Published in final edited form as:

Qual Life Res. 2015 September ; 24(9): 2289–2295. doi:10.1007/s11136-015-0949-z.

Evaluating Psychometric Properties of the Spanish-version of the Pediatric Functional Assessment of Chronic Illness Therapy-Perceived Cognitive Function (pedsFACIT-PCF)

Alex W. K. Wong,

Program in Occupational Therapy and Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA

Jin-Shei Lai,

Departments of Medical Social Sciences and Pediatrics, Northwestern University Feinberg School of Medicine, 633 N St. Clair, #19-039, Chicago, IL 60611, USA js-lai@northwestern.edu

Helena Correia, and

Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

David Cella

Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Abstract

Purpose—The pediatric Functional Assessment of Chronic Illness Therapy-Perceived Cognitive Function (pedsFACIT-PCF) is a 13-item short-form derived from the pediatric Perceived Cognitive Function item bank (pedsPCF), which was developed to measure children's daily cognitive behaviors and was validated on the US general population and children with cancer. This study evaluated the psychometric properties of Spanish language pedsFACIT-PCF and the measurement equivalence between Spanish and English versions.

Methods—pedsFACIT-PCF items were translated into Spanish using a standard iterative methodology. A total of 1358 English- and 604 Spanish-speaking children aged 8–17 years who completed English and Spanish versions of pedsFACIT-PCF, respectively, were administered through an Internet survey company. Unidimensionality was evaluated using confirmatory factor analysis. Item responses were modeled using item response theory. The presence and impact of differential item functioning (DIF) were evaluated using ordinal logistic regression.

Results—Unidimensionality of the pedsFACIT-PCF was supported. One of the 13 items demonstrated statistically significant DIF by language; however, impacts of language DIF on both individual scores and at the test level were negligible. No Spanish items showed DIF with respect to age and gender.

Correspondence to: Jin-Shei Lai.

Conflict of interest None.

Conclusions—The 13-item pedsFACIT-PCF demonstrated stable measurement properties on language, gender and age and can be used for future trials.

Keywords

Perceived Cognitive Function; Children; Differential item functioning; Language; Age; Gender

Introduction

Assessing and monitoring cognitive function for children at risk for cognitive decrements due to neurological conditions (e.g., childhood brain tumors) or neurotoxic treatments are important [1, 2], which is traditionally evaluated by neuropsychological testing. Unfortunately, such testing is not always feasible due to its length, costs, practice effects, or limited manpower in clinical settings. We thus developed the *pediatric Perceived Cognitive Function (pedsPCF) item bank*, which is intended to serve as a screening tool to identify children at risk and to facilitate timely referral for comprehensive assessments [3, 4].

The 43-item pedsPCF item bank is psychometrically sound, reliably differentiates children into clinical categories at significant accuracy rates, and is significantly correlated with computer-based neuropsychological tests [3–5]. The pedsPCF consists of both self- and proxy-reported versions. Multiple short-forms are available, and their scores are comparable between each other as well as to the full-length item bank because all items are calibrated using Item response theory (IRT) model. Of 43 items, 13 items were selected to form the *Pediatric Functional Assessment of Chronic Illness Therapy-Perceived Cognitive Function (pedsFACIT-PCF)*. These 13 items were written based on concerns raised by patients with brain tumors, their parents, teachers and clinicians and have been included in the pedsFACIT—Childhood Brain Tumor Survivor [6]. We translated the English items into Spanish and other language versions (www.facit.org) according to the FACIT translation methodology [7], a validated and widely used iterative method consisting of forward- and back-translations, multiple reviews and testing with cognitive debriefing to ensure conceptual and measurement equivalence between language versions. Although the pedsPCF and derived short-forms show promising clinical utility, their development should consider the cross-cultural content, to ensure that different language versions tapped the same cognitive domains. In the case of Spanish, this is particularly important given that the proportion of minority groups, especially Latinos born into or descended from a Spanish-speaking community, is dramatically increasing in the US general population [8, 9].

Although the translatability of the pedsPCF items was assessed during the development process, this study provided further psychometric evidence to ensure that the pedsPCF is harmonized for cross-cultural or language comparisons, by assessing differential item functioning (DIF) [10], an approach which has been used to evaluate potential item bias across different populations [11, 12]. This study compared responses with the child self-reported version of the pedsFACIT-PCF items among English- and Spanish-speaking populations to provide evidence of the cultural/linguistic equivalence. To monitor PCF from childhood through adulthood, it is necessary to have a validated measure that accounts for developmental changes and differences between boys and girls. Although we found stable

item measurement properties in the English version within sample from general population and cancer patients [5], this study further compared responses with Spanish items by age and sex to ensure that the Spanish version is equivalent in measurement across these subgroups.

