41	<.01	.928	-.029	.101	.023

Note: df = degree of freedom, R χ^2 = Robust chi squared estimator, $\Delta R\chi^2$ = increase of the Robust chi squared estimator, CFI = Comparative Fit Index, ΔCFI = increase of the Comparative Fit Index, RMSEA = Root Means Square Error Approximation, $\Delta RMSEA$ = increase of the Root Means Square Error Approximation

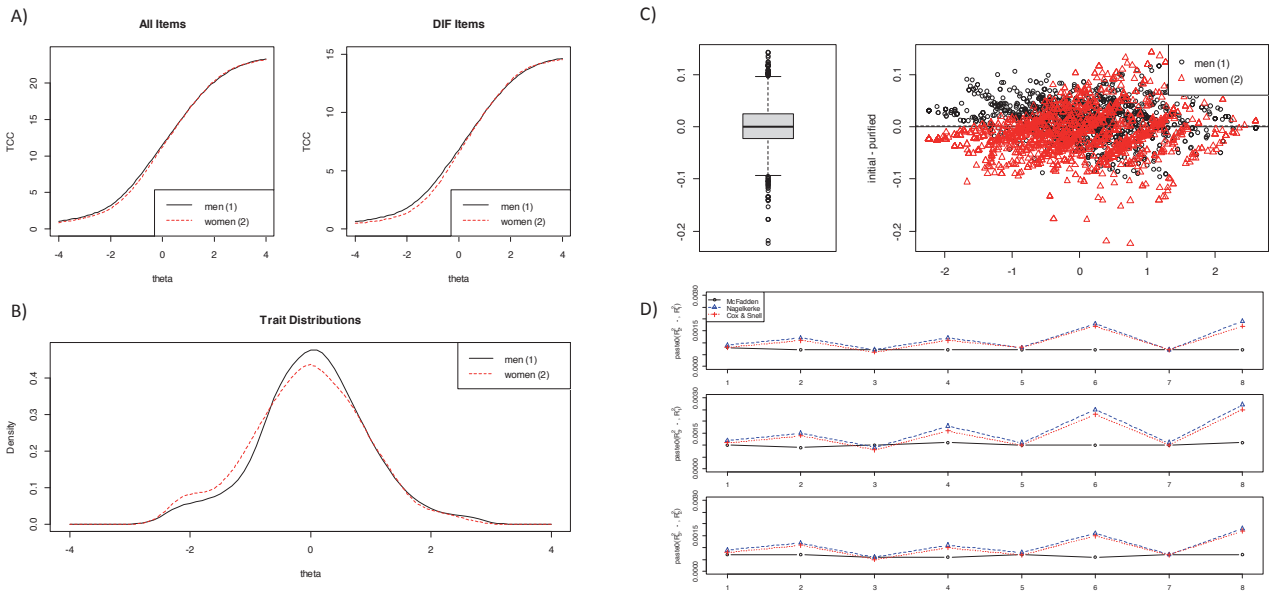


Figure 1. DIF of sub-scale Acceptance for Sexual Freedom (ASF) for the sex variable

eight items were flagged for DIF in the ASS dimension. As can be seen in Figure 2, the presence of DIF was negligible; as in the previous case. The highest observed R^2 – again, the Nagelkerke – was for the item in second position; similarly, with .020. In Figure 2D, one can observe how the effect size was achievable without problems in clean replications of DIF. It was true that the largest purification of these items appeared to have an effect, compared to the previous dimension – at least in Figure 2C two lines can be observed – but, again the ratings with DIF and purified DIF were virtually the same. In Choi et al. (2011) one can see a detailed description of the figures presented here.

With regard to education level, for ASF as well as ASS, a significant presence of DIF was observed when taking into account

the Likelihood Ratio χ^2 . Albeit, this had been amortised for all dimensions, with maximum R^2 for both sub-scales of .015. The DIF details can be seen in Figures 3 and 4, taking into account the technical descriptions of the previous paragraph.

True DIF had been observed with regard to age (see Figures 5 and 6), as far as the ASF dimension was concerned. Items 11 “A man should be sexually experienced when he gets married” and 14 “A woman should be sexually experienced when she gets married” had displayed a Nagelkerke R^2 , for the model 1-3 which jointly collected the presence of uniform and non-uniform DIF of .038 and .033 respectively. Although it was true that item 14 was below the threshold of .035 (see Figure 5D), the presence of DIF was significant. In this case, the presence of DIF was creating bias

