Equivalence and Standard Scores of the Hurlbert Index of Sexual Assertiveness Across Spanish Men and Women

Pablo Santos-Iglesias*, Pablo Vallejo-Medina, and Juan C. Sierra

Department of Personality, Assessment, and Psychological Treatment, University of Granada. Granada (Spain)

Título: Equivalencia y Baremos del Hurlbert Index of Sexual Assertiveness Entre Hombres y Mujeres Españoles.
Resumen: El objetivo de este estudio fue analizar la presencia de invarianza de medida y funcionamiento diferencial del ítem de la versión española del Hurlbert Index of Sexual Assertiveness en función del sexo. La muestra estuvo compuesta por 1.600 mujeres y 1.598 hombres de España, con edades comprendidas entre los 18 y 84 años. El Hurlbert Index of Sexual Assertiveness solo mostró invarianza débil entre hombres y mujeres. El análisis de funcionamiento diferencial del ítem mostró que únicamente el ítem 2 ("Creo que soy tímido en el ámbito sexual") mostró funcionamiento diferencial uniforme moderado. De forma más específica, las mujeres tendieron a responder "Siempre" a este ítem de forma más frecuente que los hombres. Los resultados sugirieron eliminar el ítem 2 resultando en una versión final compuesta por 18 ítems agrupadas en dos dimensiones, con buenos índices de fiabilidad tanto para hombres como para mujeres. Los baremos obtenidos para la escala de Inicio y Ausencia de timidez/Rechazo reflejaron la existencia de roles sexuales tradicionales en hombres y mujeres.
Palabras clave: Asertividad sexual; Hurlbert Index of Sexual Assertiveness; invarianza de medida; funcionamiento diferencial del ítem; baremos.

Abstract: The purpose of the present study was to analyze the measurement invariance and differential item functioning of the Spanish version of the Hurlbert Index of Sexual Assertiveness across gender. The sample was composed of 1,600 women and 1,598 men from Spain, with ages ranging from 18 to 84 years old. The Hurlbert Index of Sexual Assertiveness only showed weak invariance for men and women. The differential item functioning analysis showed that only item 2 ("I feel that I am shy when it comes to sex") flagged moderate uniform differential item functioning. More specifically, women tended to respond "Always" to this item more frequently than did men. Results strongly suggested eliminating item 2, resulting in a final version with 18 items clustered into two dimensions with good reliability values for men and women. Standard scores for both Initiation and No Shyness/Refusal reflected traditional sexual scripts for men and women.
Key words: Sexual assertiveness; Hurlbert Index of Sexual Assertiveness; measurement invariance; differential item functioning; standard scores.

Introduction

Sexual assertiveness has been defined in a variety of ways. Painter (1997) stated that sexual assertiveness is the ability to develop assertive behaviors in a sexual context. Dunn, Lloyd, and Phelps (1979) noted that it involves using "behavioral skills to obtain sexual satisfaction for yourself and your partner" (p. 294). Morokoff et al. (1997) provided a clearer picture of sexual assertiveness by stating that it embraces the ability to initiate desired sexual contacts, refuse unwanted sexual contacts, and the ability to prevent pregnancy or STIs with a regular partner. In line with this definition, several studies have explored the relevance of sexual assertiveness for human sexual life (for a review, see Santos-Iglesias & Sierra, 2010a) and concluded that it helps develop sexual healthy behaviors (*e.g.,* use of condom) and obtain greater sexual satisfaction. Finally, sexual assertiveness training programs help promote positive sexual outcomes and behaviors (Kelly, St. Lawrence, Hood, & Brasfield, 1989; Murphy, Coleman, Hoon, & Scott, 1980, St. Lawrence et al., 1995).

