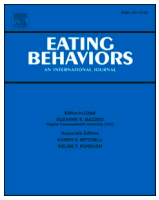
Eating Behaviors 52 (2024) 101827

Contents lists available at ScienceDirect

Eating Behaviors

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



Factor structure and measurement invariance of the Weight-Related Abuse Questionnaire (WRAQ)

Alejandro Magallares [a], Patricia Recio [b,*], Ignacio Jáuregui-Lobera [c], Pilar Benito de Valle [d], José Antonio Irles [d], Genna Hymowitz [e]

[a] School of Psychology, Social Psychology Department, Spanish Open University (UNED), Madrid, Spain
[b] School of Psychology, Methodology Department, Spanish Open University (UNED), Madrid, Spain
[c] School of Experimental Sciences, Nutrition and Bromatology, Universidad Pablo de Olavide, Sevilla, Spain
[d] Clinical Nutrition Unit, Hospital de Valme, Sevilla, Spain
[e] Bariatric and Weight Loss Psychology, Psychiatry and Behavioral Health, Stony Brook University, New York, United States

ARTICLE INFO

Keywords:
Factor structure
Measurement invariance
Obesity
Physical and verbal abuse
Weight-related abuse questionnaire

ABSTRACT

Weight-related abuse is defined as verbal or physical maltreatment specific to one's weight. The Weight-Related Abuse Questionnaire (WRAQ) is an instrument specifically designed to measure weight-related abuse. The main goal of this research was to study the factor structure and measurement invariance of the Spanish version of the WRAQ in a non-clinical and a clinical sample. The clinical sample included 150 participants with obesity (60 % women) from the (masked for blind review) Hospital. The non-clinical sample included 301 students (79 % women) from the (masked for blind review) University. Scales to measure weight self-stigma and fear of gaining weight were used to analyze the convergent validity of the WRAQ. A confirmatory factor analysis showed that a two-factor model (verbal and physical abuse) was an acceptable fit for the data in both the clinical and non-clinical samples. Multigroup Confirmatory Factor Analysis revealed scalar measurement invariance by sample and gender. Cronbach's alpha coefficients and composite reliability for both samples were found to be good, with values ranging from 0.83 to 0.96. Fear of gaining weight was correlated to verbal ($r = 0.36$, $p < .01$) and physical ($r = 0.12$, $p < .05$) abuse, and weight self-stigma was also related to physical ($r = 0.21$, $p < .01$) and verbal ($r = 0.41$, $p < .01$) abuse. These results suggest that the WRAQ can be used in clinical and non-clinical samples to assess verbal and physical abuse in both men and women.

1. Introduction

Weight-related abuse (WRA) may be defined as a "significant verbal or physical victimization or maltreatment specific to one's weight" (Salwen & Hymowitz, 2015, pp. 150–151). Previous studies identified a relationship between WRA and binge eating, emotional eating, night eating, and unhealthy weight control behaviors, and suggest that WRA may play a unique role in the development of disordered eating (Salwen et al., 2015). Research has also indicated associations between WRA and weight self-stigma, which implies that WRA may be related to personal experiences of shame, negative self-evaluations, and perceived discrimination (Castillo-Luna & Jauregui-Lobera, 2017). Thus, improving our understanding of WRA can help us better identify individuals at risk for disordered eating and weight self-stigma and assist with the refinement of interventions and prevention programs.

The Weight-Related Abuse Questionnaire (WRAQ; Salwen & Hymowitz, 2015) is a 15-item questionnaire, originally developed in the United States with a clinical sample of patients with obesity and a college sample, that retrospectively assesses the average frequency of negative weight-related events, including an 8-item verbal abuse (VA) subscale and 7-item physical abuse (PA) subscale. However, in the WRAQ development study (Salwen & Hymowitz, 2015), the authors distinguished between mild VA (6 items) and severe VA (2 items), and between mild PA (5 items) and severe PA (2 items), suggesting a four-factor structure, instead of the two-factor model initially proposed.

