Revisiting the Ambivalent Sexism Inventory's Adolescent and Brief Versions: Problems, solutions, and considerations

Joel Juarros-Basterretxea[1],*, Nadia Yazmin Ocampo[2], Jose Luis Rojas-Solis[3], Francisco Javier Rodríguez-Díaz[4], and Eduardo García-Cueto[4]

1 Universidad de Zaragoza (Spain)
2 Universidad de Guadalajara, Guadalajara (Mexico)
3 Benemérita Universidad Autónoma de Puebla, Puebla (Mexico)
4 Universidad de Oviedo, Oviedo (Spain)

Título: Revisitando la versión para adolescentes y breve del Inventario de Sexismo Ambivalente: problemas, soluciones y consideraciones.
Resumen: El inventario de sexismo ambivalente (ASI) es frecuentemente utilizado en la investigación aplicada, lo que ha llevado a que en ocasiones se propongan versiones modificadas del mismo. La presente investigación tuvo por objetivo evaluar la validez factorial y la fiabilidad de la versión para adolescentes (ASI-A) y la versión breve (B-ASI) de este inventario en población adolescente. Se utilizaron los datos de dos muestras análogas de estudiantes de secundaria de México (estudio 1: n_1 = 975; estudio 2: n_2 = 1020). El modelo de dos dimensiones omitiendo los ítems 1 a 4 resultó ser el más recomendable en el caso del ASI-A presentando un ajuste aceptable (estudio 1). El modelo bidimensional también resultó ser el más apropiado en el caso del B-ASI al obtener un ajuste adecuado a los datos sin requerir ninguna modificación (estudio 2). El ASI-A es un instrumento de medida válido cuando algunos de los ítems son omitidos, pero la demostrada validez de la versión original de la escala incluso en población adolescente cuestiona la pertinencia teórica y empírica para desarrollar una versión específica para adolescentes.
Palabras clave: Sexismo. Inventario de sexismo ambivalente. Breve. Adolescentes. Propiedades psicométricas.

Abstract: The Ambivalent Sexism Inventory (ASI) has been widely used in applied research and sometimes modified versions have been proposed. The current research aimed to test the internal structure and reliability of the ASI for adolescents (ASI-A) and Brief ASI (B-ASI) versions in the adolescent population. Two analogue samples of Mexican secondary students (Study 1: n_1 = 975; Study 2: n_2 = 1020) composed the sample. The ASI-A showed that the two-dimension model omitting items 1 to 4 is the most recommendable model and it presents an acceptable fitting to the data (Study 1). The B-ASI bidimensional model showed an appropriate fitting to the data and no modifications were required (Study 2). The ASI-A is a valid measure of ambivalent sexism when some of its items are omitted, but the demonstrated validity of the original items in the adolescent population supports that there was no theoretical or empirical reason for developing a specific version for adolescents of the ASI.
Keywords: Sexism. Ambivalent sexism inventory. Brief. Adolescent. Psychometric properties.

Introduction

The Ambivalent Sexism Inventory (ASI) was originally developed by Glick and Fiske (1996, 1997, 2001a; Connor et al., 2016) as a corresponding measure of the sexism theory based on the ambivalent attitudes toward women. In their study, Glick and Fiske (1996) proposed a "preferred model" with two positively related factors called hostile sexism (11 items) and benevolent sexism (11 items), with three subfactors for benevolent sexism (protective paternalism [4 items], complementary gender differentiation [3 items], and heterosexual intimacy [4 items]; see Figure 1), which outperformed the other tested models (one general sexism factor and two related factors of hostile and benevolent sexism without subfactors). Since its publication in 1996, the ASI has been widely adapted to different contexts and used in applied research. For example, in their study, Glick et al. (2000) evaluated 15.000 participants from 19 nations after translating the items of the ASI into the country's language. As demonstrated in the mentioned study, the preferred model tended to be significantly better than the alternative models and the scale showed acceptable-to-good psychometric properties. Never-

theless, the ASI use in the applied research has been generally reduced to the estimation of hostile and benevolent sexism factors. This general practice has been carried out under the shelter of the systematically demonstrated existence of two related but different sexist attitudes (Glick, 2005; Glick & Fiske, 2011; Glick et al., 2000) and as a cause of two manifest reasons: (1) the interpretation of the two-dimensions of ASI instead of the better fitting of the preferred model is more parsimonious and (2) the reliability of the general benevolent sexism (11 items benevolent sexism factor should be more reliable than 3-4 items subfactors).

The extended use of the ASI in the applied research is a consequence of the strong theoretical sense and its empirical support across the world (e. g., Glick et al., 2000). As in many cases, this wide use of the ASI comes with the modification proposals of the instruments, as more brief versions which facilitated its application (e. g., Bendixen & Kennair, 2017; Bonilla-Algovia & Rivas-Rivero, 2020; Glick & Whitehead, 2010; Rodríguez-Castro et al., 2009; Rollero et al., 2014). These shorter versions of measurement instruments are usually theoretically and methodologically appropriate and thus well embraced by applied researchers because of the widely known of difficulties of long measures (e. g., lower participation or cognitive fatigue). In the particular case of the ASI, its brief versions only present the limitation of benevolent sexism subfactor: As long as the original ASI is not

* Correspondence address [Dirección para correspondencia]:
Joel Juarros Basterretxea, Universidad de Zaragoza (Spain).
E-mail: joeljuarros@unizar.es
(Article received: 29-01-2022; revised: 14-05-2022; accepted: 15-08-2022)

particularly long, the brief versions have not enough items per subfactor to calculate them without (a) oversaturating the model and (b) losing reliability. Thus, the brief versions only allow researchers to estimate two-correlated factors of hostile sexism and benevolent sexism. Nonetheless, this limitation is not especially significant considering that the use of two general factors of hostile and benevolent sexism is generalized in the applied research even when the original ASI of 22 items is used and at least one of the subfactors of the long version is still oversaturated (less than 4 items).

