Brazilian Portuguese translation, cross-cultural adaptation and reproducibility assessment of the modified Bristol Stool Form Scale for children

Debora Rodrigues Jozalaa, Isabelle Stefan de Faria Oliveirae, Erika Veruska Paiva Ortolanb, Wilson Elias de Oliveira Juniorc, Giovana Tuccille Comesc, Vanessa Mello Granado Cassettaire, Mariella Marie Sef, Pedro Luiz Toledo de Arruda Lourençob,*

Abstract
Objective: To translate and culturally adapt the modified Bristol Stool Form Scale for children into Brazilian Portuguese, and to evaluate the reproducibility of the translated version. Methods: The stage of translation and cross-cultural adaptation was performed according to an internationally accepted methodology, including the translation, back-translation, and pretest application of the translated version to a sample of 74 children to evaluate the degree of understanding. The reproducibility of the translated scale was assessed by applying the final version of Brazilian Portuguese modified Bristol Stool Form Scale for children to a sample of 64 children and 25 healthcare professionals, who were asked to correlate a randomly selected description from the translated scale with the corresponding representative illustration of the stool type.

* Corresponding author.
E-mail: lourencao@fmb.unesp.br (P.L. Lourenço).

https://doi.org/10.1016/j.jped.2018.01.006
0021-7557/© 2018 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/).
Introduction

Characterizing stool consistency is of fundamental importance for diagnosis, therapeutic monitoring, and scientific research on intestinal diseases.\(^1,2\) The use of graphic scales is an effective tool during this workup.\(^3-6\) The Bristol scale for stool consistency (Bristol Stool Form Scale [BSFS]) was developed and validated in the city of Bristol, United Kingdom, about two decades ago.\(^7-11\) This scale is composed of pictures representing seven stool types, associated with accurate descriptions of their shape and consistency, using recognizable illustrations.\(^11\) In adults, the use of the BSFS to determine stool consistency is well established in clinical practice, epidemiological studies, and clinical trials.\(^11-13\)

The BSFS for children (mBSFS-C) was recently modified.\(^4\) This scale proposed reducing the number of stool types from seven to five, and adapting the language used in the descriptions to allow children to fully understand the depictions of each stool type. The use of this scale has been validated in the United States for children aged 6–8 years, when the descriptions are read aloud, and for those aged over 8 years, who read the descriptions themselves.\(^5\)

These two scales were created and have been validated in the English language. However, to allow their use in clinical practice and in scientific research in countries with other official languages, the translation and cultural adaptation is fundamental.\(^6,14-16\) To achieve this, the process must follow internationally accepted standards, i.e., the items must not only be translated properly from a linguistic point of view, but must also be culturally adapted, while maintaining the validity of the original instrument.\(^15,16\)

In Brazil, the BSFS has been translated and cross-culturally adapted for Brazilian Portuguese for adult populations (>18 years).\(^6\) However, the mBSFS-C still has not undergone this process, to allow its application in the
Brazilian pediatric population. Thus, the authors decided to undertake the translation and cultural adaptation of the mBSFS-C into Brazilian Portuguese and to evaluate the reproducibility of the translated version.

Methods

This was a single-center study, conducted at the Botucatu Medical School, São Paulo State University (UNESP), São Paulo, Brazil, between June 2015 and July 2017. The study included 138 children aged between 6 and 18 years who had no cognitive impairment and who were being treated at the General Pediatric Outpatients Clinic due to other health problems, not related to intestinal complaints. A group of 25 healthcare professionals, composed of physicians and nurses with experience in the areas of gastroenterology and pediatrics, have participated during the reproducibility assessment stage. Participants and/or their guardians were informed of the purpose of the research and signed a term of free, informed consent. This study was approved by the local Research Ethics Committee, under protocol No. 28104614.8.0000.5411.

Stage 1: Translation and cross-cultural adaptation

To ensure the quality of the adaptation process, the scale was translated and adapted according to an internationally accepted and recommended methodology, consisting of six phases.6,14–18

Phase 1: Translation into Brazilian Portuguese

Two independent translations of the mBSFS-C were made by two bilingual translators whose mother language is Brazilian Portuguese.

