Evaluation of the psychometric performance of the SCOFF questionnaire in a Mexican young adult sample

Omar Sánchez-Armass, PhD,(1) Flavia Cristina Drumond-Andrade, PhD,(2) Angela R Wiley, PhD,(2) Marcela Raffaelli, PhD,(2) Celia Aradillas-García, PhD,(1) UP AMIGOS 2008 Study Group*

(1) Facultad de Psicología, Universidad Autónoma de San Luis Potosí. San Luis Potosí, Mexico.
(2) University of Illinois at Urbana-Champaign. Champaign IL, USA.
* Study group members: Carmen Rojas, Eduardo Medina Cerda, Esperanza de la Cruz Mendoza, Juan Manuel Vargas Morales (Universidad Autónoma de San Luis Potosí, México) y Margarita Terán García (University of Illinois at Urbana-Champaign, USA).

Abstract

Objective. To examine the psychometric performance of the SCOFF, a brief screening instrument for eating disorders (ED). Materials and methods. Mexican university applicants (n = 3594, 55.7% female, M age = 18.1 years) completed self-report measures and a health screen. Results. Confirmatory factor analyses revealed one factor for females. However a bifactor model fits better for males and females. Reliability was lower for females (KR20 = .49) than males (KR20 = .59). More females (24.2%) presented risk for ED (SCOFF > 2) than males (11.2%). Nomological validity indicated that risk for ED in young women was associated with demographic (e.g., parental education), psychological (e.g., depression, weight management efficacy), physical (e.g., BMI), and social (e.g., family conflict) indicators in conceptually coherent ways. Fewer variables were significant for males. Conclusion. Although the SCOFF may be a useful ED screen in Mexico, further research must examine its criterion validity, sensitivity, and specificity.

Key words: eating disorders; psychometrics; students; Mexico

Resumen

Objetivo. Examinar el desempeño psicométrico del SCOFF, un instrumento de tamizaje para trastornos de conducta alimentaria (TCA). Material y métodos. Aspirantes mexicanos a una universidad (n = 3594, 55.7% mujeres, edad M = 18.1 años) completaron cuestionarios y una revisión médica. Resultados. Análisis factoriales confirmatorios revelaron un factor para mujeres, aunque un modelo bifactorial funcionó mejor para hombres y mujeres. La fiabilidad fue menor en mujeres (KR20 = .49) que en hombres (KR20 = .59). Las mujeres (24.2%) presentaron mayor riesgo de TCA (SCOFF > 2) que hombres (11.2%). Validez nomológica indicó que el riesgo de TCA en mujeres jóvenes estuvo asociado con indicadores demográficos (e.g., educación paterna), psicológicos (e.g., depresión, control eficaz de peso), físicos (e.g., IMC), y sociales (e.g., conflicto familiar) de forma conceptualmente coherente. Un subconjunto de estas variables fueron significativas para los hombres. Conclusión. El SCOFF podría ser un tamizaje de TCA útil en México. Sin embargo, se requiere más investigación sobre su validez de criterio, sensibilidad y especificidad.

Palabras clave: trastornos alimenticios; psicometría; estudiantes; México
Eating disorders (ED)-disruptions in normal eating patterns such as self-starvation, unhealthy dieting, or binge/purge cycles-have been reported around the world. Estimates of clinical prevalence rates worldwide range from 0.02 - 2.2% for anorexia nervosa (AN) and 0.7 - 5.4% for bulimia nervosa (BN); rates of eating disorder not otherwise specified (EDNOS; a diagnosis for people who meet some criteria for AN or BN) are between 3 - 5%. Although males and females of all ages may be affected by ED, most cases occur in young women. ED represent a threat to long-term physical and psychosocial well-being. ED are persistent, with symptoms lasting up to 10-15 years for 12 - 14% of cases, and relapse rates of 4 - 27%.