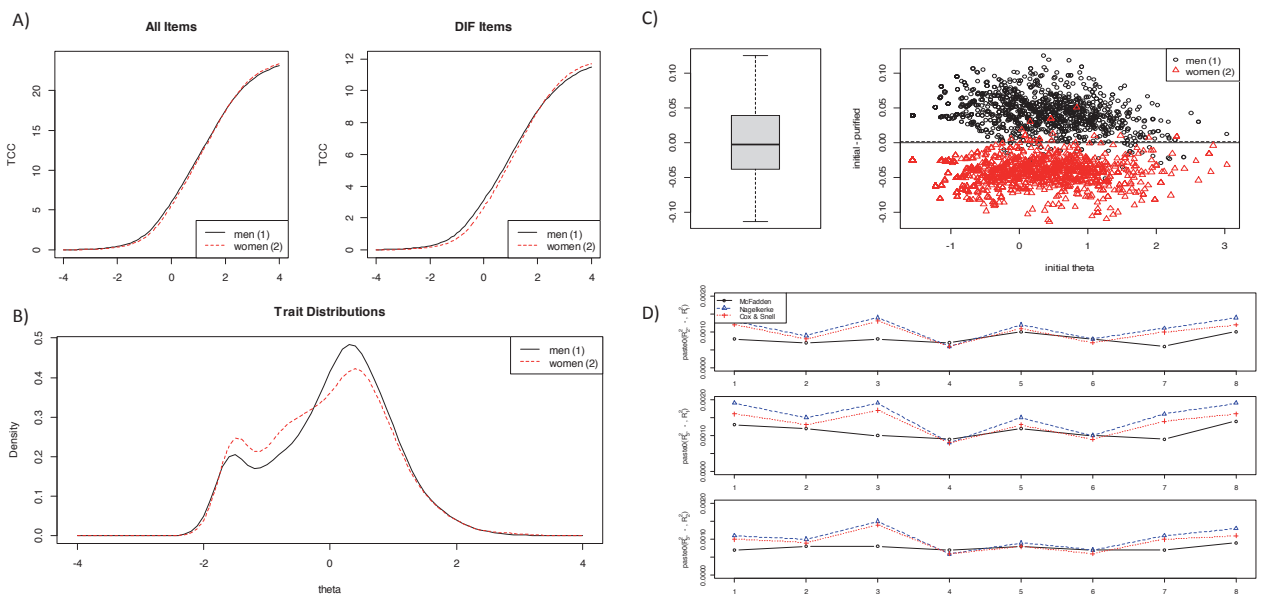


Figure 2. DIF of sub-scale Acceptance for Sexual Shyness (ASS) for the sex variable

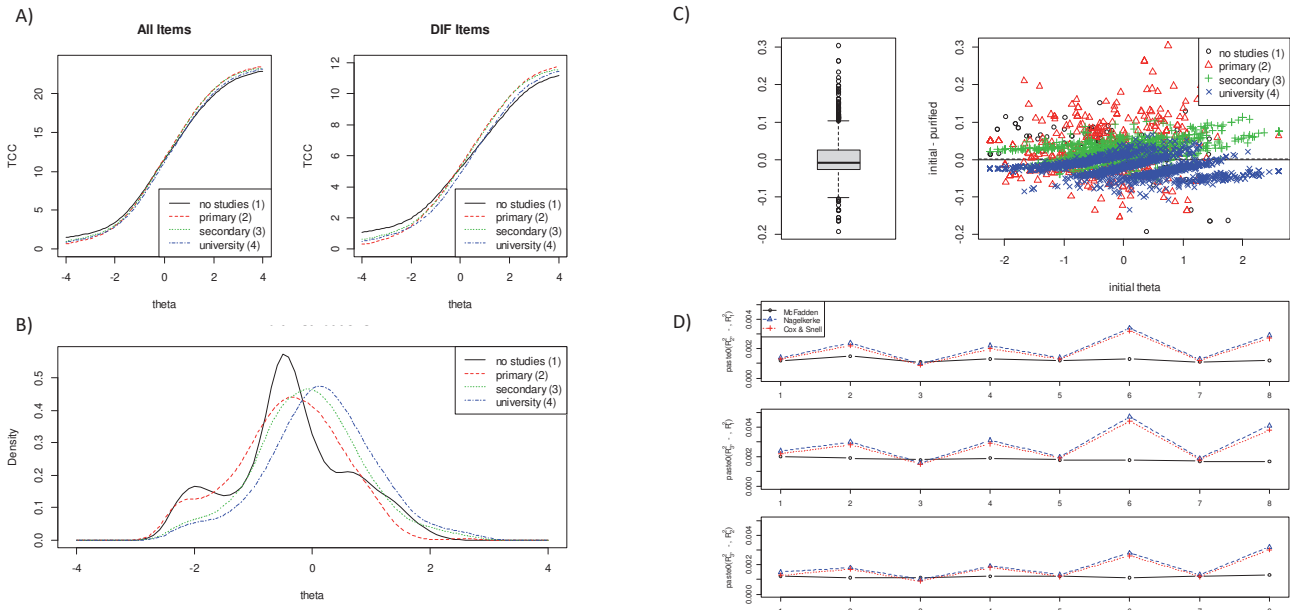


Figure 3. DIF of sub-scale Acceptance for Sexual Freedom (ASF) for the education variable

in the scale for different age ranges. As can be seen in Figure 5A, the presence of DIF appeared in the answers with high values of these items (i.e., strongly agree) as well as in the answers with low values (i.e., strongly disagree). Young participants were less likely to select the options with a high value for these items (theta). In terms of age – but with regard to the ASS dimension and according to the aforementioned interpretations – no presence of DIF was observed with a maximum R^2 of .011 for the item in first place (number 8 on the scale).

