Published in final edited form as:

J Prosthodont Res. 2018 October ; 62(4): 456–461. doi:10.1016/j.jpor.2018.05.003.

Development, Validity and Reliability of the Orofacial Esthetic Scale – Spanish Version

Miguel Simancas-Pallares, DMD, MSc¹, Mike T. John, DMD, PhD², Swati Prodduturu, BDS, MPH³, William A. Rush⁴, Christopher J. Enstad⁵, and Patricia Lenton, BSDH, MA, CCRP⁶

¹Dental Research Department. University of Cartagena. Cartagena, Colombia. **Postal Address:** 30th St # 48-152. Health Sciences Campus, Zaragocilla. Suite 301. Faculty of Dentistry. **Phone:** +57 5 6698172, line 110

²Department of Diagnostic and Biological Sciences, University of Minnesota, Minneapolis, MN, USA. **Postal Address:** 515 Delaware St SE. **Phone:** 601-896-5861

³Department of Diagnostic and Biological Sciences, University of Minnesota, Minneapolis, MN, USA. 515 Delaware St SE, Rm 7-546, Minneapolis, MN 55455 **Phone:** 601-896-5861

⁴HealthPartners Institute for Education and Research, Minneapolis, MN, USA. **Postal Address:** 3311 E Old Shakopee Rd, Bloomington, MN 55440-1524. **Phone:** 952-967-5007

⁵HealthPartners Institute for Education and Research, Minneapolis, MN, USA. **Postal Address:** 3311 E Old Shakopee Rd, Bloomington, MN 55440-1524. **Phone:** 952-967-5007

⁶Oral Health Clinical Research Clinic, University of Minnesota, Minneapolis, MN, USA. **Postal Address:** 3M Center, 270-2S-18, St. Paul, MN 55144-1000 United States **Phone:** 651-737-9541

Abstract

Purpose—To develop a Spanish version of the Orofacial Esthetic Scale (OES-Sp) and to determine its psychometric properties in dental patients.

Methods—We performed a validation study for OES development and validation into Spanish. In the development phase, eight individuals participated in the translation process (cross-cultural adaptation) following the PROMIS methodology. In the validation phase, 331 dental patients (age mean±sd: 42.9±12.3 years, 59% female) from the HealthPartners dental clinics in Minnesota, USA responded to a survey of Spanish speaking patients (N=540, response rate: 61%). We first explored dimensionality by means of exploratory factor analysis and scree plot, and then we computed reliability measures with the Cronbach's Alpha statistic. Finally, we assessed convergent validity by computing Pearson/Spearman rank correlations between OES-Sp and OHIP-based Orofacial Appearance measures. All statistical procedures were performed using Stata v.13 for Windows (StataCorp).

Corresponding Author: Miguel Simancas-Pallares. Dental Research Department. University of Cartagena. 30th St # 48-152. Health Sciences Campus, Zaragocilla. Suite 301. Cartagena, Colombia. Postal Code: 130011. Colombia. msimancasp@unicartagena.edu.co.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Results—Exploratory factor analysis provided evidence that a single factor represents the Spanish OES version. Score reliability was high with Cronbach's Alpha statistic of 0.95 (lower limit of the 95% CI: 0.94). Score validity was sufficient indicated by Pearson and Spearman rank correlations between -0.53 and -0.69 (all 95% confidence intervals: less than ± 0.10).

Conclusions—OES is a valuable instrument to measure the Orofacial Appearance construct for Spanish-speaking populations. OES was concluded to be unidimensional with excellent score reliability and sufficient convergent score validity when compared with other Orofacial Appearance measures. Our results suggest that OES could be used in Spanish-speaking patients, an important and growing population around the globe.

Keywords

validation studies; psychometrics; dental esthetics; factor analysis

1. Introduction

Orofacial Appearance (OA) has become an important component within the general (health-related) quality of life and the specific oral health-related quality of life (OHRQoL) concepts [1]. Consequently, it is one of the four OHRQoL dimensions [2,3]. The increased demand of orthodontic, prosthetic and dental cosmetic treatments suggest that patients consider these issues relevant to their psychological well-being [4]. Several studies show how important orofacial esthetics are; including, how important facial appearance and profile is from the Hispanic/Latino patient's perspective [5–7]. In addition to this, orofacial esthetics constitute one of the most common reasons for patients looking for prosthodontic treatment. Thus being correlated with patient's self-perceived satisfaction [8]. Restorative treatment is also able to change the appearance of teeth, gums and jaws [9]. Some research shows that quality of life and self-confidence assessments were greater in patients who were pleased with their restorations [10].

