UC Davis
UC Davis Previously Published Works

Title

Translation and Cross-cultural Validation of the English Young Adult Burn Outcome Questionnaire (YABOQ) in Spanish

Permalink

Journal

ISSN

Authors

Brady, Keri JS
Ni, Pengsheng
Grant, Gabrielle G
et al.

Publication Date

DOI

Peer reviewed

Translation and Cross-cultural Validation of the English Young Adult Burn Outcome Questionnaire (YABOQ) in Spanish

Keri J. S. Brady, MPH[*,†,‡], **Pengsheng Ni, MD, MPH**[*], **Gabrielle G. Grant, BA**[†,‡], **Catherine R. Thorpe, BSc**[†,‖], **Deborah Nadler, MD**[$], **Austin Lee, PhD**[‡,‖], **Gabriel D. Shapiro, PhD**[¶], **Ann Kogosov**[†], **Tina L. Palmieri, MD**[**,††], **Ronald G. Tompkins, MD, ScD**[‡‡], **Jeffrey C. Schneider, MD**[†,‖], **Lewis E. Kazis, ScD**[*], **Colleen M. Ryan, MD**[†,‡]

[*]Department of Health Law, Policy, and Management, Boston University School of Public Health, Boston, MA

[†]Shriners Hospitals for Children—Boston, Boston, MA

[‡]Department of Surgery, Massachusetts General Hospital, Boston, MA

[‖]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Charlestown, MA

[$]Department of Internal Medicine, Rhode Island Hospital, Providence, RI

[¶]McGill University, Montreal, QC, Canada

[**]Shriners Hospital for Children—Northern California, Sacramento, CA

[††]University of California Davis School of Medicine, Department of Surgery, Division of Burn Surgery and Reconstruction, Sacramento, CA

[‡‡]Harvard Medical School, Boston, MA

Abstract

The Young Adult Burn Outcome Questionnaire (YABOQ) is a validated, English-language patient-reported outcome assessment of young adults' recovery from burn injury across 15 scale domains. We evaluated the cross-cultural validity of a newly developed Spanish version of the YABOQ. Secondary data from English- and Spanish-speaking burn survivors (17 to 30 years of age) were obtained from the Multicenter Benchmarking Study. We conducted classic psychometric analyses and evaluated the measurement equivalence of the English and Spanish YABOQs in logistic and ordinal logistic regression differential item functioning analyses. All multi-item scales in the Spanish YABOQ demonstrated adequate reliability except the Pain and Itch scales. One item in the Perceived Appearance scale showed differential item functioning across English- and Spanish-speaking burn survivors, but the observed differential item functioning had no clinically significant impact on scale-level Perceived Appearance scores. Our findings support the cross-cultural validity of the YABOQ Physical Function, Perceived Appearance, Sexual Function,

Address correspondence to Colleen M. Ryan, MD, Massachusetts General Hospital, 55 Fruit Street, Bigelow 1302, Boston, MA 02114. cryan@mgh.harvard.edu.

Emotion, Family Function, Family Concern, Satisfaction with Symptom Relief, Satisfaction with Role, Work Reintegration and Religion scales among English- and Spanish-speaking young adult burn survivors. This work supports the use of these English and Spanish YABOQ scales to assess the effect of therapeutic interventions on young adults' burn outcomes in pooled analyses and to assess disparities in young adults' burn outcomes across language groups.

