



ORIGINAL ARTICLE

Cross-cultural adaptation and validation of the universal Portuguese-version of the Pediatric Functional Assessment of Chronic Illness Therapy – Fatigue (pedsFACIT-F)[☆]



Karla P. Fernandes ^{a,*}, Bruno S. Teixeira ^a, Benjamin J. Arnold ^b,
Tânia M. da S. Mendonça ^a, Sthela M. Oliveira ^a, Carlos Henrique M. da Silva ^{a,c,d}

^a Universidade Federal de Uberlândia, Faculdade de Medicina, Programa de Pós-graduação em Ciências da Saúde, Uberlândia, MG, Brazil

^b FACITtrans, Elmhurst, United States

^c Universidade de São Paulo (USP), Faculdade de Medicina, São Paulo, SP, Brazil

^d Universidade Federal de Uberlândia (UFU), Grupo de Pesquisa em Qualidade de Vida Relacionada à Saúde, Uberlândia, MG, Brazil

Received 1 August 2018; accepted 18 December 2018

Available online 26 April 2019

KEYWORDS

Fatigue;
Children;
Cross cultural
translation;
Validation;
Calibration;
Item response theory

Abstract

Objective: To cross-culturally adapt and validate the universal Portuguese version of the Pediatric Functional Assessment of Chronic Illness Therapy – Fatigue (pedsFACIT-F).

Method: The universal Portuguese version of the pedsFACIT-F was cross-culturally adapted and validated in 323 children and adolescents aged 8–18 years, 173 healthy individuals, and 150 with chronic diseases (cancer, juvenile idiopathic arthritis, and diabetes). Reliability (internal consistency and test–retest reliability) was assessed. Item response theory model assumptions were evaluated using confirmatory and exploratory factor analyses. Items were calibrated using a graded response model. Differential item functioning was assessed regarding age, gender, and clinical condition (healthy vs. chronic diseases).

Results: No major cultural adaptations were needed. Internal consistency (Cronbach's alpha = 0.84) and test–retest reliability (intraclass correlation coefficient = 0.92) were good. CFA (CFI = 0.92, TLI = 0.90, RMSEA = 0.097) and CFE analysis confirmed sufficient unidimensionality. The data also fit the GRM and demonstrated good coverage of the fatigue construct (threshold parameters range: –1.42 to 4.56). No items demonstrated significant differential item functioning.

[☆] Please cite this article as: Fernandes KP, Teixeira BS, Arnold BJ, Mendonça TM, Oliveira SM, Silva CH. Cross-cultural adaptation and validation of the universal Portuguese-version of the Pediatric Functional Assessment of Chronic Illness Therapy – Fatigue (pedsFACIT-F). J Pediatr (Rio J). 2020;96:456–63.

* Corresponding author.

E-mail: fernandeskp@uol.com.br (K.P. Fernandes).

<https://doi.org/10.1016/j.jpmed.2019.01.003>

0021-7557/© 2019 Sociedade Brasileira de Pediatria. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

PALAVRAS-CHAVE

Fadiga;
Crianças;
Tradução
transcultural;
Validação;
Calibração;
Teoria de resposta ao
item

Conclusion: The universal Portuguese version of the pedsFACIT-F provides a reliable, precise, and valid measure after being assessed by robust psychometric properties. Stability of the measurement properties of the pedsFACIT-F scale allows its use to assess fatigue in clinical research in Portuguese-speaking children and adolescents.

© 2019 Sociedade Brasileira de Pediatria. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Adaptação transcultural e validação da versão portuguesa universal da Escala Avaliação Funcional Pediátrica de Terapia de Doença Crônica – Fadiga (pedsFACIT-F)

Resumo

Objetivo: Adaptar transculturalmente e validar a versão portuguesa universal da escala Avaliação Funcional Pediátrica de Terapia de Doença Crônica – Fadiga (*pedsFACIT-F*).