Materials and methods

Data sources and measures

This study was approved by the Institutional Review Board of all participating sites. Data were collected from Spanish-speaking parents and their children in the household drawn from the US general population by an Internet survey company, Toluna (www.toluna-group.com). Procedures for data quality control are described at <http://www.tolunagroup.com/about-toluna/about/data-quality-approach>. Toluna adopted standard operating procedures to ensure that survey data are reliable, valid, de-duplicated and representative. English-speaking sample that was collected in another study [4] served as the comparison for language DIF analysis.

The company sent emails to invite potential parents of participating children from their database to participate in this study. Potential participants were screened via Internet to ensure their eligibility. Criteria for the Spanish sample were (1) Spanish being the primary language, (2) Hispanic and (3) having a child aged 7 or older. Whereas criteria for the English sample included (1) English being the primary language, (2) non-Hispanic and (3) having a child aged 7 or older. Spanish-speaking parents must also meet the eligibility criteria (i.e., <3.0 score in the Short Acculturation Scale for Hispanics) [13]. Eligible parents provided consent for their child to participate and completed a brief socio-demographic and clinical form about the child. Then they passed the computer to the child to complete the items. Recruitment was terminated when the preset accrual goal was reached ($n = 300$ aged 8–12; $n = 300$ aged 13–17).

Sample characteristics

Data were collected from a total of 1358 English- and 604 Spanish-speaking children. Demographic characteristics of children in English and Spanish samples were similar. The mean $\pm$ SD age of children was 12.5 years $\pm$ 2.9 in English and 12.3 years $\pm$ 3.1 in Spanish samples. Both English and Spanish samples consisted of 50 % children aged 8–12 years, and 50 % aged 13–17 years. In the English sample, 83 % of children were white and 57 % were male, whereas in the Spanish sample 86 % were white and 54 % were male. The mean $\pm$ SD age of Spanish parents was 40.0 years $\pm$ 7.7; all were Hispanic or Latino origin; 84 % were white, and 38 % were male; 36.9 % were high school graduates or less, 30.8 % some college, and 32.3 % had a college degree or higher. Thirty-four percent of Spanish-speaking parents had been told by a physician or a health professional that their child had the following conditions: epilepsy (3.6 %), diabetes (2.0 %), cancer (3.3 %), muscular dystrophy (2.8 %), depression (4.8 %), anxiety (3.5 %), alcohol (2.2 %), sleep disorder (6.5 %), multiple sclerosis (0.7 %), asthma (15.9 %) and arthritis (0.8 %). Demographic characteristics of English-speaking parents (3) and clinical information of their children [4] were evaluated elsewhere and are not repeated here.

Statistical analysis

We reversed the pedsFACIT-PCF scores prior to data analysis so that high scores reflected better Perceived Cognitive Functioning. We estimated descriptive statistics and reliability indexes and examined the item response theory (IRT) assumptions for the Spanish items. Unidimensionality of the items was examined by fitting a one-factor confirmatory factor analysis (CFA) model using Mplus (criteria: comparative fit index, CFI > 0.9; Tucker-Lewis index, TLI > 0.9; root-mean-squared error of approximation, RMSEA < 0.08) [14, 15]. Local independence was assessed based on residual correlations among items (criteria: residual correlation < 0.2) [16]. We estimated IRT item parameters with Samejima's graded response model (GRM) [17] using IRTPro [18].

By using an IRT-based ordinal logistic regression framework as implemented in LORDIF [11], we evaluated DIF on language (Spanish vs. English) for both samples, as well as age (8–12 vs. 13–17 years old) and gender (males vs. females) for Spanish sample only. DIF on age and gender of English sample was evaluated elsewhere [5] and is not repeated here. Zumbo [19] suggested that a change in pseudo- R^2 statistics lower than 0.13 represents negligible DIF. The current study followed Paz et al.'s studies [20, 21] to adopt a more sensitive criterion (i.e., McFadden's pseudo- R^2 change < 0.02) to identify meaningful DIF. To test the impact of DIF on individual scores, we calculated differences between scores, ignoring DIF (initial theta) and those accounting for DIF (purified), and estimated the percentages of theta differences that exceed the median standard error (SE) of initial scores and the individual's SEs of initial scores [11]. To test its impact at the test level, we compared Cohen's effect sizes for differences between the two group means, ignoring DIF and those accounting for DIF.