The Young Adult Burn Outcome Questionnaire (YABOQ) is one of the few burn-specific, patient-reported outcome measures (PROMs) designed to assess young adults' (19 to 30 years) recovery following burn injury.[1] The 48-item instrument assesses burn outcomes across 15 age-relevant domains of health-related quality of life (HRQOL), including physical function, fine motor, pain, itch, social function limited by physical function, perceived appearance, social function limited by appearance, sexual function, emotion, family function, family concern, satisfaction with symptom relief, satisfaction with role, work reintegration, and religion. The original YABOQ was developed in English and has demonstrated adequate construct validity (structural and criterion validity), reliability (test–retest and internal consistency reliability), and responsiveness to change in cross-sectional and prospective studies.[1] Normative data benchmarking burn survivors' recovery across each HRQOL domain relative to a nonburned cohort are also available.[1]

Although validity evidence supporting use of the YABOQ to assess recovery outcomes among young adult burn survivors has been established,[1] the YABOQ's use is limited by its lack of availability in Spanish. Spanish is the most common language spoken in U.S. households among non-English speakers and is the primary language of a large proportion of young adult burn survivors in the United States and other countries.[2] The goal of this study was to translate the English YABOQ to Spanish and establish the measurement equivalence of the English and Spanish forms. Measurement equivalence is a key aspect of cross-cultural validity that ensures that different translations of the same instrument measure the concepts they intend to measure in the same way. For example, a lack of measurement equivalence across English and Spanish versions of a scale could result if multiple scale items are interpreted differently across English- and Spanish-speaking burn survivors. This could be due to poor translation efforts made on behalf of scale developers or to differences in the meaning of assessed concepts across cultures and their item relationships within and between domains. When clinically meaningful group differences in observed scale scores stem from differences in cultural interpretation/translations of scale items across the groups, such differences represent a bias and are not due to true group differences in the underlying construct assessed.[3] Lack of measurement equivalence between the English and Spanish YABOQs could result in false differences in recovery levels between English and Spanish-speaking burn survivors.[4] It could also bias the true effect of a clinical intervention on burn survivors' HRQOL if data are pooled across English and Spanish forms.[4] Establishing measurement equivalence following the translation of the YABOQ into Spanish is therefore a necessary prerequisite before YABOQ scores can be compared across forms or pooled data can be used to evaluate the effect of an intervention.

METHODS

Spanish Translation of YABOQ

The English YABOQ was translated to Spanish using a forward and backward translation process. Two translators (Translator 1, who is a native English speaker and fluent in Spanish, and Translator 2, who is a native Spanish speaker and fluent in English) worked independently on two forward translations (English to Spanish). After working independently, these two translators reconciled differences through discussion, resulting in a revised forward translation. Next, backward translation (Spanish to English) was conducted independently by a third translator (Translator 3) who is a native Spanish speaker and fluent in English. The backward translation was then evaluated by Translator 1 for any discrepancies between the back translation and the original YABOQ. The forward translation was revised by Translators 1 and 3 for items that did not accurately convey their original intent in the backward translation. Translators 1 and 3 performed the final proofreading and reformatting of the final Spanish YABOQ (Supplemental Appendix 1). The Spanish spoken by all three translators was of Latin American influence. This approach was informed by the National Institutes of Health Patient Reported Outcome Measurement Information System (NIH PROMIS) and International Society for Quality of Life Research (ISOQOL) recommendations and guidelines for translating HRQOL assessments.[5,6]

Design and Sample

This study used a pooled sample of secondary, de-identified English and Spanish YABOQ response data collected as part of the Multicenter Benchmarking Study, which was a multisite prospective cohort study to evaluate the effects of burn injuries on children and young adults.[1,7] Further study design and sampling details are published in Kazis et al.[1] and Ryan et al.[7] Patients were included in the original study if they were 17 years old, were able to speak and understand English or Spanish, had a TBSA burn 5% (or a burn to critical areas of the hands, face, feet, or genitals), and were discharged from their initial inpatient stay.

