



Published in final edited form as:

*J Pediatr.* 2011 September ; 159(3): 437–441.e1. doi:10.1016/j.jpeds.2011.03.002.

## Reliability and Validity of a Modified Bristol Stool Form Scale for Children

Mariella M. Lane, Ph.D.<sup>1,3</sup>, Danita I. Czyzewski, Ph.D.<sup>1,2,3</sup>, Bruno P. Chumpitazi, M.D., M.P.H.<sup>2,3</sup>, and Robert J. Shulman, M.D.<sup>2,3,4</sup>

<sup>1</sup> Baylor College of Medicine Menninger Department of Psychiatry and Behavioral Sciences, Houston, Texas

<sup>2</sup> Department of Pediatrics, Houston, Texas

<sup>3</sup> Texas Children's Hospital, Houston, Texas

<sup>4</sup> USDA/ARS Children's Nutrition Research Center, Houston, Texas

### Abstract

**Objectives**—This study sought to: 1) evaluate the ability of children to reliably use a modified Bristol Stool Form Scale for Children (mBSFS-C), 2) evaluate criterion-related validity of the mBSFS-C, and 3) identify the lower age limit for mBSFS-C use.

**Study design**—The mBSFS-C comprises five stool form types described and depicted in drawings. Children 3–18 years rated stool form for ten stool photographs. Due to low reliability when stool form descriptors were not read aloud (n=119), a subsequent sample (n=191) rated photographs with descriptors read.

**Results**—Intraclass correlation coefficients for descriptor-unread versus -read samples were 0.62 and 0.79. Children were increasingly reliable with age. Percentage of correct ratings varied by stool form type but generally increased with age. With descriptors unread, children 8 years and older demonstrated acceptable inter-observer reliability with over 78% of ratings correct. With descriptors read, children 6 years and older demonstrated acceptable reliability and over 80% of ratings correct.

**Conclusions**—The mBSFS-C is reliable and valid for use by children, age 6 being the lower limit for scale use with descriptors read and age 8 without descriptors read. We anticipate that the mBSFS-C can be effectively used in pediatric clinical and research settings.

### Keywords

Bristol Stool Form Scale; pediatric gastrointestinal disorders

---

Alterations in stool form are associated with numerous gastrointestinal symptoms and conditions that affect children. Patient report of stool form changes are used clinically for

---

© 2011 Mosby, Inc. All rights reserved.

Corresponding/reprint request author/author of first draft: Mariella M. Lane, Ph.D., Psychiatry Service, Texas Children's Hospital, 6701 Fannin St, Suite 1740, Houston, TX 77030. mmlane@texaschildrens.org Phone: (832)822-3750 Fax: (832)825-3747.

The authors declare no conflicts of interest.

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

diagnosis and management and are measured in research settings to assess clinical outcomes<sup>1,2</sup>.

Despite the importance of accurately assessing stool form changes, a stool form scale has not been validated for this use in pediatric patients. In adults, the Bristol Stool Form Scale (BSFS) is an increasingly used self-report instrument for assessing stool form, though it was validated as a measure of gastrointestinal transit time<sup>3-8</sup> without assessment of the ability of adults to reliably and accurately identify stool form type. The scale allows classification of stool form into seven types ranging from “separate hard lumps like nuts” (type 1) to “watery, no solid pieces” (type 7)<sup>3</sup>. The BSFS has been used to evaluate stool form in a variety of clinical studies<sup>9-16</sup>, and the Rome foundation has recommended its use to assess stool form in adults with functional GI disorders<sup>17</sup>.

In children, gathering information about stool form may be challenging because caregivers may not directly observe all stools. As such, healthcare providers are often placed in a position whereby a child’s description of his/her stools is required for diagnostic and clinical decision making. However, procuring accurate descriptions from children may be particularly challenging without an objective tool. Given that no stool form scale has been validated for use in children, we sought to validate a modified Bristol Stool Form Scale for children (mBSFS-C) for this indication.