Previously developed and validated questionnaires like, the General/Geriatric Oral Health Assessment Index (GOHAI) [11], the Oral Health Impact Profile (OHIP) [12] and the Oral Impacts on Daily Performances (OIDP) [13], have been proposed to measure OHRQoL. To this end, they include items to measure dental, oral, and facial appearance, but they don't summarize the specific appearance-related impact with a score. By using GOHAI, only two out of twelve items are able to quantify Orofacial Appearance impacts. With the most used instrument, the OHIP, only 3 out of 49 questions address Orofacial Appearance directly [8], and when we use OIDP, only two items do this. Consequently, the need for instruments targeting Orofacial Appearance specifically is an obvious necessity, not only from the perspective of government public health service needs, but also for clinicians since they need valid and reliable instruments to quantify, for example, the impact of dental treatments on the Orofacial Appearance construct.

To address this gap of not having specific instruments to quantify OA, in 2010 Larsson et al. [8] developed the Orofacial Esthetic Scale (OES). OES items ask patients how they feel about their face, facial profile, mouth, tooth alignment, tooth shape, tooth color and gums and the eighth question requests an overall impression of OA. The scale was originally

developed in Swedish and was accompanied by a translation into English. Since its original development, it has been translated into several languages including Croatian, Chinese, Albanian, German and Dutch while showing consistent validity and reliability measures [14–18].

When originally tested in prosthodontic patients, OES seemed to discriminate adequately between esthetically impaired patients and healthy controls and showed good convergent validity, satisfactory internal consistency, as well as excellent test-retest reliability [19]. Factor analyses provided support that a single score can sufficiently characterize the latent trait OE. OES has also been validated in general population subjects, showing adequate psychometric properties and suggesting that the extension of its use beyond prosthodontic patients appears to be justified [9].

With more than 20 countries having primarily Spanish-speaking population and more than 414 million people speaking Spanish worldwide [20], Spanish is the second most widely spoken native language, even surpassing English as a native language. Also, because the Spanish language has been one of the most rapidly growing languages in recent years [20], there is a marked increase in the need to be able to measure OE in Spanish. Based on the research and clinical care needs that involve measurement of OE and the increase in number of multinational research projects requiring globally applicable instruments, available questionnaires that have been developed in English-speaking countries need to be accompanied by cross-culturally validated OE instruments in Spanish. Thus far, it has been needed, but has not been made available.

The aim of this study was to develop a Spanish version of the Orofacial Esthetic Scale (OES-Sp) and to determine its psychometric properties in dental patients.

2. Materials and methods

2.1 Orofacial Esthetic Scale

The scale was originally developed in Swedish (OES-S) [8]. To facilitate international application the authors provided an English translation as well. The questionnaire has seven questions assessing the appearance of the face, mouth and teeth. Item response options are available on an 11-point numeric rating scale, ranging from 0, “Very dissatisfied”, to 10, “Very satisfied.” A summary score is obtained by summing items one to seven. The lowest score is 0 (worst OA) and the highest score is 70 (best OA). In addition to the seven items, an eighth item is judged separately ranging between 0–10 on the numeric rating scale, providing a global assessment of Orofacial Appearance.

2.2 Development of a Spanish Orofacial Esthetic Scale

2.2.1 Individuals participating in the scale development—Eight individuals participated in the translation process. The majority came from diverse Spanish-speaking countries (Spain, N=3; Peru, N=1; Venezuela, N=3). One participant was a native English-speaker who was fluent in Spanish. Most of our participants involved in the translation process were dental professionals (N=5), well versed with the colloquial and formal verbiage of orofacial structures.

2.2.2 Translation Process—PROMIS (Patient-Reported Outcomes Measurement Information System) methodology provides a universal approach for scale translation and harmonization across all varieties of a particular language [21]. This results from an iterative process of forward- and back-translations, bilingual expert review, and pre-testing with cognitive debriefing (linguistic validation). Harmonization across all language varieties and a universal approach to translation are key factors guiding the process. We adopted this approach as it is considered a minimum standard for ensuring accurate and culturally appropriate translations, as well as equivalence between the languages [22].

PROMIS advocates the usage of FACIT (Functional Assessment of Chronic Illness Therapy) methodology for the translation and cross-cultural validation of health status questionnaires [23–25]. This methodology, an 11-step procedure, is also consistent with the guidelines recommended by the International Society for Pharmacoeconomic and Outcomes Research (ISPOR) for translation of Patient Reported Outcomes (PRO) instruments [24]. Table 1 presents the steps and a brief description of the methods.