Statistical Analyses

Descriptive Analyses and Assumption Assessment—Descriptive sample statistics were computed across language groups using means and SD for continuous variables and frequencies and percentages for categorical variables. Descriptive scale statistics included scale means, SD, and minimum and maximum scores across language groups. Because the evaluation of measurement equivalence is only meaningful if the YABOQ is reliable and essentially unidimensional in the reference (English) and focal (Spanish) groups, we evaluated both of these assumptions prior to evaluating measurement equivalence in differential item functioning (DIF) analyses. We computed Cronbach alpha coefficients to evaluate each scale's internal consistency reliability in the Spanish-speaking group relative to the English-speaking group. Scales showing <0.70 reliability in the Spanish-speaking group were not included in the DIF analysis. We evaluated the dimensionality of the YABOQ in the English- and Spanish-speaking groups in two correlated traits confirmatory factor analyses (CFAs) (using weighted least square mean and variance adjusted estimation and polychoric correlations on pairwise complete observations). In the correlated traits

CFA structure, each scale item was specified as loading onto a factor representing the corresponding scale (i.e., one factor was included in the model for each scale), and all factors were allowed to correlate, which is appropriate for HRQOL scales. We considered factor loadings of 0.40 and a sufficient fit to the specified CFA model supportive of essential unidimensionality of each scale. Sufficient fit was defined as a scaled root mean square error approximation (RMSEA) 0.10, a Tucker–Lewis Index (TLI) and Comparative Fit Index (CFI) of 0.90, and standardized root mean square residual (SRMSR) of 0.08.[8-13] The use of a single CFA model in each language group vs fitting separate CFA models to each scale in each language group was chosen because the majority of the YABOQ scales have 3 scale items, which results in a CFA model that is saturated and, therefore, we could not assess model fit. Furthermore, we chose to fit two separate CFA models to each language group over a multigroup CFA model due to our sample size. To further explore the psychometric quality of each scale prior to DIF analyses, we also conducted supplemental item response theory analyses (Supplemental Appendix 2).

Differential Item Functioning Analyses.—We evaluated measurement equivalence of the Spanish and English YABOQs using binary logistic and ordinal logistic regression DIF analyses for dichotomous and polytomous scales, respectively, which is suitable for brief scales and small samples.[14] DIF indicates a lack of item-level measurement equivalence and occurs when the probability of an item response differs across English-speaking (reference group) and Spanish-speaking (focal group) respondents after controlling for the two groups' underlying HRQOL levels.[15] For each scale item, DIF was detected by comparing the change in McFadden's pseudo-R^2 in three nested ordinal logistic regression DIF models (Models 1–3), where Models 1 and 2 and Models 2 and 3 were compared with each other to assess uniform DIF and nonuniform DIF, respectively (Supplemental Appendix 3).[14,16,17] Uniform DIF was defined as a change in the McFadden's pseudo-R^2 value between Models 1 and 2 of 2%,[16,18,19] indicating an item was systematically easier or harder to endorse for the Spanish-speaking group compared with the English-speaking group. Nonuniform DIF was defined as a change in the McFadden's pseudo-R^2 value between Models 2 and 3 of 2%, indicating an item was more or less strongly related to the underlying construct assessed in the Spanish-speaking group compared to the English-speaking group.[16,18,19] For any uniform DIF detected, we further evaluated the magnitude of the DIF by computing the odds ratio (OR) and 95% confidence interval (CI) for the grouping variable in Model 2. Slight to moderate DIF was defined by a CI for the OR that is outside the interval of 0.65 to 1.53; moderate to large DIF was defined by a CI for the OR that is outside the interval of 0.53 to 1.89.[20,21] If a scale item showed uniform and/or nonuniform DIF, we assessed the clinical significance of the DIF by computing mean scale score differences between language groups using all items (including DIF items) and, separately, using only non-DIF items.[22] If no statistically significant mean differences in scale scores across language groups existed using a scale score computed with all (including DIF) items, we considered the impact of DIF to be clinically insignificant.[22] Conversely, if a statistically significant mean difference in scale scores existed across groups using all scale items but did not exist when computed using only non-DIF items, we considered the impact of DIF to be clinically significant. Two-sided significance levels were set to $\alpha < 0.05$. The fine motor scale was excluded from the DIF analysis because DIF is not applicable to single-item measures. We

conducted all statistical analyses in R (v3.5.1) using the *psych* (v1.8.12), *lavaan* (v0.6-3), and *lordif* packages (v0.3-3).[16,23-25]

RESULTS

Demographic Characteristics

The overall sample included a total of 81 and 153 unique English- and Spanish-speaking burn survivors, respectively (Table 1). There were no significant differences in the sex distributions across the language groups but significant differences in the age distributions, with the English group being older than the Spanish group on average. The majority of burn survivors in the Spanish and English groups identified as Hispanic/Latino and non-Hispanic White race/ethnicity, respectively.