Table 2 displays the estimation of the reliability of the SDSS sub-scales for the different groups.

Lastly – given the fact that the scale is equivalent by sex, education level, and age – discarding items 11 and 14 in the last case, we examined the differences for the SDS indices based on these variables. Table 3 contains the descriptive data for indices IDS-SF and IDS-SS. Figure 7 shows the densiograms of these differences after ANOVA data analysis, as well as their statistical contrasts by sex, educational level, and age.

Discussion

The literature points to differences in SDS based on sex (Allison & Risman, 2013; England & Bearak, 2014; Guo, 2019;

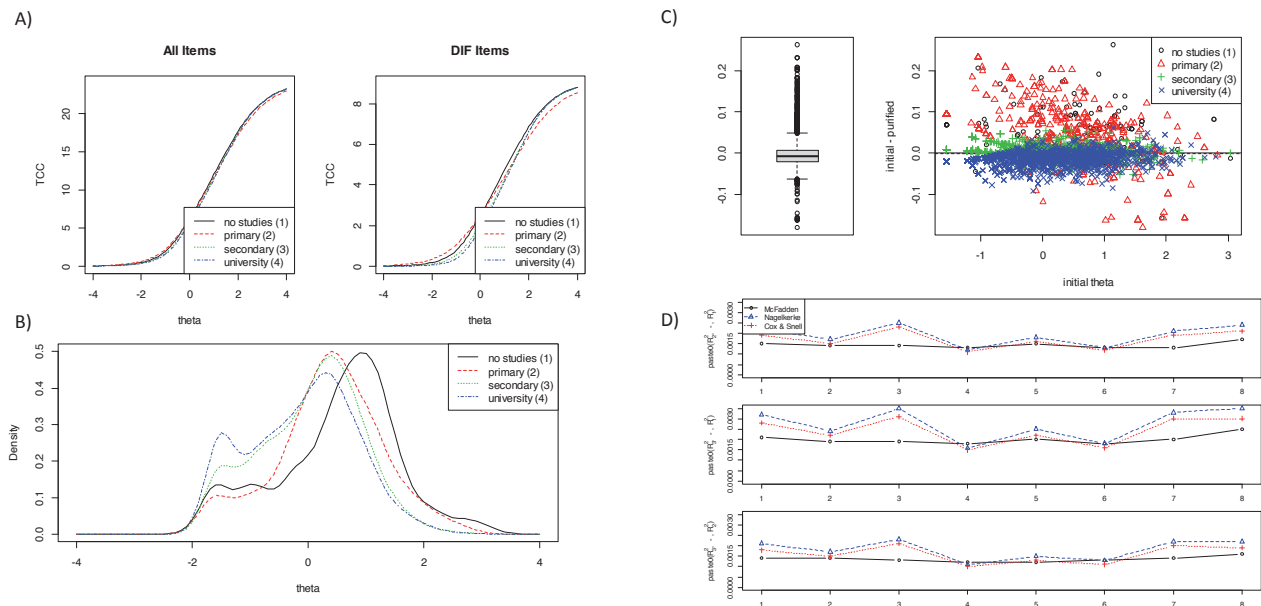


Figure 4. DIF of sub-scale Acceptance for Sexual Shyness (ASS) for the education variable

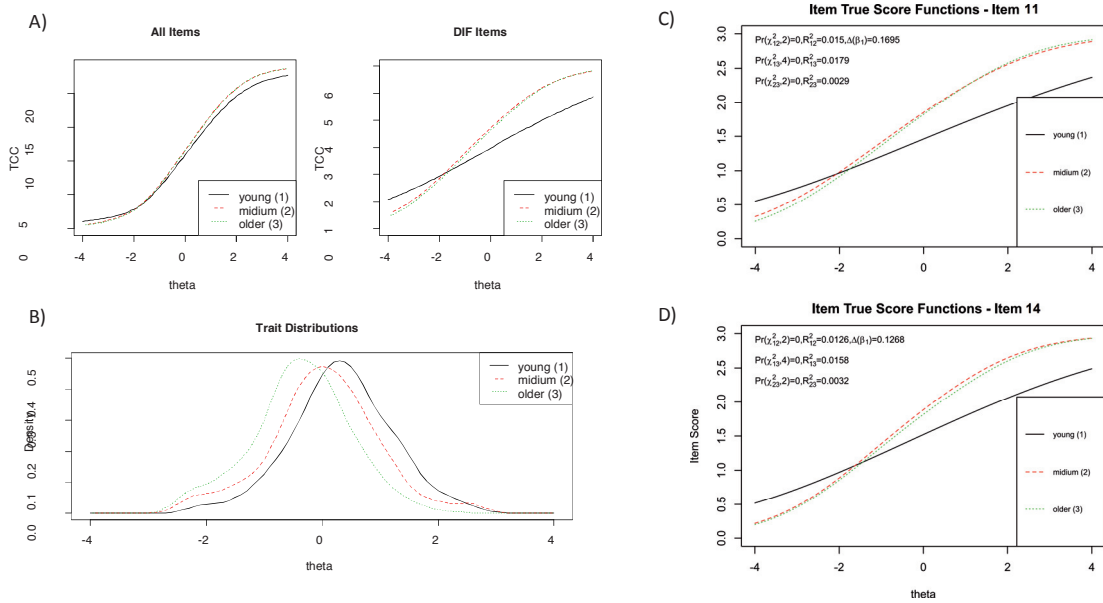


Figure 5. DIF of sub-scale Acceptance for Sexual Freedom (ASF) for the age variable