Método: A versão traduzida para o português universal e adaptada transculturalmente da escala *pedsFACIT-F* foi validada em 323 crianças (entre 8 e 18 anos), 173 saudáveis e 150 com doenças crônicas (câncer, artrite idiopática juvenil e diabetes). A confiabilidade foi avaliada pela consistência interna e confiabilidade teste-reteste. Os pressupostos do modelo da teoria da resposta ao item foram avaliados por meio da análise fatorial confirmatória e exploratória. Os itens foram calibrados segundo modelo de resposta gradual. O funcionamento diferencial do item foi examinado com respeito à idade, ao gênero e à condição de saúde (saudáveis *versus* doenças crônicas).

Resultados: A adaptação cultural não apresentou dificuldades substantivas. A confiabilidade da consistência interna (alfa-Cronbach = 0,84) e do teste-reteste (correlação intraclassa = 0,92) foram adequadas. As análises da AFC (CFI = 0,92, TLI = 0,90, RMSEA = 0,097) e AFE confirmaram suficiente unidimensionalidade. O estudo de calibração demonstrou bom ajuste do MRG e boa cobertura do construto fadiga (variação dos limiares das categorias de resposta: –1,42 a 4,56). Não foi verificada presença de funcionamento diferencial do item significante.

Conclusão: A versão portuguesa universal da escala *pedsFACIT-F* é uma medida confiável, precisa e válida, verificada após análises de propriedades psicométricas robustas. A estabilidade das propriedades de medida da escala permite seu uso para avaliação de fadiga em estudos clínicos com crianças e adolescentes em países lusófonos.

© 2019 Sociedade Brasileira de Pediatria. Publicado por Elsevier Editora Ltda. Este é um artigo Open Access sob uma licença CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Fatigue is a symptom consisting of physical, mental, and emotional components, characterized by lack of energy, decreased ability to perform daily activities, and a feeling of tiredness.¹ It is a subjective, multidimensional construct that can present as acute, episodic, or chronic, and it is an important component of health-related quality of life, as it correlates with health status, pain, cognitive function, and the familial or social role played by an individual.² It is a frequent symptom in “healthy” adolescents (30%–40%), especially females³ and in children and adolescents with chronic diseases such as cancer, diabetes, inflammatory bowel disease, and juvenile idiopathic arthritis.¹

Fatigue assessment scales in children and adolescents are important because they allow the identification of the symptom impact, differentiate subjects with greater vulnerability, and evaluate and compare results of clinical trials.³ Some scales have been developed to measure fatigue in the pediatric age group: Fatigue Scale-Child (FS-C)⁴ and Fatigue Scale-Adolescent (FS-A)⁵; PROMIS (Patient-Reported

Outcomes Measurement Information System) – Pediatric Fatigue⁶; PedsQL Multidimensional Fatigue Scale (PedsQL-MFS)⁷; and Pediatric Functional Assessment of Chronic Disease Therapy – Fatigue (pedsFACIT-F).⁸ Of these, only PedsQL-MFS has been translated into Brazilian Portuguese and used to evaluate fatigue in children with juvenile idiopathic arthritis.⁹

PedsFACIT-F was developed to be answered by children and adolescents between the ages of 8 and 18 years, and its psychometric properties are considered adequate by classical and modern tests.⁸ PedsFACIT-F has been shown to be an important scale for clinicians and researchers wishing to assess the presence of fatigue in healthy and chronically-ill children and adolescents.¹ The collection of information directly from children and adolescents through a valid and reliable instrument is essential to better understand the perceptions of fatigue in this age group. The availability of this scale for universal use in Portuguese-speaking countries will contribute to the integral health care of this group of children and allow strategies to be established for prevention and treatment of this symptom. The aim of the present

work was to carry out cross-cultural adaptation and to evaluate the psychometric properties of the universal Portuguese version of pedsFACIT-F.

Method

Study design and participants

This is an observational methodological study, approved by the Ethics Committee for Research in Human Subjects of Universidade Federal de Uberlândia (CAAE: 13446913.2.0000.5152) and formal permission was obtained by e-mail from the FACIT.org representative for use of the original version of pedsFACIT-F.