Descriptive Scale Statistics and Assumption Assessment

The overall pooled sample included a total of 142 and 368 responses from the Spanish and English groups, respectively (Table 2). Although the Pain and Itch scales showed adequate internal consistency reliability in the English group, these 11-point anchored visual analogue scales showed low internal consistency reliability among Spanish-speaking burn survivors and require future revision. Qualitative inspection of the scales themselves revealed that this may be due to nonintuitive labeling of the second items in the Pain and Itch scales (pain2, itch2), whereby a score of 0 was labeled "very severe pain"/"very severe itch" and a score of 10 was labeled "no pain"/"no itch." Evaluation of the inter-item correlations in the Social Functioning Limited by Physical Function and Social Functioning Limited by Appearance scales across language groups also revealed that a number of items were very highly correlated (0.95 to 0.99) indicating that the items function redundantly in both English- and Spanish-speaking groups and require revision (Supplemental Appendix 4). Given the need for revisions to the pain, itch, social functioning limited by physical function, and social functioning limited by appearance scales, we excluded these scales from the further analyses.

The 10-factor correlated traits CFA model demonstrated good fit in the Spanish response data (RMSEA = 0.019, CFI = 0.997, TLI = 0.997, SRMSR = 0.085) and in the English response data (RMSEA = 0.027, CFI = 0.999, TLI = 0.999, SRMR = 0.065) and factor correlation patterns were similar across both CFA models, providing support for the structural equivalency of the Spanish version compared with the English version. (See Supplemental Appendix 5 for item response theory analysis results of the English YABOQ.)

Differential Item Functioning

No DIF was detected in the physical function, sexual function, emotion, family function, family concern, satisfaction with symptom relief, satisfaction with role, and work reintegration scales across language groups (Table 3). All items in the perceived appearance (PA) scale were also free of DIF across language groups except PA3 ("I believe unsure of myself around strangers"), which displayed moderate to large uniform DIF (OR = 0.30, 95% CI = 0.19–0.46). Compared with English-speaking burn survivors, Spanish-speaking burn survivors were less likely to endorse PA3 despite both groups having the same

underlying perceived appearance levels (i.e., the item was harder to endorse for Spanish- vs English-speaking respondents). Despite this, two-sample *t*-tests evaluating mean scale score differences across language groups using all items (PA1 ["I believe the burn is unattractive to others"], PA2 ["I think people would not want to touch me"], and PA3: $t(242.10) = 0.46$, $p = .6488$) and only DIF-free items (PA1–PA2: $t(235.46) = 0.87$, $p = .3865$) revealed no significant differences when all items or only non-DIF items are used. The lack of mean differences observed among scale scores that were calculated using all items (including DIF items) suggests that there is no clinical significance of the DIF observed on item PA3