Sierra et al., 2018), sociocultural levels (Crawford & Popp, 2003) or age (Sierra et al., 2018), which is why it is necessary to have instruments that allow us to evaluate it without variance in the analysis of the construct based on population groups (Dimitrov, 2010). As such, the aim of this study was to determine whether the Spanish version of the SDSS (Sierra et al., 2018) is equivalent by sex, education level, and age. It is important to note that this version allows us to directly evaluate the subject's attitude towards freedom (ASF) or towards sexual shyness (ASS), and indirectly measures the SDS by combining the sum of the ratings of the items that refer to men and women in each sub-scale.

Strict invariance was reached with regard to sex and no differences in measurements were found through DIF analysis. It is confirmed that the scores of SDSS was equivalent among men and women; meaning that it allows for comparison with a minimum bias and affirms that it is rigorous in both groups (Dimitrov, 2010). A similar result was also obtained with the Double Standard Scale, which showed that it was invariant by sex in a Peruvian sample (Monge et al., 2013). This allows us to affirm that there is a greater presence of traditional SDS in men than in women – as has been reported by previous studies (Allison & Risman, 2013; England & Bearak, 2014; Guo, 2019; Sierra et al., 2018). It also allows us to identify and compare more modern, potential expressions of the

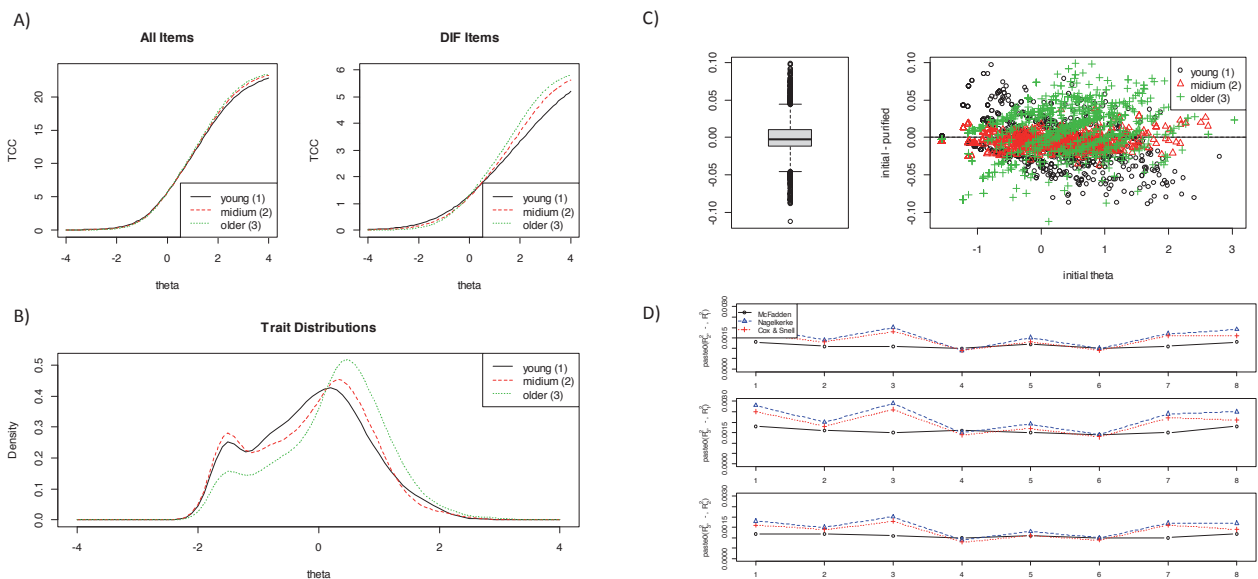


Figure 6. DIF of sub-scale Acceptance for Sexual Shyness (ASS) for the age variable

		ASF	ASS
		α	α
Sex	Male	.75	.81
	Female	.82	.85
Studies	No studies	.79	.86
	Primary	.78	.80
	Secondary	.78	.80
	University	.80	.83
Age*	Young	.79	.81
	Medium	.83	.83
	Old	.81	.84

Note: *Age data are reported without items 11 and 14 corresponding to ASF dimension. α = alpha ordinal; ASF = Acceptance for sexual freedom; ASS = Acceptance for sexual shyness