Thus, the aims of this study were to: 1) evaluate the ability of children to reliably use a modified BSFS, 2) evaluate evidence for criterion-related validity of the modified scale by assessing children’s ability to correctly identify stool form type, and 3) identify the lower limit of age for which use of the mBSFS-C is appropriate. We expected to demonstrate evidence of inter-observer reliability and criterion-related validity of the mBSFS-C, hypothesizing both to increase with age and predicting that we could identify a minimum age for appropriate scale use.

## METHODS

As described previously, the original BSFS was adapted by decreasing the number of stool categories from seven to five<sup>18</sup>. This reduces the discriminations children are required to make, with the goal of maintaining scale utility while increasing accuracy of classification by young children who may have more difficulty attending to and discriminating between the original seven categories. Reducing response categories for self-report scale use by children is common practice.<sup>19-21</sup>

The five categories were chosen to adequately describe a range of stool form from hard through watery. Type 3 (“like a sausage or snake but with cracks on its surface”) and type 5 (“soft blobs with clear cut edges”) were eliminated as response options, with the remaining types 1, 2, 4, 6, and 7 comprising the modified scale. The final scale (Figure; available at [www.jpeds.com](http://www.jpeds.com)) consists of five stool form descriptors accompanied by drawings similar to the Rome II and Rome III renderings of the BSFS.<sup>17,22</sup>

In an initial evaluation of our modification of the BSFS, 14 pediatric gastroenterologists used the scale to rate 32 color photographs of stool. Photographs were obtained from publicly accessible areas of the internet and depicted focused, close-up photographs of entire bowel movements.<sup>18</sup> The modified scale had a high degree of inter- and intra-rater reliability when used by the expert raters<sup>18</sup>.

## Selection of Stool Photographs for Rating by Children

To select which photographs would be used as stimuli for child ratings, the distributions of stool form ratings by the expert raters were evaluated for the 32 stool photographs. Two photographs of each stool form type with high absolute agreement among the 14 expert raters were selected (Table I).

## Recruitment & Data Collection

The Baylor College of Medicine Institutional Review Board approved the study, and consent/assent were obtained. Participants included patients or siblings ages 3–18 years who were attending a scheduled outpatient clinic visit at Texas Children’s Hospital. Data were collected in the waiting room or an exam room. A trained research assistant presented children with the 10 stool photographs in random order, with the mBSFS-C printed beneath each. Children were asked to use the mBSFS-C to assign a stool form category to each photograph by pointing to their selection on the scale.

In our initial sample a research assistant presented the mBSFS-C without reading descriptors aloud. When these data indicated less than acceptable reliability for the total sample and particularly for the youngest children (see Results), we sought to evaluate if reading descriptors aloud could improve children’s ability to use the scale. In a second sample, stool form descriptors were read aloud for each photograph.

## Statistical Analysis

Statistics were performed using SPSS 17.0. Single measures intraclass correlation coefficients (two-way random effects model with absolute agreement) were used to assess inter-observer reliability. For both the descriptor-read and -unread samples, these values are presented for the total sample and 5 age ranges.

Sources of variance in the ratings (i.e., variance due to raters, photographs, or the interaction of rater and photograph) were examined for the descriptor-read total sample, where high degree of variance due to photographs vs. other sources provides strongest support for the scale. We also explored if children’s rating deviations from the expert rating (i.e., rating errors) were random or systematic; difference scores were not normally distributed, so nonparametric one-sample Wilcoxon Signed Ranks Tests were used. Kruskal-Wallis tests then explored if rating errors were related to age group.

Criterion-related validity was evaluated by comparing the child stool form ratings to the correct classifications determined by the physician ratings in our previous study. Percentages of children selecting the correct stool form classification for each photograph were calculated for the total descriptor-read sample and for each age range. To compare differences in criterion-related validity between the descriptor-read vs. -unread administration methods, overall percent correct for the total unread sample and for each age range was also calculated.

