
Article:
Trujillo-ChiVacuán, E., Compte, E.J. and Waller, G. orcid.org/0000-0001-7794-9546 (2025) Validation of the Eating Disorders-15 (ED-15) in Mexican patients across levels of care: Psychometric properties in a clinical sample. Eating Behaviors, 59. 102037. ISSN: 1471-0153

https://doi.org/10.1016/j.eatbeh.2025.102037

Validation of the Eating Disorders-15 (ED-15) in Mexican Patients Across Levels of Care: Psychometric Properties in a Clinical Sample

1. Introduction

Reliable and valid tools are essential for monitoring patients and guiding treatment. Brief measures with sound psychometric properties and clear factor structures are also crucial for cross-cultural research and advancing our understanding of ED psychopathology. The Eating Disorders-15 (ED-15), developed by Tatham et al. (2015), is a short self-report instrument designed to track therapeutic progress rather than provide a formal diagnosis. It includes two subscales (i.e., Weight and Shape Concerns (WSC) and Eating Concerns (EC)) identified through principal component analysis. The final version comprises 10 attitudinal and 5 behavioral items, with strong internal consistency and excellent test-retest reliability. Convergent validity was demonstrated by strong correlations with the Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994, 2008), and divergent validity was supported by weaker associations with constructs such as anxiety and depression, as expected. The ED-15 targets core features shared across ED diagnoses (Tatham et al., 2015), aligning with transdiagnostic models (Fairburn et al., 2003), and is thus suitable for heterogeneous clinical populations.

Since its development, the ED-15 has been examined in various cultural and clinical contexts, offering valuable insights into its structure and performance across diverse populations. In terms of factorial structure, in Portugal, a clinical sample study tested five models, including one- and two-factor solutions, and identified a re-specified 2-factor model with correlated residuals as the best-fitting solution (Rodrigues et al., 2019). In Chile, a non-clinical study introduced a bi-factor model to account for high factor correlations, which showed best fit (Compte et al., 2023). In Turkey, exploratory factor analysis (EFA) in clinical and non-clinical samples confirmed the original 2-factor structure (Öngün-Yılmaz et al., 2023). In Australia, confirmatory factor analysis (CFA) in a clinical sample also supported the original model (Zhou et al., 2024). Overall, findings consistently support the 2-factor structure and its

adaptability across contexts.

Strong psychometric properties have been reported across countries. Internal consistency (Cronbach's α, McDonald's ω) exceeded .80 in Portugal, Chile, and Turkey (Rodrigues et al., 2019; Compte et al., 2023; Öngün-Yılmaz et al., 2023). Test-retest reliability, also above .70, was documented in these same studies. Convergent validity was supported by strong correlations with ED measures like the EDE-Q in Portugal and Australia (Rodrigues et al., 2019; Zhou et al., 2024), and divergent validity by weaker associations with depression, anxiety, and stress in Portugal and Chile (Rodrigues et al., 2019; Compte et al., 2023). Construct validity was further supported by associations with functional impairment in Portugal and body appreciation in Chile.

Evidence for the clinical utility and diagnostic validity of the ED-15 has been reported across multiple studies. Tatham et al. (2015) found higher scores in clinical compared to community samples and observed reductions in WSC and EC subscales over the course of a ten-session cognitive-behavioral therapy in women with EDs. Rodrigues et al. (2019) identified a cut-off score of 2.1 for the total score and calculated a Reliable Change Index (RCI) to evaluate reliable and clinically significant change during treatment. The Turkish validation similarly reported higher scores in clinical samples and proposed a cut-off of 2.5 for the total score. Finally, Zhou et al. (2024) demonstrated that ED-15 cut-offs, originally proposed by Rodrigues et al. (2019), effectively distinguished between clinical and non-clinical cases, with robust odds ratios (4.69 to 21.35), providing strong support for the scale's diagnostic validity and sensitivity to change across cultural contexts.