Similarly, adaptations of the psychological tests are common, for example, to specific samples. In this regard, the sample used for ASI development also employed late adolescents (Glick & Fiske, 2001b), it has also been modified to measure the same construct (ambivalent sexism) in adolescents. For this purpose, de Lemus et al. (2008) developed the Ambivalent Sexism Inventory for Adolescents (ASI-A). The modification of the original ASI to ASI-A was based on the presumed necessity of ambivalent sexism measure adapted to the adolescent population and thus, assuming the inappropriateness of the original ASI for this population. In the words of the authors, it was necessary to adapt the original ASI to the "adolescents' language and everyday reality's behaviors" (de Lemus et al., 2008, p. 541). Indeed, this assumption is *a priori* theoretically valid, and it could be plausible that the language of one scale validated with an adult sample would not fit the reality of the adolescents.

Based on this premise, de Lemus et al. (2008) proposed the ASI-A, which is composed of 20 items for measuring hostile sexism (10 items) and benevolent sexism's (10 items) subfactors of protective paternalism (4 items), heterosexual intimacy (3 items) and complementary gender differentiation (3 items). Of the 20 items composing the ASI-A, 15 have a direct corollary in the ASI. Eight of these 15 items have minimum changes respecting the original ASI: "women" and "men" were changed by "girls" and "boys" respectively in items 6 to 10, 12, 15, and 20 (i. e., "Girls are too easily offended" in the ASI-A and "Women are too easily offended" in the ASI). In other seven items more changes were included: the item 14 ("A good boyfriend should be willing to sacrifice things he likes in order to please his girlfriend") which have its corollary in item 20 of the ASI ("Men should be willing to sacrifice their own wellbeing in order to provide financially for the women in their lives"); the item 16 ("Girls, compared to boys, have a superior sensibility toward others' fillings") which corollary in the ASI is the item 19 ("Women, compared to men, tend to have a superior moral sensibility"); the item 17 ("Generally, girls are more intelligent than boys") which have its corollary in item 22 of ASI ("Women, as compared to men, tend to have more refined sense of culture and good taste"); the item 18 ("It is important for boys to find a girl to be romantically involved with her") comparable to the item 1 of the ASI ("No matter how accomplished he is, a man is no truly complete as a person unless he has the love of a woman"); the item 19 in the ASI-A ("Being romantically involved it is essential to reach the true

happiness in the life") is comparable to the item 6 of the ASI ("People are often truly happy in life without being romantically involved with a member of the other sex"). Finally, item 5 ("Sometimes, girls are seeking for special treatment under the guise of being girls") and item 11 ("Girls are seeking for more power than boys under the guise of equality") have the same corollary in the 2nd item of the ASI ("Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for equality) because of de Lemus et al. (2008) propose two items for the ASI-A as an equivalent of one item of ASI (see table 1 in de Lemus et al., 2008). Of the 5 items without direct corollaries in the ASI, item 13 ("Boys must protect girls") seems to be a variation of item 12 mentioned above, and the 4 remaining items are additionally added by de Lemus et al. (2008) and did not have direct corollaries on the ASI and were not similar to other items like occurs with the item 13.

Nevertheless, the main purpose when developing measure instruments must be to develop instruments as usable as possible in the intended population regardless of different characteristics (e. g., age) (APA, AERA, NCME, 2014) and only when this is not possible the researchers should try to develop specific adaptations. This kind of adaptation should be based not only on theoretical aspects but also on the empirically demonstrated lack of appropriateness of the existing instruments (Muñiz, 2018; Muñiz & Fonseca-Pedrero, 2019) and there was not any previous empirical support to do not use the original ASI in the adolescent population. On the contrary, the ASI was developed using also data from adolescents (Glick & Fiske, 2001b) and has shown good reliability in this population (see, for example, Lameiras-Fernández et al., 2001). Furthermore, the majority of the changes made in the content of the items were not significant and those which represent a significant modification have no direct corollaries in the original ASI, and thus it could be not measuring the same construct.

Considering the exposed above, the current research aimed to discuss the internal structure of two instruments for measuring ambivalent sexist attitudes toward women in comparable samples of Mexican secondary students: (1) the Ambivalent Sexism Inventory for Adolescents (ASI-A; de Lemus et al., 2008) and (2) the Brief version of the Ambivalent Sexism Inventory (B-ASI; Rodríguez-Castro et al., 2009).

Study 1: Psychometric properties of the Ambivalent Sexism Inventory for Adolescents (de Lemus et al., 2008)

Method

Participants

A total of 975 Mexican secondary students composed the sample and the mean age of the sample was 14.59 (*SD* = 1.39). The 47.3% (*n* = 461) of the participants were male and the 52.7% (*n* = 514) were females.

Instrument

Ambivalent Sexism Inventory for Adolescents (ASI-A; de Lemus et al., 2008). The ASI-A was created as an equivalent measure of the ASI (Glick & Fiske, 1996) adapted to the adolescent population (de Lemus et al., 2008). To reach this aim, the authors adapted "the items and indicators used by Glick and Fiske (1996) in the construction of the ASI to the language and behaviors of the daily reality of adolescents" (de Lemus et al., 2008, p. 541) in order to facilitate the understanding of the items. The ASI-A is composed of 20 items for measuring hostile sexism (10 items) and benevolent sexism's (10 items) subfactors of protective paternalism (4 items), heterosexual intimacy (3 items) and complementary gender differentiation (3 items). At this point, it is important to note that de Lemus et al. (2008) indicate that the complementary gender differentiation is composed of items 16 and 17, corresponding to the fifth component shown in the table 3 of their article and where the item 15 is also included in the mentioned component. Thus, we assume include the three items assuming that the text has an erratum when not mentioning the 15 items as part of the complementary gender differentiation dimension. The items response scale rated from 0 (totally disagree) to 5 (totally agree) in a six-point Likert scale.