2.2.3 Assessment of score reliability and validity of Spanish version of the Orofacial Esthetic Scale in dental patients

2.2.3.1 Study participants: Starting in July 2014, patients who were scheduled for dental appointments were approached to participate in the study. A consecutive sample of patients was targeted at each participating HealthPartners dental clinic. After contacting the patients by phone, a battery of patient-reported outcome measures, including the Orofacial Esthetic Scale, was sent to the patients. Patients filled the information at home and sent the questionnaires back to the HealthPartners Institute. They received a compensation of \$50. Recruitment finished in April 2016. HealthPartners is a large upper Midwestern integrated health system that consists of a dental group, a medical group (including hospitals), a dental plan and a medical plan. HealthPartners medical group serves more than one million patients and HealthPartners dental group serves over 200,000 patients. The language indicator in the electronic dental record or in the electronic medical record identified both English and Spanish speakers. Spanish speakers could also be bilingual patients. These bilingual patients were identified by checking their country of origin in the electronic health record. The patient was considered bilingual if their language was English or missing and their country of origin was from the list of countries we identified as Spanish: Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Spain, Uruguay, and Venezuela. Bilingual patients were offered an English version instead of Spanish if they preferred. In this manner bilinguals self-selected English or Spanish.

The study was reviewed and approved by the Institutional Review Board of the HealthPartners Institute (Study Number A11-136). To investigate the psychometric properties of the OES-Sp, only the Spanish-speaking patients were included. Patients (N=540) were invited to the study and 331 provided information (response rate 61%, age mean±sd: 42.9±12.3 years, 59% female). Not all patients provided complete information for the OES and the Oral Health Impact Profile (OHIP), an oral health-related quality of life instrument used for validation of OES scores. When patients had more than one missing

OES item (N=20), they were dropped from the analyses, leaving 311 patients for analysis. Missing OES items were imputed using item median imputation.

2.2.3.2 Dimensionality: To determine OES-Sp dimensionality, we performed a parallel analysis [26]. This technique calculated a 7 OES items x 7 OES items sample correlation matrix from the 305 patients and 7 OES items dataset. Then, we created a scree plot by plotting the eigenvalues of the sample correlation matrix against their position from largest to smallest (1, 2, ..., 7). The eigenvalue points were connected with a straight line. This technique also created a simulated dataset with 305 observations randomly sampled from 7 independent normal variates. It calculated a 7 x 7 correlation matrix for the simulated data, extracted the 7 eigenvalues, and ordered them from largest to smallest. This step was repeated k times (in our case, $k = 1000$). From 1000 eigenvalues at position 1, the median was calculated. A median of eigenvalues was also calculated for position 2, 3, ..., 7. The 7-eigenvalue medians were connected with a dashed line and overlaid on the scree plot. Finally, the intersection of the solid (actual data) and the dashed lines (simulated data) was the cutoff for determining the number of dimensions present in the data.

2.2.3.3 Reliability: To assess OES-Sp score reliability, an internal consistency analysis was performed by computing Cronbach's Alpha [27]. Alpha values were interpreted according to Kline: acceptable (0.60 – 0.70); good (0.70 – 0.90); excellent (0.90) [28]. In addition, for each item, the item-rest correlations were computed. In a reliable scale, these values should be higher than 0.30 [29].

2.2.3.4 Validity - Convergent Validity: To assess OES-Sp score convergent validity, i.e., the degree of how one measure is related to another measure of the same construct, we calculated one Pearson correlation coefficient and five Spearman's rank correlation coefficients. The first coefficient (Pearson) correlated OES summary scores with the OES' eighth item, the overall Orofacial Appearance impression. Next, OES summary scores were correlated with items of the OHRQoL Orofacial Appearance dimension [2,3]. As performed for other OES language versions [18,30] with the goal to compare our results with the literature, rank correlations between three core Orofacial Appearance indicators of OHIP or their sum on one hand and the OES-Sp summary score on the other hand were computed. These OHIP items were:

1. Item 3, "Have you noticed a tooth that does not look right?"
2. Item 22, "Have you felt uncomfortable about the appearance of teeth, mouth or dentures?"
3. Item 31, "Have you avoided smiling because of problems with your teeth, mouth or dentures?"