The WRAQ has been used in Egypt (Ahmed Mokbel et al., 2020) and Pakistan (Munir & Dawood, 2021). While these studies reported data on the PA and VA subscales, they did not include factor analysis (neither

* Corresponding author at: Departamento de Metodología de las Ciencias del Comportamiento, Facultad de Psicología UNED, C/ Juan del Rosal, 10, 28040 Madrid, Spain.
E-mail address: reciop@psi.uned.es (P. Recio).

https://doi.org/10.1016/j.eatbeh.2023.101827
Received 22 May 2023; Received in revised form 26 September 2023; Accepted 16 November 2023
Available online 22 November 2023

exploratory nor confirmatory) to determine the underlying structure of the questionnaire. Confirmatory factor analysis is an appropriate procedure for establishing the factor structure of a scale when there is a plausible hypothesis about the model (Bollen, 1989). Once the factor structure is established, factorial invariance (or multiple-group invariance) becomes essential for ensuring valid score comparisons between different groups. Therefore, these studies underscore the need for additional psychometric work, including CFA to study the scale's dimensionality and an examination of invariance, to ensure meaningful cross-group comparisons of the WRAQ.

Therefore, the main goal of the current research was to compare a two-factor model of the WRAQ that distinguished between PA and VA, and a four-factor model, that differentiated between mild and severe abuse in both PA and VA (Salwen & Hymowitz, 2015) in samples of individuals in Spain. In addition, as previous studies suggested higher WRA endorsement in clinical samples (Salwen & Hymowitz, 2015) we aimed to analyze the questionnaire's invariance to demonstrate that the WRAQ's structure was equivalent and can be used in both clinical (patients with obesity) and non-clinical (college students) samples. Furthermore, given previous research suggesting possible gender differences in weight-based stigma (Spahlholz et al., 2016) we wanted to also test gender invariance of the WRAQ. Finally, to study the construct validity of the WRAQ, weight self-stigma (Castillo-Luna & Jaureguí-Lobera, 2017) was measured for the clinical sample, and fear of gaining weight (Salwen & Hymowitz, 2015) was assessed for the non-clinical sample.

2. Method

2.1. Samples

The clinical sample comprised 150 participants (60 % women) from the (masked for blind review) Spanish Hospital with a mean age of 44.04 years ($SD = 13.58$). The mean Body Mass Index (BMI) was 43.18 kg/m^2 ($SD = 8.30$) with a range between 30.10 and 68.35.

The non-clinical sample was comprised of 301 participants (79 % women) from the (masked for blind review) Spanish University with a mean age of 28.04 years ($SD = 6.47$). The mean BMI was 22.95 kg/m^2 ($SD = 4.01$) with a range between 15.42 and 35.16. For more information about the recruitment procedure in both samples, please see Appendix A.

2.2. Instruments

To measure WRA, the Spanish version of the WRAQ (Salwen & Hymowitz, 2015) was used in both the clinical and non-clinical samples (see Appendix B for more information about the adaptation procedure and the items of the Spanish version of the WRAQ). The WRAQ is a 15-item self-report measure that retrospectively explores the average frequency of negative experiences related to one's weight that occurred prior to the age of 21. It has two subscales: VA (8 items. E.g., "Someone called you names because of your weight") and PA (7 items. E.g., "Someone grabbed you because of your weight"). These statements were rated on a 7-point Likert scale ranging from 0 (never) to 6 (more than 20 times per year). Cronbach's alpha coefficients from three separate samples included in the original psychometric paper (Salwen & Hymowitz, 2015) were found to be good, with values ranging from 0.87 to 0.93 for VA, and from 0.83 to 0.89 for the PA subscale. VA and PA scores were obtained by averaging the items of VA or PA factors respectively. Higher scores on VA reflect more verbal aggressions. Higher scores on PA reflect greater physical aggressions.

To measure weight self-stigma in the clinical sample, the Weight Self-Stigma Questionnaire (WSSQ) (Lillis et al., 2010; Spanish version: Magallares et al., 2022) was used. The WSSQ contains 12 items and was originally designed for use with populations of people with overweight or obesity (Lillis et al., 2010). The items (e.g., "I became overweight because I'm a weak person") were rated on a 7-point Likert scale from 0 ("strongly disagree") to 6 ("strongly agree"). Coefficient alpha of the WSSQ on our clinical sample was 0.80. WSSQ scores were obtained by averaging the items of the questionnaire. Higher scores on WSSQ reflect more weight self-stigma.