Procedure

First, the questionnaire's language was reviewed to check if any cultural adaptation was needed to guarantee that no systematic bias derived from the language affect the results due to the differences between Spanish from Spain and Spanish from Mexico. Spanish and Mexican expert researchers carried out this preliminary evaluation and no modifications were needed. Second, researchers contacted directive teams of different educational centers to explain to them the aim of the research and to obtain authorization to scale implementation during the class. After the study was authorized and participants signed the informed consent, the scale was applied to the students by research team members in the paper-and-pencil format during the scholar hours in the classroom. The participants completed the self-reported questionnaire individually under the supervision of the researchers.

Data analysis

A series of confirmatory factor analysis were carried out in order to the fitting of the three models classically tested in the literature regarding the ASI (one-dimension, two-dimensions, and the preferred model). The fitting was established by the χ^2 and its associated probability, the *Comparative Fit Index* (CFI), and the *Root Mean Square Error Approximation* (RMSEA) and its 90% confidence interval (CI). CFI values $\geq .90$ were considered acceptable and $\geq .95$ good fitting, and RMSEA values $\leq .08$ were considered acceptable and $\leq .05$ good. In those cases, where RMSEA was between .05 and .08, the confidence interval was analyzed. Considering the categorical nature of the items (Likert scale) the *Weighted Least Square Mean and Variance Adjusted* (WLSMV) estimator was used. A theoretical analysis of the items' meaning and content based on the original theoretical model was also carried out as complementary to the empirical analysis of the model.

Results

Using confirmatory factor analysis, the three models tested by de Lemus et al. (2008) were tested. The first model, testing a one general sexism-dimension model (χ^2 = 4542.749, *p* $\leq$ .001, df = 170; CFI = .694, RMSEA = .161, 90%CI [.158, .167]) and the second model testing a two correlated dimension (hostile and benevolent sexism) model (χ^2 = 2148.323, *p* $\leq$ .001, df = 169; CFI = .861, RMSEA = .110, 90%CI [.105, .114]) showed poor fit to the data. Third, the model proposed by de Lemus et al. (2008) emulating the original model of Glick & Fiske (1996) was tested; the fitting of the model improved but it still was poor (χ^2 = 1565.979, *p* $\leq$ .001, df = 166; CFI = .902, RMSEA = .093, 90%CI [.089, .097]).

To determine the potential causes of the poor fitting of the model to the data, an analysis of the item's content based on the original theoretical model and empirical previous results obtained by de Lemus et al. (2008) was made. In their table 3, de Lemus et al. (2008) displayed the results got by exploratory factor analyses which reveal, contrary to their interpretation, that items 1 to 10 (relative to hostile sexism) are not encompassed in the same factor, but in two different factors. Based on the results displayed by de Lemus et al. (2008) there is no empirical reason to encompass items 1 to 10 in one dimension (hostile sexism) which can be the reason for the poor fitting of the model to the data. Additionally, the theoretical analysis of the items revealed that items 5 to 10 had direct corollaries in the original ASI, but items 1 to 4 no. In-depth analysis of the items and the empirical results obtained by de Lemus et al. (2008) suggest that there are measuring a construct related to hostile sexism, but different

from it, making this four-item factor theoretically incongruent.

Following the empirical and theoretical evidence, an alternative two-dimension (Figure 1) and the preferred model (Figure 2) for adolescents omitting 1 to 4 items were tested. Both models' fitting to the data improved significantly when items 1 to 4 were omitted. Nevertheless, the two-dimension model fitting was still poor ($\chi^2 = 1282.260$, $p \leq .001$, df = 103; CFI = .902, RMSEA = .108, 90%CI [.103, .114]), while the preferred model fitted acceptable ($\chi^2 = 722.97$, $p \leq .001$, df = 100; CFI = .948, RMSEA = .080, 90%CI [.074, .085]). The analysis of the modification indices revealed that releasing the covariations between item 11, 12 and 13 error terms, which were constrained to 0 by default, could significantly improve the model fitting to the data.

Figure 1
*Standardized parameter estimates for the corrected two-dimension model of ASI-A.*
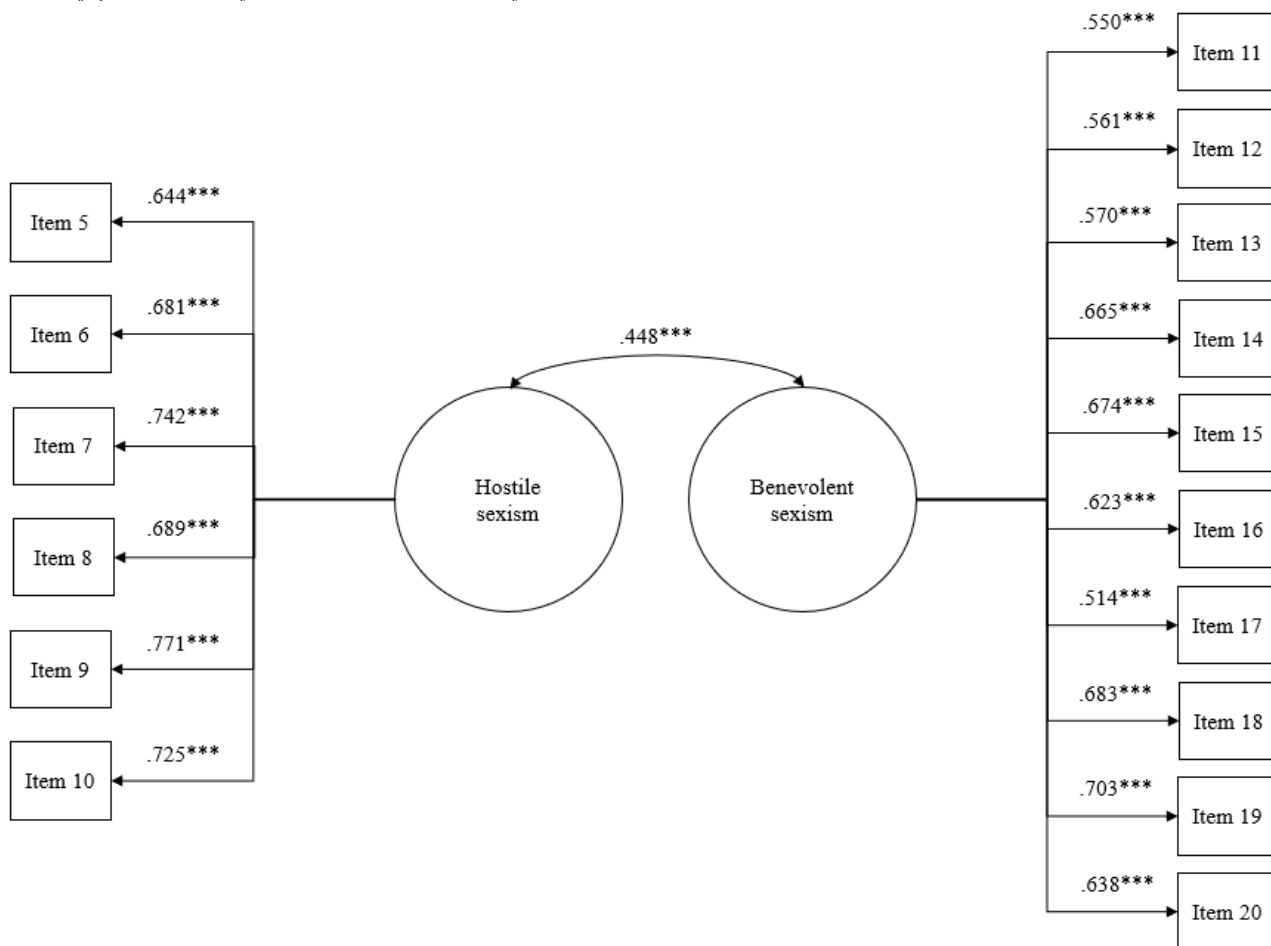
***$p \leq .001$
Note. Covariances between 11, 12 and 13 items are omitted in the figure.