The ED-15 has been translated into Portuguese, Spanish, and Turkish through rigorous procedures to ensure semantic, conceptual, and cultural equivalence. The Portuguese and Turkish versions followed forward- and back-translation processes by bilingual experts, combined by consensus, and reviewed by field specialists (Rodrigues et al., 2019; Öngün-Yılmaz et al., 2023). The Spanish version adapted in Chile, used simultaneous translation by native speakers from different Latin American countries to create a harmonized version suitable for diverse Spanish-speaking populations (Compte et al., 2023). Pretesting was conducted to

evaluate comprehension, acceptability, and emotional impact in each context.

Simultaneously, the absence of brief and validated tools continues to hinder adequate ED assessment in Latin America, where underdiagnosis and limited treatment access remain critical issues. In Mexico, where ED prevalence is rising (Villalobos-Hernández et al., 2023), clinical services urgently require efficient instruments to monitor symptoms and guide treatment. Although a Spanish version of the ED-15 was validated in a non-clinical Chilean sample (Compte et al., 2023), its performance in clinical populations has not yet been evaluated. While not a diagnostic tool per se, assessing its clinical utility in Spanish-speaking treatment settings is critical.

This study therefore aims to: (1) examine ED-15 factor models through CFA, (2) assess internal consistency, and (3) evaluate construct validity, including convergent and divergent validity, using established measures of ED symptoms, emotional distress, body image, and functional impairment.

2. Method

This study was approved by the Institutional Review Board at the Instituto Tecnológico de Monterrey (P000624-TCA 2021-CEIC-CR003). Informed written consent was obtained from all participants, as well as from their parent(s) or legal guardian(s), where applicable.

2.1 Participants

A total of 286 individuals diagnosed with EDs participated in this study, conducted at [omitted for peer review], a specialized EDs treatment center located in Monterrey, Mexico. Most participants were Mexican-born (92.1%) and female (91.1%), with a mean age of 18.5 years (SD = 6.6, range = 9 – 57) and a mean BMI of 20.2 (SD = 6.0, range = 10.3 – 44.6). At the time of data collection, approximately 70% of participants were in the early stages of treatment, 15% in intermediate stages, and 15% in advanced stages. Diagnoses included Anorexia Nervosa (54.5%), Atypical Anorexia Nervosa (13.5%), Bulimia Nervosa (5.6%), Bulimia Nervosa of Low Frequency and/or Limited Duration (9.4%), Binge Eating Disorder (3.4%), and Unspecified Feeding and Eating Disorders (13.5%). Diagnoses were established according to DSM-5 feeding and eating disorders criteria through a clinical triage process

conducted by senior clinical staff at [omitted for peer review]. Regarding treatment settings, 47.5% were enrolled in residential programs, 21.3% in partial hospitalization, and 30.5% in outpatient programs, while 0.7% had not yet been assigned to a treatment program. Participants completed a battery of questionnaires during baseline assessments or as part of their ongoing treatment.

2.2 Measures

The Eating Disorders-15 (ED-15; Tatham et al., 2015) is a brief self-report tool designed to assess session-by-session changes in core EDs attitudes and behaviors. It includes 10 attitudinal items grouped into two subscales (WSC and EC), rated on a 7-point Likert-type scale based on experiences over the past seven days. The overall score is the mean of the 10 items. The Spanish-language version (Compte et al., 2023) was developed through a multicultural adaptation process following International Test Commission (2017) guidelines. Two procedures were used: a committee-based translation by bilingual experts from various Latin American countries and a separate forward and back-translation. Both versions were reconciled to ensure conceptual and linguistic equivalence. As part of the current study, a pilot test was conducted with an additional 30 ED patients (96.7% female, Mage = 20.2, SD = 5.8, range = 12–32) to evaluate item clarity using a 5-point Likert-type scale. The mean rating was 4.46 (SD = 0.33), with minimal feedback and no semantic changes required.