To measure fear of gaining weight in the non-clinical sample, the three item Fear of Fat subscale of the Antifat Attitudes test (AFA) was used (Crandall, 1994; Spanish version: Magallares & Morales, 2014). This factor of the AFA is usually administered to university students and allows for identification of people at risk of developing disordered eating (Magallares, 2012). The items were rated on a 7-point Likert scale from 0 ("strongly disagree") to 6 ("strongly agree"). An example would be: "I feel disgusted with myself when I gain weight". Coefficient alpha of the Fear of Fat subscale on our non-clinical sample was 0.87. Fear of Fat scores were obtained by averaging the items of the questionnaire. Higher scores indicate higher fear of gaining weight.

2.3. Data analysis

Confirmatory Factor Analyses (CFA) were conducted to determine the factor structure of the questionnaire using the Maximum Likelihood Estimation with Satorra-Bentler correction. To assess the fit of the models to the data, several fit indices (Chi-square [χ^2], normed Chi-square [χ^2/df]), Comparative Fit Index [CFI], and Root Mean Square Error of Approximation [RMSEA]) were reported. χ^2 values less than 0.05 and normed Chi-square values below 3 indicate an adequate model fit (Anderson & Gerbing, 1984). CFI values between 0.90 and 0.95 indicate an acceptable model fit, with values greater than 0.95 indicating a close model fit (Hu & Bentler, 1999). Values of RMSEA below 0.08 indicate a fair fit.

Cronbach's alpha calculations were used to evaluate the internal consistency of the WRAQ, the questionnaire's factors and total score. Cronbach's alphas above 0.70 can be interpreted as acceptable (Tavakol & Dennick, 2011). Moreover, Composite Reliability (CR) was also determined. Experts suggest that CR is a more reliable measurement in structural equation modeling compared to Cronbach's alpha, as it relies on the loadings instead of the observed correlations between the variables (Brown, 2015). CR values of 0.70 or higher denote good reliability.

This study assessed measurement invariance by sample and by gender using a Multigroup Confirmatory Factor Analysis (MGCFA), following the procedure proposed by Vandenberg and Lance (2000) for model comparisons. The MGCFA procedure involves restricting different model parameters in a stepwise manner and comparing the resulting models between groups. We focused on testing configural, metric, scalar and strict invariance by gradually testing for equal form, equal factor loadings, equal intercepts, and equal residual variances. A more restrictive model was only nested if the invariance of the preceding model held. Invariance was considered to be present if parameter changes for CFI and RMSEA were no larger than ± 0.01, following the criteria suggested by Cheung and Rensvold (2002). Strict invariance is often very difficult to establish in practice, but scalar invariance (equal factor loadings and intercepts) is a prerequisite for meaningfully comparing latent factor means across groups, so this approach allowed us to determine whether observed group differences were due to true interindividual differences rather than to differences in how the construct is being measured across groups.

To analyze the convergent validity of the WRAQ, the relationships between weight self-stigma (for the clinical sample) and fear of gaining weight (for the non-clinical sample) were studied using Pearson's correlations. The CFA and the measurement invariance analyses were conducted with R (using the package Lavaan) and AMOS 27 Software. All remaining analyses were performed with SPSS Statistics 27.