ED have been diagnosed primarily in the so-called developed countries, but there is evidence that ED are increasingly common around the world. There is growing recognition of ED in Mexico, with the prevalence of risky behaviors increasing in students aged 12-19 in Mexico City from 1997 to 2003. Prevalence rates differ across studies with different samples and measures. A nationally representative study reported a BN rate of 0.6% among men and 1.8% among females aged 18-65, but no cases of AN. A review of previous studies reported ED related behaviors in 5 - 18% of Mexican female adolescents, university and high school students. Given the prevalence and related complications of ED, early detection and effective treatment can decrease severity and improve prognosis.

One barrier to early ED detection in Mexico is a lack of brief screening tools. A review of Mexican studies highlights disagreement about the instruments used to identify ED. Common diagnostic tools (e.g., EAT-26, EDI) are too lengthy for primary care settings with large, non-clinical samples considering time, convenience, and cost. Unikel and colleagues developed a brief screening instrument for assessing risky eating behaviors for the Mexican population. The 10-item instrument has high reliability and predictive validity, but has only been used in Mexico and Colombia, limiting international comparisons. Furthermore, although relatively brief, its length still may deter use in some settings.

This study focuses on the widely used SCOFF, introduced by Morgan and colleagues as a simple screening tool for ED. The SCOFF has five questions with good predictive validity. It has been translated and adapted to multiple countries, including Finland, France, Italy, Spain, China, Japan, Korea, Colombia, and Israel. These studies have shown that the SCOFF is a valuable screening tool. We expand on prior work in several ways. First, we examined not only the SCOFF’s reliability but also its factor structure with exploratory (EFA) and confirmatory factor analyses (CFA). Three prior studies examined the SCOFF’s factor structure, with inconsistent results. With a Colombian sample, Campo-Arias et al reported a single factor structure for both males and females; Muro-Sans et al found a two factor structure for females and a unifactorial solution for males in a Spanish sample. Hautala et al, using CFA to test the one and two factor models in Finland, reported good fit for both, with the two factor structure performing slightly better for females.

Second, we examined the SCOFF’s nomological validity (the extent to which it correlates in theoretically predicted ways with measures of related constructs), since the larger study from which our data were derived did not include other eating disorder measures. In these analyses, we evaluated associations between the SCOFF and a set of conceptually-identified factors related to ED in past research. These include indicators of psychological well-being (e.g., depression, stress, substance use), weight-management efficacy, conflict with parents about weight and health lifestyle, and general health (sedentary behavior and body mass index [BMI]). To our knowledge, ours is the first study to evaluate the SCOFF in a Mexican sample, representing a potential contribution to future research and practice.

Materials and methods

Sample

The sample included 3 594 young adults aged 16 to 21 years (M age = 18.1, SD = 1.1; 55.6% female) who applied to the Universidad Autónoma de San Luis Potosí in Mexico in 2008. The larger ongoing project involves annual cross-sectional data collection. The study is Institutional Review Board approved and participants provided written consent.

Measures

Respondents completed self-report questionnaires and a physical examination. Validated measures were used when available. Other measures were adapted following standard translation/back-translation procedures. This study draws on a subset of measures (Table I).

Demographic characteristics. Respondents reported their gender, age (in years), and highest level of education for either parent.

Disordered eating. The SCOFF assesses five core aspects of AN and BN (see Table II for items). Each “yes” response is given 1 point, and a total score computed by summing. Scores above 2 are considered a positive
Evaluation of the SCOFF questionnaire

screen for AN or BN. The validated Spanish version of the SCOFF was used; original wording for all items was preserved.

Psychological well-being. Respondents completed depression, stress, and substance use measures. The CESD-10 assesses depressive symptoms in the past week with 4 response options from “Rarely or none of the time (less than 1 day)” to “All of the time (5-7 days).” Positive items are reversed and their values (0 to 3) are summed. The CESD-10 had excellent psychometric properties with Mexican adolescents (α = .93; in the current sample α = .79). Respondents rated their stress in the last month in 6 areas (romantic partner, family, finances, general health, weight, and “other”) from 1 = Not at all to 6 = Extremely. An overall score was computed by summing (α = .71). There were two dichotomous (0 = no, 1 = yes) substance use items: “Do you smoke?” and “Do you drink alcohol?”