According to the sexual script theory (Simon & Gagnon, 1984, 1986, 2003), men are typically initiators of sexual encounters, while women are supposed to be restrictors of such contacts. Thus, men should score high on initiation sexual assertiveness (*i.e.,* the ability to initiate desired sexual contacts) while women should score high on refusal sexual

assertiveness (*i.e.,* the ability to refuse undesired sexual contacts). This traditional sexual script has generated some research to analyze whether men or women scored higher on sexual assertiveness. In general, results have usually found that men scored higher than women on sexual assertiveness (Haavio-Mannila & Kontula, 1997; Pierce & Hurlbert, 1999; Snell, Fisher, & Miller, 1991), although results have been mixed (Stulhofer, Graham, Bozicevic, Kufrin, & Ajdukovic, 2007). For example, Pierce and Hurlbert (1999) interviewed 54 non-clinical individuals and 46 clinical individuals attending sex therapy and showed that men in both clinical and non-clinical samples scored higher on sexual assertiveness than women. On the other hand, Sutlhofer et al. (2007) interviewed a nationally representative sample of young men and women and found that women scored higher than men on sexual assertiveness. These results can be explained by the fact that the studies by Hurlbert et al. and Snell et al. were based on sexual assertiveness scores mostly composed of initiation items, while Stulhofer et al. used refusal assertiveness items (A. Stulhofer, personal communication, March 22, 2011). Moreover, a study by Sierra, Santos-Iglesias, and Vallejo-Medina (2012) showed that, as age increased, initiation sexual assertiveness was higher in men compared to women. These authors also found that refusal sexual assertiveness was higher in women than men regardless of age. These results suggest that sexual assertiveness might follow traditional sexual scripts. They also noted that men and women have usually been compared on the basis of their sexual assertiveness. However, only the study by Sierra, Santos-Iglesias, et al. tested for measurement invariance and differential item functioning of one of these sexual assertiveness measures. They found that the Sexual Assertiveness

* **Dirección para correspondencia [Correspondence address]:**
Pablo Santos-Iglesias. Department of Psychology, University of New Brunswick, P.O. Box 4400, Fredericton, NB. E3B 5A3. Canada. E-mail: psantos@unb.ca

Scale had a strict equivalent dimensionality across sexes and only one item flagged differential item functioning, so they concluded that there is no significant bias in the scale when comparing sexual assertiveness across sexes.

Measurement invariance means that the probability of an observed score does not depend on the person's group membership (Meredith, 1993), that is: "respondents from different groups, but with the same true score, will have the same observed score" (Wu, Li, & Zumbo, 2007, p. 2). This concept implies that measuring constructs with the same instrument will reflect differences based on the performance/attribute between groups, and not differences based on confounding variables. Differential item functioning (*DIF*) is related to the conditional probability of answering an item in two or more groups after matching on the underlying ability (Hidalgo & Gómez, 2006; Zumbo, 1999). In the context of sexual assertiveness, for example, measurements should be invariant and show lack of *DIF* for comparisons between men and women to really reflect differences in sexual assertiveness and not differences based on sexist items or item comprehension, for example. Both procedures are strongly related (Dimitrov, 2010; Holland & Wainer, 1993) and are supposed to be tested together as evidence of validity, especially when test scores are used to compare groups.

The Hurlbert Index of Sexual Assertiveness (HISA; Hurlbert, 1991) is one of the instruments used most frequently to assess sexual assertiveness (Santos-Iglesias & Sierra, 2010a). In its original version, it was composed of 25 items providing an one-dimensional measure of sexual assertiveness in couples. The Spanish adaptation was shortened to a 19-item version clustered into two dimensions: (a) Initiation, which reflects the ability to begin sexual contacts and to express sexual desires and fantasies; and (b) No Shyness/Refusal, which means the difficulty starting and maintaining sexual conversations and the inability to reject undesired sexual contacts (Santos-Iglesias & Sierra, 2010b). Although the HISA has shown adequate psychometric properties (Santos-Iglesias & Sierra, 2010b; Sierra, Santos, Gutiérrez-Quintanilla, Gómez, & Maeso, 2008) and has been used to compare men and women (see Pierce & Hurlbert, 1999), no studies have tested whether its psychometric properties are the same for men and women. Thus, the main aim of the present study was to assess the measurement invariance and *DIF* of the Hurlbert Index of Sexual Assertiveness across gender using a Spanish sample. Due to the lack of normative data and its potential usefulness for clinical and epidemiological assessments, standard scores were developed for both the Initiation and No Shyness/Refusal subscales for both men and women across three different age groups (18-34, 35-49, and 50 years old or older).