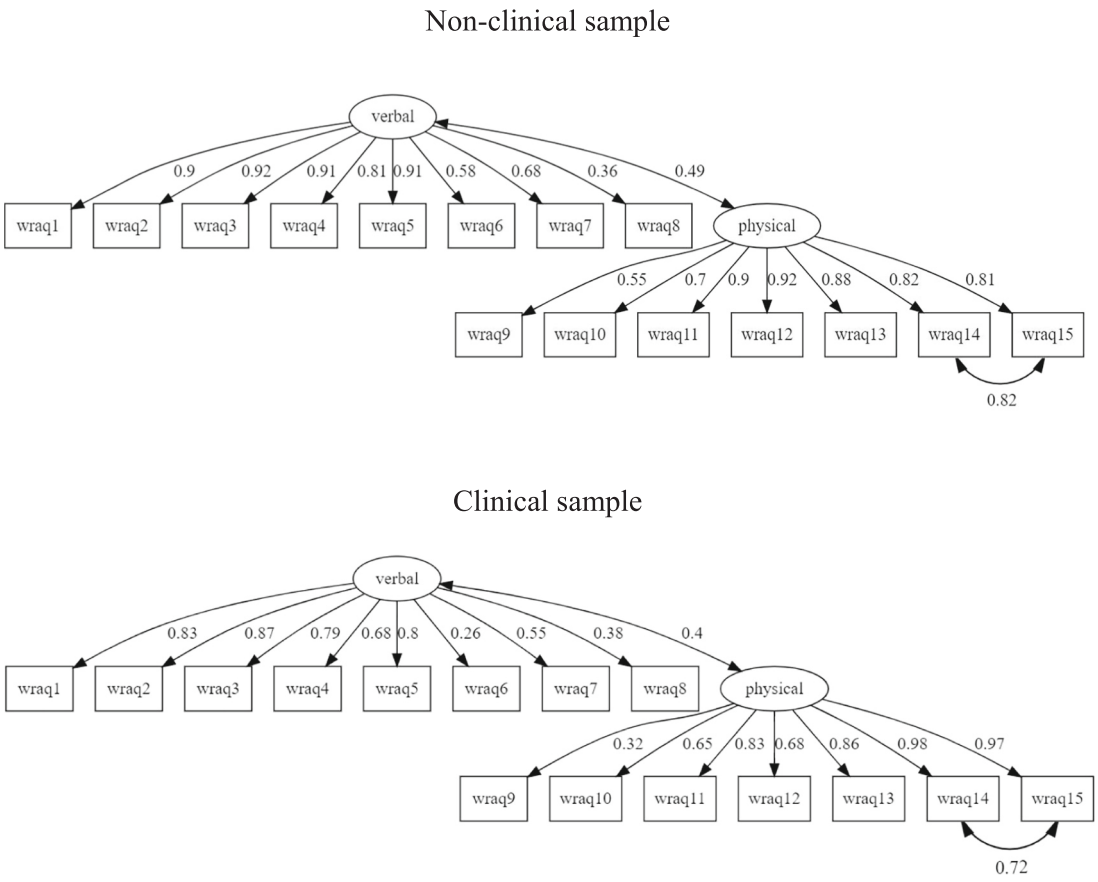


Fig. 1. Two-factor model of the Weight-Related Abuse Questionnaire in both non-clinical and clinical samples.