## RESULTS

### Inter-Observer Reliability with Unread Stool Form Descriptors

When stool form descriptors were not read to the children (n=119; 48.7% female), the single measures intra-class correlation coefficient for the total sample was 0.62 (95% CI: 0.44 – 0.85,  $P < 0.001$ ). Inter-observer reliability did not reach acceptable standards (i.e.,  $> 0.70$ ) until the children were 8–10 years of age, with the youngest children evidencing extremely low reliability (Table II). As previously mentioned, this led us to evaluate reliability and

validity of the mBSFS-C in a second sample when stool form descriptors were read aloud for each photograph.

### **Inter-Observer Reliability for the Total Sample with Stool Form Descriptors Read**

With stool form descriptors read for each photograph (n=191; 47.6% female), the single measures intraclass correlation coefficient for the total sample was 0.79 (95% CI: 0.64 – 0.93, P<0.001).

The variance in ratings due to the raters (i.e., assessment of tendency to rate all photographs more toward one end of the scale or the other) was very low at 0.01. Variance in ratings due to the interaction of raters by photographs (i.e., assessment of inconsistent choices for photographs) was also relatively low at 0.47. The variance in ratings due to the photographs (i.e., differences in stimuli are associated with different ratings) was much higher at 1.77. As such, the signal to noise ratio (variance in ratings due to photographs versus other sources of variance) was 3.77.

### **Inter-Observer Reliability by Age**

As with the descriptor-unread sample, we suspected that age would affect inter-observer reliability because children at younger ages would have more difficulty with the task. Table II provides intraclass correlation coefficient values across 5 age ranges, indeed indicating increasing reliability with age.

We next sought to explore if lower reliability in the younger children was attributable to random or systematic error. We first assessed if systematic bias (i.e., the tendency to rate as too firm or too loose) occurred across all age groups for each photograph. Results of the one-sample Wilcoxon Signed Ranks tests indicated that for the total sample, Photographs B (type 4; Z = -4.84, P<0.001), C (type 5; Z = -3.67, P<0.001), and I (type 5; Z = -3.57, P<0.001) tended to be rated as too firm, whereas Photographs F (type 2; Z = 7.17, P<.001), G (type 1; Z = 3.35, P< 0.01), and J (type 2; Z = 6.37, P<0.001) tended to be rated as too loose compared with the physician classifications. We note that for some of these photographs (types 1 and 5), this was the only direction that children could err because the correct classification was at the end of the scale.

Next, to evaluate if the younger children exhibited more systematic rating bias, a Kruskal-Wallis test was performed to determine if the difference scores (child ratings minus correct) varied according to the 5 age categories. Results indicated significant differences by age group for Photographs C (type 5;  $\chi^2 = 15.79$ , P< 0.01), F (type 2;  $\chi^2 = 26.05$ , P< 0.001), and J (type 2;  $\chi^2 = 15.97$ , P< 0.01). Examination of mean difference scores for the 5 age groups indicated that younger children were more likely than older children to make the error of rating Photograph C as too firm (the only direction they could err) and Photographs F and J as too loose, reflecting more systematic bias at the younger ages for these stimuli.

### **Criterion-Related Validity of the mBSFS-C**

The ability of the children to validly use the mBSFS-C was evaluated by comparing the child stool form ratings to the correct classifications as determined by the physician experts in the initial study.<sup>18</sup> Of the 1910 ratings made by our descriptor-read child sample, 1601 (83.8%) were in agreement with the physician expert classification. Of the 309 incorrect stool form classifications made, 139 (45%) were within one form type of the correct answer. Table I reports the percent of children selecting the correct stool form classification for each photograph for the total descriptor-read sample and each age range. For the total sample, the percent correct for each picture ranged from 61.8% to 99.0%. On average, the type 1 photographs (“hard lumps”) had the highest percent correct and type 4 (“fluffy pieces with

ragged edges”) the lowest. Overall percent correct increased with age; the 3–5 years age group had a notably lower overall percent correct, largely attributable to fewer correct ratings of stool types 2 (“sausage-shaped but lumpy”) and 4 (“fluffy pieces with ragged edges”).