For the correlation between the OES summary score and the OES' eighth item-the overall appearance impression, we expected a positive correlation. For OES, larger scores mean better Orofacial Appearance, but for OHIP, larger scores/responses indicate more problems. We expected negative correlations for the OHIP-based validity measures. For all validity analyses, we expected that the correlation coefficients should be "large", i.e., higher than 0.50, according to Cohen [31]. Patients (N=2) with more than one missing item in the OHIP

Orofacial Appearance dimension were dropped from the analysis, leaving 309 patients for the convergent validity analysis. Missing OHIP items were imputed using median imputation.

All of the statistical procedures were performed using Stata v.13 for Windows (StataCorp, College Station, TX, USA).

3. Results

3.1 Translation Process

OES items typically refer to structures of the teeth, mouth, jaws, and face. Major challenges translating the names of the structures were not present, and the 11-step translation procedure could be successfully performed without major deviations from the procedure recommended by PROMIS (Table 2). The final OES-Sp version is presented in Table 3.

3.2 Characterization of study participants' Orofacial Appearance

Study participants were typically satisfied with the appearance of their face, mouth, and teeth (Table 4). Mean values for OES items ranged from a highest value of 7.6 ("Your facial appearance") to a lowest value of 6.1 ("Color of your teeth"). Median values were even higher than mean values. Only a small proportion of study participants were "very dissatisfied", but a considerable proportion of study participants were "very satisfied" with their orofacial appearance, indicating that the item distributions were skewed to better appearance. Considerable differences among items in their distributions were not observed. When subjects rated the overall assessment of their appearance, responses were very similar compared to individual items.

3.3 Dimensionality

Only one eigenvalue of the correlation matrix was substantially larger than 1, and the scree plot presented a large difference between the first and all other eigenvalues as a sharp drop after the first factor, indicating that one dominant factor underlay OES items (Fig. 1). When observed eigenvalues were compared with simulated eigenvalues (parallel analysis), findings also suggested the presence of one dimension.

3.4 Reliability

Cronbach's alpha revealed that the scale's overall internal consistency was excellent ($\alpha=0.95$, lower limit of 95% CI: 0.94). Cronbach's alpha did not change substantially when a particular item was deleted from the scale (Table 5).

OES items correlated well with each other, indicated by an average inter-item correlation of 0.73. The individual items also correlated well with the scale. All item-rest correlations (the correlation of each item with the summed index but excluding that item from the index) were 0.79 and higher.

3.5 Convergent Validity

OES summary scores correlated in the expected direction (a positive value for the OES overall appearance impression and negative values for OHIP-based validity measures) and with the expected magnitude (all coefficients absolute magnitude 0.53) for the six validity measures (Table 6). According to Cohen, all effects were “large.”

The highest correlation (0.90) was observed for the OES overall appearance impression. OES summary scores’ correlation coefficients with the OHIP-based measures were between –0.53 and –0.69, indicating that a higher OES-Sp score (better appearance) was negatively related to more OHIP appearance problems. Confidence intervals indicated good precision of the coefficients.

4. Discussion

This is the first study that evaluated psychometric properties (validity and reliability) of the Orofacial Esthetic Scale in Spanish-speaking patients. The OES was successfully translated and validated into Spanish. This translation process was completed satisfactorily (with minor adjustments) and the convergent validity and internal consistency were excellent.

OES was originally developed for assessment of esthetical concerns in prosthodontic patients [8]. However, two years later its use was extended to the general population, showing adequate psychometric properties as well [9]. In the present study, we studied dental patients in general and results were in line with previous findings. This is partially explained due to the fact that prosthodontic patients are a large subgroup of dental patients. They usually suffer from tooth loss, resulting from the two major oral diseases, caries and periodontitis, and their treatments are not performed in isolation from other frequently performed treatments such as extractions, fillings, and endodontic treatments. Therefore, prosthodontic patients may be regarded as “typical” dental patients [32].

A fundamental psychometric property of an instrument is its dimensionality, i.e., the number of scores that represent the target construct well. Exploratory factor analysis (EFA) is often used to determine dimensionality or structural validity. It summarizes the information contained in all items with a smaller number of latent variables called “factors.” These factors are then interpreted as the latent dimensions of a test [32].

Using EFA, OES consistently demonstrated a one-dimensional construct since its development in Swedish, with later confirmation in English, Croatian, Chinese, Dutch, German, Albanian, and Arabic studies [8,14–18,33]. Our EFA agreed with these previous findings, suggesting that one OES-Sp score can be applied to assess self-perceived Orofacial Appearance in Spanish-speaking populations and further confirming OES’s cross-cultural comparability.