Method

Participants

Participants were recruited from the general population in Spain. The final sample was composed of 1,598 men and 1,600 women. Since the HISA assesses sexual assertiveness within partners, participants were required to be involved in a romantic relationship that included sexual activity at the time of the study. The mean age of men was 39.47 years (*SD* = 13.38, range 18-81), while that of women was 36.98 years (*SD* = 13.41, range 18-84). Educational level, religion, and frequency of religious practice are reported in Table 1.

Table 1. Educational level, religion, and religious practice of both men and women.

Variables	Men		Women	
	N	%	*n*	%
Educational level				
No education	25	1.6	20	1.3
Primary	274	17.1	267	16.8
Secondary	490	30.7	353	22.1
University	809	50.6	954	59.8
Religion				
Christian	1,135	71.2	1,237	77.6
Islamic	2	0.1	2	0.1
Hindu	1	0.1	3	0.2
Buddhist	2	0.1	3	0.2
None	455	28.5	350	21.9
Religious practice				
Daily	6	0.4	9	0.6
Once a week	67	4.2	119	7.4
Once a month	127	7.9	156	9.8
Once a year	754	47.2	841	52.6
Never	644	40.3	473	29.6

Measures

A background questionnaire was administered to obtain information about sex, age, whether participants were involved on a romantic relationship, whether they had sexual activity with their partners, educational level, religion, and frequency of religious practice.

Hurlbert Index of Sexual Assertiveness (HISA; Hurlbert, 1991). The Spanish version by Santos-Iglesias and Sierra (2010b) was used. It includes 19 items clustered into two factors: *Initiation* and *No Shyness/Refusal*. Participants responded using a 5-point Likert scale from 0 (*never*) to 4 (*always*). Higher scores indicated greater initiation assertiveness (Initiation subscale), and lack of shyness and greater refusal assertiveness (No Shyness/Refusal subscale). Santos-Iglesias and Sierra reported an internal consistency (McDonald's omega) of .83 for each factor and .87 for the global scale. It showed positive correlations above .10 with the Spanish version of the Sexual Assertiveness Scale (Sierra, Vallejo-Medina, & Santos-Iglesias, 2011) and above .18 with the Spanish abbreviated version of the Dyadic Adjustment Scale (Santos-Iglesias, Vallejo-Medina, & Sierra, 2009).

Procedure

Participants were recruited from the Spanish general population. A quota convenience sampling method was used to obtain the same number of men and women, distributed across different groups according to age (18-34 years old, 35-49 years old, and 50 years old or older), size of the town or city of residence (a population lesser than 50,000 and greater than 50,000), and geographical area (north and south of Spain). Participants were required to be involved in a stable heterosexual relation with sexual activity (because sexual assertiveness implies negotiation of sexual behaviors) for at least 6 months at the time of the study. Testing was conducted individually in different settings by well-trained researchers (public libraries, social centers, and public places). In university classrooms, participants were tested collectively. The purpose of the study was briefly explained to all participants. Verbal informed consent was obtained, and anonymity and confidentiality were guaranteed, as well as the exclusive use of the tests for research purposes.

Data analysis

Measurement invariance was tested using LISREL 8.51 (Jöreskog & Sörbom, 2001) following the procedure described by Wu et al. (2007) for multi-group confirmatory factor analysis (*MG-CFA*). Four models were assessed: (a) configural invariance constrained the number of factors and the pattern of free and fixed loadings across both groups; (b) weak invariance tested equality of factor loadings across groups; (c) strong invariance tested equality of intercepts for both groups; and (d) strict invariance assumed that residual variances for all items were equal across groups. These four steps were estimated using maximum likelihood. In order to avoid problems with sample size, three main indices were used to assess adjustment: the Root Mean Square Error of Approximation (*RMSEA*), Non-Normed Fit Index (*NNFI*), and Comparative Fit Index (*CFI*). In this context, *NNFI* and *CFI* values above .90 and *RMSEA* values below .05 were used as indicators of good fit (Cheung & Rensvold, 2002, Wu et al., 2007). Additionally, to assess the fit of nested models –such as the *MG-CFA*–, changes in the fit indices were examined (Cheung & Rensvold, 2002; Wu, et al., 2007). Cheung and Rensvold (2002) recommended using ΔCFI and proposed $\Delta CFI \leq -.01$ as a good indicator of measurement invariance.