### Comparison of Criterion-Related Validity for Read vs. Unread Descriptors

In an effort to evaluate differences in criterion-related validity between the descriptor-read vs. -unread administration methods, overall percent correct for the total descriptor-unread sample and for each age range was calculated. Table III presents unread sample values with those of the read sample, indicating a higher percent correct for the read sample across age ranges but most notably for the younger children.

## DISCUSSION

Despite the utility of assessing stool form in clinical practice and research, few have attempted to validate stool form scales, and such a scale has not been validated for use in children. We present the mBSFS-C as a stool form rating scale using pictorial representations and short descriptors that requires fewer discriminations than the standard BSFS while still capturing clinically relevant differentiations. We anticipate that the mBSFS-C can be effectively used in clinical and research settings by children with gastrointestinal disorders, and we hope further research will support its ability to accurately communicate stool characteristics from pediatric patients to physicians.

We did not initially intend to compare two methods of administration, but concerns about inadequate reliability prompted re-evaluation with stool form descriptors read aloud. Children 3–5 years could not use the mBSFS-C reliably under either condition. Children 6 years and older exceeded the common convention of  $\geq 0.70$  for acceptable inter-observer reliability when descriptors were read aloud, whereas those 8 years and older could reliably use the scale without descriptors read. Improvement in reliability may relate to helping child raters stay on task, a hypothesis supported by our data collection experience. Reading descriptors aloud may also prompt children to incorporate them instead of solely engaging in a visual matching task.

Considering criterion-related validity, correct ratings generally increased with age. Children 3–5 years old demonstrated markedly fewer correct classifications than older children in both samples. This and reliability data indicates that the mBSFS-C is too difficult for use by children in this age group. Overall percent correct was somewhat higher in the read versus unread sample, but discrepancies diminished with age.

Comparing child ratings with physician classifications shows that stool form type 1 was easiest for children to correctly identify, even at the youngest ages. Types 3 and 5 were also well identified, suggesting that children best discriminate these three non-adjacent and most distinct stool types. Types 2 and 4 were somewhat more difficult, but most children 6 years and older were also able to classify these correctly (descriptors read). Reducing to a three-point scale would improve young children’s reliability and accuracy but may unacceptably compromise clinical relevance and scale sensitivity.

With respect to methodological limitations, photographs of stool form type 2 that achieved high expert rater consensus were difficult to identify. We considered other potential stimuli but chose photographs as they would allow for a uniform assessment, were more feasible than actual stools, and relative to diagrams would provide a more realistic comparison to observing one’s own stool. One photograph achieved only 78.6% agreement by experts, with children varying in the same direction as physicians (i.e., deviations reflected looser

ratings). Although this stimulus may have been somewhat problematic, other photographs with over 90% expert agreement also suggest that types 2 and 4 are somewhat more difficult for children to correctly rate.

Although this study provides initial evidence for reliability and validity of the mBSFS-C, further scale evaluation would be useful. Though we previously demonstrated high intra-observer reliability in pediatric gastroenterologists<sup>18</sup>, our study in children did not include ratings at multiple time points. Further validation using actual stools would be contributive but perhaps a prohibitively onerous method of assessment. Evaluating sensitivity to change in response to treatment would also contribute support for mBSFS-C use as an outcome measure. Albeit difficult to assess, evaluating children's ability to use the mBSFS-C to accurately, retrospectively describe stool form or indicate "typical" stool form type would lend ecological validity to scale use during physician visits.

Overall, results provide evidence that our mBSFS-C is reliable and valid for children ages 6 years and older if stool form descriptors are read to them, and those 8 or more years of age could use the scale without descriptors read. That said, across all age ranges, both reliability and criterion-related validity were somewhat stronger when descriptors were read aloud. As such, if the mBSFS-C is being used for assessment at a single time point (low administration burden) and clinical decisions are being influenced by an individual measurement, such as during a physician visit, we recommend that stool form descriptors be read aloud regardless of age. However, if it is to be used for multiple ratings over time and ratings are to be aggregated as in research, we suggest that children 8 years and older can use the mBSFS-C independently with appropriate initial training in scale use.