DISCUSSION

We translated the original English YABOQ to Spanish and evaluated the cross-cultural validity of the Spanish YABOQ. The result of the CFA models for these scales was robust and indicated that the relationships of items between and within scales or domains were comparable across the English and Spanish YABOQs. This is an important finding given that the cultural differences and biases could be potentially ruled out when comparing the two YABOQ versions. Our DIF analyses supported the measurement equivalence of all 10 of the following English and Spanish YABOQ scales: physical function, perceived appearance, sexual function, emotion, family function, family concern, satisfaction with symptom relief, satisfaction with role, work reintegration, and religion. Of the 10 scales included in the DIF analyses, only the perceived appearance scale showed significant DIF in one of its three scale items, whereby Spanish-speaking respondents were less likely to endorse "feeling unsure of [themselves] around strangers" (PA3) compared with English-speaking burn survivors. This difference could have been the result of differences in the cultural relevancy of this item across Spanish and English groups, differences in the interpretation of this item across Spanish and English groups, and/or differences in Spanish-speaking burn survivors' willingness to endorse this item compared with English-speaking burn survivors. Nevertheless, at the scale level, the observed DIF in PA3 had no clinically significant impact on conclusions regarding differences in English- and Spanish-speaking burn survivors' perceived appearance outcomes. The PA3 item should not be used, however, as a single-item measure, as the observed DIF may bias cross-cultural comparisons in mean item scores.[22]

Other analyses revealed that the pain, itch, social functioning limited by physical function, and social functioning limited by appearance scales require revision. The pain and itch scales were not found to be reliable as presented among Spanish-speaking burn survivors. These symptom assessments might be readily addressable by aligning or supplementing the items with commonly used clinical or visual analogue scales. For example, these scales might be improved by aligning the direction of the scale coding (0–10) for the anchored visual analogue pain and itch scales with the direction of the scale labels (from "no" pain/ itch to "very severe" pain/itch)—such that the high and low ends of the coding and labels are not in opposing directions. The social functioning limited by physical function and social functioning limited by appearance scales were redundant in both English- and Spanish-speaking burn survivors, requiring future modifications. These scales can be improved by revising the repetitive items and assessing the validity of the revised scale.

This is the first study to develop and validate a Spanish-language YABOQ. As Spanish is the most common non-English language spoken at home in U.S. households,[2] this study takes an important step in addressing disparities in the reporting of burn outcomes among Spanish-speaking burn survivors due to the lack of validated, Spanish-language outcome assessments. However, there are several limitations to this study. Although the Spanish YABOQ was translated native Spanish speakers, all three translators resided in the United States. Therefore, the Spanish YABOQ may be less culturally relevant to Spanish speakers residing in countries outside the United States. Further qualitative and quantitative research is needed to further address this research question.[17] Qualitative research will serve to evaluate the relevance and comprehensiveness of Spanish YABOQ items among young adult burn survivors from other Spanish-speaking countries. Quantitative research administering the Spanish YABOQ to Spanish-speaking burn survivors residing in other countries will be needed to evaluate DIF across countries (e.g., United States vs Mexico). If differences in cultural relevance are identified through these methods, updates to the Spanish YABOQ would be warranted to create a revised version of the instrument that is relevant to other countries.[5] This study also had a small sample size. As a result, we were unable to specify more complex CFA models (i.e., multigroup CFA) in our dimensionality analyses and unable to test the robustness of our measurement equivalence results in sensitivity analyses using IRT-based DIF detection methods (which have shown to be more accurate than DIF detection methods that rely on the use of observed scale scores as a conditioning variable, as used in this study).[14,26] Future research that collects clinical variables in a larger samples of Spanish-speaking burn survivor in the United States and other Spanish-speaking countries will be needed to further explore these issues.

CONCLUSIONS

This work supports the use of the 10 identified English and Spanish YABOQ scales in research and clinical practice to evaluate the effectiveness of therapeutic interventions, inform clinical decision-making, and monitor patient recovery outcomes. Results support the cultural validity of these scales in English- and Spanish-speaking young adult burn survivors. Future work should consider some modifications to the four other multi-item scales.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

REFERENCES

1. Ryan CM, Schneider JC, Kazis LE, et al. ; Multi-Center Benchmarking Study Group. Benchmarks for multidimensional recovery after burn injury in young adults: the development, validation, and testing of the American Burn Association/Shriners Hospitals for Children young adult burn outcome questionnaire. J Burn Care Res 2013;34:e121–42. [PubMed: 23511284]

2. United States Census Bureau. American community survey 5-year estimates. 2018; available from https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2017/5-year.html.

