

The Ugandan version of the Pediatric Evaluation of Disability Inventory (PEDI-UG). Part II: Psychometric properties

A. Amer^{1,2*} | A. Kakooza-Mwesige^{3,4,5*} | G. Jarl^{1,2,6} | J. K. Tumwine^{3,4} | H. Forssberg⁵ | A.-C. Eliasson⁵ | L. Hermansson^{1,6}

¹Faculty of Medicine and Health, School of Health and Medical Sciences, Örebro University, Örebro, Sweden

²University Health Care Research Center, Örebro, Sweden

³Department of Paediatrics & Child Health, Makerere University College of Health Sciences, Kampala, Uganda

⁴Mulago Hospital, Kampala, Uganda

⁵Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

⁶Department of Prosthetics and Orthotics, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

Correspondence

Ahmed Amer, University Health Care Research Center (UFC), Region Örebro County, Box 1613, S-70116 Örebro, Sweden.
Email: ahmed.amer@regionorebrolan.se

Funding information

Swedish International Development Agency (SIDA); Swedish Research Council. Grant/Award Number: 5925; Frimurare Barnhus Foundation; African Population and Health Research Center (APHRC) in partnership with the International Development Research Center (IDRC) and Ford Foundation; Karolinska Institutet; Belgian Technical Cooperation (BTC), Grant/Award Number: L07UGA023

Abstract

Background: The Pediatric Evaluation of Disability Inventory (PEDI) has been recommended as a gold standard in paediatric rehabilitation. A Ugandan version of PEDI (PEDI-UG) has been developed by culturally adapting and translating the original PEDI. The aim of this study was to investigate the psychometric properties of the PEDI-UG in Ugandan children by testing the instrument's rating scale functioning, internal structure, and test-retest reliability.

Methods: Two hundred forty-nine Ugandan children (125 girls) aged 6 months to 7.5 years (Mean = 3.4, SD = 1.9) with typical development were tested using the PEDI-UG. Forty-nine children were tested twice to assess test-retest reliability. Validity was investigated by Rasch analysis and reliability by intraclass correlation coefficient.

Results: The PEDI-UG domains showed good unidimensionality based on principal component analysis of residuals. Most activities (95%) showed acceptable fit to the Rasch model. Six misfit items were deleted from the Functional Skills scales and one from the Caregiver Assistance scales. The category steps on the Caregiver Assistance scales' rating scale were reversed but functioned well when changed from a 6-point to 4-point rating scale. The reliability was excellent; intraclass correlation coefficient was 0.87–0.92 for the domains of the Functional Skills scales and 0.86–0.88 for the domains of the Caregiver Assistance scales.

Conclusion: The PEDI-UG has good to excellent psychometric properties and provides a valid measure of the functional performance of typically developing children from the age of 6 months to 7.5 years in Uganda. Further analysis of all items, including misfit and deleted items, in children with functional disability is recommended.

KEYWORDS

children, disability, PEDI, Uganda, validation studies

1 | INTRODUCTION

The Pediatric Evaluation of Disability Inventory (PEDI) is a commonly used clinical assessment instrument for children with disability in western countries. It has been recommended as a gold standard in paediatric rehabilitation (Berg, Aamodt, Stanghelle, Krumlinde-Sundholm, & Hussain, 2008; Law, 2003) and has been validated and

translated into many different languages (Berg, Dolva, Kleiven, & Krumlinde-Sundholm, 2016; Chen, Hsieh, Sheu, Hu, & Tseng, 2009; Schulze, Kottorp, Meichtry, Lilja, & Page, 2014; Stahlhut, Gard, Aadahl, & Christensen, 2011). The PEDI was developed in North America and reflects the ability and performance of children living in this cultural and social environment. In several instances, instruments from high-income countries are used indiscriminately in low- and middle-income countries, although their validity for use in a different cultural setting