## Acknowledgments

Supported by an investigator-initiated grant from Takeda Pharmaceuticals. Funding was independent of: 1) study design, 2) the collection, analysis, and interpretation of data, 3) the writing of the report, and 4) the decision to submit the paper for publication. Salary support to one or more of the authors during the conduct of this study has been provided by the National Institutes of Health (R01 NR05337, UH2 DK083990, and RC2 NR011959), the Daffy's Foundation, the USDA/ARS (under Cooperative Agreement 6250-51000-043 and P30 DK56338), which funds the Texas Medical Center Digestive Disease Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The contents do not necessarily reflect the views or policies of the USDA, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

The authors would like to express appreciation to Norah Vasen and Lesli Redenbaugh for their assistance with data collection and to Dr. Paul R. Swank for statistical consultation.

## Abbreviations

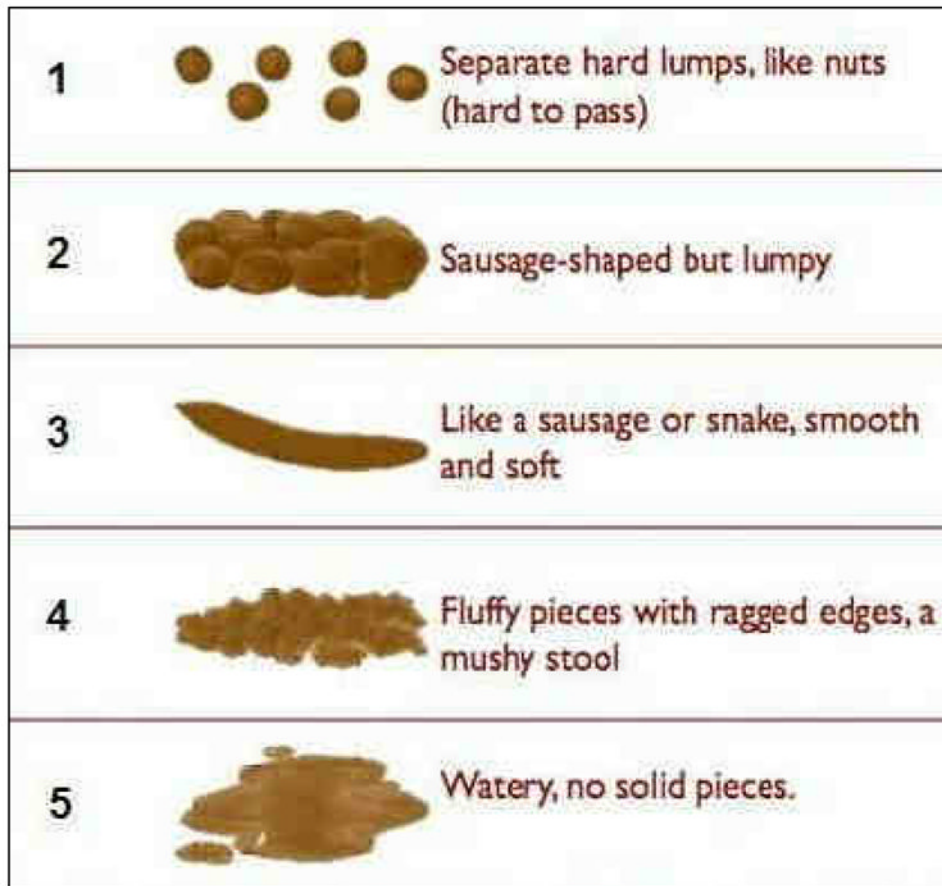
<b>BSFS</b>	Bristol Stool Form Scale
<b>mBSFS-C</b>	modified Bristol Stool Form Scale for Children