Between September 2014 and September 2016, through a convenience sample of 323 children and adolescents, aged between 8 and 18 years, consisting of 173 healthy children from local public schools and 150 with chronic diseases treated at the Hospital de Clínicas da Universidade Federal de Uberlândia, were invited to answer the pedsFACIT-F scale. The scale was applied after the informed consent was signed by the parents and a term of assent was signed by children over 12 years old, and after proof of literacy was provided, required for the reading and interpretation of two items in the tool. It was determined that participants who did not answer more than two items would be excluded from the study. A questionnaire containing sociodemographic data was answered by the parents. Regarding the sample size, to evaluate the semantic and conceptual equivalence in relation to the original items, the pre-test version was tested in ten participants, as recommended by DeWalt et al.¹⁰ To evaluate the equivalence of psychometric measures, the authors followed the guidelines by Hair et al.,¹¹ by using a sample that contemplates a minimum of five to ten participants for each item of the scale.

The pedsFACIT-F scale

The pedsFACIT-F is a scale consisting of 13 Likert-like items with five response options ("never" – score = 0 to "always" – score = 4) that measure the feeling of fatigue in the last seven days, derived from the Pediatric Fatigue Item-Bank (peds-FIB) and consisting of 51 items calibrated through item response theory (IRT) methodology. The pedsFACIT-F evaluates fatigue as a one-dimensional concept and shows satisfactory psychometric properties according to analyses by classical and modern theory.⁸

Translation and cultural adaptation

The pedsFACIT-F scale was translated into Portuguese according to the FACIT translation methodology criteria, a widely validated iterative method consisting of the stages of translation, back-translation, review of back-translation by independent reviewers, and harmonization. During the phase of translation into Portuguese, the scale was evaluated by five translators and one reviewer representing FACIT (FACIT.org). After answering the items, the children were invited, through a retrospective and cognitive interview, to comment on the understanding of each item

and its relevance to them, both personal and cultural, and were able to make suggestions about items or terms considered difficult to understand, offensive, redundant, or inappropriate. After the pre-test was applied, the answers were sent to the authors of the original scale to rephrase the terms, if necessary.¹²

Statistics and psychometric tests

Descriptive statistical analysis was used for the participants' sociodemographic data. For item analysis, the scores obtained from the reversed items were reverted, so that higher scores represented a greater perception of fatigue by the respondent.

Reliability

The scale reliability was assessed through internal consistency using Cronbach's alpha coefficient. Coefficients >0.7 were considered acceptable.¹³

To verify reproducibility, the scale was reapplied to 30 participants after 14 days and the intra-class correlation coefficient was calculated. Correlations were considered adequate when >0.90.¹⁴

Validity

To confirm the original model of the pedsFACIT scale, construct validity was evaluated through confirmatory factorial analysis (CFA).¹⁵ Additionally, the epistemological assumptions for the assessed item response theory were unidimensionality, local independence, and monotonicity.¹⁶

The unidimensionality of the construct was verified by CFA, which used inter-item polychoric correlation matrices and the mean and variance – adjusted weighted least squares (WLSMV) of the MPlus program (MPLUS, Computer Software, Version 6.12, CA, USA).¹⁷ To evaluate the adjustment indexes of the proposed model, the coefficients' comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA) were used. The items were considered one-dimensional if CFI and TLI were >0.90 and RMSEA was <0.06.¹¹ In the case of violation of one of these criteria, the unidimensionality adequacy was confirmed if the analysis of the exploratory factorial analysis (EFA) (SPE Software Performance Engineering, v. 18.0, TX, USA) of the first extracted factor explained more than 20% of the total variance in the model and if the ratio between the percentage of explanation of the first factor was 4-fold greater than the percentage of explanation of the second factor.¹⁵

Local independence was investigated through the residual matrix correlation obtained in the CFA with a general factor. Residual correlations >0.25 indicate possible local dependence. The negative impact of local dependence was assessed by observing the change in item parameters after the individual removal of the items suspected of not complying with this assumption.¹⁶

Monotonicity was verified during the calibration of the items, through the program IRTPRO (IRTPRO_ Item response theory for patient-reported outcomes, v. 4.218, WA, USA).