Eating Disorder Examination-Questionnaire 7 (EDE-Q-7; Fairburn & Beglin, 1994, 2008; Grilo et al., 2013; Trujillo-ChiVacuán et al., 2025), is a concise adaptation of the 28-item EDE-Q, originally derived from the Eating Disorder Examination (Fairburn & Cooper, 1993), a widely used semistructured interview for assessing EDs psychopathology. The EDE-Q-7, validated in Mexican patients with EDs (Trujillo-ChiVacuán et al., 2025), includes seven items organized into three subscales: Restraint, Shape/Weight Overevaluation, and Body Dissatisfaction. Responses are scored on a 7-point Likert-type scale ranging from 0 = No days to 6 = Every day, reflecting behaviors and attitudes over the past 28 days. Psychometric evaluations in Mexican samples have demonstrated strong internal consistency (Cronbach's α = .90 for the total scale; .78–.90 for subscales).

Body Shape Questionnaire-8 (Cooper et al., 1987; Franco-Paredes et al., 2020), is a self-report measure designed to assess body dissatisfaction related to size and shape. Originally consisting of 34 items, the BSQ has been adapted and validated in Mexican populations, with the 8-item version (BSQ-8) demonstrating the best psychometric properties (Franco-Paredes et al., 2020). The BSQ-8 is unidimensional, with excellent internal consistency (Cronbach's α = .89) and convergent validity with measures of eating behaviors and sociocultural influences on body image. It effectively discriminates between individuals with and without abnormal eating behaviors and has been validated for use across adolescents and young adults.

Body Appreciation Scale-2 (BAS-2; Góngora et al., 2020; Tylka & Wood-Barcalow, 2015), is a widely used measure of positive body image, assessing an individual's acceptance, respect, and appreciation of their body. The BAS-2 consists of 10 items rated on a 5-point Likert-type scale ranging from 1 = Never to 5 = Always, with higher scores indicating greater body appreciation. A Latin American Spanish translation of the BAS-2 was validated in adolescent samples from Mexico, Argentina, and Colombia (Góngora et al., 2020), demonstrating a unidimensional factor structure and measurement invariance across countries and gender groups. Psychometric evaluations in these populations have reported strong internal consistency (Cronbach's α = .93 to .94) and significant negative correlations with body dissatisfaction, thin-ideal internalization, and sociocultural pressures related to appearance.

Patient Health Questionnaire-9 (PHQ-9; Spitzer et al., 1999), is a brief self-report measure assessing depressive symptoms over the past two weeks. The nine items are rated on a 4-point Likert-type scale (0 = Not at all to 3 = Nearly every day), with total scores ranging from 0 to 27 and categories for depression severity. In Mexican women, the PHQ-9 has shown a unidimensional structure, high internal consistency (Cronbach's α = 0.89), and adequate validity in identifying moderate to severe symptoms (Familiar et al., 2015).

Clinical Impairment Assessment (CIA; Bohn et al., 2008), is a 16-item self-report questionnaire designed to assess the psychosocial impairment caused by EDs psychopathology. It evaluates three domains—personal, social, and cognitive impairment—over the past 28 days, with items rated on a 4-point Likert-type scale (0 = "Not at all" to 3 = "A lot"). A higher total

score (range 0–48) indicates greater impairment. The Spanish version of the CIA was adapted and validated in a clinical sample of 178 patients with EDs in Spain (Martín et al., 2015), maintaining the original three-factor structure. The validation demonstrated excellent internal consistency (Cronbach's α = .90 to .96) and strong construct validity, with strong correlations with related constructs such as eating attitudes and health-related quality of life.

Table 2 presents the internal consistency values for the measures used in the current sample.