		IDS-SF	IDS-SS
		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Sex	Male (<i>n</i> = 1,908)	0.46 (1.86)	0.57 (2.03)
	Female (<i>n</i> = 1,930)	-0.27 (1.40)	-0.13 (1.49)
Studies	No studies (<i>n</i> = 105)	0.77 (2.07)	0.72 (2.19)
	Primary (<i>n</i> = 435)	0.33 (1.98)	0.64 (2.26)
	Secondary (<i>n</i> = 1,051)	0.18 (1.89)	0.36 (2.05)
	University (<i>n</i> = 2,247)	-0.03 (1.47)	0.05 (1.53)
Age*	Young (<i>n</i> = 1,298)	-0.05 (1.42)	0.05 (1.68)
	Medium (<i>n</i> = 1,247)	0.17 (1.34)	0.20 (1.69)
	Old (<i>n</i> = 1,293)	0.17 (1.44)	0.41 (2.02)

Note: *Age data are reported without items 11 and 14. *M* = Mean; *SD* = Standard deviation; IDS-SF = Index of Double Standard for Sexual Freedom; IDS-SS = Index of Double Standard for Sexual Shyness

SDS by sex (Fasula et al., 2014). Likewise, for education level, the scale almost obtained a total invariance among the evaluated groups. In addition, the DIF analysis was insignificant in both dimensions of the scale. Therefore, the equivalence of the SDSS scale is confirmed, allowing for the precise and independent

evaluation and comparison by education level. For the age variable, the ASS factor did not obtain differences in measurement of its items among the three age quotas (18-34, 35-49 and older than 50 years). In contrast, for the ASF dimension, DIF was present in items 11 and 14 which are parallel items referring to sexual

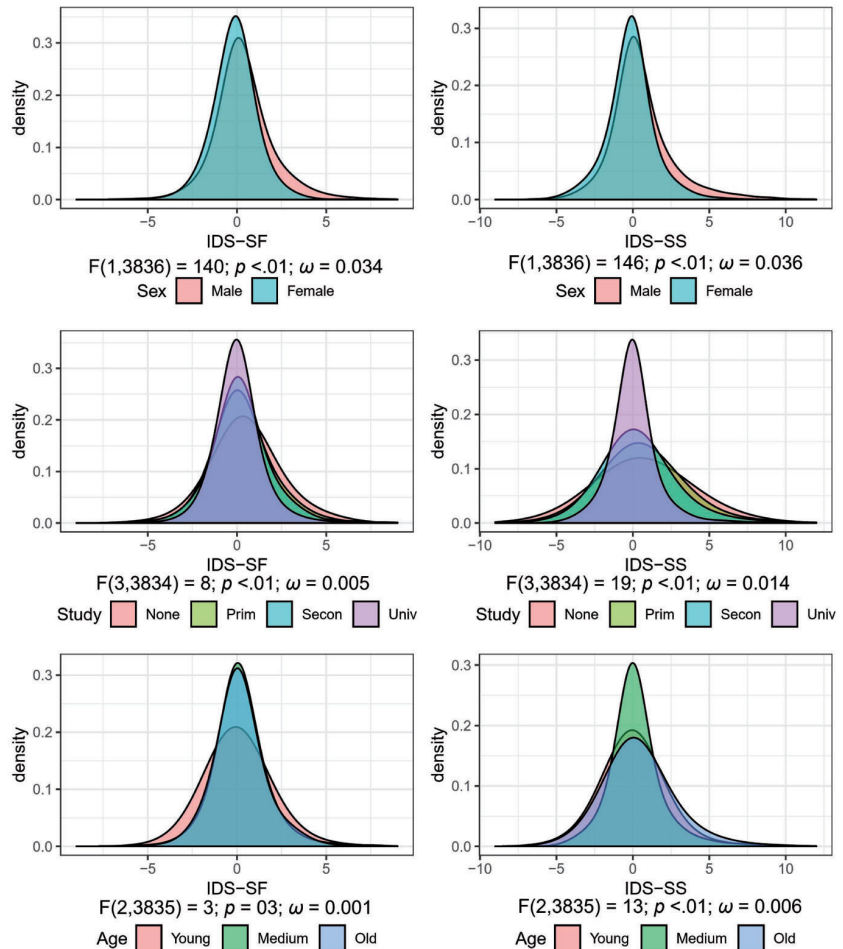


Figure 7. Densigram of Index of Double Standard for Sexual Freedom (IDS-SF) and Index of Double Standard for Sexual Shyness (IDS-SS) for the variables: sex, education level and age, along with the corresponding hypothesis contrasts

experience before marriage— creating measurement bias in the scale. The analysis suggests that it was not advisable to keep these two items for comparing SDS by age range. These results are found in the field of study of Monge et al. (2013) where the factorial equivalence and metric of the Double Standard Scale is analysed and where the variance by age was discarded. Studies found that changes social, cultural, and historical may play a role in the different life stages (Arnett, 2016). These changes allowed to identify a new period between 18 years old to late twenties that is different because of the demographic fluctuation (Arnett, 2016). Therefore, the study of both, education level and social class in this period is key (Arnett, 2016). Furthermore, this stage is characterized on one hand, by developmental identity, and on the other hand by belief and values determinations (Arnett & Mitra, 2018), this may be an explanation that justifies differences between youngers and the other age ranges.