How consistently scores measure the construct is characterized by score reliability. We used the Cronbach’s Alpha statistic to investigate the internal consistency of scores, i.e., how similar the items are for measuring OA. Computed estimates ($\alpha=0.95$) showed consistent

results compared with the previous OES validated versions. The Croatian, Chinese and Albanian versions showed very similar values of $\alpha=0.96$ [14–16].

Regarding evidence on whether the OES measures what it purports to measure, i.e., score validity, we had also observed similar findings when compared to previously validated OES versions. When comparing OES-Sp scores and OHIP items separately, we found correlation coefficients between -0.51 and -0.63 , which indicate that the higher the OES score, the lower the OHIP item score; as expected due the nature of OHIP and OES item wording. These findings are also consistent with previous results of OES Swedish, German, and Albanian validations [16,17,19]. When we generated an OHIP Orofacial Appearance score and computed its correlation with the overall OES-Sp score, we also obtained the hypothesized result ($r=-0.67$), indicating an adequate convergent validity of the proposed OES Spanish version. These results are also similar to previously validated versions [14,16–19].

This study has strengths and limitations. Our response rate was 61% and therefore, how our study participants generalize to the target population could be affected by bias. For studies that investigate correlations, such as ours, non-response bias is typically not a substantial threat because there is often no reason to assume that the correlation between, for example, the OES summary score and the OHIP Orofacial Appearance score, differs between subjects who participated in the study and subjects who refused. Our sample size is considered not small and consequently, the statistical uncertainty around our validity and reliability results is not large. In particular, for validation studies with exploratory factor analysis, an item/patient ratio of 1:10 has been suggested [34]. In our study, we included three times more patients per item, suggesting robust score dimensionality results. We focused on basic psychometric properties and other characteristics such as the instrument's responsiveness to change; a score change should reflect a change exists in the underlying construct in a longitudinal assessment [17]. Such very clinically oriented studies, including the assessment of response shift (change in the meaning of one's evaluation of a construct as a result of a change in one's internal standards of measurement, a change in one's values, or a change in one's definition of the construct) [35,36], are certainly a recommended next step in the evaluation of the OES. Similarly, targeting an improved clinical use or investigating a suggested 5-point ordinal response format as suggested by previous authors [14] could be further steps. Finally, we did not perform a confirmatory factor analysis (CFA), which tests the relationships between the observed items and the latent factors. Only two OES studies have performed CFA [9,15], and we intend to be comparable with the majority of previous dimensionality analyses.

5. Conclusion

Orofacial Appearance is a dimension of construct Oral Health-Related Quality of Life and as such an important area of patient concern. It represents a major target for dental inventions. The Spanish version of the OES questionnaire measures this construct with good psychometric properties. They are very similar compared to other language versions, supporting that OES is a suitable instrument not only applicable for Spanish-speaking

populations, but is also able to assess Orofacial Appearance in clinical and research settings globally.