2.3 Data Analysis

Continuous variables were summarized as means and standard deviations, and categorical variables as percentages. A nonparametric test indicated that item-level missing data on the ED-15 were completely at random (p = .634). Given the very low proportion (0.02%), multivariate imputation by chained equations was used to preserve statistical integrity (Azur et al., 2011). A Monte Carlo simulation was conducted to estimate the sample size required for a CFA of the original model with two latent factors (one with four and one with six indicators). This approach, following the recommendations by Muthén and Muthén (2002), served a similar purpose to an a priori power analysis by evaluating whether the planned sample size would yield stable parameter estimates and acceptable model fit. Sample sizes from 50 to 500 were simulated, and model fit was evaluated using the criteria described below. Results indicated that a minimum of 250 participants was necessary for adequate model fit, aligning with recommendations for CFA with moderate communalities (MacCallum et al., 1999). The simulation confirmed that a sample size of 286 participants provided sufficient statistical power (>80%) to detect differences in factor structure and to reliably evaluate model fit indices. Given evidence of multivariate non-normality (Henze-Zirkler test = 4.924, $p < .001$), CFAs were conducted using robust maximum likelihood estimation with Satorra-Bentler correction (Satorra & Bentler, 1994). Four models were tested: (1) the original two-factor structure (Tatham et al., 2015), (2) a one-factor solution collapsing all attitudinal items (Rodrigues et al., 2019), (3) a higher-order model accounting for a general latent construct, and (4) a bi-factor model including domain-specific factors (ED-SWC and ED-EC) and a general factor (g-factor), following

Compte et al. (2023). The inclusion of the higher-order model was theoretically justified due to the high correlation between latent factors, which may indicate the presence of a broader underlying construct. This approach aligns with psychometric recommendations suggesting that when inter-factor correlations exceed .70, a second-order model should be tested (Kline, 2011). Model fit was assessed using the normed chi-square statistic (χ^2/df), CFI, TLI, RMSEA (with 90% CI), and SRMR. Adequate fit thresholds were: χ^2/df < 3.00, CFI and TLI > .95, RMSEA < .06, and SRMR < .08 (Kline, 2011). We also examined the Akaike Information Criterion (AIC) and the Expected Cross-Validation Index (ECVI), with lower values indicating better model fit. Differences of 2 points or more in AIC are generally considered meaningful when comparing model parsimony (Burnham & Anderson, 2004).

Convergent and divergent validity were evaluated via Spearman rank-order correlations: $r_s \geq .10$–$.29$ = weak, $r_s \geq .30$–$.49$ = moderate, and $r_s \geq .50$ = strong (Cohen, 1992). Internal consistency was estimated using Cronbach's alpha (α) and the Omega (ω) coefficient with 95% confidence intervals (Dunn et al., 2014). For 2-item subscales, the Spearman-Brown coefficient was used (Eisinga et al., 2012). Values $\geq .80$ were considered acceptable (Nájera Catalán, 2019; Nunnally, 1978).

Statistical significance was set at two-tailed $p < .05$. All analyses were conducted using R (version 4.3.0) with the following packages: MVN (Korkmaz et al., 2014), Lavaan (Rosseel, 2012), semPlot (Epskamp, 2013), MBESS (Kelley & Lai, 2012), psych (Revelle, 2018), and Hmisc (Harrel, 2008).

3. Results

3.1 Confirmatory Factor Analyses

Four models were tested to evaluate the factor structure of the ED-15. As shown in Table 1, all models demonstrated acceptable fit based on conventional cutoffs. The original 2-factor model (Model 1) and the hierarchical second-order model (Model 3) yielded identical fit indices, supporting their equivalence in representing the data. The 1-factor model (Model 2) showed a comparatively poorer fit, with a higher RMSEA, as well as higher AIC and ECVI values, indicating reduced parsimony. The bi-factor model (Model 4) showed slightly superior

absolute fit indices, yet its AIC and ECVI were higher than those of Model 1. Although the bi-factor model demonstrated slightly better fit, the original 2-factor structure (Model 1) was retained due to its theoretical coherence, model parsimony, and consistency with prior research. This retained model will be used for subsequent analyses of internal consistency and construct validity. Figure 1 presents a conceptual representation of Model 1, where all factor loadings were significant ($ps < .001$) and ranged from .36 to .95.