The two sub-scales of the Spanish version of the SDSS showed good evidence of internal structure, with its coefficients varying between .75 of ASF in men and .86 of ASS in respondents with no education level – values which are very similar to those reported in the adaptation of the scale (Sierra et al., 2018).

Once the invariance of the scores was demonstrated by sex, education level and age (discarding items 11 and 14 in the last case), and once the validity of the sub-scales which make up the Spanish version of the SDSS was demonstrated, the different groups were compared in the SDS. Even though, in general, the effect size found in the differences were small, statistically significant differences were indeed found in the expected direction among the groups. Thus, men reported more traditional SDS than women within the scope of sexual liberty as well as sexual shyness – results which are in line with the previous studies (Allison & Risman, 2013; England & Bearak, 2014; Guo, 2019; Sierra et al., 2018). The novelty of the results lies within the average scores reported by women, as the negative scores were reflecting an inverse SDS tendency – a nuance which this version of the SDSS is able to capture. With regard to the association of the SDS with education level, the results were as expected: as the education level rises, the SDS indices decrease; as we have seen in previous studies (Sierra et al., 2012, 2018). In this case – even though it is very low – the tendency towards inverse SDS is only shown in people who have

university– education level, within the scope of sexual freedom. Lastly, the association between SDS and age also followed an expected pattern: higher levels were found in older people than in younger people. A slight inverse SDS was also observed in this case within the scope of sexual freedom in young people between the ages of 18 and 34. The comparison of the SDSS across age range and education level will allow us to know what the influence of sociocultural factors are on SDS and to explain the prevalence of the traditional and other demonstrations of the SDS in men and women.

The identification of differences in SDS among specific groups (such as those that we had examined in this study and using scales for analysing possible answer bias, as in the case of SDSS), is useful for adjusting the interventions on sexual health variables traditionally associated with SDS, such as the sexual victimization of women (Sierra, Bermúdez, Buela-Casal, Salinas, & Monge, 2014; Sierra, Monge, Santos-Iglesias, Bermúdez, & Salinas, 2011), male sexual aggression towards women (Kim, Lee, & Lee, 2019; Moyano, Monge, & Sierra, 2017; Sierra, Gutiérrez-Quintanilla, Bermúdez, & Buela-Casal, 2009), a greater risk of sexually transmitted diseases (Fasula et al., 2014; Ramiro-Sánchez, Ramiro, Bermúdez, & Buela-Casal, 2018) and reduced sexual satisfaction in both sexes (Haavio-Mannila & Kontula, 2003; Santos Iglesias et al., 2009).

In conclusion, the invariance attained and the absence of DIF showed that the SDSS (Sierra et al., 2018) is a scale that allows us to measure SDS in an equivalent way by sex, education level and age (eliminating items 11 and 14 in this case). Nevertheless, it is necessary to point out some limitations of the study. Despite using a large sample selected by quota, it is not a probability sample; and, as this is a transversal study, causality relationships cannot be established between the SDS and the variables that were analysed.

Acknowledgements

This study has been funded by the *Ministerio de Economía y Competitividad* through the Research Project PSI2014-58035-R and the Bursary FPU 16/04429 for University Professor Training, as part of the first author's thesis (Psychology Doctoral Programme B13 56 1; RD 99/2011).

References

- Allison, R., & Risman, B. J. (2013). A double standard for "Hooking Up": How far have we come toward gender equality? *Social Science Research, 42*, 1191-1206. doi: 10.1016/j.ssresearch.2013.04.006
- Arnett, J. J. (2016). Does emerging adulthood theory apply across social classes? National data on a persistent question. *Emerging Adulthood, 4*, 227-235. doi: 10.1177/2167696815613000
- Arnett, J. J., & Mitra, D. (2018). Are the features of emerging adulthood developmentally distinctive? A comparison of ages 18-60 in the United States. *Emerging Adulthood, 1*, 1-8. doi:10.1177/2167696818810073
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397-438. doi:10.1080/10705510903008204
- Auguie, B. (2017). gridExtra: Miscellaneous functions for "Grid" graphics [Computer software]. Retrieved from <https://CRAN.R-project.org/package=gridExtra>
- Bordini, G. S., & Sperb, T. M. (2013). Sexual double standard: A review of the literature between 2001 and 2010. *Sexuality y Culture, 17*, 686-704. doi: 10.1007/s12119-012-9163-0
- Byrne, B. M., & van de Vijver, F. J. R. (2017). The maximum likelihood alignment approach to testing for approximate measurement invariance: A paradigmatic cross-cultural application. *Psicothema, 29*, 539-551. doi: 10.7334/psicothema2017.178
- Chen F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464-504. doi: 10.1080/10705510701301834
- Chi, X., Bongardt, D. V. D., & Hawk, S. T. (2015). Intrapersonal and interpersonal sexual behaviors of Chinese university students: Gender differences in prevalence and correlates. *Journal of Sex Research, 52*, 532-542. doi: 10.1080/00224499.2014.914131
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software, 39*, 1-30. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3093114/>