Parent-child health conflicts. The Young Adult-Parent Conflicts About Lifestyle (YAPCAL) scale consists of three items assessing frequency of family conflict due to respondents’ health lifestyle (weight, exercise, and eating patterns) on a scale from 1 (Never) to 6 (Always). An overall score was created by summing; higher scores indicate a higher level of weight management self-efficacy.

Physical activity. A dichotomous item asked “Do you do regular physical activity?” (0 = no, 1 = yes). Sedentary behavior was calculated by summing the daily hours participants reported watching television and playing videogames (actual range = 0 - 14 hours).

<table>
<thead>
<tr>
<th>Table I</th>
<th>DESCRIPTIVE STATISTICS OF PARTICIPANTS AT RISK (SCOFF &gt; 2) AND NOT AT RISK (SCOFF ≤ 2) OF EATING DISORDERS OVERALL AND BY SEX. SAN LUIS POTOSI, MEXICO, 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>SCOFF ≤ 2</strong></td>
</tr>
<tr>
<td>n = 1982</td>
<td>n = 1503</td>
</tr>
<tr>
<td><strong>Continuous variables (Mean ± SD)</strong></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>18.0 ± 1.1</td>
</tr>
<tr>
<td>Depression</td>
<td>7.3 ± 5.3</td>
</tr>
<tr>
<td>Stress</td>
<td>19.8 ± 6.1</td>
</tr>
<tr>
<td>WEL – Negative emotions</td>
<td>17.3 ± 5.6</td>
</tr>
<tr>
<td>WEL – Availability</td>
<td>15.7 ± 5.1</td>
</tr>
<tr>
<td>WEL – Social pressure</td>
<td>16.5 ± 5.0</td>
</tr>
<tr>
<td>WEL – Physical discomfort</td>
<td>17.8 ± 5.3</td>
</tr>
<tr>
<td>WEL – Positive activities</td>
<td>16.5 ± 5.2</td>
</tr>
<tr>
<td>Parent-child health conflict</td>
<td>14.8 ± 3.5</td>
</tr>
<tr>
<td>Sedentary behavior (hours)</td>
<td>4.6 ± 4.4</td>
</tr>
<tr>
<td>BMI (kg/m2)</td>
<td>23.4 ± 4.5</td>
</tr>
<tr>
<td><strong>Categorical variables (percent)</strong></td>
<td></td>
</tr>
<tr>
<td>Smoke (% yes)</td>
<td>10.5</td>
</tr>
<tr>
<td>Drink (% yes)</td>
<td>18.9</td>
</tr>
<tr>
<td>Exercise (% yes)</td>
<td>57.2</td>
</tr>
<tr>
<td>Parental education – less than bachelor</td>
<td>74.4</td>
</tr>
<tr>
<td>Parental education – bachelor</td>
<td>25.6</td>
</tr>
</tbody>
</table>

Note: Significant differences between SCOFF ≤ and > 2:

* p ≤ 0.05
‡ p ≤ 0.01
BMI. Height was recorded to the nearest 0.5 centimeter and weight to the nearest 0.1 kg and BMI calculated as kg/m². Respondents were classified as underweight (<18.50), normal weight (18.50–24.99), overweight and obese (≥ 25.0).

**Data analysis**

Data analysis was conducted with R version 2.9.41**Significance level was set at 95% and kept constant across multiple comparisons with Bonferroni's correction. Item-level and factor analyses analyzed psychometric performance.**42 Point-biserial correlations were calculated to estimate each SCOFF item’s discriminatory power. Reliability was assessed with Kuder and Richardson’s formula 20 (KR20) and McDonald’s omega (ω). For the factor analyses, the male and female samples were each randomly split. A principal components exploratory factor analysis (EFA) with promax rotation was used on the first halves; the Very Simple Structure and Velicer MAP criteria were used for factor selection.43 On the second halves, maximum likelihood CFA tested previously reported models in the literature (one factor M1 and two factor solution M2, which consisted of Loss of Control Over Food [items 2, 4, and 5] and Purging Behaviors [items 1 and 3]).19,22,28

To examine nomological validity, we compared respondents with positive and negative ED screens on the conceptually-identified indicators. Univariate normality and normal distribution of residuals were assessed with the Shapiro-Wilk test.44 Residual homogeneity of variance was evaluated with the Brown-Forsythe test.45 Group mean comparisons were performed with Student’s t-test or Mann-Whitney’s U test, depending on normality and homogeneity of variance of residuals.