3.2 Construct Validity (Convergent and Divergent) and Internal Consistency

The ED-15 demonstrated strong convergent and divergent validity with measures of ED psychopathology, body dissatisfaction, body appreciation, depressive symptoms, and psychosocial impairment. As shown in Table 2, the ED-15 subscales and total score showed positive correlations with ED-related measures ($r = .50$ to $.88$), and with the CIA subscales and PHQ-9 ($r = .57$ to $.80$). A strong negative correlation was found with body appreciation, supporting divergent validity. Internal consistency for the ED-15 subscales and total score was high ($\alpha = .80$ to $.93$; $\omega = .86$ to $.93$), and related measures also showed excellent reliability (see Table 2).

4. Discussion

Validating the ED-15 in a clinical Latin American population fills a critical gap in the availability of culturally appropriate tools for ED assessment. This is the first study to validate the ED-15 in a Spanish-speaking clinical sample from Latin America, expanding its cross-cultural utility. The ED-15 facilitates timely intervention by enabling session-by-session monitoring, improving treatment outcomes (de Jong et al., 2025). Furthermore, its robust psychometric properties across levels of care underscore its potential as a standard tool for clinical and research applications in the region. This study validated the ED-15 in a clinical sample of Latin American patients with EDs, using the multicultural Spanish-language version developed by Compte et al. (2023). The findings confirmed the scale's robust psychometric properties, including a strong factorial structure, significant construct validity (including convergent and divergent validity) with established measures of EDs psychopathology and related constructs, and high internal consistency. These results support the utility of the ED-15

as a reliable tool for assessing core ED symptoms and associated impairments in Spanish-speaking populations, addressing a critical gap in the availability of culturally appropriate measures for the region. Although most participants presented with an anorexia-type diagnosis, the ED-15 was designed to assess core cognitive and behavioral features common across eating disorders (Tatham et al., 2015), in line with transdiagnostic models (Fairburn et al., 2003), supporting its use across diagnostic profiles. The decision to test a second-order model was further supported by the high correlation observed between the two latent factors (r = .78), suggesting the presence of a higher-order construct. This approach is consistent with psychometric guidelines (Kline, 2011) and strengthens the theoretical interpretation of the ED-15 as capturing a unified dimension of ED psychopathology while preserving clinically meaningful subscales.The findings extend the ED-15´s applicability to Latin America, addressing an unmet need for culturally sensitive assessment tools. Additionally, the study sets a foundation for cross-cultural comparisons in ED research.

The factorial analyses supported the original two-factor structure of the ED-15, with adequate and robust fit indices across multiple indicators. This structure includes two correlated latent factors (i.e., WSC and EC), which align with the original conceptualization proposed by Tatham et al. (2015). Although a bi-factor model demonstrated slightly better absolute fit, the 2-factor model was retained based on its theoretical coherence, parsimony, and empirical support in previous validations (Rodrigues et al., 2019; Öngün-Yılmaz et al., 2023). This decision reinforces the stability and applicability of the ED-15 structure across diverse clinical settings, including Latin American populations. Retaining the original two-factor model also offers a clear and interpretable structure that supports clinical decision-making by distinguishing between specific symptom domains, namely, weight and shape concerns and eating concerns, without requiring more complex latent structures in routine assessment contexts.

The ED-15 demonstrated strong construct validity, including expected convergent and divergent associations, with significant correlations observed between the ED-15 subscales and total score and measures of ED psychopathology, body dissatisfaction/appreciation, depressive symptoms, and psychosocial impairment. These results replicate findings from Rodrigues et al.