The gradual response model (GRM)¹⁸ was run in this program, which considers a positive value for parameter (*a*) (discrimination), resulting in increasing monotonicity. Well-adjusted items are indications that this assumption has been duly met.¹⁹

Calibration

The pedsFACIT-F items were calibrated according to the GRM proposed by Samejima,¹⁸ which estimates the degree of discrimination, parameter (*a*), and the degree of difficulty, parameter (*b*), from the answers of items.¹⁸ Parameter (*a*) represents the ability of the item responses to discriminate respondents with close latent variable magnitudes. Values of parameter (*a*) between 0.01 and 0.34 are considered to be very low; from 0.35 to 0.64, low; from 0.65 to 1.34, moderate; from 1.35 to 1.69, high; and, above 1.70, very high.²⁰ Parameter (*b*) refers to the item difficulty, *i.e.*, its location along the continuum of the latent variable (fatigue) according to the threshold of the response categories, using the marginal maximum likelihood method. Ideally, perfect items should have enough coverage to contain the degrees of difficulty of the measured construct.²¹ The SS- χ^2 index was used to evaluate the fit of the model, and items with significant values (<0.01) were considered as unadjusted. These analyses were performed using the IRTPRO program (IRTPRO. Item response theory for patient-reported outcomes, v. 4.02, WA, USA).^{18,22}

Differential item functioning (DIF)

Differential item functioning (DIF) refers to the possibility that two individuals with the same level of latent trait (fatigue) respond differently due to some particular characteristic. In the present study, different age groups (8–12 years vs. 13–18 years), genders (male vs. female), and health status (chronic diseases vs. healthy) were considered.

The magnitude and impact of DIF on fatigue items were evaluated using ordinal logistic regression analysis models, using the computer software interface R (R Foundation for Statistical Computing, lordif package, version 0.2-2, Vienna, Austria).²³

The magnitude of the DIF was assessed using the chi-squared test (significance level of 0.01) and the “pseudo- R^2 ” McFadden’s test (with a 2% change in critical values). The impact of DIF was evaluated based on the coefficient of determination (R^2). R^2 coefficients <0.13 were considered non-significant, between 0.13 and 0.26, moderately significant, and >0.26, of significant DIF impact.²³

Results

Translation and cultural adaptation

The process of translation, back-translation, and review of the back-translation did not show any difficulties. For the pre-test phase, the scale was applied to ten children, five healthy ones and five with chronic diseases. In general, participants reported good understanding and that it was easy to answer the items. The reviewers considered the items

Table 1 Sociodemographic characteristics of the participants.

Characteristics	Chronic diseases (<i>n</i> = 150)	Healthy individuals (<i>n</i> = 173)
<i>Mean age; months (SD)</i>	138.0 (28.0)	138.2 (26.8)
<i>Age range (%)</i>		
8–12 years	106 (70.6)	114 (65.9)
13–18 years	44 (29.3)	59 (34.1)
<i>Gender (%)</i>		
F	77 (51.3)	99 (57.2)
M	73 (48.7)	74 (42.8)
<i>Level of schooling (%)</i>		
Elementary school	38 (25.3)	47 (27.2)
Middle school	98 (65.3)	121 (69.9)
High school	14 (9.3)	5 (2.9)
<i>Chronic diseases (%)</i>		
Juvenile arthritis	42 (28.0)	-
Cancer	46 (30.6)	-
Diabetes	62 (41.3)	-

SD, standard deviation.

relevant for the Portuguese culture and did not suggest any changes in the scale. The time taken to complete the scale ranged from five to ten minutes.

Participants

There were 323 children and adolescents who answered the scale, 150 patients with chronic disease and 173 healthy students. The sociodemographic data are shown in Table 1. The mean age was 138 months (SD = 26.4 months) and most were females, both among the patients (51.3%) and the students (57.2%). Over 80% attended elementary school.

Reliability

The reliability of internal consistency and reproducibility were considered adequate. Cronbach’s alpha coefficient obtained for the scale was 0.84 and the intra-class correlation coefficient was 0.92 (confidence interval of 0.87–0.95).

Validity

The CFA analyses showed a good fit of the model, with TLI=0.90, CFI=0.92, and RMSEA=0.097 (90% CI: 0.08–0.11). As there was a violation of the unidimensionality (RMSEA > 0.06), EFA was performed, in which the first factor extracted explained 38.0% of the total variance of the items and the second factor explained 8.2%, indicating the unidimensionality adequacy (Fig. 1). Only one pair of items (pF11 and pF1) had a residual correlation >0.25, suggesting possible violation of local independence. After the individual removal of these items, no differences were observed in the parameters estimated by IRT. Therefore, the local independence of the items was confirmed.