- Cox, D. R., & Snell, E. J. (1989). *The Analysis of Binary Data* (2^a ed.). London: Chapman and Hall.
- Crane, P. K., Gibbons, L. E., Jolley, L., & van Belle, G. (2006). Differential item functioning analysis with ordinal logistic regression techniques: DIF detect and dif with par. *Medical Care*, *44*, S115-S123. doi: 10.1097/01mlr.0000245183.28384.ed
- Crane, P. K., Gibbons, L. E., Ocepek-Welikson, K., Cook, K., Cella, D., Narasimhalu, ...Teresi, J. A. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research*, *16*(Suppl 1), 69-84. doi: 10.1007/s11136-007-9185-5
- Crawford, M., & Popp, D. (2003). Sexual double standard: A review and methodological critique of two decades of research. *Journal of Sex Research*, *40*, 13-26. doi: 10.1080/00224490309552163
- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*, 121-149. doi: 10.1177/0748175610373459
- England, P., & Bearak, J. (2014). The sexual double standard and gender differences in attitudes toward casual sex among U.S. university students. *Demographic Research*, *30*, 1327-1338. doi: 10.4054/DemRes.2014.30.46
- Fasula, A. M., Carry, M., & Miller, K. S. (2014). A multidimensional framework for the meanings of the sexual double standard and its application for the sexual health of young black women in the U.S. *Journal of Sex Research*, *51*, 170-183. doi: 10.1080/00224499.2012.716874
- Fonseca-Pedrero, E., & Muñiz, J. (2017). Quinta evaluación de tests editados en España: mirando hacia atrás, construyendo el futuro [Fifth review of tests published in Spain: Looking back, building the future]. *Papeles del Psicólogo*, *38*, 161-168. doi: 10.23923/pap.psicol.2017.2844
- Gómez-Benito, J., Sireci, S., Padilla, J. L., Hidalgo, M. D., & Benítez, I. (2018). Differential item functioning: Beyond validity evidence bases on internal structure. *Psicothema*, *30*, 104-109. doi: 10.7334/psicothema2017.183
- Guo, Y. (2019). Sexual double standards in white and Asian Americans: Ethnicity, gender, and acculturation. *Sexuality & Culture*, *23*, 57-95. doi: 10.1007/s12119-018-9543-1
- Gutiérrez-Quintanilla, J. R., Rojas-García, A., & Sierra, J. C. (2010). Comparación transcultural de la doble moral sexual entre estudiantes universitarios salvadoreños y españoles [Cross-cultural comparison of the sexual double standard between Salvadoran and Spanish college students]. *Revista Salvadoreña de Psicología*, *1*, 31-51.
- Haavio-Mannila, E., & Kontula, O. (2003). Single and double sexual standards in Finland, Estonia, and St. Petersburg. *Journal of Sex Research*, *40*, 36-49. doi: 10.1080/00224490309552165
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55. doi: 10.1080/10705519909540118
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, *14*, 329-349. doi: 10.1207/S15324818AME1404_2
- Kim, Y., Lee, E., & Lee, H. (2019). Sexual double standard, dating violence recognition, and sexual assertiveness among university students in South Korea. *Asian Nursing Research*, *13*, 47-52. doi: 10.1016/j.anr.2019.01003
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*, 936-949. doi: 10.3758/s13428-015-0319-7
- Marks, M. J., Young, T. M., & Zaikman, Y. (2019). The sexual double standard in the real world. Evaluations of sexuality active friends and acquaintances. *Social Psychology*, *50*, 67-79. doi: 10.1027/1864-9335/a000362
- Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, *54*, 17-24. doi: 10.1080/00031305.2000.10474502
- Milhausen, R. R., & Herold, E. S. (2002). Reconceptualizing the sexual double standard. *Journal of Psychology and Human Sexuality*, *13*, 63-86. doi: 10.1300/J056v13n02_05
- Monge, F. S., Sierra, J. C., & Salinas, J. M. (2013). Equivalencia factorial y métrica de la Double Standard Scale por sexo y edad [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. doi:10.1016/j.ejpal.2016.06.001
- Muehlenhard, C. L., & Quackenbush, D. M. (2011). Sexual Double Standard Scale. In T. D. Fisher, C. M. Davis, W. L. Yarber, & S. L. Davis (Eds.), *Handbook of sexuality-related measures* (3th ed., pp. 199-201). New York, NY: Routledge.
- Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición [International Test Commission Guidelines for test translation and adaptation: Second edition]. *Psicothema*, *25*, 151-157. doi: 10.7334/psicothema2013.24
- Muñiz, J., & Fonseca-Pedrero. (2019). Diez pasos para la construcción de un test [Ten steps for test development]. *Psicothema*, *31*, 7-16. doi: 10.7334/psicothema2018.291
- Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*, 691-692. doi: 10.1093/biomet/78.3.691
- Ortiz, A. P., Soto-Salgado, M., Suárez, E., Santos-Ortiz, M. C., Tortolero-Luna, G., & Pérez, C. M. (2011). Sexual behaviors among adults in Puerto Rico: A population-based study. *Journal of Sexual Medicine*, *8*, 2439-2449. doi:10.1111/j.1743-6109.2011.02329.x
- Peixoto, C., Botelho, F., Tomada, I., & Tomada, N. (2016). Comportamento sexual de estudantes de medicina portuguesas e seus fatores preditivos [Sexual patterns in Portuguese medical students and its predictive factors]. *Revista Internacional de Andrología*, *14*, 53-68. doi: 10.1016/j.androl.2015.02.001 1698-031X
- Ramiro-Sánchez, T., Ramiro, M. T., Bermúdez, M. P., & Buela-Casal, G. (2018). Sexism and sexual risk behavior in adolescents: Gender differences. *International Journal of Clinical and Health Psychology*, *18*, 245-253. doi: 10.1016/j.ijchp.2018.04.002
- R Core Team. (2017). R: A language and environment for statistical computing Vienna: R Foundation for Statistical Computing [Computer software]. Retrieved from <https://www.R-project.org/>
- RStudio Team. (2016). RStudio: Integrated Development for R Boston, MA: RStudio, Inc. [Computer software]. Retrieved from <http://www.rstudio.com/>
- Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1-36. doi: 10.18637/jss.v018.i02
- Sakaluk, J. K., Todd, L. M., Milhausen, R., & Lachowsky, N. J. (2014). Dominant heterosexual sexual scripts in emerging adulthood: Conceptualization and measurement. *Journal of Sex Research*, *51*, 516-531. doi: 10.1080/00224499.2012.745473
- Santos Iglesias, P., Sierra, J. C., García, M., Martínez, A., Sánchez, A., & Tapia, M. I. (2009). Índice de Satisfacción Sexual: un estudio sobre su fiabilidad y validez [Index of Sexual Satisfaction (ISS): A study on the reliability and validity]. *International Journal of Psychology and Psychological Therapy*, *9*, 259-273.
- semTools Contributors. (2016). semTools: Useful tools for structural equation modeling [Computer software]. Retrieved from <https://CRAN.R-project.org/package=semTools>
- Sierra, J. C., Bermúdez, M. P., Buela-Casal G., Salinas, J. M., & Monge, F. S. (2014). Variables asociadas a la experiencia de abuso en la pareja y su denuncia en una muestra de mujeres [Factors associated with the intimate partner violence and their complaint in a sample of women]. *Universitas Psychologica*, *13*, 37-46. doi: 10.11144/Javeriana. UPSY13-1.vaea
- Sierra, J. C., Costa, N., & Monge, F. S. (2012, octubre). *Actitudes sexuales machistas en mujeres: análisis transcultural entre Brasil y Perú* [Sexist sexual attitudes in women: cross-cultural analysis between Brazil and Peru]. Poster session presented at meeting of XVI Congreso Latinoamericano de Sexología y Educación Sexual. Medellín, Colombia.
- Sierra, J. C., Gutiérrez-Quintanilla, J. R., Bermúdez, M. P., & Buela-Casal, G. (2009). Male sexual coercion: Analysis of a few associated factors. *Psychological Reports*, *105*, 69-79. doi: 10.2466/PRO.105.1.69-79
- Sierra, J. C., Monge, F. S., Santos-Iglesias, P., Bermúdez, M. P., & Salinas, J. M. (2011). Validation of a reduced Spanish version of the Index of Spouse Abuse. *International Journal of Clinical and Health Psychology*, *11*, 363-383.