References

1. Isiekwe GI, Sofola OO, Onigbogi OO, Utomi IL, Sanu OO, daCosta OO. Dental esthetics and oral health-related quality of life in young adults. *Am J Orthod Dentofacial Orthop*. 2016; 150:627–36. [PubMed: 27692420]
2. John MT, Reissmann DR, Feuerstahler L, Waller N, Baba K, Larsson P, et al. Exploratory factor analysis of the Oral Health Impact Profile. *J Oral Rehabil*. 2014; 41:635–43. [PubMed: 24909881]
3. John MT, Feuerstahler L, Waller N, Baba K, Larsson P, elebi A, et al. Confirmatory factor analysis of the Oral Health Impact Profile. *J Oral Rehabil*. 2014; 41:644–52. [PubMed: 24909797]
4. Perši S, elebi A. Influence of different prosthodontic rehabilitation options on oral health-related quality of life, orofacial esthetics and chewing function based on patient-reported outcomes. *Qual Life Res*. 2015; 24:919–26. [PubMed: 25294397]
5. Mejia-Maidl M, Evans CA, Viana G, Anderson NK, Giddon DB. Preferences for facial profiles between Mexican Americans and Caucasians. *Angle Orthod*. 2005; 75:953–8. [PubMed: 16448237]
6. Sanders AE. A Latino advantage in oral health-related quality of life is modified by nativity status. *Soc Sci Med*. 2010; 71:205–11. [PubMed: 20434250]
7. Born CD, Divaris K, Zeldin LP, Rozier RG. Influences on preschool children's oral health-related quality of life as reported by English and Spanish-speaking parents and caregivers. *J Public Health Dent*. 2016; 76:276–86. [PubMed: 26990804]
8. Larsson P, John MT, Nilner K, Bondemark L, List T. Development of an Orofacial Esthetic Scale in prosthodontic patients. *Int J Prosthodont*. 2010; 23:249–56. [PubMed: 20552092]
9. John MT, Larsson P, Nilner K, Bandyopadhyay D, List T. Validation of the Orofacial Esthetic Scale in the general population. *Health Qual Life Outcomes*. 2012; 10:135. [PubMed: 23158767]
10. Ellis JS, Pelekis ND, Thomason JM. Conventional rehabilitation of edentulous patients: the impact on oral health-related quality of life and patient satisfaction. *J Prosthodont*. 2007; 16:37–42. [PubMed: 17244306]
11. Atchison KA, Dolan TA. Development of the Geriatric Oral Health Assessment Index. *J Dent Educ*. 1990; 54:680–7. [PubMed: 2229624]
12. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health*. 1994; 11:3–11. [PubMed: 8193981]
13. Adulyanon S, Sheiham A. Oral Impacts on Daily Performances. In: Slade GD, editor *Measuring Oral Health and Quality of Life*. 1. Chapel Hill, N.C: The University of North Carolina at Chapel Hill; 1997. 151–8.
14. Perši S, Milardovic S, Mehulic K, elebi A. Psychometric properties of the Croatian version of the Orofacial Esthetic Scale and suggestions for modification. *Int J Prosthodont*. 2011; 24:523–33. [PubMed: 22146251]
15. Zhao Y, He SL. Development of the Chinese version of the Oro-facial Esthetic Scale. *J Oral Rehabil*. 2013; 40:670–7. [PubMed: 23829233]
16. Bimbashi V, elebi A, Staka G, Hoxha F, Perši S, Petricevic N. Psychometric properties of the Albanian version of the Orofacial Esthetic Scale: OES-ALB. *BMC Oral Health*. 2015; 15:97. [PubMed: 26306391]
17. Reißmann DR, Benecke AW, Aarabi G, Sierwald I. Development and validation of the German version of the Orofacial Esthetic Scale. *Clin Oral Investig*. 2015; 19:1443–50.
18. Wetselaar P, Koutris M, Visscher CM, Larsson P, John MT, Lobbezoo F. Psychometric properties of the Dutch version of the Orofacial Esthetic Scale (OES-NL) in dental patients with and without self-reported tooth wear. *J Oral Rehabil*. 2015; 42:803–9. [PubMed: 26037598]
19. Larsson P, John MT, Nilner K, List T. Reliability and validity of the Orofacial Esthetic Scale in prosthodontic patients. *Int J Prosthodont*. 2010; 23:257–62. [PubMed: 20552093]
20. Lewis MP, Simons GF, Fennig CD. *Ethnologue: Languages of the World*. 20. Dallas, TX: SIL International, Global Publishing; 2017. 505

21. PROMIS. Instrument development and psychometric evaluation scientific standards. Bethesda: National Institutes of Health; 2012.
22. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care*. 2007; 45:S3–S11.
23. Bonomi AE, Cella DF, Hahn EA, Bjordal K, Sperner-Unterweger B, Gangeri L, et al. Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. *Qual Life Res*. 1996; 5:309–20. [PubMed: 8763799]
24. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005; 8:94–104. [PubMed: 15804318]
25. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof*. 2005; 28:212–32. [PubMed: 15851774]
26. Zwick WR, Velicer WF. Factors Influencing Four Rules For Determining The Number Of Components To Retain. *Multivariate Behav Res*. 1982; 17:253–69. [PubMed: 26810950]
27. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951; 16:297–334.
28. Kline P. The handbook of psychological testing. 2. New York, NY: Routledge; 2000. 744
29. Field A. Discovering statistics using IBM SPSS Statistics. 4. London, London: Sage publications; 2009. 952
30. Larsson P, John MT, Nilner K, List T. Normative values for the Oro-facial Esthetic Scale in Sweden. *J Oral Rehabil*. 2014; 41:148–54. [PubMed: 24372184]
31. Cohen J. Statistical power analysis for the behavioral sciences. 2. Hillsdale, NJ: Erlbaum; 1988. 381–9.
32. John MT, Reissmann DR, Feuerstahler L, Waller N, Baba K, Larsson P, et al. Factor analyses of the Oral Health Impact Profile - overview and studied population. *J Prosthodont Res*. 2014; 58:26–34. [PubMed: 24440578]
33. Alhajj MN, Amran AG, Halboub E, Al-Basmi AA, Al-Ghabri FA. Development, validation and psychometric properties of the Arabic version of the Orofacial Esthetic Scale: OES-Ar. *J Prosthodont Res*. 2016
34. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*. 2005; 10:1–9.
35. Reißmann DR, Remmler A, John MT, Schierz O, Hirsch C. Impact of response shift on the assessment of treatment effects using the Oral Health Impact Profile. *Eur J Oral Sci*. 2012; 120:520–5. [PubMed: 23167468]
36. Reißmann DR, John MT, Feuerstahler L, Baba K, Szabó G, elebi A, et al. Longitudinal measurement invariance in prospective oral health-related quality of life assessment. *Health Qual Life Outcomes*. 2016; 14:88. [PubMed: 27267885]