*Ahmed Amer and Angelina Kakooza-Mwesige shared first authorship

can be questioned (Schulze, Page, Kottorp, & Lilja, 2013). Typically developing children in Norwegian, Slovenian, Turkish, and Saudi studies have reported significantly different normative values than the original American values (Al-Khudair & Al-Eisa, 2014; Berg, Frooslie, & Hussain, 2003; Erkin, Elhan, Aybay, Sirzai, & Ozel, 2007; Srsen, Vidmar, & Zupan, 2005). The living context of a child, for example, in Uganda, contrasts markedly from a child growing up in North America. Many children live in rural areas with no access to transportation and hygiene facilities. Consequently, instruments developed in high-income countries may have inappropriate or unfamiliar terms, irrelevant items, or reference values that subsequently lead to unreliable results. There is thus a need to adapt and validate instruments prior to using them in new settings, even when transferred between various high-income countries. In a German version, several adaptations of the items were made and new items added (Schulze et al., 2013; Schulze et al., 2014), and in Norway, the norms for achieving age-related activities differed from the original PEDI norms (Berg et al., 2016).

In a separate report, we have described the development of a Ugandan version of PEDI (PEDI-UG) by culturally adapting the original version for children living in Uganda. This was accomplished through a series of steps using successive technical advisory group meetings and field testing (Beaton, Bombardier, Guillemin, & Ferraz, 2000; Wild et al., 2005; World Health Organization, 2015). The processes composed of translation to Luganda, alteration in the wording, incorporation of local examples, insertion of seven new items and deletion of 19 irrelevant items, and accommodation for major differences in living conditions between rural and urban areas in Uganda (type of toilet or bed) by providing 10 alternative items. The adapted version was subsequently tested in the field, together with feedback from in-depth interviews with primary caregivers validating the test content (linguistic and cultural validity) to make sure that the items were understandable and relevant for the caregivers (Kakooza-Mwesige, Tumwine, Forssberg, & Eliasson, 2018). However, in order to ensure that PEDI-UG is a reliable and valid measurement of what Ugandan children actually do in everyday routine activities, psychometric analysis is required on response processes and internal structure, as well as on the stability of the measure over time (Miller, McIntire, & Lovler, 2013, pp. 157–189). The aim of this study is, therefore, to investigate the psychometric properties of the PEDI-UG by testing the instrument's rating scale functioning, internal structure, and test-retest reliability.

2 | METHODS

This was an observational cross-sectional study, conducted in 11 districts from all regions of the Republic of Uganda. English is the official language of the country, and *Luganda* is the most widely spoken local language.

2.1 | Participants

Two hundred forty-nine typically developing children between 6 months and 7.5 years were included in this study. Inclusion criteria were no signs of a disability as assessed by the Ten Question Screen (Durkin, Zaman, Thorburn, Hasan, & Davidson, 1991), no signs of malnutrition, and the presence of a primary caregiver able to provide the history of the child. Sixty-one percent of the caregivers were mothers;

Key messages

- This study demonstrated that the Ugandan version of the Pediatric Evaluation of Disability Inventory is a valid and reliable instrument for assessment of ability and performance of typically developing children in Uganda.
- In low- and middle-income countries where there is a shortage of assessment instruments to measure functional performance in children, the Ugandan version of the Pediatric Evaluation of Disability Inventory could potentially be used to fill this gap.

there were also fathers (23.3%), grandparents (11.2%), other caregivers responding (2.4%), and missing data (2.0%). Children were selected from different socio-economic categories, from both urban and rural areas. A majority of the respondents answered the English version of PEDI-UG (65%), with a higher proportion from the urban areas (89%) than from the rural areas (43%). Each 6-month age interval included a minimum of 10 children, with equal representation of boys and girls and children from both urban and rural areas (Table 1). Prior to testing, the caregivers received both verbal and written information and gave written or thumb prints consent. Ethical approval was obtained from the Makerere University Research and Ethics Committee, Mulago National Referral Hospital, and the Uganda National Council of Science and Technology (Reference HS 628).