Differential item functioning was tested using discriminant logistic analysis (Hidalgo & Gómez, 2006; Hidalgo & Gómez-Benito, 2010) through SPSS 17.0. A 3-step hierarchical procedure was followed. Step 1 tested the contribution of each subscale score (Initiation and No shyness/Refusal). Step 2 tested whether item score significantly contributed to differences between men and women (dependent variable), and Step 3 tested the interaction between subscale score and item score. Significance of Step 2 - Step 1 (Step 2 itself) indicated uniform DIF, while significance of

Step 3 - Step 2 (Step 3 itself) was considered evidence of non-uniform DIF. Effect size was tested through the increase in Nagelkerke's R^2, so that values up to .035 indicated negligible DIF, values between .035 and .070 showed moderate DIF, and values above .070 indicated large DIF (Jodoin & Gierl, 2001). A stepwise purification procedure was performed for all the items showing DIF. Finally, to analyze the response category in which the DIF did exist, a discriminant logistic analysis using a cumulative probability model was performed on each item showing DIF (Mellenberg, 1995).

Results

Measurement invariance

Measurement invariance started by testing configural invariance. Results showed that the model was the same for men and women (see Table 2). Although the χ^2 value was extremely high due to the large sample size, *CFI*, and *RMSEA* showed adequate fit. Step 2 involved testing whether weak invariance, or factor loading equivalence, was supported. The *NNFI*, *CFI* showed good fit, and *RMSEA* were close to a good fit, and the increase in *CFI* was -.002, indicating good fit for nested models between model 1 and model 2. Step 3 tested strong invariance or equivalence of intercepts across groups. Results showed an increase in the *RMSEA* and a decrease in the GFI, *NNFI*, and *CFI*, so that strong invariance is not supported. Furthermore, changes in the *CFI* reached .023, which meant that this nested model did not fit the data and therefore that strong invariance was not supported. At this point, the modification indices in the Tau-x matrix were assessed and revealed that items 2, 9, and 13 had large modification values (110.19, 62.58, and 57.39, respectively) and large expected change values too (.149; -.152; and -.149, respectively). This suggested testing strong invariance again without restrictions for these three items. Although results showed a slight non-significant decrease in the *CFI* (ΔCFI = -.01), values of the *NNFI*, *CFI*, and *RMSEA* did not supported partial strong invariance either. Since strong invariance and partial strong invariance are not supported, we did not check for strict invariance across gender (Byrne & van de Vijver, 2010; Wu et al., 2007).

Differential item functioning

As shown in Table 3, the only item flagging moderate uniform DIF across gender was item 2 (ΔR^2 = .059; "I feel that I am shy when it comes to sex"). The purification process showed that, after deleting item 2 from the matching score, uniform DIF was still moderate (χ^2 = 144.55, $p < .001$, ΔR^2 = .059). Results of the discriminant logistic analysis performed on response scale categories revealed moderate uniform DIF in response category 4 (*always*) for item 2 (χ^2 = 158.56, $p < .001$, ΔR^2 = .065), indicating that women chose this anchor more frequently than men (OR = 0.37).

Table 2. Goodness-of-fit indices for measurement invariance models.

Model	χ^2	df	NNFI	CFI	ΔCFI	RMSEA
1. Configural invariance	2,006.75***	305	.890	.902		.059
2. Weak invariance	2,063.30***	322	.894	.900	-.002	.058
3. Strong invariance	2,482.13***	341	.876	.877	-.023	.062
3a. Partial strong invariance[a]	2,231.16***	338	.889	.890	-.01	.059
4. Strict invariance[a]	2,379.52***	357	.885	.880	-.01	.059

[a] Without restrictions on intercepts in items 2, 9, and 13.   *** $p < .001$.

Table 3. Differential item functioning of the Initiation and No Shyness/Refusal subscales.