- Sierra, J. C., Monge, F. S., Santos-Iglesias, P., Rodríguez, K., & Aparicio, D. L. (2010). Propiedades psicométricas de las versiones en español de la Double Standard Scale (DSS) y de la Rape Supportive Scale (RSAS) en mujeres peruanas [Psychometric properties of Spanish version of Double Standard Scale (DSS) and of Rape Supportive Attitude Scale (RSAS) in Peruvian women]. *Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace*, *95*, 57-66.
- 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. doi: 10.1016/j.ijchp.2017.05.003
- Soller, B., & Haynie, D. L. (2017). Variation in sexual double standards across schools: How do they matter of adolescent sexual behavior? *Sociological Perspectives*, *60*, 702-721. doi: 10.1177/0731121416668865
- Velten, J., Scholten, S., & Margraf, J. (2018). Psychometric properties of the sexual excitation/sexual inhibition inventory for women and men (SESI-W/M) and the Sexual Excitation Scales/Sexual Inhibition Scales Short Form (SIS/SES-SF) in a population based sample in Germany. *Plos ONE*, *13*, e0193080. doi: 10.1371/journal.pone.0193080
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis* [Computer software]. New York: Springer-Verlag. Retrieved from <https://cran.r-project.org/web/packages/ggplot2/index.html>