Phase 2: Producing a synthesized forward translation

A meeting was held between the two translators who participated in Phase 1 and a team of specialists, composed of professionals with experience in child healthcare (two doctors, one nurse, one psychologist, and one nutritionist), and a university professor with extensive experience in cross-cultural adaptation of health assessment tools. A synthesized forward translation was produced based on the evaluation, reflection, and discussion of this group of professionals.

Phase 3: Back-translation

The synthesized translation was then back-translated into English, independently, by two bilingual translators whose mother language is English; these translators did not participate in the first phase and are not healthcare professionals. These translators were not informed of the concepts explored by the instrument, and both translations were made without prior knowledge of the original version of the scale.

Phase 4: Preparing the pre-final version of the translated scale

The pre-final version of the translated scale was produced after joint discussion and assessment between the four translators and the team of specialists. At this stage, the back-translations were compared with the original version of the mBSFS-C and all four forward-translated versions were analyzed, edited, and consolidated in the joint development of the pre-final version of the Brazilian Portuguese scale.

Phase 5: Pretest and evaluation of the degree of understanding

The pretest was applied to a sample of 74 children (21 aged 6–8 years and 53 aged 8–18 years). A five-point verbal numerical scale was applied to assess the degree of understanding of the translated version of the mBSFS-C as a whole and each of its components, in reference to the translated descriptions of each of the five stool types. The guiding question for evaluating the translated scale as a whole was "Did you understand what was asked and the differences between these types of stools?" and, for evaluating each of the individual items of the scale, "Do you understand what this type of stool is?" The minimum value was 0 ("I don’t understand anything") and the maximum was 5 ("I understand perfectly and have no doubts"). Values less than 3 were considered to indicate insufficient understanding.17,19 In conformity with the validation conducted in the United States, the translated scale descriptions were read aloud to the group of children aged 6–8 years, while the group of children over 8 years read the descriptions by themselves.5

Phase 6: Evaluating the results and producing the final version

This phase consisted of analyzing the results obtained in the pretest by the team of specialists and making minor modifications to the pre-final version based on difficulties in understanding observed in the population evaluated, in order to produce the final version of the Brazilian Portuguese mBSFS-C (Fig. 1).

Stage 2: Reproducibility assessment of the translated scale

This stage was conducted by five different examiners, all doctors, working in the area of pediatrics. The participants in this stage were 64 children (28 aged 6–8 years, and 36 over 8 years), and a group of 25 health care professionals with expertise in the areas of gastroenterology and pediatrics. The scale’s descriptions were only read aloud to children aged between the ages of 6 and 8 years.5

Instrument reproducibility was investigated by applying the final version of Brazilian Portuguese mBSFS-C for the 64 children and for the 25 health professionals, who were asked to correlate a randomly selected description from the translated scale with the corresponding representative illustration of the stool type. With this test, it was possible to compare the correlations obtained by the group of children with those obtained by health care professionals group.

Furthermore, the authors evaluated the potential influence of scale interpretation by different individuals (inter-observer reliability), by the same individual in different moments (intra-observer reliability), by different examiners (inter-examiners reliability), and by the same examiner in different moments (intra-examiner reliability).
To evaluate inter-observer reliability, a test was elaborated in which each of the five examiners applied the scale, with each asking five different children to correlate a randomly selected description and the corresponding representative illustration of the stool type; 25 of the 64 children participated, ten aged 6–8 years and 15 aged over 8 years. To evaluate intra-observer reliability, one of these examiners applied the scale five times to the same child, requesting that they correlate a randomly selected description with the corresponding representative illustration of the stool type, at 60-min intervals; ten of the 64 children participated, five aged 6–8 years and five over 8 years. Inter-examiner reliability was assessed by a test in which five different examiners asked the question “Which of these types of stool is most like yours most of the time?” to the same child, at 60-min intervals; 11 of the 64 children participated, five aged 6–8 years and six aged over 8 years. To evaluate intra-examiner reliability, an examiner posed this same question five times to the same child, at 60-min intervals; ten of the 64 children participated, five aged 6–8 years and five aged over 8 years.