Scale	Item	Step 2 - Step 1			Step 3 - Step 2		
		χ^2	p	ΔR^2	χ^2	p	ΔR^2
Initiation	3	8.75	.003	.003	22.79	< .001	.009
	4	2.53	.11	.001	15.14	< .001	.006
	5	0.35	.55	.000	1.60	.20	.000
	7	1.83	.17	.000	11.19	< .001	.005
	16	7.87	.005	.003	1.20	.27	.000
	17	35.54	< .001	.014	1.81	.17	.001
	18	23.78	< .001	.009	29.81	< .001	.012
	21	2.40	.12	.001	2.74	.10	.001
No Shyness/	1	2.72	.09	.001	1.88	.17	.001
Refusal	2	145.49	< .001	.059	5.29	.02	.002
	6	1.12	.28	.001	1.16	.28	.000
	9	57.65	< .001	.024	2.09	.14	.001
	10	4.42	.03	.002	4.18	.04	.002
	11	4.79	.02	.002	2.36	.12	.001
	13	44.09	< .001	.018	12.88	< .001	.006
	14	.02	.88	.000	3.55	.06	.002
	19	.30	.58	.000	2.23	.13	.001
	24	.56	.45	.000	6.97	.008	.003
	25	.06	.80	.000	6.10	.013	.003

Standard scores and internal consistency

Standard scores for Initiation and No Shyness/Refusal were created from z score transformations due to the violation of normality (see Table 4 and Table 5, respectively). Cronbach's alpha values for both subscales are shown in parenthesis in Table 4 and Table 5. It must be noted that, according to the DIF results, item 2 was eliminated from the

Table 4. Standard scores of the Initiation subscale.

Initiation	Men (α = .74)			Women (α = .81)		
	18-34	35-49	50-	18-34	35-49	50-
M	24.40	22.53	21.62	22.84	21.20	16.73
SD	5.68	6.25	6.96	6.35	7.35	7.86
Cent						
99	32	32	32	32	32	32
95	32	32	32	32	32	30
85	30	29	29	30	30	25
75	29	27	27	28	28	23
65	28	26	25	26	25	21
55	26	24	23	24	23	18
50	25.50	23	23	24	22	16
45	25	22	22	23	20	15
35	23	21	19	21	18	13
25	21	18	17	19	16	10
15	18	16	14	16	13	8
5	13	11	9	10	7	4
1	8.06	6	4	6.08	5	0

No Shyness/Refusal subscale before calculating standard scores. Results showed that men tended to report higher scores than women on initiation assertiveness across all ages. Regarding the No Shyness/Refusal subscale, young women scored slightly higher than younger men, but older women scored higher than younger men (see Figure 1 and Figure 2).

Table 5. Standard scores of the No Shyness/Refusal subscale.

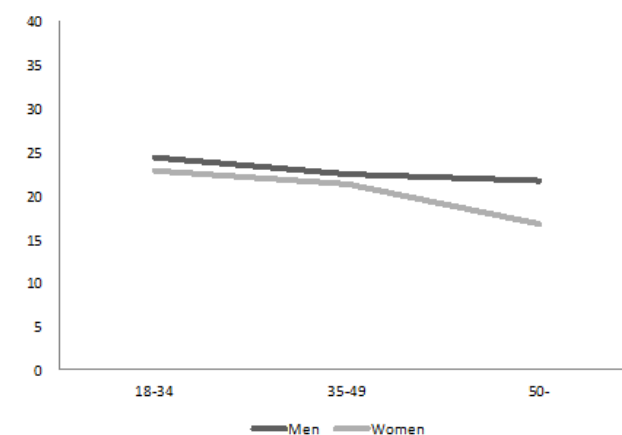| No Shyness	Men (α = .78)			Women (α = .83)		
	18-34	35-49	50-	18-34	35-49	50-
M	32.97	32.28	31.72	34.25	32.63	29.47
SD	5.31	5.89	5.91	4.87	6.04	7.28
Cent						
99	40	40	40	40	40	40
95	40	40	40	40	40	39
85	38	38	38	39	39	37
75	37	37	36	38	37	35
65	35	36	34	37	36	33
55	34	34	33	36	35	32
50	34	33	33	36	34	31
45	33	32	32	35	33	30
35	32	31	31	34	31	28
25	30	29	29	32	30	26
15	28	27	26	30	26	21.30
5	23	21	21	24	20	15.10
1	16	10.05	12	18	13.28	7