Figure 1. OES scores' dimensionality (scree plot and parallel analysis with actual eigenvalues (●-) and simulated eigenvalues (---) from the sample OES item correlation matrix).

Table 1

Methods of PROMIS-recommended translation and harmonization process across Spanish language varieties.

Step	Description
1) Two simultaneous forward-translations (2 Fwd)	Source items in English are translated into Spanish by two independent professional translators who are native Spanish speakers.
2) Reconciled single target language translation (1 Rec)	A third independent translator, also a native Spanish speaker, reconciles the two forward translations by selecting one of the forward translations, creating a hybrid version, or providing a new version. This translator must also note the reasons why the reconciled version is the best way to convey the meaning of the source.
3) Back-translation (1 BT)	This reconciled version is then back-translated by a native English-speaking translator who is fluent in the target language. The translator does not see the original English source items or item definitions. The back-translation into English must reflect what the target language translation says, without embellishing it.
4) Back-translation review	The Translation Project Manager (TPM) compares source and back-translated English versions and identifies discrepancies in the back translations and provides clarification to the reviewers on the intent behind the items. This step also results in a preliminary assessment of harmonization between the languages.
5) Expert reviews (3 Revs)	Three experts, who are native Spanish speakers, independently examine all of the preceding steps and select the most appropriate translation for each item or provide alternate translations, if the previous translations are not acceptable.
6) Pre-finalization review	The Translation Project Manager evaluates the merit of the reviewer's comments, identifies potential problems in their recommended translations, and formulates questions and comments to guide the Language Coordinator (LC) for the target language.
7) Finalization (1 LC)	The Language Coordinator (LC), a native Spanish speaker, determines the final translation by reviewing all the information in the Item History and addressing the TPM's comments. Along with the final translation, the LC also provides the respective literal back-translation and polished back-translation for each item. The LC explains the choice of final translation and offers justification for the decision if the final translation is different from the reconciled version or from what reviewers individually recommended.
8) Harmonization and quality assurance	The Translation Project Manager makes a preliminary assessment of the accuracy and equivalence of the final translation by comparing the final back-translations with the source, and verifying that documentation of the decision making process is complete. The Language Coordinator may be consulted again for additional input.
9) Formatting, typesetting and proofreading of final questionnaire	Item forms by two proofreaders working independently, and reconciliation of the proofreading comments.
10) Cognitive testing and linguistic validation	The Spanish version is pretested with participants who are native Spanish speakers. The goal is to have each new item debriefed in the target country by at least 5 participants in a cognitive debriefing interview to verify the meaning of the item is equivalent to the English source after translation.
11) Analysis of participants' comments and finalization of translation	The Translation Project Manager compiles participants' comments (back-translated into English) and summarizes the issues. The Language Coordinator reviews the issues and proposes translation solutions. The TPM verifies that solutions proposed by the LC harmonize with the source and with other languages.

Table 2

Results of PROMIS-recommended translation and harmonization process across the Spanish language varieties