Results

The range for item means $\pm$ SDs was 3.69 ± 1.09 – 3.95 ± 1.13 . Frequencies for the items are provided in Table 1. No items had any categories with <10 responses. Coefficient alpha was 0.97, and item–total correlation ranged from 0.79 to 0.86, suggesting adequate scale reliability. Unidimensionality of items was supported by the acceptable fit from one-factor CFA analysis: CFI = 0.997, TLI = 0.996 and RMSEA = 0.067. Factor loadings of all items were greater than 0.4, and no item-pairs had a residual correlation of 0.2 or higher. Thus, all items were retained for IRT calibration and DIF analyses.

Item parameters

Table 2 shows slope and threshold estimates from GRM for all 13 items. The slope estimates ranged from 3.01 to 4.12, suggesting that this scale showed high discrimination properties (expected values between 1 and 5). The threshold parameters ranged from –2.13 to 0.72. The IRT-scaled scores ranged from –2.207 to 1.659.

Identification and impact of DIF

One of the 13 items exhibited statistically significant language DIF (Table 3). Item “Your child reacts slower than most people his/her age when he/she plays games” was found to have uniform DIF based on the pseudo- R^2 change of 0.02 criterion. The same DIF result

was obtained when using the Zumbo's criterion. The impact of language DIF on test characteristic curves (TCCs) is shown in Fig. 1. The graph on the left shows the TCC for all 13 items, while the graph on the right shows the TCC for an item with DIF. These curves indicated that the expected total score is higher for the English- than the Spanish-speaking sample, controlling for underlying Perceived Cognitive Function. Figure 2 shows the difference between scores ignoring DIF and those that account for DIF. The difference ranged from -0.064 to 0.048 with a mean of -0.003 as shown in the right panel. No subjects (0 %) experienced salient DIF as their score changes were smaller than the median SE of initial scores (0.206) and their initial SE estimates (ranged from 0.166 to 0.547), indicating that the language DIF effect on individual scores was negligible. Cohen's effect size for the difference between two group means was small (0.019) after identifying and accounting for DIF (from 0.339 to 0.320), indicating that the impact of language DIF at the test level was also minimal.

No items had meaningful DIF with respect to age and gender within the Spanish-speaking sample, indicating stable parameters across age-groups (8–12 vs. 13–17 years old) and gender (male vs. female).

Discussion and conclusion

Lifelong decrements in cognitive function are among the most prominent concerns for children with neurological conditions, which negatively impact their quality of life and family well-being. As the proportion of non-English-speaking minority groups in the USA increases, it appears important to have adequate outcome measures to assess them equivalently [22, 23]. Results of this study support the satisfactory psychometric properties of the Spanish version of the pedsFACIT-PCF. DIF analysis showed that both individual- and group-level impact of DIF related to language of administration was minimal, producing only a small effect and clinically insignificant difference in the total score. Thus, we recommend that English version-based parameters could be used on the Spanish-speaking population, which allows for including Spanish samples in the future when evaluating Perceived Cognitive Function.

The limitations of this study are worth noting. We evaluated item bias using the pseudo- R^2 change criterion. However, there is no universal agreement on DIF detection methods. Different methods may have produced different results. Another limitation is that the sample was recruited online, and parents provided consent for their children. Thus, computer illiterate parents and their children were excluded. We did not screen children for Spanish language proficiency. Hence, we were not able to verify whether all children understood the items. Further, our study had two-third respondents who have completed some college or higher education. Although an Internet-based sample with high education status is not uncommon, further studies are necessary for validation of the pedsFACIT-PCF by using respondents with a wider range of education and technology competence that may better represent the US general population. Future studies may collect and evaluate additional information such as families' socioeconomic status as these factors may impact children's cognitive behaviors. Future research may include Perceived Cognitive Function as a

predictor of real-world outcomes, such as academic attainment and employment, to understand its predictive value in children's real life.