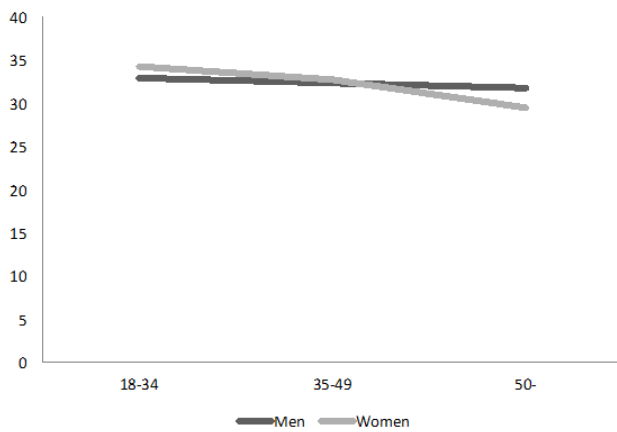**Figure 1.** Means of Initiation assertiveness for men and women.

Figure 2. Means of No Shyness/ Refusal assertiveness for men and women.

Discussion

When assessment instruments are used to compare groups (*i.e.,* cultures, gender, etc.) it is essential for such instruments to operate in the same way for each group (Dimitrov, 2010). The main purpose of the present study was to analyze the measurement invariance and differential item functioning of the Spanish version of the Hurlbert Index of Sexual Assertiveness (HISA; Santos-Iglesias & Sierra, 2010b), because it is a construct that has typically been compared across men and women. The Spanish version of the HISA only showed weak measurement invariance. Only the item 2 ("I feel that I am shy when it comes to sex") flagged moderate uniform DIF. Thus, we highly recommend eliminating this item from the scale.

Regarding measurement invariance, results show that the model proposed by Santos-Iglesias and Sierra (2010b) is the same for men and women, as proven by the configural invariance test. Further, not only is the structure the same, but factor loadings are also equivalent across gender. On the other hand, results failed to support both strong and strict invariance. Thus, since strong and, specially, strict invariance are not satisfied (Wu et al., 2007) we can not assume the Spanish version of the HISA to be invariant across sexes. This result is particularly relevant when we want to compare the sexual assertiveness scores of men and women using this scale. In such a case, we deeply encourage to use the Spanish version of the Sexual Assertiveness Scale (Sierra et al., 2011), which has been demonstrated to be invariant between men and women (Sierra et al., 2012).

Differential item functioning revealed that item 2 showed uniform differential item functioning, which means that men and women have different probabilities of endorsing a response even if they belong to the same attribute level. More specifically, women have a greater probability of responding "Always" to item 2 ("I feel that I am shy when it comes to sex") compared to men. These results are related with traditional sexual scripts and gender-role stereotypes, in

which women are supposed to follow traditionally feminine attributes like being sympathetic or shy (Bem, 1974; Holt & Ellis, 1998) and are encouraged not to talk overtly about sex (Quina, Harlow, Morokoff, Burkholder, & Deiter, 2000). Based on these results, we propose that item 2 should be eliminated from the Spanish version of the HISA, resulting in a 18-item questionnaire, clustered into two different factors: Initation (8 items) and No shyness / Refusal (10 items).

Finally, standard scores are provided. Results of mean scores reveal that assertiveness still follows traditional sexual scripts and gender-role stereotypes, especially among older participants. According to this, men assertively initiate sexual contacts more frequently than women (Haavio-Mannila & Kontula, 1997; López, Carcedo, Fernández-Rouco, Blázquez, & Kilani, 2011; Pierce & Hurlbert, 1999; Snell, et al., 1991) because they are supposed to initiate sexual contacts while women are supposed to act as restrictors of such contacts (Simon & Gagnon, 1984, 1986, 2003). In addition, young women scored slightly higher than young men on the No Shyness/Refusal subscale, which indicates that young women are less shy and refuse sexual contacts more often than young men. Regarding older men and women, results reveal that older women are shyer and less able to refuse undesired sexual contacts than older men. These results, although contrary to traditional sexual scripts, are consistent with some gender stereotypes, such as shyness in women (Bem, 1974; Holt & Ellis, 1998) actually show that sexual assertive skills were not traditionally taught to women (Muehlenhard & McCoy, 1991). This is particularly true in the case of Spanish women, who were taught to be "good wives" and comply with their partners' sexual desires in the past (Vázquez García & Moreno Mengíbar, 1997).