## References

1. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. *Gastroenterology*. 2006; 130:1480–91. [PubMed: 16678561]
2. Rasquin A, Di Lorenzo C, Forbes D, Guiraldes E, Hyams JS, Staiano A, et al. Childhood functional gastrointestinal disorders: child/adolescent. *Gastroenterology*. 2006; 130:1527–37. [PubMed: 16678566]
3. O'Donnell LJD, Virjee J, Heaton KW. Detection of pseudodiarrhoea by simple clinical assessment of intestinal transit rate. *BMJ*. 1990; 300:439–440. [PubMed: 2107897]



4. Probert CJ, Emmett PM, Heaton KW. Intestinal transit time in the population calculated from self made observations of defecation. *J Epidemiol Community Health*. 1993; 47:331–3. [PubMed: 8228773]
5. Heaton KW, O'Donnell LJD. An office guide to whole-gut transit time: patients recollection of their stool form. *J Clin Gastroenterol*. 1994; 19:28–30.
6. Degen LP, Phillips SF. How well does tool form reflect colonic transit? *Gut*. 1996; 39:109–113. [PubMed: 8881820]
7. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol*. 1997; 32:920–4. [PubMed: 9299672]
8. Saad RJ, Rao SS, Koch KL, Kuo B, Parkman HP, McCallum RW, et al. Do stool form and frequency correlate with whole-gut colonic transit? Results from a multicenter study in constipated individuals and healthy controls. *Am J Gastroenterol*. 2010; 105:403–411.
9. Heaton KW, Radvan J, Cripps H, Mountford RA, Braddon FE, Hughes AO. Defecation frequency and timing, and stool form in the general population: a prospective study. *Gut*. 1992; 33:818–24. [PubMed: 1624166]
10. Timmouth J, Tomlinson G, Kandel G, Walmsley S, Steinhart HA, Glazier R. Evaluation of stool frequency and stool form as measures of HIV-related diarrhea. *HIV Clin Trials*. 2007; 8:421–8. [PubMed: 18042507]
11. Giralt J, Regadera JP, Verges R, Romero J, de la Fuente I, Biete A, et al. Effects of probiotic *Lactobacillus casei* DN-114 001 in prevention of radiation-induced diarrhea: results from multicenter, randomized, placebo-controlled nutritional trial. *Int J Radiat Oncol Biol Phys*. 2008; 71:1213–1229. [PubMed: 18243569]
12. Konning CJ, Jonkers DM, Stobberingh EE. The effect of a multispecies probiotic on the intestinal microbiota and bowel movements in healthy volunteers taking the antibiotic amoxicillin. *Am J Gastroenterol*. 2008; 103:178–189.
13. Austin GL, Dalton CB, Hu Y, Morris CB, Hankins J, Weinland SR, et al. A very low-carbohydrate diet improves symptoms and quality of life in diarrhea-predominant irritable bowel syndrome. *Clin Gastroenterol Hepatol*. 2009; 7:706–708. e1. [PubMed: 19281859]
14. Silk DB, Davis A, Vulevic J, Tzortis G, Gibson GR. Clinical trial: the effects of a trans-galactooligosaccharide prebiotic on faecal microbiota and symptoms in irritable bowel syndrome. *Aliment Pharmacol Ther*. 2009; 29:508–518. [PubMed: 19053980]
15. Valerio F, Russo F, de Candia S, Riezzo G, Orlando A, Lonigro SL, et al. Effects of probiotic *Lactobacillus paracasei*-enriched artichokes on constipated patients: a pilot study. *J Clin Gastroenterol*. 2010; 44 (Suppl1):S49–S53. [PubMed: 20495470]
16. Koning CJ, Jonkers D, Smidt H, Rombouts F, Penings HJ, Wouters E, et al. The effect of a multispecies probiotic on the composition of the faecal microbiota and bowel habits in chronic obstructive pulmonary disease. *Br J Nutr*. 2010; 103:1452–1460. [PubMed: 20021703]
17. Drossman, DA.; Corazziari, E.; Delvaux, M.; Spiller, RC.; Talley, NJ.; Thompson, WG.; Whitehead, WE., editors. *The Functional Gastrointestinal Disorders*. 3. Degnon Associates, Inc; McLean, VA: 2006. ROME III.
18. Chumpitazi BP, Lane MM, Czyzewski DI, Weidler EM, Swank PR, Shulman RJ. Creation and initial evaluation of a stool form scale for children. *J Pediatr*. 2010; 157:594–597. [PubMed: 20826285]
19. Kovacs, M. *Children's Depression Inventory*. Multi-Health Systems; North Tonawanda, NY: 1992.
20. Spielberger, CD. *State-Trait Anxiety Inventory for Children*. Mind Garden; Redwood City, CA: 1973.
21. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: Reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Cores Scales in Healthy and Patient Populations. *Med Care*. 2001; 39:800–812. [PubMed: 11468499]
22. Drossman, DA.; Corazziari, E.; Talley, NJ.; Thompson, WG.; Whitehead, WE., editors. *Diagnosis, Pathophysiology and Treatment: A Multinational Consensus*. 2. Degnon Associates; McLean, VA: 2000. ROME II. *The Functional Gastrointestinal Disorders*.