(2019) and Zhou et al. (2024), which demonstrated comparable associations with ED-related measures. Additionally, the negative correlation observed with body appreciation, as reported by Compte et al. (2023), further supports the ED-15's ability to capture the broader emotional and psychological dimensions of EDs. The strong correlations with established ED-related measures, including the EDE-Q and BSQ-8, reinforce the scale's alignment with validated tools used in clinical practice. These findings highlight the ED-15's utility in capturing not only core ED symptoms but also associated emotional and psychosocial impairments, making it a comprehensive tool for clinical assessment and research. Furthermore, the associations with depressive symptoms and body dissatisfaction align with previous studies, confirming the scale's sensitivity to comorbidities and related constructs, which are critical in understanding and treating EDs. These results are consistent with the ED-15's theoretical underpinnings and prior validation efforts, reinforcing its construct validity.

The use of previously validated scales in Mexico, such as the EDE-Q-7, BSQ-8, PHQ-9, and BAS-2, strengthens the reliability and cultural relevance of the present findings. By employing these measures, the study ensures that comparisons are grounded in instruments that have demonstrated strong psychometric properties within the target population. This approach enhances the validity of the ED-15's evaluation by reducing potential biases that may arise from using non-validated tools. Additionally, incorporating well-established measures facilitates consistency across studies, allowing for more accurate cross-cultural and longitudinal comparisons in ED research within Latin American contexts.

The ED-15 showed excellent internal consistency for both subscales and the total score, with reliability coefficients exceeding .80. These results are consistent with prior validations. The high reliability observed across diverse cultural and clinical contexts underscores the robustness of the ED-15 as a psychometrically sound measure. This consistency enhances its applicability in both research and clinical settings, ensuring that the scale provides reliable and reproducible results across populations. A notable strength of this study is the inclusion of participants from diverse treatment settings, including residential, partial hospitalization, and outpatient programs. This heterogeneity reflects a wide range of symptom severity and

281 treatment intensities, broadening the applicability of the findings. The ED-15 demonstrated
282 robust psychometric performance across these varied settings, suggesting that it is a reliable tool
283 for assessing ED symptoms regardless of the treatment context. This is particularly important
284 for clinical settings, where assessment tools must function effectively across different levels of
285 care. Future studies could explore whether specific treatment settings influence the scale's
286 sensitivity to symptom changes or its associations with related constructs.
287 Despite its strengths, this study has several limitations. Test-retest reliability was not
288 assessed, which limits the ability to determine the scale's stability over time. This decision was
289 based on the intensive nature of the treatment programs at [omitted for peer review], which aim
290 to facilitate early therapeutic changes that could significantly alter scale scores within a short
291 timeframe. As a result, conducting a reassessment within this context would not have provided a
292 meaningful evaluation of temporal stability. Future research should address this limitation by
293 assessing test-retest reliability in settings with less intensive treatment approaches, where
294 stability over time can be more accurately measured. Additionally, the sample was
295 predominantly female and Mexican-born, which may limit the generalizability of the findings to
296 other genders or Latin American subpopulations. Moreover, the predominance of anorexia-type
297 presentations may also restrict generalizability to other ED diagnoses. However, the ED-15
298 captures transdiagnostic features, supporting its applicability across different clinical profiles
299 (Tatham et al., 2015; Fairburn et al., 2003). Nevertheless, the large sample size, rigorous data
300 collection procedures, and inclusion of participants across diverse treatment settings represent
301 significant strengths that enhance the study's validity and applicability. Furthermore, this study
302 was not pre-registered, which may affect the transparency of analytical decisions and hypothesis
303 testing. Despite these limitations, the findings contribute valuable evidence supporting the ED-
304 15 as a reliable and valid measure for assessing eating disorder symptoms in diverse clinical
305 settings in Mexico.
306 This study provides robust evidence for the validity and reliability of the ED-15 in a
307 clinical Latin American population, confirming its factorial structure and psychometric
308 strengths. The findings highlight the scale's potential for clinical assessment and research on

ED psychopathology in Spanish-speaking populations. Future studies should explore the ED-15´s sensitivity to therapeutic changes across treatment modalities and its utility in predicting long-term recovery outcomes. Expanding validation efforts to include test-retest reliability and diverse cultural settings will further solidify its role as a comprehensive tool for monitoring progress, guiding interventions, and advancing the understanding of EDs in underserved regions.