Some implications of these results must be noted. First, the factor structure found by Santos-Iglesias and Sierra (2010b) has been replicated in a sample of Spanish men and women, which is an indicator of construct validity of the scale. Second, although the factor structure is replicated the results from this study demonstrate that this scale has no measurement invariance between men and women. Thus, it is not recommended to use this scale when the purpose of the study is to compare male and female scores. In such a case, it is possible to use the Spanish version of the Sexual Assertiveness Scale (Sierra et al., 2011), whose equivalence across gender has been proven. Third, standard scores provided here are useful tools for clinicians and applied psychologists who want to assess individuals' sexual assertiveness. Finally, some limitations must be noted. For example, results are based on a non-representative sample with a large proportion of participants with high educational level, which implies that these results cannot be generalized to the entire Spanish population. Second, such results only aply to the Spanish version of the HISA, so no inferences can be made about the original English version.

References

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 47*, 155-162. doi: 10.1037/h0036215

Byrne, B. M., & van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*, 107-132. doi: 10.1080/15305051003637306

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 235-255. doi: 10.1207/S15328007SEM0902_5

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

Dunn, M., Lloyd, E. E., & Phelps, G. H. (1979). Sexual assertiveness in spinal cord injury. *Sexuality and Disability, 2*, 293-300. doi: 10.1007/BF01101395

Haavio-Mannila, E., & Kontula, O. (1997). Correlates of increased sexual satisfaction. *Archives of Sexual Behavior, 26*, 399-419. doi: 10.1023/A:1024591318836

Hidalgo, M. D., & Gómez, J. (2006). Nonuniform DIF detection using discriminant logistic analysis and multinomial logistic regression: a comparison for polytomous items. *Quality & Quantity, 40*, 805-823. doi: 10.1007/s11135-005-3964-2

Hidalgo, M. D., & Gómez-Benito, J. (2010). Education measurement: Differential item functioning. In P. Peterson, E. Baker & B. McGraw (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 36-44). Oxford, UK: Elsevier - Science & Technology.

Holt, C. L., & Ellis, J. B. (1998). Assessing the current validity of the Bem Sex-Role Inventory. *Sex Roles, 39*, 929-941. doi: 10.1023/A:1018836923919

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum.

Hurlbert, D. F. (1991). The role of assertiveness in female sexuality: a comparative study between sexually assertive and sexually nonassertive women. *Journal of Sex and Marital Therapy, 17*, 183-190. doi: 10.1080/00926239108404342

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

Jöreskog, K. G., & Sörbom, D. (2001). *LISREL 8.51*. Chicago, IL: Scientific Software International.

Kelly, J. A., St. Lawrence, J. S., Hood, H. V., & Brasfield, T. L. (1989). Behavioral intervention to reduce AIDS risk activities. *Journal of Consulting and Clinical Psychology, 57*, 60-67.

López, F., Carcedo, R., Fernández-Rouco, N., Blázquez, M.I., & Kilani, A. (2011). Diferencias sexuales en la sexualidad adolescente: afectos y conductas. *Anales de Psicología, 27*, 791-799.