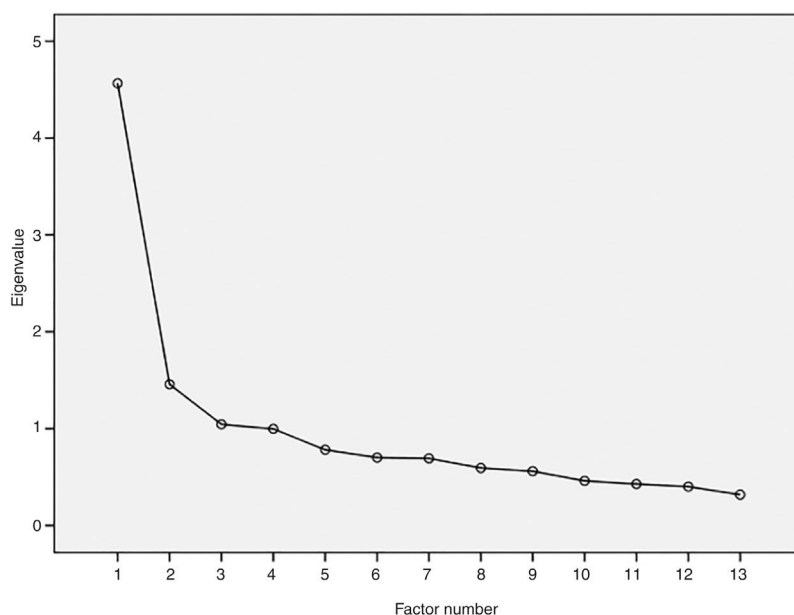


Figure 1 Scree plot extracted from the exploratory factor analysis of the pedsFACIT-F scale.

Calibration

The adjustment of the categories (item parameters) in the GRM is shown in Table 2.

The values of the discrimination parameter (a) ranged from 0.84 (pF3) to 2.28 (pF7). Five items (pF2, pF3, pF6, pF9, and pF11) showed values of the discrimination parameter (a) considered to be moderate, five (pF1, pF5, pF10, pF12, and pF13) showed high discriminating power, and three (pF4, pF7, and pF8), very high. The values of parameter (b) for difficulty ranged from -1.42 to 4.56 , which shows adequate coverage of latent trait variation.

The test information curve of the universal Portuguese version of the pedsFACIT-F scale that associates the test information with the observed measurement error is shown in Fig. 2.

DIF

There was no DIF between groups of participants with chronic diseases and healthy individuals. Regarding age, there was a uniform DIF in pF9 and pF11, and regarding gender, only item pF11 showed uniform DIF. The impact of DIF was considered irrelevant, since the pseudo- R^2 values obtained were <0.13 .

Discussion

The final version translated into universal Portuguese of the pedsFACIT-F scale was obtained after a strict methodology that guaranteed semantic, conceptual, and cultural equivalence. In addition to the classic psychometric measures, the validation process included IRT principles, which offer advantages in the evaluation of patient-reported outcomes (PROs), since they allow the creation of computerized adaptive tests or individualized scales with few items (short

forms), because they estimate the parameters of difficulty and discrimination of the item alone.²⁴ The universal Portuguese version of the pedsFACIT-F scale showed acceptable reliability and unidimensionality. Item calibration showed satisfactory model adjustments and consistent measurement properties regarding age and gender assessed by DIF.

It was observed that, during the translation and cultural adaptation phase, the scale was answered quickly and there were no suggestions for alterations in items or terms considered inadequate or unnecessary, which shows a good understanding and acceptance of the items by the respondents. Of the available scales for the evaluation of the fatigue symptom in children and adolescents, only PedsQL-MFS has been translated and culturally adapted to Brazilian Portuguese.⁹ It is interesting to note that the term “fatigue” (*fadiga*, in Portuguese) is not mentioned in the items of the universal Portuguese version of the pedsFACIT-F scale, but the construct is guaranteed by easy-to-understand expressions such as “lack of energy,” “tiredness,” and “need for rest.” In fact, the term “fatigue” is not always easy to understand, especially in the pediatric age group. Therefore, it is desirable to provide a scale for the evaluation of this symptom that translates the complexity and multidimensionality of the construct.^{25,26} The universal Portuguese version of the pedsFACIT-F scale is an important tool for health care providers and researchers from Portuguese-speaking countries who treat mainly children and adolescents with chronic diseases. The validation of a tool for the context in which it will be used guarantees the reliability of the collected data and its application in multicenter studies.¹² The reliability of the internal consistency and reproducibility of the universal Portuguese version of the pedsFACIT-F scale were considered adequate.