**Results**

Nearly one quarter (24.2%) of female respondents, and 11.2% of male respondents, scored above 2 on the SCOFF, indicating a positive screen for ED. Table I presents the descriptive statistics of participants at risk and not at risk of ED overall and by sex.

**Female participants**

Psychometric analyses. The average score on the SCOFF was 0.87 (SD = 0.99). The corrected item-total point-biserial correlations are presented in Table II. The Kaiser-Meyer-Olkin measure of sample adequacy test was 0.61, and Bartlett’s test of sphericity was significant, χ²(10) = 303.7, p < .001, indicating that the items shared enough common variance to attempt a factor analysis. EFA with 983 participants yielded a unifactorial solution with an eigenvalue of 1.68 (34% of variance explained). KR20 was 0.48 (ω = 0.62). Inter-item tetrachoric correlations ranged from 0.04 to 0.58. Results of the CFA for the second half of female participants (n = 999) are presented in Table III. Inter-item tetrachoric correlations ranged from 0.22 to 0.58. Goodness-of-fit statistics indicated a good fit for both M1 and M2. Reliability for M1 were KR20 = .48 and ω = 0.49. KR20 for M2 factors were .48 and .17 respectively, with an ω = .61.

**SCOFF Validity.** Comparisons of females with positive versus negative screens revealed multiple differences (Table I). Those with positive screens reported higher psychological distress (depressive symptoms, stress, and substance use), lower weight management efficacy in two contexts (when experiencing negative emotions and when food is available), and higher BMI.

### Table II

<table>
<thead>
<tr>
<th>SCOFF Item</th>
<th>Males rpbi</th>
<th>Females rpbi</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOFF 1. Do you make yourself sick because you feel uncomfortably full? (¿Usted se provoca el vómito porque se siente lleno(a)?)</td>
<td>0.18</td>
<td>0.23</td>
</tr>
<tr>
<td>SCOFF 2. Do you worry you have lost control over how much you eat? (¿Le preocupa que haya perdido el control sobre la cantidad de comida que ingiere?)</td>
<td>0.27</td>
<td>0.37</td>
</tr>
<tr>
<td>SCOFF 3. Do you believe yourself to be fat when others say you are too thin? (¿Cree que está gordo(a) aunque los demás digan que está demasiado delgado(a)?)</td>
<td>0.16</td>
<td>0.09</td>
</tr>
<tr>
<td>SCOFF 4. Have you recently lost more than fourteen pounds in a 3-month period? (¿Ha perdido recientemente más de 7kg en un periodo de 3 meses?)</td>
<td>0.33</td>
<td>0.36</td>
</tr>
<tr>
<td>SCOFF 5. Would you say that food dominates your life? (¿Usted diría que la comida domina su vida?)</td>
<td>0.20</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Note: Original items in Spanish in parenthesis.
They endorsed higher weight management efficacy when experiencing physical discomfort and lower levels of health-related conflict with parents. They had less educated parents than those with a negative screen. Being overweight or obese was associated with higher likelihood of endorsing items 1, 2, 3 and 5, whereas being underweight was associated with lower likelihood of endorsing items 2, 4 and 5 compared to those of normal weight (available upon request).