Statistical analysis

The reproducibility of the translated scale was assessed by the agreement between the randomly selected translated descriptions and the corresponding illustrations of the stool types. The sample size was estimated in 73 agreement tests, which considered the expected discriminant capacity of the scale of 95%, with 95% confidence interval and 5% error. The agreement values were determined by Kappa with quadratic weights (Fleiss-Cohen), considering the ordinal character of the mBSFS-C. The tests conducted to evaluate the reliability of the translated scale in different times or occasions were analyzed by percentage rates of agreement between the descriptions and illustrations of the stool types and among the responses given by children when asked on more than one occasion. The influence of age in these scenarios was determined using a simple logistic regression model. To assess inter-examiner reliability, Cohen’s kappa index was used to determine the agreement between the children’s responses with different examiners, as this is a dichotomous variable. Comparative analyses between two groups were performed using Fisher’s exact test for categorical variables and Mann–Whitney test for numerical variables with abnormal distribution. The level of significance considered was 5%, and the analyses were performed using SPSS software (IBM SPSS Statistics for Windows, Version 22.0. NY, USA).

Results

Stage 1: Translation and cross-cultural adaptation

The synthesized forward-translation was based on the two previous translations of the scale from the source language to the target language, Brazilian Portuguese, followed by the assessment, reflection and discussion of the team of specialists, together with the Phase 1 bilingual translators. This version was independently back-translated into the source language of the questionnaire by two translators who did not participate in the first phase and who are not healthcare professionals. These translators are native English speakers who were not informed of the concepts explored by the instrument. The pre-final version was created following the assessment and discussion of the team of specialists, in conjunction with the four translators involved. At this stage, the back-translations were compared with the original version of the scale. The purpose of the team was to consolidate the translated versions and create the pre-final version of the Brazilian Portuguese mBSFS-C, having reviewed all translations to reach a consensus on any disagreement observed.

This pre-final version was assessed by a pretest (Table 1). The translated descriptions achieved a maximum of 15% of the values considered to indicate insufficient understanding, with median values of understanding higher than 3.00 obtained on the numerical verbal scale. No significant differences were observed when comparing the two age groups regarding the values of understanding obtained for the translated pre-final version and each of its descriptions. These results were discussed at a subsequent meeting of the team of specialists, when certain items that continued to present comprehension difficulties among the population tested were reviewed. The final version of the Brazilian Portuguese mBSFS-C is shown in Fig. 1.

Stage 2: Reproducibility assessment of the translated scale

Agreements between randomly selected descriptions from the translated scale and corresponding representative illustrations of the stool types, as determined by children and health professionals, are presented in Tables 2 and 3.
Table 1  Pretest results: degree of understanding of the translated version of the mBSFS-C as a whole and each of its components, in reference to the translated descriptions of each of the five stool types.

<table>
<thead>
<tr>
<th></th>
<th>6–7 years (n=21)</th>
<th>&gt;8 years (n=53)</th>
<th>p&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of individuals with insufficient understanding</td>
<td>Median&lt;sup&gt;a&lt;/sup&gt;</td>
<td>% of individuals with insufficient understanding</td>
</tr>
<tr>
<td>Pre-final version</td>
<td>10%</td>
<td>5 (1–5)</td>
<td>0%</td>
</tr>
<tr>
<td>Description 1</td>
<td>15%</td>
<td>5 (1–5)</td>
<td>11%</td>
</tr>
<tr>
<td>Description 2</td>
<td>15%</td>
<td>4 (1–5)</td>
<td>11%</td>
</tr>
<tr>
<td>Description 3</td>
<td>5%</td>
<td>5 (2–5)</td>
<td>7.5%</td>
</tr>
<tr>
<td>Description 4</td>
<td>15%</td>
<td>4 (1–5)</td>
<td>15%</td>
</tr>
<tr>
<td>Description 5</td>
<td>5%</td>
<td>5 (2–5)</td>
<td>5.6%</td>
</tr>
</tbody>
</table>

<sup>a</sup> Median (min/max) values based on the results obtained through the application of the numerical verbal scale.