3. Teresi JA, Fleishman JA. Differential item functioning and health assessment. Qual Life Res 2007;16(Suppl 1):33–42. [PubMed: 17443420]

4. 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]

5. Health Measures. PROMIS instrument development and validation scientific standards version 2.0. 2013; available from http://www.healthmeasures.net/explore-measurement-systems/promis/measure-development-research.

6. Koller M, Kantzer V, Mear I, et al. ; ISOQOL TCA-SIG. The process of reconciliation: evaluation of guidelines for translating quality-of-life questionnaires. Expert Rev Pharmacoecon Outcomes Res 2012;12:189–97. [PubMed: 22458620]

7. Kazis LE, Lee AF, Hinson M, et al. ; Multi-Center Benchmarking Study Working Group. Methods for assessment of health outcomes in children with burn injury: the multi-center benchmarking study. J Trauma Acute Care Surg 2012;73(3 Suppl 2):S179–88. [PubMed: 22929545]

8. Lt Hu, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equat Model 1999;6:1–55.

9. Steiger JH. Structural model evaluation and modification: an interval estimation approach. Multivariate Behav Res 1990;25:173–80. [PubMed: 26794479]

10. Chen F, Curran PJ, Bollen KA, Kirby J, Paxton P. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. Sociol Methods Res 2008;36:462–94. [PubMed: 19756246]

11. Muthén BO. Robust inference using weighted least squares and quadratic estimating in latent variable modeling with categorical and continuous outcomes. Psychometrika 1997;75:1–45.

12. Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. Psychometrika 2001;66:507–14.

13. Xia Y Investigating the chi-square-based model-fit indexes for WLSMV and ULSMV estimators [doctoral dissertation]. Tallahassee, Florida: Florida State University; 2016.

14. Teresi JA, Jones RN. Methodological issues in examining measurement equivalence in patient reported outcomes measures: methods overview to the two-part series, "measurement equivalence of the patient reported outcomes measurement information system® (PROMIS®) short forms". Psychol Test Assess Model 2016;58:37–78. [PubMed: 28983448]

15. Teresi JA. Overview of quantitative measurement methods. Equivalence, invariance, and differential item functioning in health applications. Med Care 2006;44(11 Suppl 3):S39–M9. [PubMed: 17060834]

16. Choi SW, Gibbons LE, Crane PK. Lordif: an r package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte carlo simulations. J Stat Softw 2011;39:1–30.

17. Measures Health. PROMIS minimum requirements for the release of PROMIS instruments after translations and recommendations for further psychometric evaluation. 2014; available from http://www.healthmeasures.net/images/PROMIS/Standards_for_release_of_PROMIS_instruments_after_translation_v8.pdf.

18. Liu H, Cella D, Gershon R, et al. representativeness of the patient-reported outcomes measurement information system internet panel. J Clin Epidemiol 2010;63:1169–78. [PubMed: 20688473]

19. Bjorner JB, Kosinski M, Ware JE Jr. Calibration of an item pool for assessing the burden of headaches: an application of item response theory to the Headache Impact Test (HIT™). Qual Life Res 2003;12:913–33. [PubMed: 14651412]

20. Scott NW, Fayers PM, Aaronson NK, et al. ; EORTC Quality of Life Group and the Quality of Life Cross-Cultural Meta-Analysis Group. Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. Health Qual Life Outcomes 2010;8:81. [PubMed: 20684767]

21. Petersen MA, Groenvold M, Bjorner JB, et al. ; European Organisation for Research and Treatment of Cancer Quality of Life Group. Use of differential item functioning analysis to

assess the equivalence of translations of a questionnaire. Qual Life Res 2003;12:373–85. [PubMed: 12797710]

22. Bjorner JB, Kreiner S, Ware JE, Damsgaard MT, Bech P. Differential item functioning in the Danish translation of the SF-36. J Clin Epidemiol 1998;51:1189–202. [PubMed: 9817137]

23. Revelle W Psych: procedures for personality and psychological research. 2018; available from https://CRAN.R-project.org/package=psych.