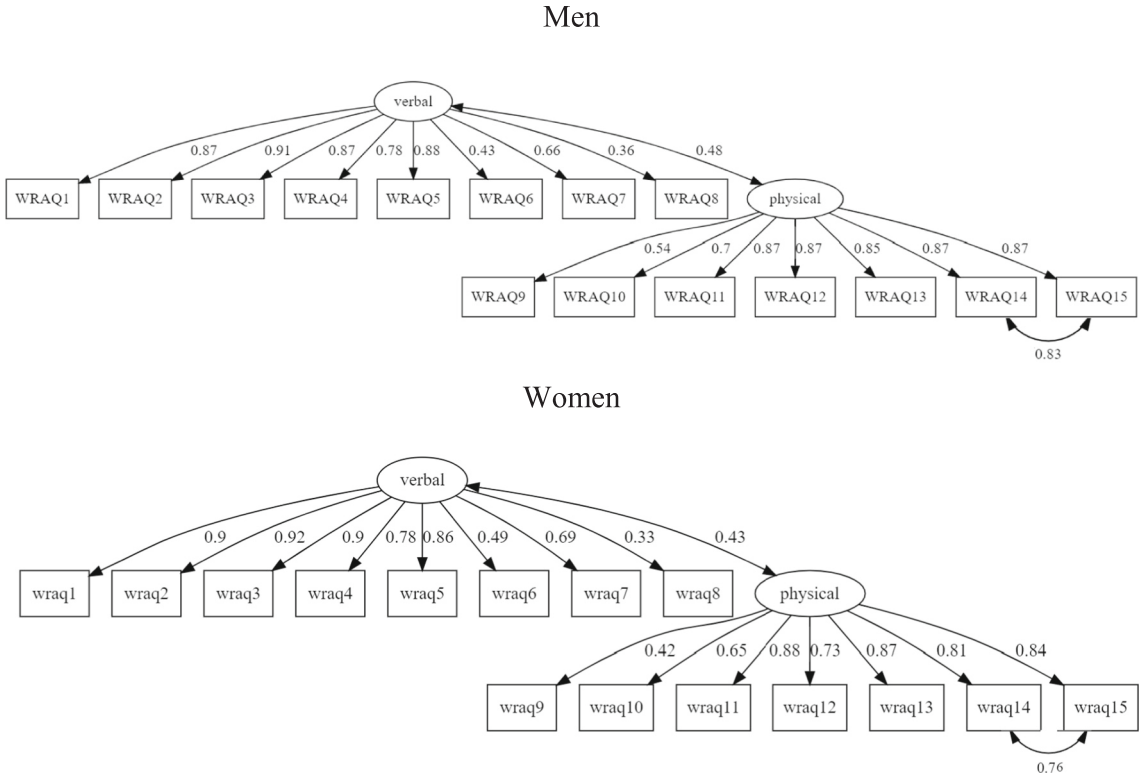


Fig. 2. Two-factor model of the Weight-Related Abuse Questionnaire in both men and women.

Table 1
Descriptive statistics and reliability for non-clinical and clinical samples and for men and women.

Variable	Non-clinical sample				Clinical sample			
	Mean	*SD*	α	CR	Mean	*SD*	α	CR
Verbal abuse	1.14	1.42	0.92	0.92	1.02	1.13	0.83	0.86
Physical abuse	0.24	0.77	0.92	0.92	0.15	0.46	0.89	0.91
Total scale	0.72	1.00	0.92	0.96	0.62	0.73	0.85	0.94
	Men				Women			
Verbal abuse	1.13	1.27	0.90	0.92	1.09	1.38	0.91	0.93
Physical abuse	0.28	0.87	0.91	0.91	0.17	0.58	0.88	0.89
Total scale	0.74	0.98	0.93	0.96	0.66	0.89	0.92	0.95

Note. CR = Composite reliability. Scores range from 0 (never) to 6 (more than 20 times per year).

3. Results

3.1. Confirmatory factor analysis

Initially, we conducted CFA in both samples to assess the two-factor and four-factor models. The results of fit indices indicated that the models were not acceptable. In the non-clinical sample, both the CFI and the Root Mean Square Error of Approximation (RMSEA) were outside the recommended range ($\chi^2 = 234.39$, $p = .00$, CFI = 0.86, RMSEA = 0.14). In the clinical sample, RMSEA exceeded the cutoff value of 0.08, indicating inadequate fit ($\chi^2 = 121.31$, $p = .01$, CFI = 0.92, RMSEA = 0.083). Furthermore, when examining the four-factor models, we found that the correlation between severe physical abuse and mild physical abuse was very high (0.95 in the clinical sample and 0.80 in the non-clinical sample). This suggests that it is not appropriate to distinguish between these factors.

Modification indexes were checked for significant residual correlations that were theoretically meaningful, and as a result, the residuals of items 14 and 15 were allowed to correlate. The content of items 14 ("Someone hit you because of your weight") and 15 ("Someone kicked you because of your weight") was very similar, justifying this correlation (Gerbing & Anderson, 1984). After these adjustments, all models improved their fit and appeared to be good.

Regarding the clinical sample, the two-factor model showed good fit ($\chi^2 = 110.57$, $p = .00$, $\chi^2/df = 1.26$, CFI = 0.94, RMSEA = 0.04) and distinguished between PA and VA, which were moderately correlated ($r = 0.40$, $p < .01$). The four-factor model also showed adequate fit ($\chi^2 = 104.62$, $p = .06$, $\chi^2/df = 1.25$, CFI = 0.95, RMSEA = 0.04), but the correlations between the mild and severe abuse factors of the VA ($r = 0.89$, $p < .01$) and PA ($r = 0.95$; $p < .01$) subscales were too high to consider them as separate dimensions. Therefore, for the clinical sample, the two-factor structure was deemed a better fit.

For the non-clinical sample, we found a similar scenario: the two-factor ($\chi^2 = 155.41$, $p = .00$, $\chi^2/df = 1.77$, CFI = 0.94, RMSEA = 0.05) and four-factor ($\chi^2 = 146.84$, $p = .00$, $\chi^2/df = 1.75$, CFI = 0.94, RMSEA = 0.05) structures showed acceptable fit. However, the correlations between the two factors (mild and severe) of VA ($r = 0.73$; $p < .01$) and the two factors of PA (mild and severe) ($r = 0.80$; $p < .01$) were too high once again, so the two-factor structure was also deemed a better fit for the non-clinical sample. Item loading ranged from 0.36 to 0.92 (non-clinical) and 0.26 to 0.98 (clinical).