**Figure.**  
The modified Bristol Stool Form Scale for Children (mBSFS-C)



**Table 1**

Percent (%) Correct Ratings for Each Stool Photograph (Descriptors Read)

	Age Range (years)					
	Total sample	3-5 (n = 34)	6-7 (n = 36)	8-10 (n = 43)	11-13 (n = 36)	14-18 (n = 42)
Photograph D (Type 1)	99.0	100	97.2	97.7	100	100
Photograph G (Type 1)	92.7	91.2	94.4	90.7	91.7	95.2
Photograph F (Type 2)	68.6	35.3	66.7	76.7	86.1	73.8
Photograph J (Type 2)	74.9	44.1	77.8	74.4	86.1	88.1
Photograph A (Type 3)	96.9	97.1	88.9	100	97.2	100
Photograph H (Type 3)	89.5	76.5	91.7	93.0	94.4	90.5
Photograph B (Type 4)	71.7	44.1	52.8	81.4	86.1	88.1
Photograph E (Type 4)	61.8	35.3	61.1	62.8	77.8	69.0
Photograph C (Type 5)	91.1	76.5	86.1	93.0	97.2	100
Photograph I (Type 5)	92.1	91.2	86.1	95.3	91.7	95.2
Overall Percent Correct	83.8	69.1	80.3	86.5	90.8	90.0

Note. Expert rater agreement for each photograph was as follows: Photograph A 100%, Photograph B 100%, Photograph C 100%, Photograph D 100%, Photograph E 92.9%, Photograph F 78.6%, Photograph G 100%, Photograph H 100%, Photograph I 92.9%, Photograph J 92.9%.

**Table II**

Intra-Class Correlation Coefficients by Age Range for Descriptors-Unread and -Read

Age Range (years)	Stool Form Descriptors Unread (n = 119)	Stool Form Descriptors Read (n = 191)
3–5	0.16 (n=22)	0.65 (n = 34)
6–7	0.61 (n=18)	0.74 (n = 36)
8–10	0.72 (n=25)	0.82 (n= 43)
11–13	0.81 (n=28)	0.86 (n = 36)
14–18	0.86 (n=26)	0.90 (n = 42)

**Table 3**  
Overall Percent (%) Correct Ratings for Stool Form Descriptors Unread vs. Read

	Total sample	3-5 years	6-7 years	8-10 years	11-13 years	14-18 years
Descriptors Unread (n = 119)	72.5	37.7	71.7	78.8	82.9	85.4
Descriptors Read (n = 191)	83.8	69.1	80.3	86.5	90.8	90.0