2.2 | Data collection procedure

One author (A. K.-M.) collected the data together with a team of 11 trained research assistants. Each research assistant interviewed a minimum of 22 caregivers within the home environment and in the caregiver's language of preference. To assess test-retest reliability, 49 caregivers answered PEDI-UG twice, the second time 1 week after the first interview. The same assistant interviewed the same caregiver in both interviews. Each interview took between 30 and 45 min.

TABLE 1 Characteristics of the participants in the Ugandan version of Pediatric Evaluation of Disability Inventory (PEDI-UG) interviews

	Boys n = 124	Girls n = 125	All n = 249
Age, mean	3 years 4.5 months, SD 1 year 8.6 months	3 years 4.6 months, SD 1 year 9.5 months	3 years 4.6 months, SD 1 year 9.1 months
Age grouping, n (%)			
<2 years	31 (25%)	31 (25%)	62 (25%)
2–5 years	65 (52%)	62 (50%)	127 (51%)
>5 years	28 (23%)	32 (25%)	60 (24%)
Language version, n (%)			
English	77 (62%)	86 (69%)	163 (65%)
Luganda	47 (38%)	39 (31%)	86 (35%)
Community, n (%)			
Rural	65 (52%)	62 (50%)	127 (51%)
Urban	59 (48%)	63 (50%)	122 (49%)

2.3 | Instruments

The PEDI-UG is available in English and Luganda (see Kakooza-Mwesige et al., 2018).

Same as the original PEDI, it includes two main scales: Functional Skills scales and Caregiver Assistance scales, which measure the child's current capability in three domains: Self-care, Mobility, and Social function. In the preliminary version of PEDI-UG (see Table 2a), the Functional Skills scales contain 185 dichotomous items (i.e., with two response alternatives; 0 = unable and 1 = able). If the respondent finds the item is not applicable due to age or previous experience, the answer should be 0. The Caregiver Assistance scales include 20 items rated on a six-category rating scale ranging from 0 (*total assistance*) to 5 (*independent*). A sum score is calculated for each domain, with higher numbers indicating better ability.

Many items in PEDI-UG use culturally adapted words and local examples. In some cases, the cultural adaptation of items in PEDI-UG includes alternative items (Kakooza-Mwesige et al., 2018; Table 2b). The alternative questions were answered depending on the type of toilet (pit latrine or flushing toilet) or bed (floor bed or actual bed) the child used.

2.4 | Statistical analysis

We used Rasch analysis to investigate the validity of PEDI-UG (Bond & Fox, 2001). The dichotomous model was used for the Functional Skills scales as these items were rated by a two-category rating scale, and the rating scale model was used for the Caregiver Assistance scales, because these items were rated by a six-category rating scale (Wright & Mok, 2004). Each domain was analysed separately, resulting in six Rasch analyses using the Rasch program WINSTEPS 3.81 (Linacre, 2014b).

The validity of evidence for PEDI-UG was investigated based on six aspects:

1. Item and person fit: The data were considered to usefully fit the Rasch model if at least 95% of the items and persons had an infit mean square within 0.5–1.5 (Linacre, 2002; Linacre, 2003; Wright & Linacre, 1994)