Acknowledgments

This study was supported in part by a grant from the National Cancer Institute (#R01CA125671; PI: Jin-Shei Lai). First author's efforts were supported in part by grants from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitation Services, US Department of Education (#H133F140037) and Craig H. Neilsen Foundation (#290474). We certify that no party having a direct interest in the results of the research supporting this article has or will confer a benefit on us or on any organization with which we are associated AND, if applicable, we certify that all financial and material support for this research and work are clearly identified in the manuscript. The content of this article is solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

Abbreviations

CFA	Confirmatory factor analysis
DIF	Differential item functioning
GRM	Graded response model
IRT	Item response theory
PRO	Patient-reported outcomes
pedsFACIT	Pediatric Functional Assessment of Chronic Illness Therapy
PCF	Perceived Cognitive Function
SE	Standard error
TCC	Test characteristic curve

References

1. Mulhern RK, Merchant TE, Gajjar A, Reddick WE, Kun LE. Late neurocognitive sequelae in survivors of brain tumours in childhood. *Lancet Oncology*. 2004; 5(7):399–408. [PubMed: 15231246]
2. Tonning Olsson I, Perrin S, Lundgren J, Hjorth L, Johanson A. Long-term cognitive sequelae following pediatric brain tumor related to medical risk factors, age and gender. *Pediatric Neurology*. (in press).
3. Lai J-S, Butt Z, Zelko F, Cella D, Kieran MW, Krull KR, et al. Development of a patient-report cognitive function item bank using item response theory and exploration of its clinical utility in computerized adaptive testing. *Journal of Pediatric Psychology*. 2011; 27:285–293.
4. Lai J-S, Zelko F, Butt Z, Cella D, Kieran MW, Krull KR, et al. Parent-perceived child cognitive function: Results from a sample drawn from the US general population. *Child's Nervous System*. 2011; 27(2):285–293.
5. Lai J-S, Zelko F, Krull KR, Cella D, Nowinski C, Manley PE, et al. Parent-reported cognition of children with cancer and its potential clinical usefulness. *Quality of Life Research*. 2014; 23(4): 1049–1058. [PubMed: 24197478]
6. Lai JS, Cella D, Tomita T, Bode RK, Newmark M, Goldman S. Developing a health-related quality of life instrument for childhood brain tumor survivors. *Child's Nervous System*. 2007; 23(1):47–57.
7. Eremenco S, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaire. *Evaluation and the Health Professions*. 2005; 28(2):212–232. [PubMed: 15851774]

8. Bureau, USC. [Accessed July 23, 2014] 2010 census shows nation's Hispanic population grew four times faster than total US population. 2011. http://www.census.gov/newsroom/releases/archives/2010_census/cb11-cn146.html.
9. Bureau, USC. [Accessed July 23, 2014] The Hispanic Population: 2010. 2011. <http://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf>.
10. Teresi JA, Stewart AL, Morales LS, Stahl SM. Measurement in a multi-ethnic society: Overview to the special issue. *Medical Care*. 2006; 44(11 Suppl 3):S3–S4. [PubMed: 17060831]
11. 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. *Journal of Statistical Software*. 2011; 39(8):1–30. [PubMed: 21572908]
12. Crane PK, Gibbons LE, Ocepek-Welikson K, Cook K, Cella D, Narasimhalu K, et al. A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research*. 2007; 16(1):69–84. [PubMed: 17554640]
13. Marín G, Sabogal F, VanOss Marín B, Otero-Sabogal F, Pérez-Stable EJ. Development of a short acculturation scale for Hispanics. *Hispanic Journal of Behavioral Sciences*. 1987; 9:183–205.
14. Muthén LK, Muthén BO. Mplus. 2008 www.statmodel.com.
15. Lai J-S, Crane PK, Cella D. Factor analysis techniques for assessing sufficient unidimensionality of cancer related fatigue. *Quality of Life Research*. 2006; 15(7):1179–1190. [PubMed: 17001438]
16. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the patient-reported outcomes measurement information system (PROMIS). *Medical Care*. 2007; 45(5):S22–S31. [PubMed: 17443115]
17. Samejima, F. Estimation of latent ability using a response pattern of graded scores (Psychometrika monograph no. 17). Richmond: Psychometric Society; 1969.
18. Cai L, Thissen D, du Toit S. IRTPRO 2.1 for windows. 2011 <http://www.ssicentral.com/irt/>.
19. Zumbo, BD. A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.
20. Paz SH, Spritzer KL, Morales LS, Hays RD. Evaluation of the patient-reported outcomes information system (PROMIS[®]) Spanish-language physical functioning items. *Quality of Life Research*. 2013; 22(7):1819–1830. [PubMed: 23124505]
21. Paz SH, Spritzer KL, Morales LS, Hays RD. Age-related differential item functioning for the patient-reported outcomes information system (PROMIS[®]) physical functioning items. *Primary Health Care*. 2013; 3(131):1–4.
22. Cella D, Hernández L, Bonomi AE, Corona M, Vaquero M, Shiimoto G, et al. Spanish language translation and initial validation of the functional assessment of cancer therapy quality-of-life instrument. *Medical Care*. 1998; 36(9):1407–1418. [PubMed: 9749663]
23. Lent L, Hahn E, Eremenco S, Webster K, Cella D. Using cross-cultural input to adapt the functional assessment of chronic illness therapy (FACIT) scales. *Acta Oncologica*. 1999; 38(6): 695–702. [PubMed: 10522759]