The pedsFACIT-F scale was originally validated for children and adolescents with cancer. As in the present study, the unidimensionality of the fatigue construct was also confirmed by CFA. However, it is not yet clear whether

Table 2 Parameters of IRT items for the pedsFACIT-F scale.

Items	<i>a</i>	Thresholds of response categories				SS- χ^2
		<i>b</i> ₁	<i>b</i> ₂	<i>b</i> ₃	<i>b</i> ₄	
I feel tired pF1	1.61	−1.42	−0.22	1.40	2.22	0.39
I have energy (or strength) pF2	1.07	0.01	1.75	3.20	4.56	0.03
I managed to do my usual things at home pF3	0.84	0.45	1.83	3.05	4.37	0.24
I had difficulty getting things done because I was too tired pF4	2.10	−0.39	0.46	1.43	2.15	0.11
I had difficulty finishing things because I was too tired pF5	1.62	−0.34	0.64	1.75	2.69	0.17
I had to sleep during the day pF6	0.95	−0.51	0.67	1.86	2.46	0.06
I was irritated because I was too tired to do the things I wanted to pF7	2.28	−0.09	0.48	1.04	1.57	0.10
I had difficulty playing or going out with my friends as much as I wanted to because I was tired pF8	1.76	0.54	1.15	1.77	2.27	0.01
I needed help to do my usual things at home pF9	0.96	−0.07	1.11	2.35	2.96	0.02
I feel weak pF10	1.52	0.05	0.98	2.05	2.93	0.20
I was too tired to eat pF11	1.19	1.26	2.05	2.92	4.42	0.00
Being tired makes me feel sad pF12	1.56	0.52	1.31	1.82	2.41	0.37
Being tired makes me feel angry pF13	1.68	0.08	0.78	1.32	1.81	0.22

a, parameter (*a*) of item discrimination; *b*_{1–4}, categories of the parameter answers.
b, of item difficulty.

fatigue should be measured by a one-dimensional or multidimensional approach.⁸ In the process of evaluation of the psychometric properties of the PROMIS[®] pediatric fatigue item bank, two dimensions were identified (“fatigue” and “lack of energy”).⁶ In this study, only 23% of the participants had chronic diseases. Although fatigue occurs in both healthy children and those with diseases, this symptom has been studied predominantly in children with cancer²⁶ and is related both with “tiredness” and “lack of energy,” but in the context of healthy children only the dimension “tiredness” has been used.³ Therefore, perhaps the unidimensionality or multidimensionality of the concept depends on the assessed population, whether it consists of healthy individuals or those with chronic clinical conditions.

As identified in the validation of the original scale, no items were observed with DIF of significant magnitude in the present study, regarding age, gender, or health status. This means that the items in the universal Portuguese version of the pedsFACIT-F scale were equally understood regardless of age, gender, or presence of chronic diseases. Lai et al.,⁶ however, identified DIF in more than 50% of the items of the PROMIS[®] pediatric fatigue item bank when analyzing a

cohort of 3048 children and adolescents aged between 8 and 17 years. According to these authors, the understanding of the concept “tired” may vary with age. The pedsFACIT-F scale should be applied in future studies with a larger number of participants to confirm whether the understanding of the concept of fatigue remains stable throughout the pediatric age group.

Local dependence was verified only in two items: pF1: “I feel tired” and pF11: “I was too tired to eat.” Local dependence means that the response given to an item facilitates or interferes with responses to other items. However, after the individual exclusion of these items, no differences were observed in the parameters estimated by IRT. Therefore, the authors believe that, for the moment, an item should not be removed from a pair of items with local dependence, until this trend is proven in studies with healthy children and adolescents, as well as those with chronic diseases.