Once these covariations were freely estimated the fitting of both models, two-dimension model ($\chi^2 = 609.781$, $p \leq .001$, df = 100; CFI = .958, RMSEA = .072, 90%CI [.067, .078]) and the preferred model ($\chi^2 = 477.969$, $p \leq .001$, df = 97; CFI = .968, RMSEA = .063, 90%CI [.058, .069]) was acceptable. The one-dimension model was not tested considering the so-demonstrated lack of fitting the literature as well as in previous results obtained in the current research.

Figure 2
Standardized parameter estimates for the corrected preferred model of ASI-A.
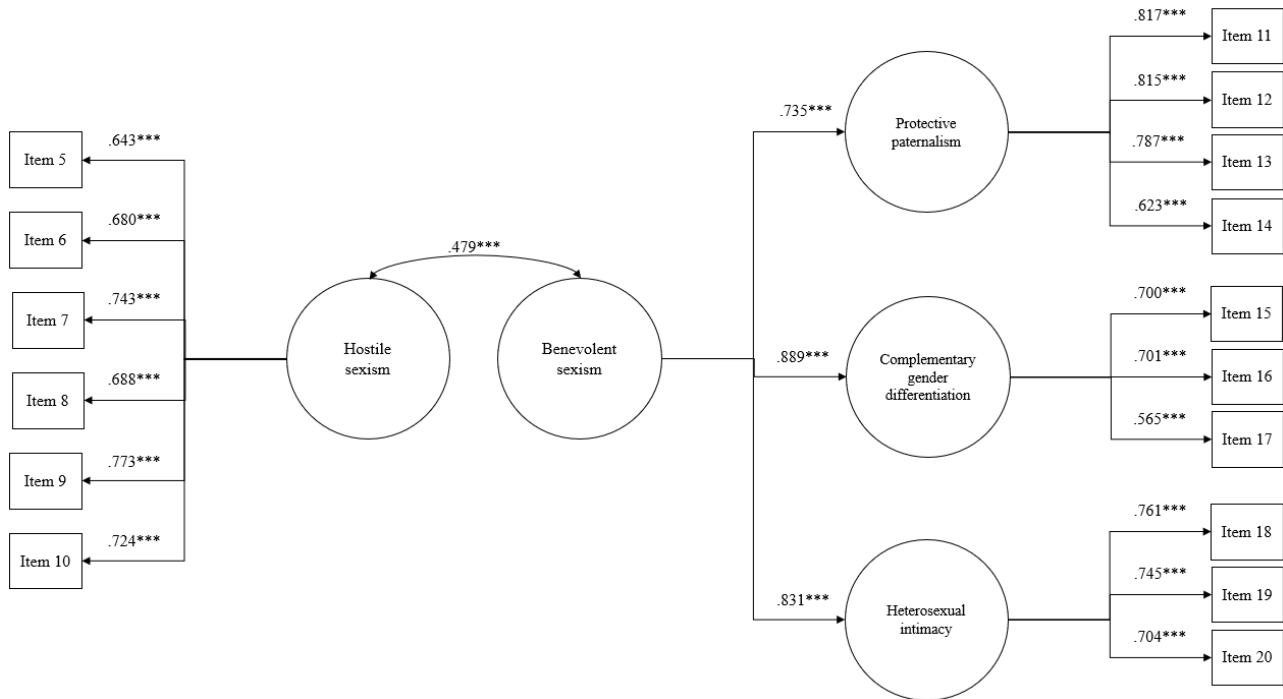
***p ≤ .001
Note. Covariances between 11, 12 and 13 items are omitted in the figure.

The factor reliabilities for the two-dimension model and the preferred model are displayed in table 1. The hostile sexism (ω = .86) showed good reliability. Regarding the benevolent sexism, the reliability was good for the two-dimension model (ω = .86) and acceptable for the preferred model (ω = .75). Nonetheless, the benevolent sexism subfactors of protective paternalism (ω = .69), complementary gender differentiation (ω = .22) and heterosexual intimacy (ω = .48) showed low reliability.