**Male participants**

Psychometric analyses. The average score on the SCOFF was 0.49 (SD = 0.79). Corrected item-total point-biserial correlations are presented in Table II. KMO was 0.62, and Bartlett’s test was significant, \( \chi^2(10) = 200.2, p < .001 \). EFA with 806 participants yielded a unifactorial solution with an eigenvalue of 1.64 (33% of variance explained); KR20 was 0.62 (\( \omega = 0.49 \)). Inter-item tetrachoric correlations ranged from .10 to .47. Results of the CFA for the second half of male participants (n = 806) are presented in Table III. Goodness-of-fit statistics indicated a good fit only for M2; factor loadings ranged from .30 to .46. Factor reliability coefficients were .40 and .18, respectively, with an \( \omega = .48 \).

SCOFF Validity. Males with a positive ED screen had higher levels of depression and stress, higher BMI, and were less likely to exercise than their counterparts (Table I). They reported lower levels of weight and lifestyle conflicts with their parents. The probability of endorsing items 1, 2 and 3 was higher among those with BMI > 25; whereas endorsing items 2 and 4 was lower among underweight than normal weight males (available upon request).

### Discussion

As ED become more common worldwide, effective yet brief screening tools are needed for use across countries. The SCOFF has been used widely to assess ED in developed countries, but its use in Latin America has been limited.25,26,28 To our knowledge, ours is the first study to evaluate the SCOFF’s psychometric performance in Mexico. We evaluated the internal consistency, factor structure, and nomological validity of the SCOFF in a sample of Mexican young adults. Findings provide information on the utility of the SCOFF as a screening measure in Mexico, and allow identification of future research directions.

The internal consistency of the SCOFF questionnaire was moderately low for both males (KR20 = .62) and females (KR20 = .49). Although these internal consistency coefficients are lower than generally deemed desirable, they are similar to those reported in other

<table>
<thead>
<tr>
<th>Table III</th>
<th>CONFIRMATORY FACTOR ANALYSES FOR TWO MODELS OF THE SCOFF. SAN LUIS POTOSI, MEXICO, 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>Females (n = 999)</td>
</tr>
<tr>
<td>Model</td>
<td>M1</td>
</tr>
<tr>
<td>Factor</td>
<td>F1</td>
</tr>
<tr>
<td>Item 1</td>
<td>.31</td>
</tr>
<tr>
<td>Item 2</td>
<td>.56</td>
</tr>
<tr>
<td>Item 3</td>
<td>.18</td>
</tr>
<tr>
<td>Item 4</td>
<td>.46</td>
</tr>
<tr>
<td>Item 5</td>
<td>.48</td>
</tr>
<tr>
<td>r</td>
<td>.73</td>
</tr>
<tr>
<td>( \chi^2 )</td>
<td>7.05</td>
</tr>
<tr>
<td>df</td>
<td>5</td>
</tr>
<tr>
<td>p-value</td>
<td>.22</td>
</tr>
<tr>
<td>( \chi^2/df )</td>
<td>1.41</td>
</tr>
<tr>
<td>RMSEA</td>
<td>.02</td>
</tr>
<tr>
<td>AIC</td>
<td>27.05</td>
</tr>
<tr>
<td>AGFI</td>
<td>.99</td>
</tr>
</tbody>
</table>

Note: RMSEA = root mean square error of approximation; AIC = Akaike’s information criterion; AGFI = adjusted goodness of fit index
studies (.44 -.58) that concluded the SCOFF is a valuable screening tool. 21-23,25,26 Only two studies that found lower internal consistencies (0.2 - 0.4) have concluded that the SCOFF has limited use as a screening instrument. 24,26 Because the SCOFF is short (just five items) with a simple response format (yes/no) and focuses on diverse behaviors indicative of distinct types of ED, modest reliability estimates are expected. 46

In terms of factor structure, most studies assume the SCOFF has a unidimensional factor structure. However, two recent studies 19,22 raise the possibility that the SCOFF may yield two factors reflecting different aspects of ED (control over food vs. purging behavior). Building on this work, we conducted both EFA and CFA on random halves of the male and female samples. EFA indicated only one factor but CFA indicated that a two factor solution may be preferable for males. In EFA, the explained variance for the males (33%) was similar to Muro-Sans and colleagues 22 (35%) and Campo-Arias and colleagues 28 (35%). However, explained variance for females (34%) was lower than they reported in these two studies (55 and 37, respectively). Future testing of the two models is needed to establish clinical utility of the SCOFF in Mexico.