Regarding gender, the two-factor model was estimated separately for men and women using the adjustments previously described (residuals of items 14 and 15 were allowed to correlate). This model had an acceptable fit for men ($\chi^2 = 168.16$, $p = .00$, $\chi^2/df = 1.91$, CFI = 0.95, RMSEA = 0.05) and women ($\chi^2 = 163.32$, $p = .00$, $\chi^2/df = 1.86$, CFI = 0.93, RMSEA = 0.05). Item loadings ranged from 0.36 to 0.91 (men) and 0.33 to 0.92 (women).

The diagrams of the two-factor solutions for both clinical and non-clinical samples are presented in Fig. 1, while Fig. 2 displays the corresponding solutions for men and women.

Table 2
Results of the invariance analysis by sample and gender.

Model	CFI	RMSEA	ΔCFI	ΔRMSEA
Non-clinical vs clinical sample				
Configural model	0.91	0.07	–	–
Metric model	0.90	0.08	0.01	0.01
Scalar model	0.89	0.08	0.01	0.00
Strict model	0.84	0.09	0.05	0.01
Gender				
Configural model	0.90	0.08	0.00	0.00
Metric model	0.90	0.08	0.00	0.00
Scalar model	0.90	0.08	0.00	0.00
Strict model	0.86	0.09	0.04	0.01

3.2. Reliability

Cronbach's alpha coefficients and CR for both non-clinical and clinical samples, and for men and women, were good, with values ranging from 0.83 to 0.96. No significant differences were found between the clinical and non-clinical samples in WRA measured with the WRAQ as a whole ($F_{1,449} = 1.29$, $p = .26$), or in its subscales of VA ($F_{1,449} = 1.40$, $p = .78$) and PA ($F_{1,449} = 1.62$, $p = .20$). Furthermore, no significant gender-based differences were observed in either the total WRAQ score ($F_{1,446} = 0.68$, $p = .41$) or the subscales ($F_{1,446} = 0.11$, $p = .75$ for VA subscale, and $F_{1,446} = 2.66$, $p = .10$ for PA subscale). The means, standard deviations and reliability are shown in Table 1, broken down by sample and gender.

3.3. Measurement invariance

Table 2 presents the results of testing the invariance of the WRAQ by sample and gender. Multigroup two-factor models were estimated with the same adjustments as in the prior CFA, where residuals of items 14 and 15 were correlated.

The next step was to test the metric invariance model, which involved equalizing the factor loads of the models in the two groups. When the fit indexes were examined, they were found to be acceptable, and the difference in CFI between the two models was equal to 0.01. Thus, nesting the metric model into the configural model did not result in a worse fit, indicating that the relationship between items and factors was equal across both populations. The results of the analysis to test the scalar invariance model showed that the difference in CFI was also equal to 0.01, similar to the second model. Finally, the error variances of the scale items were compared for both samples after being equalized (strict model). In this case, although the increment in RMSEA did not exceed the established threshold, the ΔCFI value (ΔCFI = 0.06) was much higher than the maximum recommended (Chen, 2007). The model showed a good fit in the restriction of factor loadings (metric invariance) and intercepts (scalar invariance). Thus, the WRAQ demonstrated scalar invariance across both non-clinical and clinical samples.

Regarding gender, the multigroup CFA model was estimated simultaneously for men and women, representing the invariance configural model, which showed an adequate fit with RMSEA = 0.08 and CFI = 0.90. Based on the evidence of a 0.01 increase in CFI and RMSEA, it can be assumed that the factor loadings and intercepts were invariant by gender (see Table 2). Therefore, scalar invariance by gender was established. When assessing strict invariance it was observed that ΔCFI value (ΔCFI = 0.04) exceeded the recommended threshold, suggesting a lack of strict invariance by gender.

3.4. Convergent validity

As expected, in the non-clinical sample fear of gaining weight was

Table 3
Correlations for non-clinical and clinical samples.

Non-clinical sample	1	2	3	4	Clinical sample	1	2	3	4
1. Total WRAQ		0.95**	0.75**	0.32**	Total WRAQ		0.96**	0.66**	0.41**
2. Verbal abuse			0.51**	0.36**	Verbal abuse			0.44**	0.41**
3. Physical abuse				0.12*	Physical abuse				0.21*
4. Fear of Fat					WSSQ				

Note. WRAQ = Weight-Related Abuse Questionnaire, WSSQ = Weight Self-Stigma Questionnaire, Verbal abuse and Physical abuse are the WRAQ subscales, *$p < .05$, **$p < .01$.