2. Unidimensionality by using principal component analysis of residuals and point-biserial correlation. It is desirable that measures explain >60% of variance, that the first contrast explains <5% of variance, and that the eigenvalue of the first contrast is <2.0 (Fisher, 2007; Linacre, 2014a). A positive point-biserial correlation indicates that items contribute positively to the total raw score (Haley, Coster, Ludlow, Haltiwanger, & Andrellos, 1992; Schumacker, 2004)
3. Reliability and separation of persons and items where the reliability index should be ≥ 0.80 and separation be ≥ 2 (Fisher, 2007; Schumacker, 2004)
4. Targeting between each domain and the respondents by investigating the distance between item and person means, and ceiling and floor effects where it is desirable that the proportion of respondents who get the maximum or minimum measure is $\leq 2\%$ (Fisher, 2007; Linacre, 2014c)
5. Rating scale functioning (for the Caregiver Assistance scales). It is desirable that each rating scale category includes ≥ 10 observations and has an outfit mean square <2.0 and that average measures and step difficulties increase monotonically (Linacre, 2004)
6. Analysis of differential item functioning (DIF) was used to investigate the stability of item difficulty between living areas (urban vs. rural). To be considered a statistically significant DIF, the difference in item measure between the two groups should exceed 1.0 logit with probability <0.05 (Linacre, 2014c). It is desirable that the proportion of items demonstrating DIF should be less than 5%.

The alternative items, for example, pit latrine or flushing toilet (Table 2b), were analysed as independent items under the same domain.

Calculation of test-retest reliability was performed in two steps. First, estimates of item measures and rating scale structures from the first response occasion were used as anchors for the estimation of person measures on the second response occasion. Hence, both estimations of children's ability were based on the same references. Next, intraclass correlation coefficient (ICC) was calculated based on a two-way mixed model with absolute agreement. An ICC value >0.75 was

TABLE 2 Structure of the Ugandan version of Pediatric Evaluation of Disability Inventory (PEDI-UG)

(a) Structure of preliminary version of PEDI-UG					
Scales	Rating scale	Domains (no. of items)			Total no. of items
		Self-care	Mobility	Social function	
Functional Skills	0,1	73	48	64	185
Caregiver Assistance	0–5	8	7	5	20
(b) Alternative items created in the PEDI-UG for Functional Skills scales					
	Original items	Alternative items		No. of items	
Self-care domain					
• Toileting tasks	(a) Flushing toilet	(b) Pit latrine		1	
Mobility domain					
• Toilet transfers	(a) Flushing toilet	(b) Pit latrine		5	
• Bed mobility/transfers	(a) Actual bed	(b) Floor bed		4	
(c) Final structure of PEDI-UG					
Scales	Rating scale	Domains (no. of items)			Total no. of items
		Self-care	Mobility	Social function	
Functional Skills	0.1	70	45	64	179
Caregiver Assistance	0–3	8	6	5	19

considered excellent, 0.60–0.75 good, 0.40–0.59 moderate, and <0.40 poor (Fleiss, Levin, & Cho, 2003, pp. 598–626). SPSS was used for the analyses.

3 | RESULTS

3.1 | Functional Skills scales

The first analysis showed that the number of misfit items exceeded the recommendation in two domains accounting for 7% in the Self-care domain and 13% in the Mobility domain but was acceptable with 3% in the Social function domain. Based on these results and guided by clinical relevance, the three least relevant misfit items were deleted from the Self-care domain and from the Mobility domain in order to reduce the misfit percentage to the recommended value (i.e., $\leq 5\%$; Table 3a). Subsequent analyses were run without these six misfit items. The final analyses for the Functional Skills scales indicated three misfit items (4%) in the Self-care domain, four misfit items (7%) in the Mobility domain, and two misfit items (3%) in the Social function domain. These were all retained (Table 3b). Person misfits were 7% in the Self-care domain, 9% in the Mobility domain, and 9% in the Social function domain. The point-biserial correlation values were positive for all items including the misfits, indicating that they all contribute positively to the domain raw sum score. For the Self-care and Social function domains, the variance explained by the first dimension was above the recommended value; in the Mobility domain, measures explained less than 60% of the variance (Table 4). In all domains, the first contrast had an eigenvalue above 2 but it explained less than 5% of the variance, thus supporting unidimensionality. All domains fulfilled the criteria for person and item reliability. There was more than 1 logit distance between item and person mean measures in all domains, a ceiling effect in the Mobility domain, and a floor effect in the Social function domain, indicating that the sample ability was mistargeted to the item difficulty (Table 4). Between 20% and 29% of the items in the domains of the Functional Skills scales demonstrated DIF (Table 4) for rural versus urban areas; for details, see Table S1. Analysis of test-retest reliability for the Functional Skills scales yielded excellent ICC values, ranging between 0.87 and 0.92 (Table 4).

3.2 | Caregiver Assistance scales

The rating scales were analysed before the internal structure of the Caregiver Assistance scales was studied. In all domains, the rating scales fulfilled three of the criteria, namely, there were ≥ 10 observations per category, average measures increased monotonically across categories, and the outfit mean square was < 2 for each category (Table 5a). However, in all domains, the rating scale category steps were located too close to each other, and in the Mobility domain, the order of the step difficulties was reversed (Table 5a, Figure 1a). Combining Category 2 with 3 and Category 4 with 5 improved the rating scale functioning for all three domains (Table 5b, Figure 1b). Further analyses of the Caregiver Assistance scales were thus performed based on the four-category rating scale (Figure 2).

Ninety-five percent of the items had acceptable fit in the final Rasch model; one item in the Self-care domain was misfit, and one item

in the Mobility domain was overfit to the model (Table 3). The latter item, Tub transfer, was the only item with a nonpositive point-biserial correlation and was answered by only one out of 249 respondents. This item was judged to be irrelevant for the target population and was removed from further analyses, resulting in 19 items for these scales. Person misfits were 10% in the Self-care domain, 13% in the Mobility domain, and 13% in the Social function domain. In all three domains, the first contrast either explained less than 5% of the variance or had an eigenvalue less than 2 (Table 4), supporting unidimensionality. All domains fulfilled the criteria for item reliability. Only the Mobility domain did not fulfil the criteria for person reliability. However, when we excluded children with extreme scores (in other words, without the impact of ceiling and floor effects), person reliability and separation for this domain were acceptable: 0.81 and 2.01, respectively. In terms of targeting, all domains demonstrated higher mean person ability than item mean difficulty and had a ceiling effect. In addition, the Mobility and Social function domains also had a floor effect higher than the cut-off point (Table 4). Between 17% and 40% of the items in the domains of the Caregiver Assistance scales demonstrated DIF (Table 4) for rural versus urban areas. Test-retest reliability analysis for the Caregiver Assistance scales yielded excellent ICC values, ranging between 0.86 and 0.88 (Table 4).

4 | DISCUSSION

The psychometric analyses in the present study support the validity of the culturally adapted PEDI-UG (Kakooza-Mwesige et al., 2018). The analyses demonstrated that the PEDI-UG fulfilled most of the recommended criteria for validity after some adjustments and that it showed excellent test-retest reliability. We, therefore, suggest that the PEDI-UG consists of 179 items in the Functional Skills scales (three items removed from the Self-care and Mobility domains, respectively) and 19 items in the Caregiver Assistance scales (one item removed from the Mobility domain) and that the PEDI-UG Caregiver Assistance scales have a 4-point rating scale. The results from this study are encouraging and indicate that the validity of PEDI-UG should also be studied in children with disabilities who do not follow the typical developmental trajectory. The PEDI-UG may become one of the first instruments for the assessment of children's ability in activities of daily living in a sub-Saharan country, and it can possibly be used in neighbouring countries with a similar cultural and social context.

The results from the principal component analysis of residuals of PEDI-UG show that the three domains of the Functional Skills scales form a unidimensional measure, indicating that they measure the same underlying skill or construct as the original version. Some uncertainty still remains in the Mobility domain, in which less than 60% of the variance was explained by measures and more than 5% of items were misfit. At this stage, we have chosen to keep some of these items for further analyses in a clinical population because they may be clinically important, see Kakooza-Mwesige et al. (2018). In addition, there was a positive point-biserial correlation for all the items and a low effect of other dimensions (less than 5% of variance was explained by a second dimension), which indicates satisfactory unidimensionality (Haley et al., 1992; Linacre, 2014a; Schumacker, 2004). Still, we suggest further