Table 1
Reliability for ASI factor in the two-dimension model and the preferred model.

	ω	
	Two-dimension model	Preferred model
Hostile sexism	.86	.86
Benevolent sexism	.86	.75
Protective paternalism	-	.69
Complementary gender differentiation	-	.22
Heterosexual intimacy	-	.48

Discussion

Using the data of 975 Mexican secondary students, study 1 aimed to analyze the internal structure of the Ambivalent Sexism Inventory for Adolescents (de Lemus et al., 2008). The tested models with the original ASI-A generally show poor fitting to the data. First, the same models tested by de Lemus et al. (2008) were tested and congruent results were found. As in their study, the model with two related factors of hostile and benevolent sexism and three sub-factors of benevolent sexism over-performed the unidimensional model and the two related general factors model. Nevertheless, the fitting of the models still was poor.

In-depth analysis of the items showed empirical and theoretical incongruences: In the regard to empirical problems, the results got by de Lemus et al. (2008) suggest two factors or two subfactors of hostile sexism (components 1 and 3 in de Lemus et al. [2008]), but not a general hostile sexism factor. It could be argued that the estimation of the unique dimension of hostile sexism is based on theoretical aspects which make the adolescents model congruent to the preferred model of the original ASI (Glick & Fiske, 1996) and thus based on theoretical aspects more than in empirical results. Nevertheless, the analysis of the items also revealed theoretical problems: Items 5 to 10, relative to the hostile sexism (component 1), are measuring hostile sexism, but it was not clear in the case of items 1 to 4 (component 3). Even the ASI-A is supposedly created based on the ASI and as an adaptation for adolescents of that, items 1 to 4 seem not to be in line with the original ASI. As in a private conversation via email with one of the original ASI authors Peter Glick (personal communication, August 7, 2020) was concluded, these four items are not equivalent to the original scale's items and seem to be a mix of "belief in traditional gender roles and that boys should dominate girls" (P. Glick, personal communication, August 7, 2020). In this regard,

items 1 to 4 did not fit with the theoretical basis of the hostile sexism, which in his original version tapped on the notion that men and women are locked into a competitive struggle with women trying to exert control over men (P. Glick, personal communication, August 7, 2020), and thus these four items proposed by de Lemus et al. (2008) as part of the hostile sexism factor seem to be assessing "something" related to hostile sexism, but different from it.

Following this reasoning, items 1 to 4 were excluded and models estimated again excepting the one-dimension model due to its demonstrated lack of appropriateness. The results obtained a significant improvement of the model fitting to the data and are congruent to the idea about there was not an empirical reason to consider only one general dimension of hostile sexism as made by de Lemus et al. (2008), but it is reasonable to omit items 1 to 4 to measure hostile sexism appropriately.

At this point, it is necessary to make some considerations beyond the fitting of the model. According to the results get in the current research, the preferred model for adolescents, which is equivalent to the preferred model (Glick & Fiske, 1996) but for the ASI-A (de Lemus et al., 2008), is the model with the best fitting to the data. Nevertheless, it is important to note the potential limitations of this model and the possibility that fitting could be a statistical artefact. Specifically, it is not recommended to estimate factors with less than four items because of the tendency to over-saturate the model. As seen (Figure 2), the preferred model presents a benevolent sexism second-order factor with three subfactors or first-order factors: one of four items and two of three items each. Additionally, it represents an important limitation in terms of reliability; as seen, the reliabilities of the three subfactors are low, and thus, these are not reliable measures of the construct. On the contrary, the benevolent sexism factor showed acceptable reliability, but it also improves when for the two-dimension model in comparison with the preferred model, suggesting again the appropriateness of the more parsimonious model. Finally, it is important to note that the models showed acceptable fitting to the data (good CFI and acceptable RMSEA) after consideration of covariances between items 11, 12 and 13 error terms. These covariations indicate common error sources between the items which are necessary to control better estimation of the factor. This common error sources are attributable to the similar meaning of the three items which refer to the protective role of men regarding women (see Appendix 1).

Conclusion

Considering all the results available and taking into account all the theoretical and practical implications, the two-dimension model omitting items 1 to 4 seem to be the most recommendable model to use in applied research. This model presents an acceptable fitting to the data and is significantly better than proposed in previous versions (de Lemus et al., 2008) with the added value of being more parsimonious than the preferred model and an appropriate number of items per factor which allow researchers to make more reliable measures of the benevolent sexism comparing with the three sub-factor structure.

Study 2: Psychometric properties of the Brief Ambivalent Sexism Inventory (Rodríguez-Castro et al., 2009)

Method

Participants

A total of 1020 Mexican secondary students composed the sample and the mean age of the sample was 16.56 ($SD = $ 1.39). The 62.7% ($n = 640$) of the participants were male and the 37.3% ($n = 380$) were female.

Instrument

Brief version of Ambivalent Sexism Inventory (B-ASI; Rodríguez-Castro et al., 2009). The B-ASI is the brief version of the ASI (Glick & Fiske, 1996) and is composed of 12 items for measuring hostile sexism (6 items) and benevolent sexism (6 items). Due to the brief nature of the instrument, the first-order factors of benevolent sexism protective paternalism, heterosexual intimacy and complementary gender differentiation cannot be estimated due to each of them would be composed of two items. The items' response scale rated from 0 (totally disagree) to 5 (totally agree) on a six-point Likert scale.

Procedure

The same as in study 1.

Data analysis

The same as in study 1.

Results

Using confirmatory factor analysis, the sexism one-dimension model and the two-dimension (hostile and benevolent) models were tested in the study 2. The first model, testing a one general sexism-dimension model ($\chi^2 = $ 2297.857, $p \leq .001$, df = 54; CFI = .736, RMSEA = .202, 90%CI [.195, .209]) and the second model testing a two correlated hostile and benevolent sexism dimensions model ($\chi^2 = 602.292$, $p \leq .001$, df = 53; CFI = .935, RMSEA = .101, 90%CI [.094, .108]) showed poor fit to the data.