Table 4
Correlations for men and women.

Men	1	2	3	4	5	Women	1	2	3	4	5
1. Total WRAQ		0.95**	0.83**	0.18	0.46**	Total WRAQ		0.96**	0.67**	0.38**	0.36**
2. Verbal abuse			0.59**	0.25	0.43**	Verbal abuse			0.44**	0.39**	0.39**
3. Physical abuse				0.10	0.44**	Physical abuse				0.18*	0.10
4. Fear of Fat				4		Fear of Fat				4	
5. WSSQ						WSSQ					

Note. WRAQ = Weight-Related Abuse Questionnaire, WSSQ = Weight Self-Stigma Questionnaire, Verbal abuse and Physical abuse are the WRAQ subscales, the correlation between Fear of Fat and WSSQ is not provided as the non-clinical sample completed Fear of Fat, while the clinical sample completed WSSQ. *$p < .05$, **$p < .01$.

both correlated to the VA ($r = 0.36$, $p < .01$) and the PA ($r = 0.12$, $p < .05$) subscales of the WRAQ. In the case of the clinical sample, weight self-stigma was related to the PA ($r = 0.21$, $p < .01$) and the VA ($r = 0.41$, $p < .01$) factors of the WRAQ (see Tables 3 and 4).

For men, we did not observe significant correlations between the Verbal Abuse (VA) subscale ($r = 0.25$, $p = .06$) or the Physical Abuse (PA) subscale ($r = 0.10$, $p = .63$) of the WRAQ with the Fear of Fat subscale. However, for women, these relationships were statistically significant (VA $r = 0.39$, $p < .01$, and PA $r = 0.18$, $p = .01$). Regarding the Weight Self-Stigma Questionnaire (WSSQ), there were significant positive correlations with subscales of the WRAQ in both men and women (as shown in Table 4).

4. Discussion

Overall, the results of the current study indicate that the two-factor model of the WRAQ possesses a more satisfactory factor validity than the four-factor model. In addition, this study also analyzed the scale's invariance using multi-group CFA models. The results demonstrate that the questionnaire can be utilized in clinical and non-clinical samples for both men and women, given evidence of scalar invariance, which allows for meaningful comparisons of mean scores on the WRAQ across clinical and non-clinical samples and gender. No evidence of strict invariance was found, which is common, as it represents a highly restrictive model, too strict to fit data from real participants in real settings (Brown, 2015; Van De Schoot et al., 2015). Furthermore, this research shows that the translated version of the WRAQ is a psychometrically sound questionnaire for the retrospective assessment of WRA for Spanish populations. Both the VA and PA subscales of the WRAQ have good internal consistency and convergent validity (weight self-stigma and fear of gaining weight). Additionally, the results highlight the cross-cultural nature of WRA and the relationships between WRA and negative self-perceptions related to weight.

It should be noted that as the measure used in this study was an adapted and translated version of the WRAQ, and the samples evaluated in this study were from Spain it is possible that different cultural norms may have impacted study findings, as such, research including multinational samples may be warranted to further confirm the cross-cultural validity of the two-factor model. Additionally, although we used a sample of participants with obesity, evaluation of the psychometrics of the WRAQ in other clinical populations is suggested. For example, one population of particular interest is individuals diagnosed with eating disorders, as WRA may play a role in the development of disordered eating (Salwen & Hymowitz, 2015).

Although this study has several strengths, there are some limitations to the present research that should be considered. Firstly, the clinical sample only included 150 participants. Future studies should consider recruiting more participants. Secondly, we did not find differences between the clinical and the college samples. We believe that the low means of the WRAQ seems to indicate a floor effect, but further research is needed to clarify this aspect. Thirdly, the correlations were significant but not very high, especially with PA. Again, we think that because most of the participants scored low on this factor of the WRAQ, there was not much variation and, therefore, the correlations were low. Fourthly, we found low factor loadings in both of our samples. The items with the lowest factor loadings in the current study were items that can be considered more severe forms of VA (e.g., "Someone threatened to abandon you because of your weight"). This suggests that while for the Spanish translation of the WRAQ subscales distinguishing mild and severe forms of abuse may not be warranted some items might better identify more severe forms of VA and might not be as useful for measuring milder forms of abuse. Finally, no test-retest was conducted to assess the stability of the WRAQ. Future studies should evaluate the test-retest reliability of the Spanish translation of the WRAQ.