Mellenbergh, G. J. (1995). Conceptual notes on models for discrete polytomous item responses. *Applied Psychological Measurement, 19*, 91-100. doi: 10.1177/014662169501900110

Meredith, W. (1993). MI, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543. doi: 10.1007/BF02294825

Morokoff, P. J., Quina, K., Harlow, L. L., Whitmire, L., Grimley, D. M., & Gibson, P. R. (1997). Sexual Assertiveness Scale (SAS) for women: development and validation. *Journal of Personality and Social Psychology, 73*, 790-804. doi: 10.1037/0022-3514.73.4.790

Muehlenhard, C. L., & McCoy, M. L. (1991). Double standard/double bind: The sexual double standard and women's communication about sex. *Psychology of Women Quarterly, 15*, 447-461. doi: 10.1111/j.1471-6402.1991.tb00420.x

Murphy, W. D., Coleman, E., Hoon, E., & Scott, C. (1980). Sexual dysfunction and treatment in alcoholic women. *Sexuality and Disability, 3*, 240-255. doi: 10.1007/BF01207674

Painter, C. (1997). *Sexual health, assertiveness and HIV*. Cambridge, UK: Daniels Publishing.

Pierce, A. P., & Hurlbert, D. F. (1999). Test-retest reliability of the Hurlbert Index of Sexual Assertiveness. *Perceptual and Motor Skills, 88*, 31-34.

Quina, K., Harlow, L. L., Morokoff, P. J., Burkholder, G. J., & Deiter, P. J. (2000). Sexual communication in relationships: when words speak louder than actions. *Sex Roles, 42*, 523-549. doi: 10.1023/A:1007043205155

Santos-Iglesias, P., & Sierra, J. C. (2010a). El papel de la asertividad sexual en la sexualidad humana: una revisión sistemática. *International Journal of Clinical and Health Psychology, 10*, 553-577.

Santos-Iglesias, P., & Sierra, J. C. (2010b). Hurlbert Index of Sexual Assertiveness: A study of psychometric properties in a Spanish sample. *Psychological Reports, 107*, 39-57. doi: 10.2466/02.03.07.17.21.PR0.107.4.39-57

Santos-Iglesias, P., Vallejo-Medina, P., & Sierra, J. C. (2009). Propiedades psicométricas de una versión breve de la Escala de Ajuste Diádico en muestras españolas. *International Journal of Clinical and Health Psychology, 9*, 501-517.

Sierra, J. C., Santos, P., Gutiérrez-Quintanilla, J. R., Gómez, P., & Maeso, M. D. (2008). Un estudio psicométrico del Hurlbert Index of Sexual Assertiveness en mujeres hispanas. *Terapia Psicológica, 26*, 117-123.

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. (2011). Propiedades psicométricas de la versión española de la Sexual Assertiveness Scale (SAS). *Anales de Psicología, 27*, 17-26.

Simon, W., & Gagnon, J. H. (1984). Sexual scripts. *Society, 22*, 53-60. doi: 10.1007/BF02701260

Simon, W., & Gagnon, J. H. (1986). Sexual scripts: permanence and change. *Archives of Sexual Behavior, 15*, 97-120. doi: 10.1007/BF01542219

Simon, W., & Gagnon, J. H. (2003). Sexual scripts: origin, influences, and change. *Qualitative Sociology, 26*, 491-497. doi: 10.1023/B:QUAS.0000005053.99846.e5

Snell, W. E., Fisher, T. D., & Miller, R. S. (1991). Development of the Sexual Awareness Questionnaire: components, reliability, and validity. *Annals of Sex Research, 4*, 65-92. doi: 10.1007/BF00850140

St. Lawrence, J. S., Brasfield, T. L., Jefferson, K. W., Alleyne, E., O'Bannon, R. E., & Shirley, A. (1995). Cognitive-behavioral intervention to reduce African-American adolescents' risk for HIV infection. *Journal of Consulting and Clinical Psychology, 63*, 221-237. doi: 10.1037/0022-006X.63.2.221

Stulhofer, A., Graham, C., Bozicevic, I., Kufrin, K., & Ajdukovic, D. (2007). An assessment of HIV/STI vulnerability and related sexual risk-taking in a nationally representative sample of young Croatian adults. *Archives of Sexual Behavior, 38*, 209-225. doi: 10.1007/s10508-007-9234-8

Vázquez García, F., & Moreno Mengíbar, A. (1997). *Sexo y razón. Una genealogía de la moral sexual en España (siglos XVI-XX)*. Madrid, Spain: Akal.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation, 12*, 1-26.

Zumbo, B. D. (1999). *A Handbook on the theory and methods of differential item functioning (DIF): Logistic Regression Modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense.