

The toileting habit profile questionnaire: Examining construct validity using the Rasch model

Isabelle Beaudry-Bellefeuille^{1,2} , Anita Bundy³, Alison Lane⁴,
Eduardo Ramos Polo⁵ and Shelly J Lane⁶

British Journal of Occupational Therapy
1–13

© The Author(s) 2018

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0308022618813266

journals.sagepub.com/home/bjot



Abstract

Introduction: Preliminary reports support the hypothesis that sensory issues may be related to atypical defecation habits in children. Clinical practice in this area is limited by the lack of validated measures. The toileting habit profile questionnaire was designed to address this gap.

Methods: This study included two phases of validity testing. In phase 1, we used Rasch analysis of existing data to assess item structural validity, directed content analysis of recent literature to determine the extent to which items capture clinical concerns, and expert review to validate the toileting habit profile questionnaire. Based on phase 1 outcomes, we made adjustments to toileting habit profile questionnaire items. In phase 2, we examined the item structural validity of the revised toileting habit profile questionnaire.

Results: Phase 1 resulted in a 17-item questionnaire: 15 items designed to identify habits linked to sensory over-reactivity and two designed to identify sensory under-reactivity and/or poor perception items. The analysis carried out in phase 2 supported the use of the sensory over-reactivity items. Remaining items can be used as clinical observations.

Conclusion: Caregiver report of behaviour using the revised toileting habit profile questionnaire appears to adequately capture challenging defecation behaviours related to sensory over-reactivity. Identifying challenging behaviours related to sensory under-reactivity and/or perception issues using exclusively the revised toileting habit profile questionnaire is not recommended.

Keywords

Child, constipation, faecal incontinence, occupational therapy, sensation disorders, occupational therapy

Received: 14 May 2018; accepted: 22 October 2018

Background

Occupational therapists have developed expertise in understanding and supporting participation and engagement in everyday activities (World Federation of Occupational Therapists, 2012). Bowel management, considered an important activity of daily living (American Occupational Therapy Association, 2014), is key to an individual's independence and successful social participation. Additionally, acquiring continence of the bowel is considered an important milestone of childhood. As such, addressing issues related to bowel management, such as disorders of defecation and atypical defecation habits, is an important component of occupational therapy practice.

Children with bowel management concerns often show atypical defecation habits that lead to defecation disorders. In many cases the defecation disorder is considered functional, as no known organic underlying factor can be identified (Tabbers et al., 2014). Childhood functional defecation disorders (FDD) such as functional

constipation (FC) have a high prevalence worldwide (5.3–17.4%; Van Den Berg et al., 2006) and are considered a public health problem (Rajindrajith et al., 2016). Although the Rome IV diagnostic criteria (The Rome Foundation, 2016) is considered the gold standard for identification of FDD, and there exists an extensive body of research describing a range of medical and behavioural interventions (Pijpers et al., 2010; Tabbers et al.,

¹PhD candidate, University of Newcastle, Callaghan, Australia

²Occupational therapist, Private practice, Oviedo, Spain

³Professor and Department Head (Occupational Therapy), Colorado State University, Fort Collins, USA

⁴Associate Professor (Occupational Therapy, School of Health Sciences), University of Newcastle, Callaghan, Australia

⁵Paediatric gastroenterologist, private practice, Oviedo, Spain

⁶Professor and Discipline Lead (Occupational Therapy, School of Health Sciences), University of Newcastle, Callaghan, Australia

Corresponding author:

Isabelle Beaudry-Bellefeuille, Marques de Santa Cruz, 7, 1E, 33007, Oviedo, Spain.

Email: ibbergo@gmail.com

2014; Van Ginkel et al., 2003), approaches to identification and treatment are inconsistent and the precise mechanisms of childhood FDD are not well understood (Beaudry-Bellefeuille et al., 2017; Koppen et al., 2018; Rajindrajith et al., 2016). As a result, treatment effectiveness remains limited and sound comprehension of all factors involved in the emergence of the disorder, along with greater understanding of treatment elements, are needed to optimize outcomes (Freeman et al., 2014).

Preliminary reports support the hypothesis that concerns about sensory reactivity (the process of modulating neuronal activity in response to sensory stimuli) and perception (the ability to recognize and interpret sensory stimuli) may be related to atypical defecation habits (Beaudry-Bellefeuille and Lane, 2017; Beaudry-Bellefeuille and Ramos Polo, 2011; Beaudry et al., 2013). For instance, using the short sensory profile (McIntosh et al., 1999), Beaudry-Bellefeuille and Lane (2017) reported significantly more sensory over-reactivity in children with FC than in typically developing children. Furthermore, preliminary reports of the effectiveness of intervention programmes that consider the sensory issues of children with FDD are promising (Beaudry-Bellefeuille and Ramos Polo, 2011; Beaudry et al., 2013). The field of occupational therapy has developed a strong expertise in the assessment of and intervention for sensory issues affecting participation in daily occupations. These skills can be valuable in identifying underlying sensory issues that may impact defecation habits, and in adapting the toileting environment to better fit the child's ability. Clinical practice in this area, however, is limited by the lack of validated measures that can clearly distinguish children with sensory related FDD from those without such problems.

The toileting habit profile questionnaire (THPQ) (Beaudry-Bellefeuille et al., 2016) is a screening questionnaire designed to: (1) differentiate children with typical toileting habits from those with atypical toileting habits and (2) identify toileting habits potentially related to sensory concerns in children with FDD (such as constipation, faecal incontinence, and stool toileting refusal). This tool was developed by an occupational therapist (IBB) in collaboration with a gastroenterologist (ERP). Based on available literature, clinical experience, and descriptions by parents of behaviours common in children referred to occupational therapy for difficulties establishing healthy, age-appropriate defecation habits, this team identified children with FDD that had not responded to usual behavioural and/or medical management and where the core problem appeared to be sensory-based. Scored on a five-point Likert scale (almost always to never), the THPQ includes items such as *My child refuses to go to the toilet outside of the home* or *My child hides while defecating*, with low scores corresponding to more frequent problematic behaviours and habits. Subsequently, the THPQ was shown to have face and preliminary content validity (Beaudry-Bellefeuille et al., 2016).

The THPQ has been found to be clinically useful in defining sensory concerns relative to FDD (Beaudry-

Bellefeuille and Lane, 2017; Beaudry et al., 2013). A recent pilot study using the THPQ demonstrated that children with FC that had not responded to first-line medical management ($n = 16$) demonstrated a significantly higher frequency of habits hypothesized to be linked to sensory over-reactivity than typically developing children ($n = 27$) (Beaudry-Bellefeuille and Lane, 2017). While preliminary studies regarding the THPQ were promising, we recognized that the original item set, developed more than 10 years ago, required updating and that further examination of the psychometric characteristics of the THPQ was warranted before recommending its use in clinical practice. For the occupational therapy practitioner, the use of psychometrically sound assessment tools, which accurately reflect a person's level of skill, is essential in understanding the potential factors linked to occupational challenges (Schaaf, 2015). Furthermore, psychometrically sound assessment tools allow occupational therapists to defend and substantiate the need, impact, and efficacy of their interventions (Brown and Bourke-Taylor, 2014).

Recent views in psychometrics highlight the need to carefully examine the quality of patient and caregiver questionnaires before they can be used in research or clinical practice (Mokkink et al., 2010). Given that the constructs measured by these instruments are subjective, evaluating whether the instruments measure these constructs in a valid and reliable way is crucial (Mokkink et al., 2010). The COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) initiative recommends evaluating content validity of measurements before testing construct validity (Terwee et al., 2018). As such, this study describes two phases of validity testing, addressing both content and construct validity. In both phases we chose to use the Rasch measurement model (Rasch, 1960). This model allows for the construction of robust instruments that aim to measure human traits in such a way that each individual is characterized separately and independently of which instruments have been used (Rasch, 1960). Rasch models are mathematical models that require unidimensionality (a single construct measured by a set of items) and result in additivity (measurement units are the same size along the continuum) (Smith et al., 2002). Data collected from questionnaires are tested against the expectations of the model, and when the data fit the model we obtain a clear impression of the relative difficulty of items (Smith et al., 2002). Furthermore, when the data fit the Rasch model, the analysis can provide a linear transformation of ordinal raw scores, which can then confidently be used with parametric statistical tests (Boone et al., 2013). Thus, the Rasch model offers a comprehensive approach to addressing several aspects of scale development and construct validation, as well as providing a transformation of ordinal raw scores into linear scale scores (Boone et al., 2013; Pallant and Tennant, 2007; Smith et al., 2002).

In phase 1, we extend previous work with the THPQ that used expert review of test content and hypothesis testing (known group comparisons studies) to assess content and construct validity (Beaudry-Bellefeuille and Lane,

2017; Beaudry-Bellefeuille et al., 2016). Here we used Rasch analysis of existing data to assess item structural validity. To further validate the content of the THPQ, we used directed content analysis of recent literature to determine the extent to which items capture clinical concerns as well as expert review of items. Based on phase 1 outcomes, we made adjustments to THPQ items. In phase 2, we examined the structural validity of the revised THPQ (THPQ-R).

Method - Phase 1

We subjected anonymized data from a pilot study (Beaudry-Bellefeuille and Lane, 2017) to Rasch analysis using Winsteps Version 4.0.1 (Linacre, 2017a). Specifically, we asked the following question: *How well do the items of the original version of the THPQ contribute to productive measurement of sensory reactivity and perception of defecation related sensations?* Ethics approval was obtained from the University of Newcastle (#H-2016-0282). Written informed consent was not considered necessary given that this study dealt with existing, anonymized data.

Participants

Participants included parents of children with FC and no other diagnosis ($n = 16$) and parents of typically developing children ($n = 27$). Diagnosis of FC and screening for medical conditions were done by the child's referring physician as part of standard medical management of FDD. Parents of children with organic causes of defecation disorders were excluded. For both groups, parents of children with intellectual disability, neurological conditions, or psychiatric disorders were excluded.

Measure

The THPQ (Beaudry-Bellefeuille et al., 2016) is a bilingual (English–Spanish) parent-report screening tool to help differentiate typical defecation behaviours and habits from those that are associated with FDD potentially related to sensory concerns.

Data analysis

Rasch analysis transforms ordinal raw data into interval data expressed in log odds probability units (logits); both item difficulty and person scores are expressed in this unit. In the THPQ, high scores are associated with a lower frequency of problematic habits. Based on Rasch measurement assumptions, all respondents are more likely to award high scores on easy items. Likewise, respondents whose children do not have FDD are more likely to award higher scores on the more difficult items than those who have children with FDD (Wright and Stone, 1979). In the case of the THPQ, easy items are those on which respondents will award the majority of responses the maximum score ('never'; score 5), such as item 7 (see

Table 1). Likewise, difficult items are those on which respondents will award the majority of responses the minimum score ('almost always'; score 1), such as item 6. Using Rasch analysis, we examined score patterns for evidence of internal reliability and structural validity. Any concerns identified could then be addressed by means of supplemental items or adjustments to existing items before collecting new data.

We used several sources of evidence obtained from the Rasch calculations to examine item structural validity; these are as follows. (For an introduction to Rasch calculations refer to Boone et al., 2013.)

Item correlation. Winsteps provides a calculation of the Pearson correlations between each item and the overall test. Correlations were expected to be positive; a negative correlation between an item and the overall measure indicates that the item is not part of the construct. The size of a positive correlation is of less importance than the fit of the responses to the Rasch model (see below); as such, no specific value was set for the correlations (Linacre, 2017b).

Rating scale category structure. Scale categories are expected to be evenly distributed and ordered from 1 to 5, and the distance between categories should be at least 1.4 logits (Linacre, 2002). When categories are not ordered as expected or are too closely grouped together, scoring of items should be reconsidered (Andrich et al., 1997).

Goodness of item fit statistics. We examined goodness of item fit statistics (ratio of observed to expected scores), expressed as mean square (MnSq) values, to identify items that conform to the expectations of the Rasch model (MnSq of 1.0); we accepted values up to 1.5, as suggested by Linacre (2002). Both infit and outfit statistics were considered. Infit statistics are sensitive to patterns of inlying observations such as persons responding to most of the easy items incorrectly and most of the hard items correctly (Linacre and Wright, 1994). Outfit statistics are sensitive to outliers and pick up unexpected events such as a difficult item that a low-performing person responded to correctly. When MnSq values were greater than 1.5, we reconsidered the item, the respondent's interpretation of the item, and the theory underlying the item as possible sources of divergence from the Rasch model (Boone et al., 2013). Items with unacceptably large fit values should be carefully considered to determine if they are part of the construct. Items deemed theoretically sound in spite of unacceptably large fit statistics must be carefully reviewed for clarity of expression. Additionally, items with unacceptably large fit values can be examined for unexpected responses; removal of these responses to observe the impact on the fit statistics is another way of exploring divergent fit.

Construct representation. Rasch arranges item difficulty and person ability on a single hierarchy. We examined the spread in item difficulty, looking for regions suggesting the construct being examined was not well represented (no

Table 1. Items of the toileting habit profile questionnaire – revised.

Type of Sensory issue	THPQ-R Item#	THPQ Item#	Challenging defecation behaviours potentially related to sensory processing
1	1	1	My child hides to poop.
1	2	2	My child asks for a diaper when he/she feels the need to poop.
1	3	0	My child prefers to poop in his/her clothing although the potty or toilet is nearby.
1	4	3	My child refuses to sit on the potty or the toilet to poop, but will accept peeing in the potty or toilet.
1	5	0	My child refuses to sit or seems uncomfortable sitting on the toilet or potty for both peeing and pooping, even at home.
1	6	0	My child withholds poop or resists the urge to poop.
1	7	4*	My child follows an unusual ritual when pooping which involves actions or places not typically associated with pooping or with the age of the child.
1	8	5	My child seems to feel pain when pooping, even if the poop is soft.
1	9	6	My child refuses to poop outside of the home.
1	10	7	My child shows exaggerated disgust at the smell of his/her poop.
1	11	8	My child refuses to wipe or be wiped after pooping.
1	12	0	My child shows fear or refusal related to certain features of the bathroom, such as fear of flushing the toilet.
1	13	0	My child needs to pay attention to something else while pooping (a book, a game); this seems to help him/her tolerate the sensation of pooping.
1	14	0	My child is sensitive to taste and/or food textures, making it difficult for him/her to accept laxative medicine or high fibre foods.
1	15	0	My child felt the urge to poop very early (younger than 12 months). My child would grunt in a certain way and I would sit him/her on the potty to poop.
2	16	9	My child does not seem to feel the urge to poop.
2	17	10*	My child does not realize he/she has soiled (pooped in) his/her clothes or is not upset by soiling.

Note: Type of sensory issue: Type 1: sensory over-reactivity; Type 2: sensory under-reactivity and/or issues with perception.

Item # THPQ: numbers refer to the item number on the original THPQ for those items that are common to both questionnaires; 0 indicates that this item was not present in the original THPQ; * indicates the wording from the original THPQ was modified for the THPQ-R (see text for original wording); items in Spanish are available from the authors.

THPQ: toileting habit profile questionnaire; THPQ-R: THPQ – revised

items), or over-represented (items clustering at the same level of difficulty). We also looked to see where children were and were not being tested clearly by the items along the hierarchy. Having items along the entire continuum of the construct ensures the content domain is well represented and also improves measurement precision (Smith et al., 2002). Clinically, this means the measurement tool can be used with children with a wide variety of abilities.

Logic of the item hierarchy. This non-statistical procedure consists of comparing the order of the items produced by the Rasch analysis to the expected order from a theoretical or clinical perspective. Clinical experience has shown that the behaviours reflected in the items of the THPQ are very infrequently observed in children without FDD. A panel of experts reviewed the items of the THPQ and reached similar conclusions (Beaudry-Bellefeuille et al., 2016). This led us to anticipate that control children would have no items at their level, with all items being relatively easy for this group. We also anticipated that THPQ items 3 and 6 (see Table 1) would be considered easy given that these behaviours are frequent in children with FDD and are also seen, from time to time, in typical children without FDD. THPQ items 1 and 5 were expected to be of middle

difficulty as they are somewhat frequent in children with FDD. Finally, THPQ items 8 and 9 were expected to be in the more difficult range as they are not as frequently observed clinically. The remaining items were expected to distribute throughout the middle to difficult range. We then checked the congruence of the Rasch-generated hierarchy with that of the expected hierarchy.

Principal components analysis (PCA). Winsteps provides a PCA of residuals as a means of examining evidence that a construct is *not* unidimensional (Linacre, 1998). When unexplained variance is significant (>40%), the possibility of additional dimensions is suggested, particularly when the Eigenvalue associated with the first contrast is >3 (Tennant and Pallant, 2006).

Differential item functioning (DIF). We examined the possible systematic differences in item scores, also known as item bias, between English and Spanish speakers across all the items, aiming for the same amount of difficulty regardless of language. If a DIF above 0.64 was identified, the threshold value suggested by Linacre based on the work of Zwick et al. (1999), a *t*-test with a significance level set at $p < 0.05$ was then used to identify significant pairwise

differences between the language groups. If items differed systematically between groups, that would suggest that the hierarchy of difficulty of items differs between Spanish- and English-speaking participants, possibly suggesting the need for different scoring procedures.

Internal reliability. We examined the person separation index and the discernible strata calculated from this index to assess how well the items differentiated children with varying levels of toileting habits (Wright and Masters, 2002). Person separation values have a minimum value of 0; values above 1.5 are considered acceptable (Wright and Masters, 1982, cited in Boone et al., 2013). We sought to create an instrument that would separate people into at least two groups of toileting ability: those with healthy and age-appropriate toileting habits vs those with unhealthy and/or inappropriate habits. We also examined the person reliability statistic (a measure of internal consistency conceptually similar to Cronbach's α), aiming for a value of .70 or higher, considered acceptable in the early stages of research (Nunnally, 1978).

Directed content analysis and expert review

Using directed content analysis, we sought to answer the following question: *How completely do the items of the THPQ represent the behavioural and sensory processing characteristics of children with FDD reported in the literature?* Directed content analysis is the method of choice when prior research exists about a phenomenon but the information available is incomplete (Hsieh and Shannon, 2005). We used this method to compare the items of the THPQ with two recent reviews: a systematic review identifying behaviours associated with FDD (Beaudry-Bellefeuille et al., 2017) and a scoping review identifying sensory concerns associated with FDD (Beaudry-Bellefeuille et al., in press). The items of the THPQ provided the initial framework to categorize behavioural and sensory concerns identified in the literature. Any behavioural or sensory concern identified in the reviews that could not be categorized within the THPQ-generated coding scheme was given a new category. Given that the aim of the THPQ is to document defecation behaviours potentially related to sensory issues, we retained only challenging defecation behaviours that reflected potential sensory issues. Likewise, we retained only sensory concerns related to toileting and defecation.

We further separated all behavioural categories as follows. Type 1: sensory over-reactivity, including avoiding, withdrawing, feeling pain, experiencing negative emotional reactions, expressing dislike or repulsion towards sensory stimuli and/or following rituals in personal hygiene (Dunn, 2014; Schaaf and Lane, 2015). Type 2: sensory under-reactivity and/or issues with perception, including difficulties perceiving and interpreting the qualities of sensory information, and/or diminished awareness or lack of reaction to sensory stimuli (Schaaf and Lane, 2015). We grouped under-reactivity and poor perception

together (Type 2) because it can be difficult to separate these issues based exclusively on descriptions of behaviour. Reactivity is assessed mostly with self- or proxy-reports of behavioural responses to sensation; however, use of standardized testing and/or specialized techniques is needed to assess perception (Schaaf and Lane, 2015). Three experienced occupational therapists with advanced training in the assessment of sensory concerns reviewed the categorization; all had between five and 10 years of clinical experience in the treatment of children with atypical defecation habits. We asked if they could think of any additional challenging defecation behaviours related to sensory concerns. Then we asked them to separate the categories according to our established subtypes.

Results - Phase 1

Both the Winsteps analysis and the directed content analysis revealed strengths and flaws in the THPQ. Based on these results, we revised the questionnaire.

Structural validity

Item correlation. All point measure correlations (PMC) were positive, indicating the desired direction to support the construct (see Table 2).

Rating scale category structure. Scale order was distorted for three of the items and distance between response categories was inadequate for several items (see Figure 1). Examination of the probabilities of rating scale selection showed that all items had disordered thresholds and that parents most often used the extremes of the scale, suggesting that a dichotomous scale was more appropriate. This finding reflected clinical experience in that atypical defecation behaviours are either present most of the time or, on the contrary, never or rarely observed. One item (item 4) did not adequately discriminate between typically developing children and children with FDD. Only 14% of the sample responded that their child never or rarely had a ritual for defecation.

Goodness of fit statistics. With one exception (original THPQ Item 10: *My child does not realize he has soiled [faeces] his clothes*), all item infit statistics had acceptable values (<1.5). Relative to outfit statistics, items 4 and 10 were both above acceptable values (>1.5 ; see Table 2), indicating data from 80% of the items met the assumptions of the Rasch model (Wright and Linacre, 1994).

Construct distribution. The analysis revealed gaps in item distribution, especially in the difficult and middle range (see Figure 2), with item difficulty ranging from -0.72 to 1.68 logits. The children with FDD had 90% of the items against them and their scores were evenly distributed over a range of -0.91 to 1.03 logits. The typically developing children had only one item against them (original THPQ Item 4: *My child always follows the same ritual when*

Table 2. Item measure and fit statistics of the THPQ, THPQ-R (17 items), and THPQ-R (15 items).

Phase 1: original THPQ						Phase 2: THPQ-R (17 items first analysis)						Phase 2: THPQ-R (15 items final analysis)					
Item	Meas	SE	Infit	Outfit	PMC	Item	Meas	SE	Infit	Outfit	PMC	Item	Meas	SE	Infit	Outfit	PMC
4	1.68	0.15	1.15	2	0.46	6	1.65	0.3	0.89	0.83	0.56	6	1.92	0.31	0.98	0.9	0.57
6	0.44	0.14	0.77	0.95	0.75	4	1.5	0.27	0.71	0.63	0.69	4	1.65	0.29	0.75	0.64	0.7
3	0.33	0.13	0.36	0.21	0.83	8	0.7	0.27	0.93	0.87	0.55	13	0.83	0.32	0.95	0.86	0.53
1	0.09	0.14	0.7	0.43	0.73	13	0.67	0.31	0.91	0.81	0.5	8	0.74	0.29	0.99	0.99	0.58
5	0.05	0.17	0.74	0.82	0.74	1	0.62	0.27	0.82	0.72	0.61	1	0.66	0.29	0.88	0.74	0.64
10	-0.03	0.16	1.65	2.6	0.36	9	0.62	0.27	0.93	0.9	0.54	9	0.66	0.29	0.99	1	0.58
8	-0.47	0.17	1.19	0.93	0.52	17	0.62	0.27	1.55	1.94	0.14	3	-0.51	0.34	1.05	1.11	0.45
9	-0.67	0.19	0.55	0.34	0.69	3	0.37	0.32	0.9	0.75	0.5	14	0.51	0.34	1.15	1.2	0.39
2	-0.71	0.22	1.42	0.39	0.38	14	0.37	0.32	1.1	1.17	0.33	7	0.03	0.31	0.66	0.56	0.69
7	-0.72	0.22	1.36	1.27	0.34	7	0.07	0.29	0.68	0.62	0.65	5	-0.27	0.39	0.95	0.7	0.47
						16	-0.19	0.3	1.27	1.29	0.29	11	-0.47	0.33	1.32	2.16	0.33
						5	-0.34	0.38	0.88	0.66	0.45	12	-1.24	0.51	1.31	1.2	0.2
						11	-0.38	0.31	1.16	1.79	0.31	2	-1.52	0.41	1.1	1.02	0.38
						12	-1.24	0.49	1.15	0.91	0.2	15	-1.90	0.64	0.93	3.3	0.21
						2	-1.34	0.39	1.1	0.85	0.34	10	-2.11	0.48	1.04	2.12	0.3
						15	-1.84	0.62	0.94	3.03	0.14						
						10	-1.88	0.46	1	1.25	0.29						

Meas: measure in logits; SE: standard error; PMC: point measure correlation; THPQ: toileting habit profile questionnaire; THPQ-R: THPQ - revised

defecating) and their scores were evenly distributed over a range of 0.58 to 2.17 logits.

PCA. The explained variance was 63.7% and the Eigenvalue associated with the first contrast was 2.18. Both values are within the accepted range.

Logic of hierarchy. Only one item (Item 4) appeared in an unexpected order. It was located as a difficult item when we expected it to appear around the middle of the hierarchy.

DIF. All participants in phase 1 were Spanish-speaking, thus no DIF analysis could be completed.

Internal reliability. The person separation index was 1.71, with a calculated strata of 2.61, indicating that the instrument was able to separate people into >2 levels of ability. The reliability index (similar to Cronbach's α) was .93.

Directed content analysis and expert review

The analysis revealed that several challenging defecation behaviours hypothesized to be related to sensory concerns and commonly reported in children with FDD were not represented in the THPQ. Seven behaviours were identified from the literature review that were not captured by items in the original THPQ. The identified behaviours served as a basis to generate new items (see Table 1). A panel of experts reviewed the new, 17-item version of the THPQ and agreed with the categorization of behaviours and their potential relationship to sensory processing issues. They did not suggest any additional items and

there was 100% agreement with the classification of the items.

Discussion - Phase 1

In response to the findings of phase 1, the authors made the following changes to the THPQ: (1) new items were created; (2) the questionnaire's scale was modified to a dichotomous scale (1: frequently or always; 2: never or rarely); (3) item 4, which did not adequately discriminate, was revised based on the narrative responses of participants in the pilot study (original THPQ item 4: *My child always follows the same ritual when defecating*; THPQ-R item 4: *My child follows an unusual ritual when pooping which involves actions or places not typically associated with pooping or with the age of the child*).

The large outfit MnSq (2.60) of Item 10 pointed to a possible degradation of measurement from this item (Wright and Linacre, 1994). We judged that the item was a part of the construct and that erratic scoring likely was the result of lack of clarity of the item. Thus, we reworded this item to better fit the descriptions of this behaviour found in the literature and chose to maintain it in its revised version (THPQ-R item 17; see Table 1) to reconsider its validity once more data was gathered. Finally, no ceiling or floor effect was observed.

Phase 2

Once the items of THPQ were revised (THPQ-R) in phase 1, we engaged in phase 2 with the aim of examining the structural validity of the THPQ-R. We proceeded to collect data from parents of children with FC aged 3–6 years,

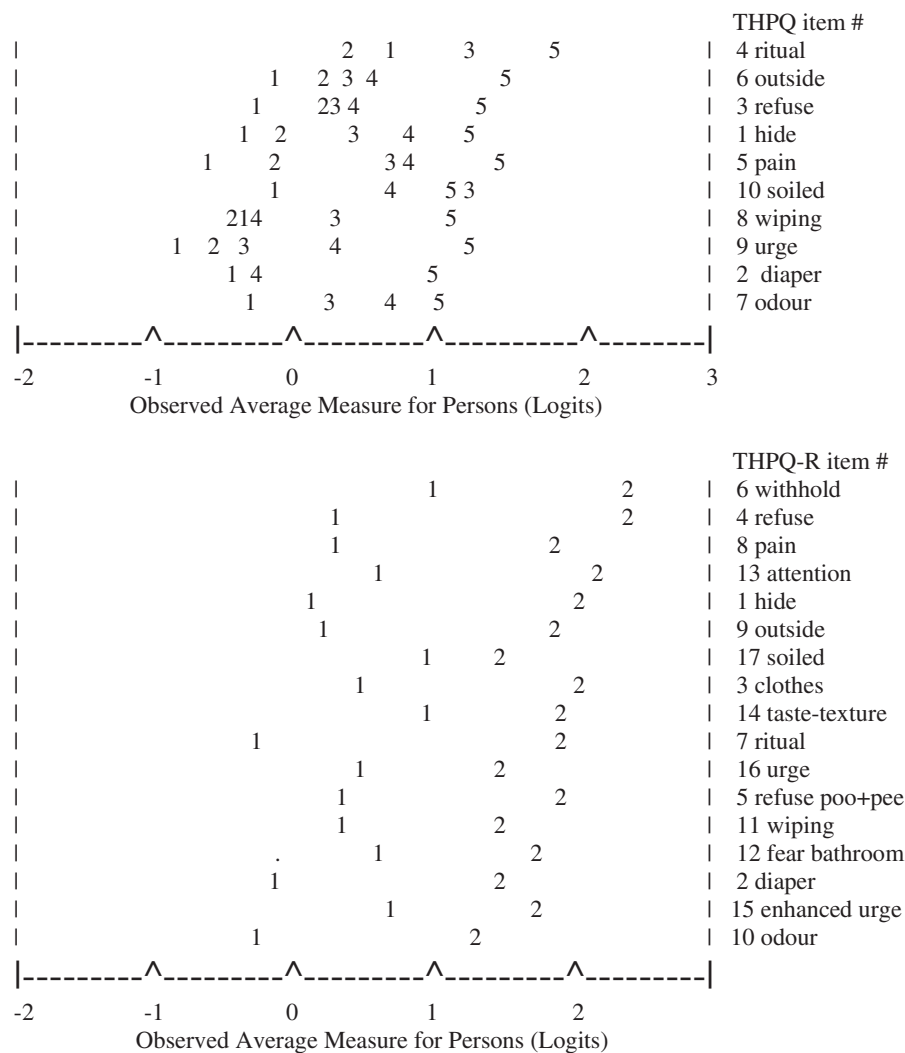


Figure 1. Observed average person measures for categories of THPQ (before collapsing categories) and of THPQ-R (after collapsing categories).

Note: numbers within the graph represent the score categories for the original (top) and revised (bottom) THPQ. The original THPQ had five score categories ranging from almost always (1) to never (5); the THPQ-R has two categories: frequently or always (1) and never or rarely (2). The categories are aligned with the horizontal axis at the point of the observed average person measure in logits. Each row of categories corresponds to a specific item labelled on the right side of the figure. Top of figure: original THPQ score category distribution for items; note some item score categories show disorder, and categories are not well distributed. Bottom of figure: THPQ-R, with score categories now dichotomous; score categories are now well ordered and scores are better distributed.

THPQ: toileting habit profile questionnaire; THPQ-R: THPQ - revised

the targeted population for this tool. This age range was chosen for two reasons. First, ongoing defecation concerns become apparent in this age range given that children typically acquire faecal continence by approximately three years (Schum et al., 2002). Second, symptoms such as feeling pain upon defecation or toilet refusal appear to be more significant during the preschool period, which is when most FDD develops (Borowitz et al., 1999).

To broaden the characteristics of the sample, we included the data from the FC group of the THPQ-pilot study for those items common to both tools. Moreover, this phase included children with autism, given that current literature reports a higher prevalence of FC in children with autism than in the typical population (Ibrahim et al., 2009).

Method - Phase 2

Using Rasch analysis of responses gathered with the THPQ-R, we examined the structural validity of our tool. We sought to verify that the changes made following phase 1 had improved the validity of the questionnaire.

Participants

Participants were parents of children aged 3–6 years with FC and no other diagnosis ($n = 55$; three years $n = 24$, four years $n = 16$, five years $n = 7$, six years $n = 8$) and parents of children aged 3–6 years with FC and autism ($n = 21$; three years $n = 6$, four years $n = 7$, five years $n = 6$, six years $n = 2$), identified by parent report of diagnosis.

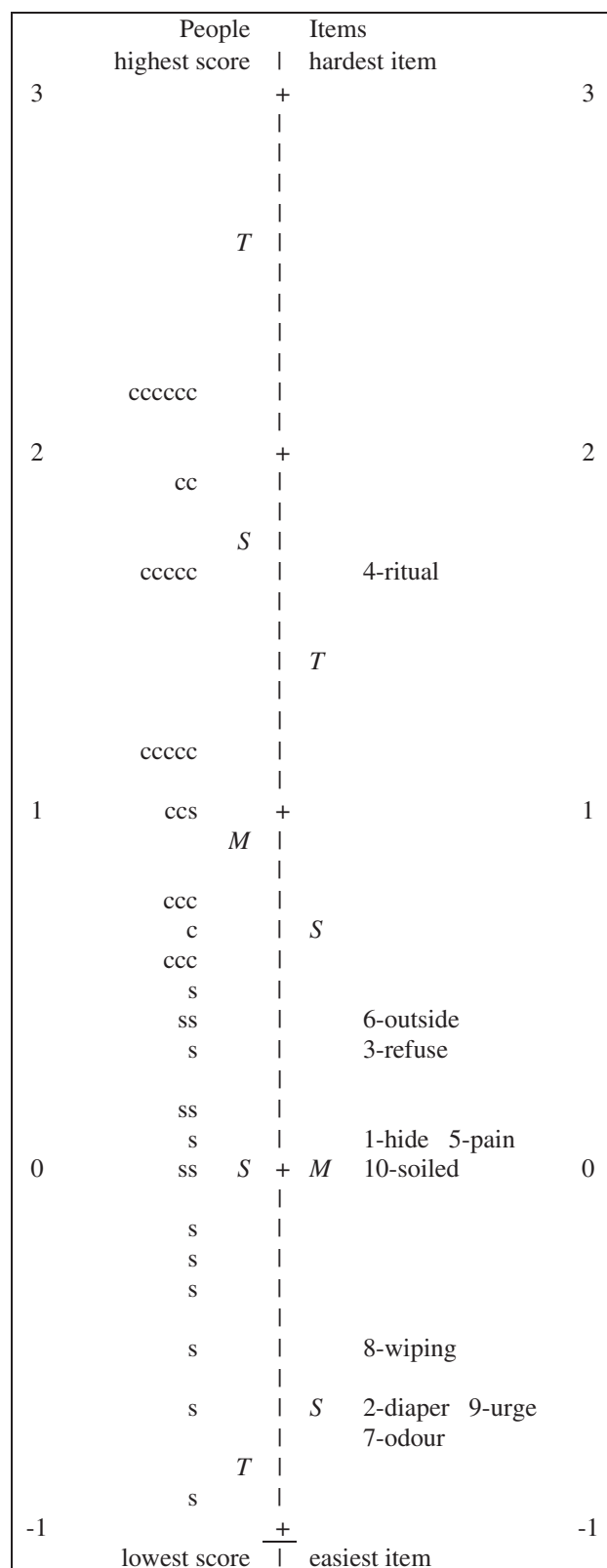


Figure 2. Hierarchy of people and items.

Note: The letter groupings on the left side of the map reflect individuals: c=control or s=subject with constipation. The right side of the map represents items; the terms are abbreviated item descriptions. The bold and italicized letters along the central line indicate the following: M: mean; S: 1 standard deviation; T: 2 standard deviations. Logits for items and persons are shown in the far left (persons) and far right (items) columns. As can be seen here, the

In this phase, 15 participants were English-speaking and 61 were Spanish-speaking.

Parents of children with organic causes of defecation issues were excluded. Parents of children with intellectual disability, neurological conditions or psychiatric disorders were excluded. Apart from the children with a diagnosis of autism, children who qualified for their school's special needs programme or who had been referred to early intervention programmes were excluded. Participants whose children had not yet initiated toilet training were also excluded given that the THPQ-R is concerned with toiletting. Initiation of toilet training was defined as asking the child to use the potty or toilet at least three times a day regardless of continence or diaper use.

Parents were recruited through parent support groups and social media. Paediatric gastroenterologists and occupational therapists were also contacted for recruitment of parents of children diagnosed with FDD and/or children with autism. Snowball recruitment was also used.

Measures

Toileting habit profile questionnaire revised (THPQ-R). The THPQ-R has 17 items organized into two sections: (a) sensory over-reactivity and (b) sensory under-reactivity and/or problems in perception. Each item is scored using a dichotomous scale (1: frequently or always; 2: never or rarely; see Table 1). All of the items were developed simultaneously in English and in Spanish following accepted guidelines for cross-cultural adaptation and validation of health questionnaires (Ramada-Rodilla et al., 2013).

Probe questions based on Rome IV diagnostic criteria of FC. The Rome Foundation is a non-profit organization that supports the creation of scientific data to assist in the diagnosis and treatment of functional gastrointestinal disorders. Questions based on Rome IV criteria (The Rome Foundation, 2016) were used to verify a diagnosis of FC.

Data collection procedure

We collected responses using a web-based survey tool (Qualtrics®; Qualtrics, 2017). Quality control was implemented to identify and exclude multiple entries and inconsistent reporting: (1) internet protocol address check; (2) email invitation only after interested participants contacted the researcher; (3) exclusion of respondents who were not consistent on Rome Foundation questions that are repeated or showed other evidence of indiscriminate responding such as answering opposite responses on similar questions. Written informed consent was not

original toileting habit profile questionnaire had items clustered in the easy range. In this instance, control participants subscribed to only one of the items, while participants identified as having FC subscribed to many or most of the items. There are no items in the middle to difficult range, suggesting a need to add items.

considered necessary given that participants were 18 years or older and no identifying data were collected. This study was approved by the ethics committee of the University of Newcastle (#H-2017-0079).

Data analysis

The Rasch procedures described for phase 1 were used again in phase 2.

Results - Phase 2

Initial analyses were conducted with the 17-item version of the THPQ-R. Iterative analyses indicated that the sensory under-reactivity / poor perception items appeared to measure a separate construct. These items ($n = 2$) were removed and analyses repeated with a 15-item version. Results for both sets of analyses are presented below.

Structural validity (THPQ-R; 17 items)

Item correlation. Examination of individual items using PMC showed positive values for all items (see Table 2).

Rating scale category structure. The scale for each item progressed as expected (Figure 1). However, a positive response to item 15 was rare; only three participants indicated that this behaviour was frequent for their child. Such a rare occurrence may distort the overall score (Linacre, 2018).

Goodness of fit statistics. Examination of fit statistics indicated that THPQ-R items 15, 17, and 11 (see Table 1) had large outfit MnSq (3.03, 1.94, and 1.79 respectively). The remaining items showed adequate infit/outfit MnSq (< 1.5 ; see Table 2).

Construct distribution. Examination of the item hierarchy showed a relatively even distribution over the linear measure (-1.88 to 1.65 logits), covering a wider item difficulty than the THPQ, although more difficult items (> 2 logits) continued to be missing (see Table 2).

Logic of hierarchy. The hierarchy of the items common to the THPQ-R and the THPQ was very similar, except for item 7 (item 4 of the THPQ; see Table 1). As expected, the adjustments made to this item left it in the middle range of item difficulty. The hierarchy of the new items was as expected.

PCA. The explained variance was 29.8 %. On the other hand, the Eigenvalue associated with the first contrast was 2.27.

Internal reliability (THPQ-R; 17 items)

The person separation index was 1.35 and the calculated strata was 2.13, the former being below desired values. The reliability index (similar to Cronbach's α) was .87.

Given the initial analysis of the 17 items, we chose to remove THPQ-R item 17 (see Table 1). Because THPQ-R items 16 and 17 were both designed to reflect sensory under-reactivity and/or poor perception, a potentially different construct from the first 15 items, we removed both items and conducted a second analysis.

Structural validity (THPQ-R; 15 items)

Rating scale category structure. Examination of item rating scale categories showed improvement, with a distance between categories near or above 1.4 logits for all items except two (item 12, 1.07; item 14, 1.28). This suggests that there is clear discrimination between children who display each behaviour and those who do not.

Goodness of fit statistics. Examination of fit statistics showed high outfit MnSq values for THPQ-R items 10 (2.12), 11 (2.16), and 15 (3.30). When we examined item scores for the four most misfitting children, we found unexpected responses on items 10, 11, and 15. Removal of these responses resulted in a reduction of outfit, with revised MnSq outfit values of 0.31 (item 10), 1.88 (item 11), and 0.61 (item 15).

Construct distribution. Examination of the variable map again showed relatively even distribution of items over the linear measure and with a wider spread of item difficulty (-2.11 to 1.92 logits).

PCA. After removal of items 16 and 17, total explained variance was 36.3%, with an Eigenvalue in the first contrast of 2.23, indicating that the removal of these items has improved the evidence for unidimensionality of the questionnaire.

DIF. The items exhibited no DIF between English and Spanish participants ($p < .05$).

Internal reliability (THPQ-R; 15 items)

The person separation index was 1.42 and the calculated strata were 2.23. The reliability index (similar to Cronbach's α) was .89. All values were acceptable and higher than in the previous analysis.

Discussion - Phase 2

Overall, the THPQ-R shows improved validity over the THPQ. While the 17 item and the 15 item versions both had merit, we removed THPQ-R items 16 and 17 for our final analyses because they appeared to be tapping a separate construct. In contrast, we kept THPQ-R items 15 and 11 as they were considered theoretically sound and part of the sensory over-reactivity construct represented by other items.

Relative to the PCA, the low explained variance could be a reflection of our small and homogeneous sample. The relatively low person separation index could also be

related to the homogeneity of the group. Nonetheless, an Eigenvalue in the first contrast of < 3 is considered acceptable. Finally, the reliability index (similar to Cronbach's α) was adequate.

Discussion

The purpose of this study was to examine the item structural and content validity of the THPQ, create the THPQ-R, and also examine the item structural validity and internal reliability of this revised tool. Both the THPQ (Beaudry-Bellefeuille et al., 2016) and the current revision, the THPQ-R, were designed as caregiver questionnaires to identify atypical defecation habits potentially related to sensory reactivity issues. The construct of sensory reactivity is well established (Ayres & Tickle, 1980; Dunn, 2014; Parham and Ecker, 2007; Su and Parham, 2014) and the use of caregiver questionnaires has become an accepted way to document its presence. However, available tools do not address issues with toileting, a crucial childhood occupation.

Directed content analysis and examination of item spread and hierarchy led us to develop additional items to better reflect the range of challenging defecation behaviours encountered in children with FDD. Examination of item structural validity of the THPQ-R using Rasch analysis identified items that needed to be reconsidered as they appeared to diverge from our goal of creating an instrument designed to measure different levels of ability along a single construct.

Examination of THPQ-R item 15 yielded a very large outfit MnSq (3.30), likely related to the fact that this behaviour is very difficult to observe. The relative rarity of individuals subscribing to this item seemed to cause distortion at the extremes of the measurement scale (Linacre, 2018). Data from two other items (THPQ-R item 10 and THPQ-R item 11) failed to conform to the expectations of the Rasch model, with outfit MnSq values above the accepted values. However, because we considered these items to be part of the construct, we examined individual person responses for evidence that a small number of participants were responsible for the failure to fit. Using the process suggested by Boone et al. (2013) for test development phases, we identified and removed individual person responses on these items which were the most misfitting. This approach brought the MnSq values to acceptable levels, supporting continued inclusion of these items. There did not appear to be a common pattern among the respondents who had the most misfitting responses.

We removed two items (THPQ-R items 16 and 17) from the analysis, representing the totality of the under-reactivity and/or poor perception section; in retrospect we recognized that they were measuring a separate construct. While sensory over- and under-reactivity are often presented as a continuum along a single construct, there is no clear data supporting this conceptualization. In this study, we considered the large misfit on item 17 to indicate

that the item is measuring a different problem, violating the principle of unidimensionality of the Rasch model. Although item 16 showed adequate fit statistics, clinical experience shows that the sensory under-reactivity / poor perception items are more difficult for parents to respond to; many parents appear to confound behavioural issues with sensory issues when responding to these items. For example, children will often deny they have soiled, apparently to avoid being scolded. Similarly, when asked if they need to defecate, many deny the urge in order to avoid being forced to sit on the toilet, a situation which many children with FDD seem to fear. Furthermore, it is well recognized that chronic stool retention leads to a reduced perception of the urge to defecate (Gladman et al., 2006), although initially over-reactivity to defecation may have been the cause of stool retention. Accordingly, we removed both items 16 and 17 and the analysis showed improved evidence of internal reliability and unidimensionality.

In considering the construct of sensory under-reactivity and/or poor perception related to defecation, we endeavoured to identify additional items using clinical observation and two literature reviews (Beaudry-Bellefeuille et al., 2017, in press). We were unable to define additional items that could adequately measure these issues. However, when behaviours such as those defined by THPQ-R items 16 and 17 are identified, they are clinically important. As such, we have maintained these items as part of the THPQ-R and recommend that they be used to provide clinical insight into sensory reactivity concerns. However, as we look to developing this tool further, we will not include scores on these two items in a THPQ-R total score. With all items on the THPQ-R we also recommend that parents be interviewed to check for a potential relationship with sensory issues; in the absence of sensory over-reactivity, behaviours reflected in items THPQ-R 16 and 17 may provide useful insight into possible sensory under-reactivity / poor perception issues. Thus, at present the THPQ-R cannot be used to clearly identify sensory under-reactivity and/or poor perception and its relationship to challenging defecation behaviour. This is something to consider in future work.

All remaining items (1–15) on the THPQ-R, designed to measure sensory over-reactivity, proved to be effective for productive measurement. These findings, and the development of the THPQ-R, build on previous work documenting the THPQ's ability to identify children with sensory over-reactivity and FC (Beaudry-Bellefeuille and Lane, 2017). Understanding underlying sensory issues related to difficulties participating in healthy age-appropriate defecation routines may lead to better treatment programmes. In order to fully implement evidenced-based practice, occupational therapists need assessments that accurately identify and characterize clients' challenges (Schaaf, 2015). Furthermore, as our profession strives to produce rigorous effectiveness studies, ensuring that the intervention under examination is appropriate for the participant is of key importance. Failure to characterize

participants adequately produces research results that lack generalizability and meaning (Pfeiffer et al., 2018).

Limitations and future research

Participants in this study represented a small and homogeneous group of children with a single type (FC) of FDD. Further research is needed to determine if findings may be generalized to younger and/or older children with other types of FDD. A less homogeneous sample is especially needed to explore the explained and unexplained variance, in order to better understand the unidimensional nature of the construct. The homogenous sample may also be one factor contributing to the poor performance of the sensory under-responsive items. Language homogeneity, with Spanish-speaking participants making up over 80% of the sample, must also be considered relative to our finding that items exhibited no DIF between English and Spanish participants. A strength in our findings was a similarity in the means of both groups; however, a more culturally diverse sample is needed to better understand differential item functioning.

Future research is needed to strengthen the overall validity of the THPQ-R for identifying defecation behaviours related to sensory reactivity issues. Finally, criterion validity as well as other aspects of construct validity, such as hypothesis testing and cross-cultural validity, also need to be explored. This study is in fact part of a larger international study which will yield data using other defecation behaviour tools as well as information relative to sensory issues in order to examine relationships with the information collected using the THPQ-R.

Conclusion

The THPQ-R is a caregiver questionnaire designed to identify challenging defecation behaviours potentially related to sensory issues. The 15 items designed to identify sensory over-reactivity were found to be adequate for measurement. The two items proposed for measurement of sensory under-reactivity and/or perception issues appear to reflect a separate construct; we recommend they be used as a source of complementary clinical information but not as part of the total score.

Key findings

- Caregiver report of behaviour using the toileting habit profile questionnaire-revised (THPQ-R) appears to adequately capture challenging defecation behaviours related to sensory over-reactivity.
- Identifying challenging defecation behaviours related to sensory under-reactivity and/or perception issues based exclusively on caregiver report of behaviour using the THPQ-R is not recommended.

What the study has added

This study has provided some evidence of the content and construct validity of the THPQ-R.

Acknowledgements

Our gratitude to the parents who kindly took the time to answer the questionnaires. Many thanks to Miguel Sanz and Hugo Sanz, linguistic and translation consultants, for their review of the additional items of the THPQ-R. Finally, many thanks to our experts, Tania Moriyón-Iglesias, Berta Gándara-Gafo, and Paula González-Martín.

Research ethics

Human research ethics approval was obtained from the University of Newcastle (#H-2016-0282).

Consent

Phase 1 (2016): #H-2016-0282; written informed consent was not considered necessary given that this study dealt with existing, anonymized data.

Phase 2 (2017): #H-2017-0079; written informed consent was not considered necessary given that participants were 18 years or older and no identifying data were collected.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Contributorship

Isabelle Beaudry-Bellefeuille conducted the Rasch analysis and the directed content analysis, and also wrote the manuscript.

Anita Bundy consulted on and reviewed the Rasch analysis, and reviewed and edited each draft of the manuscript.

Alison Lane reviewed and edited each draft of the manuscript.

Eduardo Ramos Polo reviewed and edited each draft of the manuscript.

Shelly J Lane reviewed the Rasch analysis, conducted the content analysis, and reviewed and edited each draft of the manuscript.

ORCID iD

Isabelle Beaudry-Bellefeuille  <http://orcid.org/0000-0002-2807-3688>

References

- American Occupational Therapy Association. (2014) Occupational therapy practice framework: Domain and process (3rd edition). *American Journal of Occupational Therapy* 68(Suppl. 1): S1–S48.
- Andrich D, De Jong JHAL and Sheridan BE (1997) Diagnostic opportunities with the Rasch model for ordered response categories. In: Rost J and Langeheine R (eds) *Applications of Latent Trait and Latent Class Models in the Social Sciences*. New York: Waxmann, 59–70.
- Ayres AJ and Tickle LS (1980) Hyper-responsivity to touch and vestibular stimuli as a predictor of positive response to sensory integration procedures by autistic children. *American Journal of Occupational Therapy* 34(6): 375–381.
- Beaudry IB, Schaaf RC and Ramos EP (2013) Brief report—Occupational therapy based on Ayres sensory integration in the treatment of retentive fecal incontinence in a 3-year-old boy. *American Journal of Occupational Therapy* 67(5): 601–606.
- Beaudry-Bellefeuille I and Lane SJ (2017) Examining sensory over-responsiveness in preschool children with retentive fecal incontinence. *American Journal of Occupational Therapy* 71(5): 7105220020.

- Beaudry Bellefeuille I and Ramos Polo E (2011) Tratamiento combinado de la retención voluntaria de heces mediante fármacos y terapia ocupacional [Combined treatment of voluntary stool retention with medication and occupational therapy]. *Boletín de la Sociedad de Pediatría de Asturias, Cantabria, Castilla y León* 51(217): 169–176.
- Beaudry-Bellefeuille I, Booth D and Lane SJ (2017) Defecation-specific behavior in children with functional defecation issues: A systematic review. *The Permanente Journal* 21(4): 80–87.
- Beaudry-Bellefeuille I, Lane SJ and Lane A (in press) Sensory integration concerns in children with functional defecation disorders: A scoping review. *American Journal of Occupational Therapy* 73(3).
- Beaudry-Bellefeuille I, Lane SJ and Ramos-Polo E (2016) The toileting habit profile questionnaire: Screening for sensory-based toileting difficulties in young children with constipation and retentive fecal incontinence. *Journal of Occupational Therapy, Schools, & Early Intervention* 9(2): 163–175.
- Boone WJ, Staver JR and Yale MS (2013) *Rasch Analysis in the Human Sciences*. Dordrecht, Netherlands: Springer Science & Business Media.
- Borowitz SM, Cox DJ and Sutphen JL (1999) Differences in toileting habits between children with chronic encopresis, asymptomatic siblings, and asymptomatic nonsiblings. *Journal of Developmental and Behavioral Pediatrics* 20(3): 145–149.
- Brown T and Bourke-Taylor H (2014) Children and youth instrument development and testing articles published in the American Journal of Occupational Therapy, 2009–2013: A content, methodology, and instrument design review. *American Journal of Occupational Therapy* 68(5): e154–e216.
- Dunn W (2014) *Sensory Profile 2*.
- Freeman KA, Riley A, Duke DC, et al. (2014) Systematic review and meta-analysis of behavioral interventions for fecal incontinence with constipation. *Journal of Pediatric Psychology* 39(8): 887–902.
- Gladman MA, Lunniss PJ, Scott SM, et al. (2006) Rectal hypo-sensitivity. *The American Journal of Gastroenterology* 101(5): 1140–1151.
- Hsieh HF and Shannon SE (2005) Three approaches to qualitative content analysis. *Qualitative Health Research* 15(9): 1277–1288.
- Ibrahim SH, Voigt RG, Katusic SK, et al. (2009) Incidence of gastrointestinal symptoms in children with autism: A population-based study. *Pediatrics* 124(2): 680–686.
- Koppen IJ, Vriesman MH, Tabbers MM, et al. (2018) Awareness and implementation of the 2014 ESPGHAN/NASPGHAN guideline for childhood functional constipation. *Journal of Pediatric Gastroenterology and Nutrition* 66(5): 732–737.
- Linacre JM (1998) Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement* 2(3): 266–283.
- Linacre JM (2002) Optimizing rating scale category effectiveness. *Journal of Applied Measurement* 3(1): 85–106.
- Linacre JM (2017a) *Winsteps® Rasch Measurement Computer Program*.
- Linacre JM (2017b) *Winsteps® Rasch Measurement Computer Program User's Guide*.
- Linacre JM (2018) Rare occurring response on an item [Online forum comment]. Available at: [to raschforum.boards.net/thread/873/rare-occurring-response-on-item](https://www.raschforum.com/boards.net/thread/873/rare-occurring-response-on-item) (accessed 11 March 2018).
- Linacre JM and Wright BD (1994) Chi-square fit statistics. *Rasch Measurement Transactions* 8(2): 360.
- McIntosh DN, Miller LJ, Shyu V, et al. (1999) Overview of the short sensory profile. In: Dunn W (ed.) *The Sensory Profile: User's Manual*. San Antonio, TX: The Psychological Corporation, 59–73.
- Mokkink LB, Terwee CB, Patrick DL, et al. (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research* 19(4): 539–549.
- Nunnally JC (1978) *Psychometric Theory*. 2nd ed. Hillsdale, NJ: McGraw-Hill.
- Pallant JF and Tennant A (2007) An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology* 46(1): 1–18.
- Parham LD and Ecker C (2007) *Sensory Processing Measure (SPM)*. Torrance, CA: Western Psychological Services.
- Pfeiffer B, May-Benson TA and Bodison SC (2018) Guest editorial—State of the science of sensory integration research with children and youth. *American Journal of Occupational Therapy* 72(1): 7201170010.
- Pijpers MA, Bongers ME, Benninga MA, et al. (2010) Functional constipation in children: A systematic review on prognosis and predictive factors. *Journal of Pediatric Gastroenterology and Nutrition* 50(3): 256–268.
- Qualtrics (2017) Provo, UT. Available at: www.qualtrics.com (accessed 2 March 2017).
- Rajindrajith S, Devanarayana NM, Perera BJC, et al. (2016) Childhood constipation as an emerging public health problem. *World Journal of Gastroenterology* 22(30): 6864–6875.
- Ramada-Rodilla JM, Serra-Pujadas C and Delclós-Clanchet GL (2013) Cross-cultural adaptation and health questionnaires validation: Revision and methodological recommendations. *Salud publica de Mexico* 55(1): 57–66.
- Rasch G (1960) *Probabilistic Models for Some Intelligence and Attainment Tests*.
- Schaaf RC (2015) The issue is—Creating evidence for practice using data-driven decision making. *American Journal of Occupational Therapy* 69(2): 6902360010.
- Schaaf RC and Lane AE (2015) Toward a best-practice protocol for assessment of sensory features in ASD. *Journal of Autism and Developmental Disorders* 45(5): 1380–1395.
- Schum TR, Kolb TM, McAuliffe TL, et al. (2002) Sequential acquisition of toilet-training skills: A descriptive study of gender and age differences in normal children. *Pediatrics* 109(3): E48.
- Smith EV, Conrad KM, Chang K, et al. (2002) An introduction to Rasch measurement for scale development and person assessment. *Journal of Nursing Measurement* 10(3): 189–206.
- Su CT and Parham LD (2014) Validity of sensory systems as distinct constructs. *American Journal of Occupational Therapy* 68(5): 546–554.
- Tabbers MM, Di Lorenzo C, Berger MY, et al. (2014) Evaluation and treatment of functional constipation in infants and children: Evidence-based recommendations from ESPGHAN and NASPGHAN. *Journal of Pediatric Gastroenterology and Nutrition* 58(2): 258–274.
- Tennant A and Pallant JF (2006) Unidimensionality matters! (A tale of two Smiths?). *Rasch Measurement Transactions* 20(1): 1048–1051. Available at: www.rasch.org/rmt/rmt201c.htm accessed 10 March 2018).
- Terwee CB, Prinsen CAC, Chiarotto A, et al. (2018) COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research* 27(5): 1159–1170.
- The Rome Foundation *Rome IV Pediatric Functional Gastrointestinal Disorders: Disorders of Gut-Brain Interaction*.
- Van Den Berg MM, Benninga MA and Di Lorenzo C (2006) Epidemiology of childhood constipation: A systematic

- review. *The American Journal of Gastroenterology* 101(10): 2401–2409.
- Van Ginkel R, Reitsma JB, Büller HA, et al. (2003) Childhood constipation: Longitudinal follow-up beyond puberty. *Gastroenterology* 125(2): 357–363.
- World Federation of Occupational Therapists [WFOT] *Position Statement. Activities of Daily Living*.
- Wright BD and Linacre JM (1994) Reasonable mean-square fit values. *Rasch Measurement Transactions* 8(3): 370.
- Wright BD and Masters GN (2002) Number of person or item strata. *Rasch Measurement Transactions* 16(3): 888.
- Wright BD and Stone MB (1979) *Best Test Design*.
- Zwick R, Thayer DT and Lewis C (1999) An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement* 36(1): 1–28.