Despite moderate reliability, the SCOFF has shown good predictive validity in studies conducted by the original developers and by other teams. 17-22,25,28 In analyses to evaluate nomological validity, respondents with positive vs. negative ED screens were found to differ in psychosocial and physical wellbeing, consistent with past research. 7,30-33 As expected, depression, stress and BMI were higher among both males and females with positive screens. Females with positive ED screens evidenced lower weight management efficacy, except in the presence of physical discomfort when they were more efficacious. This latter association may be related to the nature of ED; for example, those with more ED symptoms may have successfully endured physical discomfort while controlling their intake or purging, thus increasing their efficacy in this realm. Counter to expectation, both males and females with positive screens reported lower levels of lifestyle and weight conflict with parents. This may reflect heightened secrecy and disengagement experienced by young adults with disordered eating. 47 Taken together, these patterns of association provide preliminary evidence for the nomological validity of the SCOFF in our sample of Mexican young adults. Future studies that evaluate criterion validity (e.g., by comparing the SCOFF to established measures of ED) would provide a stronger test of the SCOFF’s utility for use in Mexico.

Although not a central goal of the study, findings confirm that Mexican women are more likely to have ED than men, consistent with gender differences reported in other populations. 17,29 Twice as many females than males had a positive screen for ED (24.2% vs. 11.2%). The ratio was more balanced than may be expected, given published gender ratios that vary from 2:1 - 4:1 in early adolescence 25,48 to 10:1 - 20:1 in adulthood 20; these findings, however, are consistent with other studies involving university students. 49,50 Prevalence of positive ED screens among female applicants to UASLP were lower than those reported in SCOFF studies conducted in Korea (34.8%) 24 and Colombia (39.4%). 25 However, rates were considerably higher than reported among 10-19 year olds in the 2006 Mexican National Health and Nutrition Survey (0.4% among males and 1.0% among females), 51 and in a study of Mexican female university students. 52 These differences likely reflect variations in age groups, measurement approaches (e.g., use of screeners vs. diagnostic instruments, instrument sensitivity and specificity), and criteria for classifying risk level for ED. Nevertheless, findings are consistent with the 2006 Mexican National Survey on Health and Nutrition, which reported that 17.3% of adolescent males and 25.8% of adolescent females reported at least one of the following: worrying about gaining weight, binge eating, loss of control, dieting, fasting, or excessive exercising. Considering the content of the SCOFF items, it is understandable that the rates of risky eating behaviors are similar. 53

This paper adds to the literature assessing the psychometric performance of the SCOFF in different populations. The study had a number of strengths (e.g., large sample size, multiple correlates of ED). Future research can address study limitations. Established ED measures (e.g., EAT-26, 13 EDI 14 ) should be used to assess the SCOFF’s criterion validity, specificity and sensitivity in clinical and non-clinical samples. In addition, country-specific research is needed to establish whether the specified SCOFF cut-off of 2 is appropriate for the Mexican context. Finally, future examinations of correlates of the SCOFF should utilize more nuanced and validated measures of health behaviors.

ED are recognized as a growing public health concern in Mexico. Given the potentially serious health consequences of ED, there is a need for brief screening measures for use in general populations. The current study of Mexican college applicants revealed psychometric properties that were largely comparable to published reports from other countries. 21-23,25,28 Given the inconsistency of some findings, additional examination of the SCOFF’s psychometric properties and predictive utility is needed to establish its usefulness as a general screening tool. In the meantime, the use of instruments, such as the Brief Disordered Eating Questionnaire, 9...
which has been shown good reliability and predictive validity in Mexican samples, should be used whenever time permits. Ultimately, a better understanding of ED in Mexico could inform the development of health education programs.

Declaration of conflict of interests: The authors declare no conflict of interests.

References