Figure 1.

Impact of language DIF on test characteristics curves. *Note* The graph on the left created by the Lordif program (by default setting) shows two test characteristics curves (TCCs) (*black*: English sample; *red*: Spanish sample) for all of the items (both items with and without DIF), while the *graph* on the right shows two TCCs for the subtest of items found to have DIF. However, in this study, the *graph* on the right shows the item characteristics curves (ICC) for the item five because only this item was found to have DIF. The *Y*-axis in the graph on the right represents the item score, and the *X*-axis represents the theta (level of perceived cognitive function)

Figure 2.

Impact of language DIF at individual level. *Note* Both *graphs* show the difference between scores that ignore DIF and those that accounted for DIF. The *graph* on the left shows a *box plot* of these differences. The interquartile range, representing the middle 50 % of the differences (bound between the bottom and top of the *shaded box*), range roughly from -0.01 to 0.005 with a median of approximately -0.003 . The *graph* on the right shows the same difference scores (in *Y-axis*) that are plotted against the initial scores ignoring DIF (in *X-axis*), separately for English and Spanish samples. Guidelines are placed at 0.00 (*solid line*) (i.e., no difference) and the mean of the differences (*dotted line*). The positive values indicate that for majority of English individuals, accounting for DIF leads to slightly lower scores (i.e., initial score ignoring DIF minus score account for DIF > 0). Therefore, accounting for DIF score is less than the initial score for English-speaking sample, whereas the negative values indicate that for majority of Spanish individuals, accounting for DIF leads to slightly higher scores. Therefore, accounting for DIF score is higher than the initial score for Spanish-speaking sample

Table 1

Frequencies of pedsFACIT-PCF for Spanish items

Item	Spanish version	Corresponding English version	All of the time		Most of the time		Some of the time		A little of the time		None of the time	
			N	%	N	%	N	%	N	%	N	%
1	Olvido las cosas con facilidad	I forget things easily	19	3.1	78	13	140	23	145	24	222	37
2	Se me olvida traer a la escuela y a casa lo que necesito para hacer las tareas escolares (los deberes)	I forget to bring things to and from school that I need for homework	20	3.3	64	11	135	23	185	31	197	33
3	Se me olvida lo que voy a decir	I forget what I am going to say	12	2	55	9.1	155	26	176	29	206	34
4	Me cuesta mucho esfuerzo estar al tanto de mis tareas escolares (deberes)	I have a hard time keeping track of my homework	17	2.8	53	8.8	155	26	222	37	157	26
5	Tengo que leer las cosas varias veces para entenderlas	I have to read things several times to understand them	15	2.5	57	9.4	145	24	153	25	234	39
6	Tengo que usar listas escritas más a menudo que otras personas de mi edad para no olvidarme de las cosas	I have to use written lists more often than other people my age so I will not forget things	16	2.6	77	13	151	25	192	32	168	28
7	Tengo que hacer mucho esfuerzo para prestar atención, o de lo contrario cometo errores	I have to work really hard to pay attention or I make mistakes	18	3	75	12	158	26	169	28	184	31
8	Tengo problemas para prestar atención a los maestros	I have trouble paying attention to the teacher	15	2.5	60	9.9	135	22	186	31	208	34
9	Tengo problemas para acordarme de hacer las cosas, como proyectos escolares o tareas en casa	I have trouble remembering to do things like school projects or chores	16	2.6	56	9.3	136	23	131	22	265	44
10	Reacciono más despacio que la mayoría de las personas de mi edad cuando participo en juegos	I react slower than most people my age when I play games	21	3.5	75	13	148	25	156	26	202	34
11	Me resulta difícil concentrarme en la escuela	It is hard for me to concentrate in school	16	2.6	62	10	161	27	173	29	192	32
12	Me resulta difícil encontrar las palabras adecuadas para expresar lo que quiero decir	It is hard for me to find the right words to say what I mean	24	4	76	13	139	23	148	25	211	35
13	Tardo más que otras personas en terminar mis tareas escolares	It takes me longer than other people to get my school work done	14	2.3	62	10	159	26	164	27	205	34