<sup>b</sup> p-Value on the Mann–Whitney test.

Table 2  Number of agreements between randomly selected descriptions from the translated scale and illustrations of the stool types: children’s responses.<sup>a</sup>

<table>
<thead>
<tr>
<th>Randomly selected descriptions</th>
<th>Illustrations of the stool types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type 1</td>
</tr>
<tr>
<td>Description 1</td>
<td>15</td>
</tr>
<tr>
<td>Description 2</td>
<td>1</td>
</tr>
<tr>
<td>Description 3</td>
<td>0</td>
</tr>
<tr>
<td>Description 4</td>
<td>0</td>
</tr>
<tr>
<td>Description 5</td>
<td>0</td>
</tr>
</tbody>
</table>

<sup>a</sup> n = 87 agreements, determined by 64 children.

Table 3  Number of agreements between randomly selected descriptions from the translated scale and illustrations of the stool types: healthcare professionals’ responses.<sup>a</sup>

<table>
<thead>
<tr>
<th>Randomly selected descriptions</th>
<th>Illustrations of the stool types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type 1</td>
</tr>
<tr>
<td>Description 1</td>
<td>7</td>
</tr>
<tr>
<td>Description 2</td>
<td>0</td>
</tr>
<tr>
<td>Description 3</td>
<td>0</td>
</tr>
<tr>
<td>Description 4</td>
<td>0</td>
</tr>
<tr>
<td>Description 5</td>
<td>0</td>
</tr>
</tbody>
</table>

<sup>a</sup> n = 25 agreements, determined by 25 health professionals.

Almost complete agreement (k > 0.8)<sup>20</sup> was observed for both the children (k = 0.933) and the healthcare professionals (k = 0.990). The descriptors that presented the least number of errors in the correspondences performed by the children were types 1 (two errors), 3 (one errors), and 5 (no errors). Stratification according to age group showed that significant values of agreement were achieved by children aged between 6 and 8 years (k = 0.950), and those aged over 8 years (k = 0.975).

When assessing the reliability among the observers, it was observed that, on average, 88% (95% CI = 69–100%) of the children correctly matched the randomly selected translated description to the corresponding representative illustration of the stool type, when applied by the same examiner. A simple logistic regression model showed no significant association between the child’s age and the chance of error in matching the translated descriptions and stool types (OR = 1.31, 95% CI = 0.86–1.99; p = 0.200).

The test conducted to assess intra-observer reliability showed that 80% of the children correctly matched the translated descriptions to the corresponding representative illustration of the stool type in at least four out of five applications performed by the same examiner for the same child, at 60-min intervals. Half of the children correctly matched the items in all five applications. In this test, no significant association was observed between the child’s age and the chance of correctly matching the translated descriptions and stool types (OR = 0.47, 95% CI = 0.15–1.51; p = 0.206).

In the test conducted to assess inter-examiner reliability (Table 4), agreement was at least substantial (0.60–0.79)<sup>20</sup> among the children’s responses to the questions asked by five different examiners; the children were asked to point to
The most common stool type they evacuated after seeing the illustrations and interpreting the translated descriptions. Comparing the percentage of perfect agreement (100% agreement) among the responses obtained by five examiners, no statistically significant differences were verified between the group of children over 8 years and the group aged 6–8 years (66% vs. 100%, $p=0.454$, by Fisher’s exact test).

The test used to assess intra-examiner reliability showed variation in the response in three of the ten children (30%, 95% CI = 1.5–58%) who indicated the most common stool type they evacuated, after seeing the illustrations and interpreting the translated descriptions five times for the same examiner, at 60-min intervals. In these three cases of variation in the response, the maximum variation was between two stool types (types 2 and 4 twice, and types 1 and 3, once). The simple logistic regression model showed no significant association between the child’s age and the chance of variation in the responses obtained by the same examiner (OR = 0.47, 95% CI = 0.15–1.51; $p = 0.206$).