Data availability statement:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments:

The authors acknowledge the Comenzar de Nuevo´s clinical team for their involvement in the translation and recruitment process.

References

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*(1), 40. https://doi.org/10.1002/MPR.329

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy*, *46*(10), 1105–1110. https://doi.org/10.1016/J.BRAT.2008.06.012

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. https://doi.org/10.1177/0049124104268644

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Compte, E. J., Cruzat-Mandich, C., Pérez Ríos, L., Vásquez, S., Gómez, N., & Trujillo-ChiVacuan, E. (2023). Factor structure and psychometric properties of a multicultural Spanish-language version of the Eating Disorder-15 in a nonclinical sample of Chilean women. *International Journal of Eating Disorders*, *56*(4), 727–735. https://doi.org/10.1002/EAT.23860

Cooper, P. J., Taylor, M. J., Cooper, Z., & Fairbum, C. G. (1987). The development and validation of the body shape questionnaire. *International Journal of Eating Disorders*, *6*(4), 485–494. https://doi.org/10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Eisinga, R., Grotenhuis, M. Te, & Pelzer, B. (2012). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health 2012 58:4*, *58*(4), 637–642. https://doi.org/10.1007/S00038-012-0416-3

Epskamp, S. (2013). *semPlot: Path diagrams and visual analysis of various SEM packages' output. R package version 0.3.2.*

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363–370. https://doi.org/10.1002/1098-108X(199412).

Fairburn, C. G., & Beglin, S. J. (2008). Eating Disorder Examination Questionnaire (EDE-Q 6.0). In C. G. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 309 – 314). Guilford Press

Fairburn, C. G., & Cooper, Z. (1993). The Eating Disorder Examination (12th edition). In C. G. Fairburn & G. T. Wilson (Eds.), *Binge Eating: Nature, Assessment and Treatment* (pp. 317–360). Guilford Press.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A transdiagnostic theory and treatment. *Behaviour Research and Therapy*, 41(5), 509–528. https://doi.org/10.1016/S0005-7967(02)00088-8

Familiar, I., Ortiz-Panozo, E., Hall, B., Vieitez, I., Romieu, I., Lopez-Ridaura, R., & Lajous, M. (2015). Factor structure of the Spanish version of the Patient Health Questionnaire-9 in Mexican women. *International Journal of Methods in Psychiatric Research*, *24*(1), 74–82. https://doi.org/10.1002/mpr.1461

Franco-Paredes, K., Díaz-Reséndiz, F. J., González-Betanzos, F., & Hidalgo-Rasmussen, C. (2020). Evaluating different versions of the Body Shape Questionnaire and invariance across age in Mexican women. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity, 26*(2), 515–523. https://doi.org/10.1007/S40519-020-00874-Y

Góngora, V. C., Cruz Licea, V., Mebarak Chams, M. R., & Thornborrow, T. (2020). Assessing the measurement invariance of a Latin-American Spanish translation of the Body

Appreciation Scale-2 in Mexican, Argentinean, and Colombian adolescents. *Body Image*, *32*, 180–189. https://doi.org/10.1016/J.BODYIM.2020.01.004

Grilo, C. M., Henderson, K. E., Bell, R. L., & Crosby, R. D. (2013). Eating disorder examination-questionnaire factor structure and construct validity in bariatric surgery candidates. *Obesity Surgery*, 23, 657–662.https://doi.org/10.1007/s11695-012-0840-8

Harrel, F. (2008). *Hmisc: Harrell miscellaneous.* https://doi.org/10.1177/2158244015621113.