24. Rosseel Y lavaan: an R package for structural equation modeling. J Stat Softw 2012;48:1–36.

25. R Core Team. R: a language and environment for statistical computing. 2018; available from https://www.R-project.org/.

26. Teresi JA. Different approaches to differential item functioning in health applications. Advantages, disadvantages and some neglected topics. Med Care 2006;44(11 Suppl 3):S152–70. [PubMed: 17060822]

Table 1.

Demographic of unique sample respondents [a]

	Spanish (n = 81)	English (n = 153)	P-value [b]
Age at first survey (years), mean (SD)	20.3 (3.6)	24.7 (3.6)	<.0001
Sex, n (%)			.4117
Male	53 (67)	112 (73)	
Female	26 (33)	41 (27)	
Race/ethnicity, n (%)			<.0001
Non-Hispanic White	5 (7)	94 (62)	
Non-Hispanic Black	0 (0)	18 (12)	
Hispanic or Latino	60 (86)	15 (10)	
Other	5 (7)	24 (16)	

[a] frequencies may not add to total sample size due to missing data.

[b] Two-sample t-test for continuous and chi-square tests for categorical variables.

Table 2.

Descriptive Statistics and Classic Psychometric Analyses of pooled Spanish (n = 142) and English (n = 368) YABOQ response data [a]

Scale (number of items)	Spanish					English				
	n	Mean (SD)	Minimum	Maximum	α[b]	n	Mean (SD)	Minimum	Maximum	α[b]
Pain (2)[c]	140	3.76 (0.91)	2	5	0.05	368	3.78 (0.98)	1	5	0.71
Itch (2)[c]	140	3.58 (0.94)	1	5	0.12	368	3.37 (1.14)	1	5	0.77
Physical Function (5)	138	3.79 (0.46)	1	4	0.89	367	3.57 (0.62)	1	4	0.88
Fine Motor (1)	138	3.85 (0.48)	1	4	—[d]	367	3.74 (0.61)	1	4	—[c]
Social Function Limited by Physical Function (5)	139	0.83 (0.32)	0	1	0.88	367	0.75 (0.35)	0	1	0.88
Perceived Appearance (3)	139	3.36 (1.24)	1	5	0.89	367	3.21 (1.18)	1	5	0.81
Social Function Limited by Appearance (4)	139	0.78 (0.37)	0	1	0.92	366	0.79 (0.35)	0	1	0.90
Sexual Function (5)	134	4.25 (0.78)	1	5	0.86	362	4.16 (0.92)	1	5	0.90
Emotion (2)	138	4.26 (0.92)	1	5	0.90	365	3.85 (1.07)	1	5	0.81
Family Function (3)	140	4.56 (0.74)	1	5	0.87	362	4.26 (0.92)	1	5	0.86
Family Concern (3)	139	3.37 (1.30)	1	5	0.89	361	2.85 (1.36)	1	5	0.92
Satisfaction with Symptom Relief (3)	135	3.85 (1.21)	1	5	0.87	364	3.70 (1.14)	1	5	0.78
Satisfaction with Role (3)	135	3.92 (1.21)	1	5	0.93	364	3.39 (1.38)	1	5	0.90
Work Reintegration (3)	123	3.59 (0.97)	1	5	0.92	368	3.15 (0.75)	1	5	0.88
Religion (4)	138	2.72 (0.97)	1	4	0.84	361	2.74 (0.91)	1	4	0.91

[a] All scales are scored such that higher scores equate to better recovery outcomes.

[b] α = standardized Cronbach alpha coefficient.

[c] Items PAIN2 and ITCH2 in the original scale was a visual analogue scale coded 0 (very severe symptoms) to 10 (no symptoms) but collapsed to a 5-point scale coded 1 (very severe symptoms) to 5 (no symptoms) in the analysis.