Step	Result(s)
1) Two simultaneous forward translations (2 Fwd)	Items in English were translated into Spanish by two independent professional translators who were native Spanish speakers.
2) Reconciled single target language translation (1 Rec)	A third independent translator (native Spanish speaker), chose the version provided by one of the forward translators and made modifications to it. "Tu cara" instead of "Su aspecto facial"; "Tu" instead of "Su" and "tus" instead of "sus".
3) Back-translation (1 BT)	The English back-translator, a native English speaker and fully-fluent Spanish speaker, translated the Spanish versions with the following changes from the English version: "About" as "with" and "appearance" as "looks".
4) Back-translation review	Comparison by Translation Project Manager (TPM) resulted in the decision of reviewers in the next step to use formal "Usted" version of Spanish speech when making their choices. Use "se" instead of "te".
5) Expert reviews (3 Revs)	One of the reviewers chose the version previously translated. The word "apariencia" was chosen over "aspecto" for appearance. Also the word "filas" was selected instead of "alineacion" for rows of teeth.
6) Pre-finalization review	TPM reviewed the comments and provided the versions to the Language Coordinator (LC).
7) Finalization (1 LC)	LC determined final Spanish translation and provided respective English back-translation and explanation for choice of final translation.
8) Harmonization and quality assurance	TPM did not find the need to do the translatability review due to the survey's easily definable concepts/words across the cultures/languages in question. Furthermore, they did not enforce PROMIS categories for classification of issues due to the non-existence of translatability issues encountered by reviewers as communicated by their ongoing correspondence with TPM.
9) Formatting, typesetting and proofreading of final questionnaire	TPM reviewed LC's final Spanish translation, specifically comparing the final English back-translation to the initial English source. TPM completed documentation of translation decision-making.
10) Cognitive testing and linguistic validation	The translation team did not find the need to administer cognitive debriefing interviews. Rather, requested feedback if participants had any concerns regarding survey question's cognitive/linguistic meanings.
11) Analysis of participants' comments and finalization of translation	The participants' comments were reviewed by TPM, synthesized and sent to LC for review and final version was generated.

Table 3

The English and Spanish (showed in *italics*) versions of the Orofacial Esthetic Scale.

Orofacial Esthetic Scale (OES)	Escala de Estética Orofacial (EEO)
English (Original) Version	Spanish Version
<i>How do you feel about the appearance of your face, your mouth, your teeth and your replacements (prostheses, crowns, bridges and implants)?</i>	<i>¿Cómo se siente sobre la apariencia de su cara, su boca, sus dientes, y los reemplazos de sus dientes (coronas, puentes e implantes)?</i>
0: Very dissatisfied - 10: Very satisfied	0: Muy insatisfecho - 10: Muy satisfecho
1. Your facial appearance	1. Su apariencia facial
2. Appearance of your facial profile	2. Apariencia de su perfil facial
3. Your mouth's appearance (smile, lips, and visible teeth)	3. Apariencia de su boca (sonrisa, labios y dientes visibles)
4. Appearance of your rows of teeth	4. Apariencia de la alineación de sus dientes
5. Shape/form of your teeth	5. Forma de sus dientes
6. Color of your teeth	6. Color de sus dientes
7. Your gum's appearance	7. Apariencia de sus encías
8. Overall, how do you feel about your face, your mouth and your teeth?	8. En general, ¿cómo se siente sobre la apariencia de su cara, su boca y sus dientes?

Table 4

Descriptive statistics of OES-Sp items.

OES-Sp Item	Mean (SD)	Median (IQR)	Min-Max	% of subjects who were “very dissatisfied”	% of subjects who were “very satisfied”
Face	7.6 (3.1)	9 (6–10)	0–10	9	42
Profile	7.5 (3.2)	9 (6–10)	0–10	11	39
Mouth	6.9 (3.3)	8 (5–10)	0–10	11	33
Tooth alignment	6.5 (3.5)	8 (4–10)	0–10	16	28
Tooth shape	6.8 (3.3)	8 (5–10)	0–10	10	28
Tooth color	6.1 (3.2)	7 (4–9)	0–10	12	17
Gingiva	7.1 (3.3)	8 (5–10)	0–10	11	31
Overall impression *	7.3 (3.0)	8 (5–10)	0–10	8	28

* N=310

Table 5

OES-Sp scores' internal consistency.

OES-Sp Item	Alpha when item is deleted	Item-rest correlation
Face	0.94	0.89
Profile	0.94	0.89
Mouth	0.94	0.93
Tooth alignment	0.94	0.90
Tooth shape	0.94	0.89
Tooth color	0.95	0.79
Gingiva	0.95	0.86

Table 6

OES-Sp scores validity

Validity measure	Correlation coefficient [*]	95% confidence interval
OES overall appearance impression score	0.90	0.88 to 0.92
OHIP dimension Orofacial Appearance score	−0.69	−0.75 to −0.63
OHIP sum score of items 3, 22, and 31	−0.67	−0.73 to −0.60
OHIP item 3: Have you noticed a tooth that does not look right?	−0.53	−0.61 to −0.44
OHIP item 22: Have you felt uncomfortable about the appearance of teeth, mouth, dentures or jaws?	−0.67	−0.73 to −0.60
OHIP item 31: Have you avoided smiling because of problems with your teeth, mouth, dentures or jaws?	−0.64	−0.70 to −0.57

* all coefficients are Spearman's rank coefficients, except for first correlation which is a Pearson correlation coefficient