Despite the model one was far from acceptable fitting, the analysis of the modification indices of the bidimensional model revealed that freely estimating the covariances between items 11-12 and 5-6 would significantly improve the

model fitting to the data. After including these covariations in the model (Figure 3), a good fitting to the data was observed basing on CFI, but still only acceptable if RMSEA

was considered ($\chi^2 = 320.403$, $p \leq .001$, df = 51; CFI = .968, RMSEA = .072, 90%CI [.065, .080]).

Figure 3
Standardized parameter estimates for the preferred model adapted to B-ASI.
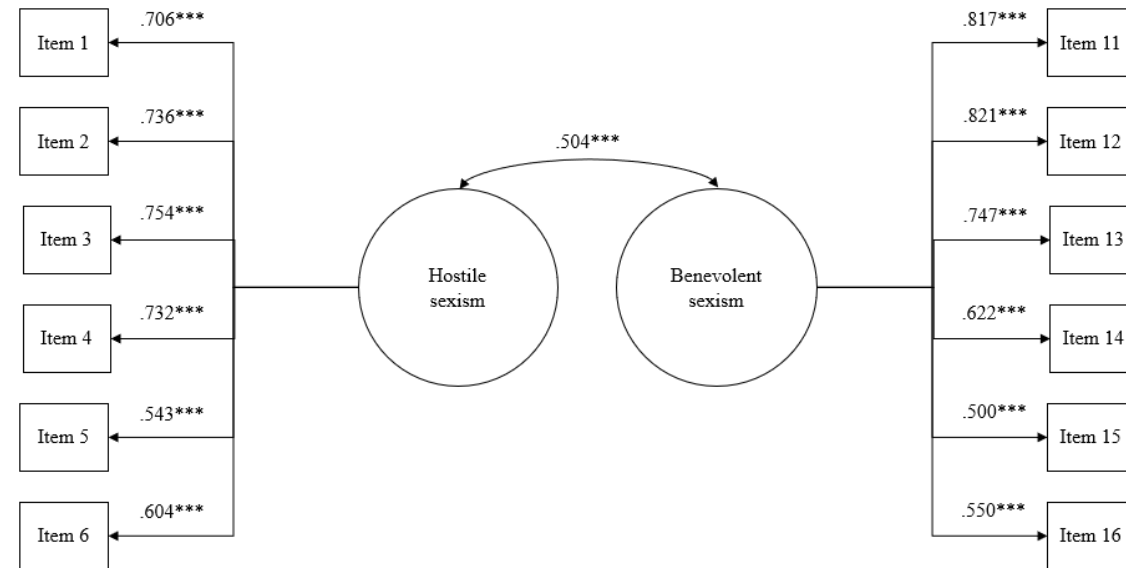
***$p \leq .001$
Note. Covariances between 11-12 and 5-6 items are omitted in the figure.

Regarding the reliability, both factors of hostile sexism (ω = .84) and benevolent sexism (ω = .84) showed good reliability for the two-dimension model of the B-ASI.

Discussion

Using the data of 1020 Mexican secondary students, the internal structure of the Brief-Ambivalent Sexism Inventory (Rodríguez-Castro et al., 2009) was tested. As in previous research, the two-dimension model over-performed the one-dimension model.

Despite the preferred model tend to be the better model (Glick & Fiske, 1996; Glick et al., 2000), it could not be estimated using the B-ASI due to the low number of items per each first-order factor of benevolent sexism. The second first-order factors of benevolent sexism (second-order factor) presented in the preferred model are theoretically identifiable in the B-ASI because benevolent sexism is composed of six items: two of protective paternalism, two of heterosexual intimacy, and two of complementary gender differentiation. Nevertheless, it is important to remember that it is recommendable to use factors composed of four or more items: first, the model could show good fitting to the data with factors composed of less than four items, but it is attributable to the oversaturation of the model and thus to a statistical artefact. Second, as mentioned in the previous study, factors with a low number of items present reliability problems.

Anyway, the two-dimension model is thoroughly used (probably the most used) in applied research to analyze the ASI results and good internal structure has been found in other brief versions and contexts (Bendixen & Kennair, 2017; Bonilla-Algovia & Rivas-Rivero, 2020; Rollero et al., 2014). The massive use of the presumably factorially worse model can be explained by two aspects: First, some researchers have found appropriate fitting to the data of this model or, at least, similar to the preferred model (Bendixen & Kennair, 2017; Bonilla-Algovia & Rivas-Rivero, 2020; Glick et al., 2000; Rollero et al., 2014). Second, this model is much more parsimonious than the preferred model, which makes it easier to calculate and interpret. It could be argued that the use of the two-dimension model implicates a loss of information and lack of representativity of the original theory prosed by Glick and Fiske (1996) because of the omission of benevolent sexism subfactors (protective paternalism, complementary gender differentiation, and heterosexual intimacy), but it is also true that B-ASI permits a very brief reliable and valid measure of the ambivalent sexism toward women in the same way that the ASI is used. Finally, it is important to note that the model showed appropriate fitting to the data (good CFI and acceptable RMSEA) after consideration of covariances between items 5-6 and 11-12 error terms. These covariations indicate common error sources between the items which are necessary to control better estimation of the factor. This common error sources are attributable to the similarity of the items (see appendix 2).

Conclusion

Considering the results obtained in this study, the B-ASI is an appropriate instrument for the ambivalent sexism toward women measure in the adolescent population and future research should evaluate the invariance of the measure in adolescents and adults to confirm if it is also an equivalent measure. Considering the benefit of using the same instrument in different populations and that the B-ASI is identical to the original ASI excepting the number of items, it would be recommendable to use it in the adolescent population as well as in the adult population.

General discussion

The question that arises here is why researchers often prefer to create and validate relatively new instruments rather than test or adapt existing reliable and valid tests which measure the same construct? Scale construction is, a *priori*, arduous, hard, and demanding work which can be a reason for trying to use existing valid measurement instruments, but it is also known that scale-development papers can be profitable for a researcher with moderate-to-good methodological knowledge.