**Discussion**

The present study was the first to translate and cross-culturally adapt the mBSFS-C into a language other than English and to evaluate the reproducibility of this translated version. The translation and cultural adaptation are key steps. In addition to language, cultural aspects considerably influence the understanding of an instrument; therefore, the cultural adaptation of the original components is required. The sequential phases of translation, back-translation, and the meetings between the translators and the team of specialists led to the development of descriptions adapted to the understanding of Brazilian children. For example, the food items “nuts” and “sausage” do not form part of the daily lives of the majority of Brazilian children and therefore were adapted to “marbles” and “banana”, respectively. The satisfactory values of understanding achieved during the pretest applications of the translated mBSFS-C among children in both age groups indicate that the translated scale has been properly adapted to Brazilian culture.

The Brazilian Portuguese mBSFS-C was evidently reproducible, since almost complete agreement was obtained among the translated descriptions and illustrations of the stool types, both among the children and the group of specialists from the fields of gastroenterology and pediatrics. All the stool types showed percentages of correct matches considered appropriate. This was true for children aged 6–8 years and those over 8 years. As in the original study of mBSFS-C, the types of stool with the highest percentage of correct responses were types 1, 3, and 5. Although types 2 and 4 presented a greater number of disagreements, the majority of children who had these randomly selected descriptors were able to correlate them appropriately. The authors believe that this finding can be explained by the ordinal character of the scale, justifying the greater easiness in identifying extremes and greater difficulty in identifying intervals.

Moreover, the Brazilian Portuguese mBSFS-C produced very similar results for the same respondents at different times, characterizing stability, and for different examiners, characterizing equivalence, which constitute the two axes of external reliability. The present results are similar to those reported in the original validation study of the mBSFS-C, which obtained values of intra-class correlation coefficient greater than 0.70 for inter-observer reliability in the correlations made by children between stool photographs and the descriptors of the scale.

The tests for reliability focused on the examiner (between different examiners and between the same examiner at different times) were based on clinical application of the scale, and asked the children point to the most common type of stool they evacuated. In these scenarios, the way the examiner asks the question can influence the results obtained; however, the children’s responses were consistent. Internal reliability was not tested because the instrument is a scale, rather than a questionnaire with numerous questions. The reliability of the translated scale was maintained regardless of the children’s age.

Two main limitations of this study must be highlighted. The first is that it was conducted in a single center, which limits the generalizations and may bring strong bias

**Table 4** Inter-examiner reliability: agreement between the children’s responses to the questions asked by five different examiners.

<table>
<thead>
<tr>
<th>Comparison between the responses obtained by different examiners</th>
<th>% Of agreement</th>
<th>Kappa index</th>
<th>$p^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1-E2</td>
<td>91</td>
<td>0.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E1-E3</td>
<td>91</td>
<td>0.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E1-E4</td>
<td>82</td>
<td>0.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E1-E5</td>
<td>82</td>
<td>0.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E2-E3</td>
<td>100</td>
<td>1.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E2-E4</td>
<td>91</td>
<td>0.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E2-E5</td>
<td>91</td>
<td>0.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E3-E4</td>
<td>91</td>
<td>0.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E3-E5</td>
<td>91</td>
<td>0.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>E4-E5</td>
<td>100</td>
<td>1.00</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

E1, Examiner 1; E2, Examiner 2; E3, Examiner 3; E4, Examiner 4; E5, Examiner 5.

$^a$ $p$-value on kappa index.
related to the social, economic, and cultural background of the sample. The second limitation is related to the fact that children who participated in the study were waiting for consultation in the general pediatrics outpatient clinic. For this reason, the time interval for applying tests to the same individual had to be limited to 60 min. However, although this time interval can be considered limited, it can be emphasized that it was the same in all situations analyzed, minimizing potentially related memory bias.

Thus, the translation and adaptation of the mBSFS-C for Brazilian Portuguese is reliable for use when reading the descriptions for children aged between 6 and 8 years, and without assistance for children over 8 years old. The authors hope this version will be useful both in clinical practice and in scientific research in Brazil.

Funding

Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP. Processos n° 2014/04271-1 e 2015/03649-3.

Conflicts of interest

The authors declare no conflicts of interest.

References