TABLE 3 Misfit^a analysis of the Ugandan version of Pediatric Evaluation of Disability Inventory (PEDI-UG)

(a) Removed misfit ^a items based on first fit analysis		Count	Measure	SE	Infit MnSq	PtBi
Domain						
Functional Skills scales						
SC	4. Does child eat all textures of table food such as adult bite-sized unground meat, chicken, cassava, sweet potatoes, millet bread, or posho, with no spilling out of mouth?	249	0.99	0.24	1.72	0.66
SC	35. Does child try to wash parts of the body such as arms or legs with the caregiver totally redoing washing?	249	-0.42	0.23	1.86	0.63
SC	68. Does child consistently stay dry during both night and day time?	249	2.42	0.19	1.68	0.56
M	19. Does child sit supported by the caregiver at a washbasin?	249	1.58	0.29	2.61	0.27
M	25. Does child walk in the home environments but must hold onto the caregiver, walls, or furniture or use devices for support?	249	0.36	0.24	1.54	0.58
M	37. Does child walk, but holding onto objects, caregiver, or devices for support?	249	0.75	0.20	1.61	0.50
Caregiver Assistance scales						
M	D. Tub transfer: Getting in and out of adult-sized tub.	1	0.91	0.82	0.00	0.00
(b) Retained misfit ^a items after the final fit analysis		Count	Measure	SE	Infit MnSq	PtBi
Domain						
Functional Skills scales						
SC	3. Does child eat cut up/chunky/diced foods such as carrots, beans, peas, cabbage, chips, groundnuts, or minced meat, with no spilling out of mouth?	249	-2.78	0.28	1.53	0.58
SC	15. Does child open mouth for teeth to be brushed?	249	-0.30	0.24	1.82	0.65
SC	16. Does child hold the toothbrush but does not functionally clean the teeth?	249	0.76	0.21	1.71	0.65
M	1. (Flushing toilet). Does child sit on potty or on a toilet supported by equipment or the caregiver?	94	-1.65	0.40	2.06	0.45
M	6. Does child sit if supported by equipment or by caregiver?	249	-2.14	0.32	2.25	0.39
M	32. Does child move in any manner within the home environment, from place to place, but does not have the capability to push or pull objects on the floor or to carry objects?	249	0.43	0.22	1.73	0.52
M	27. Does child move within the home environment with difficulty with a speed slower as compared to his normal peers?	249	0.65	0.21	1.81	0.48
SF	22. When a problem occurs, does child need to be helped immediately or else the behaviour deteriorates? (e.g., if child becomes frustrated while trying to take off shirt, you must immediately go to help child undress, and child can calm down)?	249	0.32	0.23	3.03	0.45
SF	61. Can child go about familiar environment outside of home with periodic monitoring for safety (e.g., the child can function safely in one or two familiar contexts away from home, e.g., sitting room in the neighbour's house, or the neighbour's compound)?	249	-0.88	0.25	1.69	0.66
Caregiver Assistance scales						
SC	A. Eating: Eating and drinking regular meal; does not include cutting steak, opening containers or serving food from serving dishes.	244	-2.28	0.13	2.79	0.73

Note. M = Mobility; MnSq = Mean square; PtBi = point-biserial correlation; SE = standard error; SC = Self-care; SF = Social function.

^aMisfit was defined as a mean square outside the range 0.5–1.5. Bold font indicates misfit or a nonpositive point-biserial correlation.

investigation on the unidimensionality of the Mobility scales of the Functional Skills scales.