[d] Not applicable for single-item scales.

Table 3.

Detection of differential item functioning in the YABOQ across English- and Spanish-speaking young adult burn survivors[a]

Scale/item	Scale/item description	McFadden pseudo-R^2 change from Model 1 to 2 (uniform DIF)[b]	McFadden pseudo-R^2 change from Model 2 to 3 (nonuniform DIF)
Physical Function	During the past week, has it been easy or hard for you to…		
PHY1	Climb three flights of stairs.	0.0004	0.0013
PHY2	Climb one flight of stairs.	0.0035	0.0091
PHY3	Walk three blocks.	0.0062	0.0032
PHY5	Get in and out of bed.	0.0067	0.0074
PHY6	Bend over from a standing position and pick up something off the floor.	0.0000	0.0005
Perceived Appearance			
PA1	I believe that the burn is unattractive to others	0.0152	0.0005
PA2	I think people would not want to touch me	0.0013	0.0030
PA3	I believe unsure of myself among strangers	0.0200[**]	0.0019
Sexual Function			
SEX1	I believe frustrated because I cannot be sexually aroused as well as before the burn injury.	0.0000	0.0002
SEX2	Since my burn injury I am simply not interested in sex anymore.	0.0021	0.0002
SEX3	How much of a problem was each of the following during the past 4 weeks because of the burn injury…lack of sexual interest.	0.0008	0.0018
SEX4	How much of a problem was each of the following during the past 4 weeks because of the burn injury…having difficulty in becoming sexually aroused.	0.0023	0.0002
SEX5	How much of a problem was each of the following during the past 4 weeks because of the burn injury…having difficulty in having an orgasm.	0.0000	0.0059
Emotion			
EMOT1	Over the past month, how often has your burn injury…	0.0005	0.0000
EMOT2		0.0013	0.0009
Family Function	Over the past month, how often has your burn injury…		
FAM1	Limited family's ability to have time for themselves or time with friends.	0.0010	0.0001
FAM2	Interrupted simple activities like meals.	0.0052	0.0012
FAM3	Over the past month, how often has your burn injury…limited your ability to spend time with other family members.	0.0015	0.0006
Family Concern			
CON1	Over the past month, how much worry or concern has for your family/significant other expressed out…		
	Your recovery from the effects of the burn injury.	0.0001	0.0000

Scale/item	Scale/item description	McFadden pseudo-R^2 change from Model 1 to 2 (uniform DIF)[b]	McFadden pseudo-R^2 change from Model 2 to 3 (nonuniform DIF)
CON2	Your level of pain and suffering.	0.0040	0.0002
CON3	Your future health.	0.0041	0.0008
Satisfaction with Symptom	Relief How satisfied are you now with your…		
SYM1	Pain relief.	0.0035	0.0024
SYM2	Itch relief.	0.0022	0.0005
SYM3	Sleep.	0.0000	0.0001
Satisfaction with Role	How satisfied are you now with your…		
ROLE1	Ability to do chores.	0.0095	0.0000
ROLE2	School work/job.	0.0001	0.0008
ROLE3	Ability to play and have fun.	0.0004	0.0028
Work Reintegration	Following your return to job or school after the burn injury, how would you rate your…		
WORK1	Acceptance by peers.	0.0003	0.0012
WORK2	Acceptance by teachers/boss.	0.0000	0.0000
WORK3	Ability to perform.	0.0004	0.0038
Religion			
REL1	How much is religion (and/or God), a source of strength and comfort to you?	0.0002	0.0051
REL2	How strongly religious (or spiritually oriented) do you consider yourself to be?	0.0033	0.0035
REL3	How close do you feel to God?	0.0066	0.0069
REL4	God dwells within you.	0.0000	0.0008

[a] Bolded values are 0.01 threshold for significant DIF.

[b] *indicates Slight to moderate uniform DIF magnitude; **Moderate to large uniform DIF magnitude.