The calibration analyses indicate that the items in the pedsFACIT-F scale show a good fit in the IRT model. Most items showed good discrimination parameters and good coverage of the construct, since the variation of the item response thresholds was comprehensive.

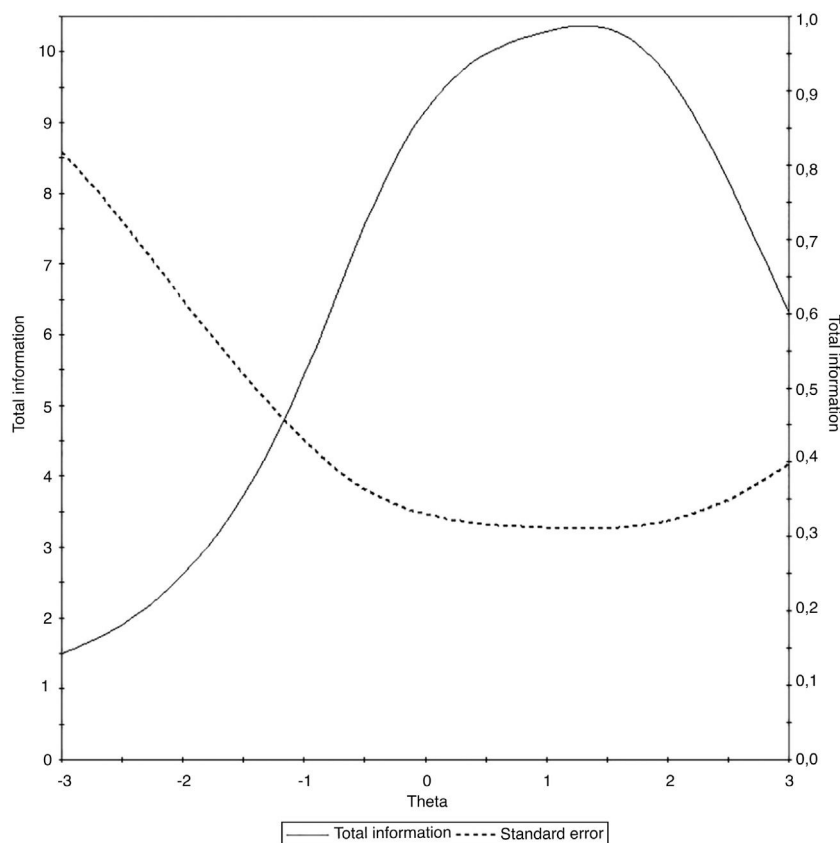


Figure 2 Test information curve of the universal Portuguese version of the pedsFACIT-F scale.

It is worth noting that, as in the validation of the original scale for children with cancer, the pedsFACIT-F scale was ideal for measuring fatigue in healthy children and adolescents and those with chronic diseases, with higher values of latent trait (theta between -1.4 and 3), *i.e.*, with moderate or severe symptom intensity. These results indicate the need to include items that can evaluate individuals with lower fatigue intensity.

Some limitations of the study can be verified regarding the sample composition and size for the psychometric analyses. Children and adolescents from a Brazilian city were recruited for the study. Given the territorial extension and Brazilian cultural diversity, as well as the possible variation in the understanding of the concept of “fatigue,” it is important to verify whether the Portuguese-language version of the pedsFACIT-F scale maintains adequate psychometric properties in other regions of the country or other Portuguese-speaking countries. The participation of healthy children and adolescents and those with chronic diseases (cancer, juvenile idiopathic arthritis, and diabetes) made it possible to validate the scale in the main clinical conditions associated with the fatigue symptom. However, it will be necessary to verify whether the absence of DIF for age, gender, and diagnosis remains in studies with a larger number of participants.