Basing on the results of the current research, the appropriateness to develop new versions of the same scale could be questioned. Using the Ambivalent Sexism Inventory, which is probably the most widely used scale for sexism measure, the correctness of developing the ASI modification for adolescents was analyzed in the current research basing on two comparable samples of Mexican secondary students who complete the Ambivalent Sexism Inventory for Adolescents (sample 1) and the Brief-Ambivalent Sexism Inventory (sample 2). Following the results obtained in the current research, the answer to the fundamental question "it was really necessary to develop the Ambivalent Sexism Inventory for Adolescents?" seems to be clear: No.

Among the different aspects of test development, the universality of them is one of the most important. Universal designs of measurement instruments imply that the measurement is precise and that the scores on the measured construct are not affected or can be differentiated for the irrelevant characteristics for the construct and do not interfere with the ability of participants to respond (APA, AERA, NCME, 2014). Following this approach, the modified versions of one instrument to adapt it to the age, as in this case, should be justified in empirical demonstration of the inappropriateness of the original scale in the intended population. On the contrary, the results obtained here indicate as in previous research (e. g., de Lemus et al., 2008) do not support the preference for the adaptation for adolescents. First, there was no theoretical nor empirical evidence of a violation of the ASI's universality among adolescents and adults which justified the modification of the ASI to the adolescent population. Second, the B-ASI, which is a brief version of the original ASI and did not change the items, is a valid instru-

ment even in the adolescent population. Third, even when an acceptable model of ASI-A has been found in the current research, it does not permit researchers to compare adults and adolescents as could be made using the B-ASI. In this regard, using the original version of the questionnaire allow researchers to make analogue inferences without the inherent problems of accommodated instruments (APA, AERA, NCME, 2014).

These results are only one of the different examples that researchers could find in the scientific literature of the unnecessary development of "new" measures. Probably the most relevant consequence of this over-abundance of scales is the promotion of confusion and the loss of information. For example, some researchers can use the traditional ASI or its brief version considering the demonstration of the applicability of ASI to the adolescent population and its advantages: it allows researchers to test invariance and make comparisons between them and adults, and it also permits longitudinal research of sexist attitudes which can be questionable with different measures. On the contrary, other researchers could use the ASI-A as originally proposed and consider that their results are comparable to those obtained with the ASI or the B-ASI but it is not true considering that they are technically different instruments. Furthermore, in the way to creating an ASI version for adolescents the inclusion of new items derived from the bias measure because they were not congruent with the original construct and only the items which had direct corollaries in the original ASI were well defined. In this regard, it could be argued that in the ASI for adolescents the good things were not new, and the new things were not good, so it would be better to test the appropriateness of the original ASI in the adolescent population before developing a new version of the instrument.

Contemporary researchers in social, educational and health sciences should consider the pros and cons of creating new scales when existing measures are potentially usable, but also should test the existing ones' psychometric properties, for example, in the samples of interest (e. g. adolescents, inmates, clinical population) before proposing and developing adaptations to these populations (APA, AERA, NCME, 2014). Nevertheless, there is important to note that there are situations where the necessity to adapt a scale's language, as in the current case, is reasonable even if the psychometric properties are not tested before. For example, Hammond and Cimpian (2020) aimed to study the ambivalent attitude toward women in 5-to-11 years old children who objectively needed a language adaptation.

References

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. American Psychological Association

Bendixen, M., & Kennair, L. E. O. (2017). When less is more: Psychometric properties of Norwegian short-forms of the Ambivalent Sexism Scales (ASI and AMI) and the Illinois Rape Myth Acceptance (IRMA) Scale. *Scandinavian Journal of Psychology, 58*, 541 – 550. https://doi.org/10.1111/sjop.12392

Bonilla-Algovia, E., & Rivas-Rivero, E. (2020). Validez de la versión reducida de la escala de sexismo ambivalente en El Salvador y México [Validity of th reduced versión of the ambivalent sexism invetory in El Salvador and Mexico]. *Revista Mexicana de Psicología, 37*(2), 92-103.

Connor, R. A., Glick, P., & Fiske, S. T. (2016). Ambivalent sexism in the 21st century. In C. Sibley & F. Barlow (Eds.), *Cambridge Handbook of the Psychology of Prejudice* (Cambridge Handbooks in Psychology, pp. 295 – 320). Cambridge University Press. https://doi.org/10.1017/9781316161579.013

de Lemus, S., Castillo, M., Moya, M., Padilla, J. L., & Ryan, E. (2008). Elaboración y validación del Inventario de Sexismo Ambivalente para adolescentes [Creation and validation of the Ambivalent Sexism Inventory for adolescents]. *International Journal of Clinical and Health Psychology, 8*, 537-562. https://www.redalyc.org/pdf/337/33712001013.pdf

Glick, P. (2005). Ambivalent sexism, power distance, and gender inequality across cultures. In S. Guimond (Ed.), *Social Comparison and Social Psychology: Understanding Cognition, Intergroup Relations, and Culture* (pp. 283 – 302). Cambridge University Press. https://doi.org/10.1017/CBO9780511584329.015

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*, 491 – 512. https://doi.org/10.1037/0022-3514.70.3.491

Glick, P., & Fiske, S. T. (1997). Hostile and benevolent sexism: Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly, 21*(1), 119 – 135. https://doi.org/10.1111/j.1471-6402.1997.tb00104.x

Glick, P., & Fiske, S. T. (2001a). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*(2), 109 – 118. https://doi.org/10.1037/0003-066X.56.2.109

Glick, P., & Fiske, S. T. (2001b). Ambivalent sexism. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 115–188). Academic Press.