Table 2

Slope and threshold parameters for the pedsFACIT-PCF 13 items in Spanish sample

Item	Slope	T1	T2	T3	T4
1. I forget things easily	3.82	-1.91	-1.12	-0.35	0.35
2. I forget to bring things to and from school that I need for homework	3.90	-1.88	-1.19	-0.39	0.48
3. I forget what I am going to say	3.99	-2.06	-1.34	-0.40	0.44
4. I have a hard time keeping track of my homework	3.21	-2.05	-1.36	-0.41	0.72
5. I have to read things several times to understand them	3.13	-2.13	-1.37	-0.46	0.30
6. I have to use written lists more often than other people my age so I will not forget things	3.01	-2.12	-1.20	-0.31	0.66
7. I have to work really hard to pay attention or I make mistakes	3.30	-2.00	-1.18	-0.28	0.57
8. I have trouble paying attention to the teacher	3.97	-1.96	-1.27	-0.46	0.42
9. I have trouble remembering to do things like school projects or chores	3.97	-1.94	-1.28	-0.46	0.15
10. I react slower than most people my age when I play games	4.08	-1.82	-1.10	-0.30	0.45
11. It is hard for me to concentrate in school	3.84	-1.95	-1.25	-0.33	0.51
12. It is hard for me to find the right words to say what I mean	3.94	-1.79	-1.08	-0.32	0.40
13. It takes me longer than other people to get my school work done	4.12	-1.98	-1.26	-0.34	0.43

T1 Threshold 1, T2 Threshold 2, T3 Threshold 3, T4 Threshold 4

Table 3

Item-level findings of DIF related to language of administration, age and gender

Item ^d	Language ^b			Age ^c			Gender ^d		
	Uniform	Non-uniform	Total	Uniform	Non-uniform	Total	Uniform	Non-uniform	Total
1	0.0002	0.0013	0.0015	0.0001	0.0029	0.0030	0.0012	0.0000	0.0012
2	0.0001	0.0015	0.0016	0.0005	0.0012	0.0017	0.0012	0.0011	0.0023
3	0.0042	0.0010	0.0052	0.0000	0.0000	0.0000	0.0000	0.0004	0.0004
4	0.0004	0.0026	0.0030	0.0015	0.0002	0.0017	0.0010	0.0003	0.0013
5	0.0246	0.0004	0.0251	0.0014	0.0002	0.0016	0.0000	0.0007	0.0007
6	0.0080	0.0002	0.0082	0.0012	0.0013	0.0025	0.0015	0.0004	0.0019
7	0.0012	0.0033	0.0046	0.0002	0.0004	0.0005	0.0000	0.0001	0.0002
8	0.0010	0.0001	0.0012	0.0001	0.0008	0.0009	0.0003	0.0027	0.0030
9	0.0060	0.0001	0.0061	0.0020	0.0009	0.0029	0.0004	0.0002	0.0005
10	0.0004	0.0022	0.0026	0.0020	0.0010	0.0030	0.0001	0.0000	0.0001
11	0.0046	0.0017	0.0063	0.0000	0.0000	0.0000	0.0029	0.0003	0.0032
12	0.0026	0.0009	0.0036	0.0002	0.0001	0.0003	0.0001	0.0000	0.0002
13	0.0059	0.0018	0.0077	0.0007	0.0004	0.0011	0.0024	0.0017	0.0041

^aThe wording of the items can be seen in Table 1^bLanguage subgroups were English ($n = 1358$) and Spanish ($n = 604$)^cAge subgroups were 8–12 years ($n = 298$) and 13–17 years ($n = 306$)^dGender subgroups were females ($n = 278$) and males ($n = 326$); DIF is present (*bold*) if McFadden's pseudo-R² change 0.02