International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests (Second edition)*. [Www.InTestCom.Org] .

Kelley, K., & Lai, K. (2012). *MBESS: MBESS. R package version 3.3.2.* http://CRAN.Rproject.org/package=MBESS

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling.* (R. B. Kline, Ed.; Third). Guilford Press.

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *R Journal*, *6*(2), 151–162. https://doi.org/10.32614/RJ-2014-031

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

McKay, R. B., Breslow, M. J., Sangster, R. L., Gabbard, S. M., Reynolds, R. W., Nakamoto, J. M., & Tarnai, J. (1996). Translating survey questionnaires: Lessons learned. *New Directions for Evaluation*, *1996*(70), 93–104. https://doi.org/10.1002/ev.1037

Muñiz, J., & Bartram, D. (2007). Improving international tests and testing. *European Psychologist*, *12*(3), 206–219. https://doi.org/10.1027/1016-9040.12.3.206

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

Nájera Catalán, H. E. (2019). Reliability, population classification and weighting in multidimensional poverty measurement: A Monte Carlo study. *Social Indicators Research*, *142*(3), 887–910. https://doi.org/10.1007/S11205-018-1950-Z/FIGURES/6

Nunnally, J. (1978). *Psychometric theory* (2nd ed). McGraw-Hill.

411 Revelle, W. (2018). *Psych: Procedures for personality and psychological research.* https://
412 CRAN.R-project.org/package=psych.
413 Rodrigues, T., Vaz, A. R., Silva, C., Conceição, E., & Machado, P. P. P. (2019). Eating
414 Disorder-15 (ED-15): Factor structure, psychometric properties, and clinical validation.
415 *European Eating Disorders Review*, *27*(6), 682–691. https://doi.org/10.1002/ERV.2694
416 Rosseel, Y. (2012). lavaan: An *R* Package for Structural Equation Modeling. *Journal of*
417 *Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02
418 Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in
419 covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables*
420 *analysis: Applications for developmental research* (pp. 399–419).
421 Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report
422 version of PRIME-MD: The PHQ Primary Care Study. *Journal of the American Medical*
423 *Association*, *282*(18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737
424 Swami, V., & Barron, D. (2019). Translation and validation of body image instruments:
425 Challenges, good practice guidelines, and reporting recommendations for test adaptation.
426 *Body Image*, *31,* 204–220. https://doi.org/10.1016/j.bodyim.2018.08.014
427 Tatham, M., Turner, H., Mountford, V. A., Tritt, A., Dyas, R., & Waller, G. (2015).
428 Development, psychometric properties and preliminary clinical validation of a brief,
429 session-by-session measure of eating disorder cognitions and behaviors: The ED-15.
430 *International Journal of Eating Disorders*, *48*(7), 1005–1015.
431 https://doi.org/10.1002/eat.22430
432 Trujillo-ChiVacuán, E. M., Pérez, L., Hutchinson-Segura, K., Leal-Alanis, E., González-Garza,
433 P., Guzmán-Arámburo, A., … Compte, E. J. (2025). Investigating the factor structure of
434 the Eating Disorder Examination Questionnaire (EDE-Q) in a clinical sample of patients
435 with eating disorders in Mexico. *Eating Disorders*, 1–17.
436 https://doi.org/10.1080/10640266.2025.2475264

Tylka, T. L., & Wood-Barcalow, N. L. (2015). The Body Appreciation Scale-2: Item refinement and psychometric evaluation. *Body Image*, *12*, 53–67. https://doi.org/10.1016/J.BODYIM.2014.09.006

Villalobos-Hernández, A., Bojórquez-Chapela, I., Hernández-Serrato, M. I., & Unikel-Santoncini, C. (2023). Prevalencia de conductas alimentarias de riesgo en adolescentes mexicanos: Ensanut Continua 2022. *Salud Publica de Mexico*, *65*. https://doi.org/10.21149/14800