Glick, P., & Fiske, S. T. (2011). Ambivalent sexism revisited. *Psychology of women quarterly, 35*(3), 530 – 535. https://doi.org/10.1177/0361684311414832

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., Adetoun, B., Osagie, J. E., Akande, A., Alao, A., Brunner, A., Willemsen, T. M., Chipeta, K., Dardenne, B., Dijksterhuis, A., Wigboldus, D., Eckes, T., Six-Materna, I., Expósito, F., Moya, M., Foddy, M., Kim, H. J., Lameiras, M., Sotelo, M. J., Mucchi-Faina, A., Romani, M., Sakalli, N., Udegbe, B., Yamamoto, M., Ui, M., Ferreira, M. C., & López-López, W. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*, 763 – 775. https://doi.org/10.1037//0022-3514.79.5.763

Glick, P., & Whitehead, J. (2010). Hostility toward men and the perceived stability of male dominance. *Social Psychology, 41*, 177 – 185. https://doi.org/10.1027/1864-9335/a000025

Hammond, M., & Cimpian, A. (2020). "Wonderful but weak": Children's ambivalent attitudes toward women. *Sex Roles, 84*, 76 – 90. https://doi.org/10.1007/s11199-020-01150-0

Holtz, P., Deutschmann, E., & Dobewall, H. (2017). Cross-cultural psychology and the rise of academic capitalism: Linguistic changes in CCR and JCCP articles, 10970-2014. *Journal of Cross-Cultural Psychology, 48*, 1410 – 1431. https://doi.org/10.1177/0022022117724902

Ibabe, I., Arnoso, A., & Elorriaga, E. (2016). Ambivalent Sexism Inventory: Adaptation to Basque population and sexism as a risk factor of dating violence. *The Spanish Journal of Psychology, 19*, 1 – 9. https://doi.org/10.1017/sjp.2016.80

Lameiras-Fernández, M., Rodríguez-Castro, Y., & Sotelo-Torrejón, J. (2001). Sexism and racism in a Spanish sample of secondary school students. *Social Indicators Research, 54*, 309 – 328. https://doi.org/10.1023/A:1010871706454

Muñiz, J. (2018). *Introducción a la psicometría*. Pirámide.

Muñiz, J., & Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test [Ten steps for test construction]. *Psicothema, 31*, 7 – 16. https://doi.org/10.7334/psicothema2018.291

Rodríguez-Castro, Y., Lameiras-Fernández, M., & Carrera-Fernández, M. V. (2009). Validación de la versión reducida de las escalas ASI y AMI en una muestra de estudiantes españoles [Validation of scales shortened versión of ASI and AMI in a sample of spanish]. *Psicogente, 12*, 284 – 295. https://dialnet.unirioja.es/servlet/articulo?codigo=3265018

Rollero, C., Glick, P. & Tartaglia, S. (2014). Psychometric properties of short versions of the ambivalent sexism inventory and ambivalence toward men inventory. *TPM – Testing, Psychometrics, Methodology in Applied Psychology, 21*, 149 – 159. https://doi.org/10.4473/TPM21.2.3

Appendix I.
Items of the ambivalent sexism inventory for adolescents. Final version.

	Totalmente en desacuerdo	Bastante en desacuerdo	Un poco el desacuerdo	Un poco de acuerdo	Bastante de acuerdo	Totalmente de acuerdo
1. A veces las chicas utilizan lo de ser chicas para que las traten de manera especial.						
2. Cuando las chicas son vencidas por losmchicos en una competición justa, generalmente, ellas se quejan de haber sido discriminadas.						
3. Las chicas se ofenden muy fácilmente.						
4. Las chicas suelen interpretar comentarios inocentes como sexistas.						
5. Las chicas suelen exagerar sus problemas.						
6. Las chicas con la excusa de la igualdad pretenden tener más poder que los chicos.						
7. Por las noches los chicos deben acompañar a las chicas ahasta su casa para que no les ocurra nada malo.						
8. Las chicas deben ser queridas y protegidas por los chicos.						

	Totalmente en desacuerdo	Bastante en desacuerdo	Un poco el desacuerdo	Un poco de acuerdo	Bastante de acuerdo	Totalmente de acuerdo
9. Los chicos deben cuidas a las chicas.						
10. Un buen novio debe estar dispuesto a sacrificar cosas que le gustan para agradar a su chica.						
11. En caso de una catástrofe las chicas deben ser salvadas antes que los chicos.						
12. las chicas tienen una mayor sensibilidad hacia los sentimientos de los demás que los chicos.						
13. Las chicas en general son más inteligentes que los chicos.						
14. Para los chicos es importante encontrar a una chica con quien salir.						
15. Las relaciones de pareja son esenciales para alcanzar la verdadera felicidad en la vida.						
16. Un chico puede sentirse incompleto si no sale con una chica.						

Note: The numeration is maintained as in the original article of De Lemus et al. (2008).

Appendix II
Items of the brief ambivalent sexism inventory. Final version.

ÍTEMS	Totalmente en desacuerdo	Bastante en desacuerdo	Un poco en desacuerdo	Un poco de acuerdo	Bastante de acuerdo	Totalmente de acuerdo
1. Las mujeres intentan ganar poder controlando a los hombres.						
2. Cuando las mujeres son vencidas por los hombres en una competencia justa, generalmente ellas se quejan de haber sido discriminadas.						
3. Una vez que una mujer logra que un hombre se comprometa con ella, por lo general intenta controlarle estrechamente.						
4. Las mujeres exageran los problemas que tienen en el trabajo.						
5. Las mujeres feministas están haciendo demandas completamente irracionales a los hombres.						
6. Existen muchas mujeres que para burlarse de los hombres, primero se insinúan sexualmente a ellos y luego rechazan los avances de estos.						
7. Todo hombre debe tener una mujer a quien amar.						
8. El hombre está incompleto sin la mujer.						
9. Los hombres deberían de estar dispuestos a sacrificar su propio bienestar con el fin de proveer seguridad económica a las mujeres.						
10. Las mujeres deben ser queridas y protegidas por los hombres.						
11. Las mujeres en comparación con los hombres, tienden a tener una mayor sensibilidad moral.						
12. Muchas mujeres se caracterizan por una pureza que pocos hombres poseen.						