5. Conclusion

To the best of our knowledge, this is the first documented work that conducts a CFA to analyze the factor structure of the WRAQ and tests the measurement invariance of this questionnaire by sample and gender. Results indicate that the translated version of the WRAQ is a reliable and valid measure of VA and PA in clinical and non-clinical samples for both men and women for the Spanish population. Future studies should evaluate the validity of the WRAQ in eating disorders patients and

researchers should continue to evaluate the role of WRA in the development and maintenance of obesity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eatbeh.2023.101827.

References

Ahmed Mokbel, R., Attia Abd Elghaffar Moustafa, A., & Salah Yonis Mahrous, E. (2020). Relation between weight-related bullying and emotional reactions among adolescents and their coping strategies. *Egyptian Journal of Health Care, 11*(3), 892–913. https://doi.org/10.21608/ejhc.2020.203203

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49*(2), 155–173. https://doi.org/10.1007/BF02294170

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17*, 303–316. https://doi.org/10.1177/0049124189017003004

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* Guilford Press.

Castillo-Luna, A., & Jauregui-Lobera, I. (2017). Weight-related experiences of social rejection. *Trastornos de la Conducta Alimentaria, 25*, 2637–2682.

Chen, F. F. (2007). Sensitivity of goodness of ft indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Crandall, C. S. (1994). Prejudice against fat people: Ideology and self-interest. *Journal of Personality and Social Psychology, 66*(5), 882–894. https://doi.org/10.1037/0022-3514.66.5.882

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research, 11*(1), 572–580. https://doi.org/10.1086/208993

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Lillis, J., Luoma, J. B., Levin, M. E., & Hayes, S. C. (2010). Measuring weight self-stigma: The weight self-stigma questionnaire. *Obesity, 18*, 971–976. https://doi.org/10.1038/oby.2009.353

Magallares, A. (2012). Well-being and prejudice toward obese people in women in risk to develop eating disorders. *Spanish Journal of Psychology, 15*(3), 1293–1302. https://doi.org/10.5209/rev_SJOP.2012.v15.n3.39415

Magallares, A., Benito de Valle, P., Irles, J. A., Recio, P., & Jauregui-Lobera, I. (2022). Psychometric properties of the Spanish version of the Weight Self-Stigma Questionnaire (S-WSSQ) in a sample of participants with obesity seeking weight loss treatment. *Eating and Weight Disorders, 27*(8), 3685–3693. https://doi.org/10.1007/s40519-022-01511-6

Magallares, A., & Morales, J. F. (2014). Spanish adaptation of the Antifat Attitudes Scale. *International Journal of Social Psychology, 29*(3), 563–588. https://doi.org/10.1080/02134748.2014.972707

Munir, M., & Dawood, S. (2021). Weight stigma and disordered eating behaviors in Pakistani overweight adolescents: The mediating role of body esteem. *Eating and Weight Disorders, 26*(6), 1939–1948. https://doi.org/10.1007/s40519-020-01038-8

Salwen, J. K., & Hymowitz, G. F. (2015). The weight-related abuse questionnaire (WRAQ): Reliability, validity, and clinical utility. *Eating Behaviors, 19*, 150–154. https://doi.org/10.1016/j.eatbeh.2015.09.001

Salwen, J. K., Hymowitz, G. F., Bannon, S. M., & O'Leary, K. D. (2015). Weight-related abuse: Perceived emotional impact and the effect on disordered eating. *Child Abuse & Neglect, 45*, 163–171. https://doi.org/10.1016/j.chiabu.2014.12.005

Spahlholz, J., Baer, N., König, H. H., Riedel-Heller, S. G., & Luck-Sikorski, C. (2016). Obesity and discrimination - A systematic review and meta-analysis of observational studies. *Obesity Reviews, 17*(1), 43–55. https://doi.org/10.1111/obr.12343

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology, 6*, 1–4. https://doi.org/10.3389/fpsyg.2015.01064

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–69. https://doi.org/10.1177/109442810031002