In conclusion, the universal Portuguese version of the pedsFACIT-F scale showed adequate psychometric properties evaluated by both the classical theory and IRT. This scale, which consists of a few items, is accurate, reliable,

and valid and, given the stability of its measurement properties regarding age, gender, and health status, it can be used for the evaluation of fatigue in clinical studies of children and adolescents, both healthy and with chronic diseases.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Crichton A, Knight S, Oakley E, Babl FE, Anderson V. Fatigue in child health conditions: a systematic review of assessment instruments. *Pediatrics*. 2015;135:e1015–31.
2. O’Connell C, Stokes EK. Fatigue: concepts for physiotherapy management and measurement. *Phys Ther Rev*. 2007;12:314–23.
3. Ter Wolbeek M, van Doornen LJ, Kavelaars A, Heijnen CJ. Predictors of persistent and new-onset fatigue in adolescent girls. *Pediatrics*. 2008;21:e449–57.
4. Hockenberry MJ, Hinds PS, Barrera P, Bryant R, Adams-McNeill J, Hooke C, et al. Three instruments to assess fatigue in children with cancer. *J Pain Symptom Manag*. 2003;25:319–28.
5. Hinds PS, Hockenberry M, Tong X, Rai SN, Gattuso JS, McCarthy K, et al. Validity and reliability of a new instrument to measure cancer-related fatigue in adolescents. *J Pain Symptom Manag*. 2007;34:607–18.
6. Lai JS, Stucky BD, Thissen D, Varni JW, DeWitt EM, Irwin DE, et al. Development and psychometric properties of

- the PROMIS[®]) pediatric fatigue item banks. *Qual Life Res.* 2013;22:2417–27.
7. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer reliability and validity of the pediatric quality of life inventoryTM generic core scales, multidimensional fatigue scale, and cancer module. *Cancer.* 2002;94:2090–106.
 8. Lai JS, Cella D, Kupst MJ, Holm S, Kelly ME, Bode RK, et al. Measuring fatigue for children with cancer: development and validation of the Pediatric Functional Assessment of Chronic Illness Therapy-Fatigue (pedsFACIT-F). *J Pediatr Hematol Oncol.* 2007;29:471–9.
 9. Paulo LT, Len CA, Hilario MO, Pedroso AS, Vitale MS, Terreri MT. Fatigue in children with juvenile idiopathic arthritis: reliability of the ‘‘Pediatric Quality of Life Inventory-Multidimensional Fatigue Scale’’. *Rheumatol Int.* 2015;35:133–9.
 10. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates. The PROMIS qualitative item review. *Med Care.* 2007;45:S12–21.
 11. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate data analysis.* 6th ed. New Jersey: Prentice Hall; 2006.
 12. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaire. *Eval Health Prof.* 2005;28:212–32.
 13. Nunally JC, Bernstein IH. *Psychometric theory.* 3rd ed. New York: McGraw-Hill; 1994.
 14. Fayers PM, Machin D. *Quality of life.* 2nd ed. Chichester: John Wiley & Sons; 2007.
 15. Kahn JH. *Factor analysis in counseling psychology research, training and practice: principles, advances and applications.* *Couns Psychol.* 2006;34:684–718.
 16. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Qual Life Res.* 2007;16:5–18.
 17. Muthen LK, Muthen BO. *Mplus: statistical analysis with latent variables.* Los Angeles, CA: Muthen and Muthen; 2004.
 18. Samejima F. Estimation of latent ability using a response pattern of graded scores. Chicago: Psychometrika Society; 1969 [Series Psychometrika Monograph Supplement, 17].
 19. Thissen DCL, du Toit SHC. *IRTpro for windows.* Lincolnwood, IL: Scientific Software International; 2017.
 20. Baker FB. *The basics of item response theory.* 2nd ed. Washington: Eric Clearinghouse on Assessment and Evaluation; 2001.
 21. Bortolotti SL, Tezza R, Andrade DF, Bornia AC. Relevance and advantages of using the item response theory. *Qual Quant.* 2013;47:2341–60.
 22. Orlando M, Thissen D. Further examination of the performance of $S-X^2$, an item fit index for dichotomous item response theory models. *Appl Psychol Meas.* 2003;27:289–98.
 23. Choi SW, Gibbons LE, Crane PK. 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.
 24. DeVellis RF. *Classical test theory.* *Med Care.* 2006;44:S50–9.
 25. Tomlinson D, Zupanec S, Jones H, O’Sullivan C, Hesser T, Sung L. The lived experience of fatigue in children and adolescents with cancer: a systematic review. *Support Care Cancer.* 2016;24:3623–31.
 26. McCabe M. Fatigue in children with long-term conditions: an evolutionary concept analysis. *J Adv Nurs.* 2009;65:1735–45.