Person and item reliability and separation in all three domains of the Functional Skills scales were excellent, indicating that high confidence can be placed in the order of items and persons along the latent variable. This means that PEDI-UG can distinguish children of different levels of self-care, mobility, and social function. Targeting indicated that this sample of children had higher ability than what PEDI-UG can measure, especially for the Caregiver Assistance scales, but this is expected as there were only typically developing children included in this study. Some floor effects were seen, and this might be explained by parenting styles including low expectations of independence. Similar results have been seen in previous studies in Saudi Arabia, Turkey, and Puerto Rico (Al-Khudair & Al-Eisa, 2014; Erkin et al., 2007; Gannotti & Cruz, 2001). More research is needed to get insight in the influence of underestimation or overestimation of children's ability by parents and the influence of different parenting styles in different cultural contexts. The PEDI-UG showed excellent test-retest reliability

in both Functional Skills scales and Caregiver Assistance scales. This is an important instrument attribute as it defines whether any small difference can be detected following an intervention.

The person fit in both Functional Skills scales and Caregiver Assistance scales was somewhat below the recommended values, meaning that some children received unexpectedly high or low scores on the items, given their ability level as measured by the Rasch model. This might be explained by the fact that these children live in widely varying environments with different possibilities especially regarding hygiene and indoor activities, as further described by Kakoza-Mwesige et al. (unpubl. observ.). The unexpected response pattern might depend on the fact that some older children with a high ability had not been exposed to certain activities in their living environment. For example, several children in rural areas are not exposed to cars and do not use shoes with laces, and therefore, were scored "0" on these items. This interpretation is supported by the fact that a substantial proportion of items demonstrated DIF, which reflects the differences between living in rural and urban areas of Uganda. Even within countries in Europe,

TABLE 4 Recommended and observed values of validity criteria for the Ugandan version of Pediatric Evaluation of Disability Inventory (PEDI-UG)

Criterion	Recommended value	Functional Skills scales			Caregiver Assistance scales		
		Self-care	Mobility	Social function	Self-care	Mobility	Social function
Unidimensionality (PCA of residuals)							
Variance explained by measures	>60%	67.5	54.8	68.7	81.0	74.0	72.9
Unexplained variance in first contrast	<5%	1.9	3.7	2.2	4.7	6.1	8.7
Eigenvalue	<2.0	4.3	4.4	4.5	2.0	1.4	1.6
Reliability							
Person reliability	>0.80	0.97	0.91	0.97	0.89	0.71	0.80
Item reliability	>0.80	0.99	0.97	0.99	0.99	0.99	0.98
Person separation	≥2	5.52	3.15	5.41	2.82	1.56	2.01
Item separation	≥2	10.00	5.69	10.10	8.86	8.94	6.57
Targeting							
Mean person ability		1.49	2.27	1.45	3.78	2.95	1.38
Ceiling effect	≤2%	2.0	9.6	1.2	32.4	38.5	16.8
Floor effect	≤2%	1.6	1.6	2.8	2.0	2.5	2.5
DIF (percentage of DIF items)							
Living area (rural vs. urban)	<5%	28.6	20.0	21.9	25.0	16.7	40.0
Test-retest reliability							
ICC (95% CI)	≥0.75	0.87 [0.77, 0.93]	0.88 [0.80, 0.93]	0.92 [0.86, 0.96]	0.87 [0.67, 0.92]	0.86 [0.71, 0.92]	0.88 [0.73, 0.94]

Note. ICC = intraclass correlation coefficient; PCA = principal component analysis; DIF = differential item functioning. Bold font indicates deviation from criteria.

TABLE 5 The Ugandan version of Pediatric Evaluation of Disability Inventory (PEDI-UG) Caregiver Assistance scales: Fulfilment of criteria for rating scale analyses

Criterion	Self-care	Mobility	Social function
(a) Original six-category scale			
Observations	≥10 Obs./category	Yes	Yes
Average measure	Increase monotonically	Yes	Yes
Step difficulties	Increase monotonically	Yes	No, reversal of Steps 1/2 (0.43) and 2/3 (-0.55)
Distance between adjacent steps	1.4–5.0 logits	No, distance between Steps 1/2 and 2/3: 0.8 ; between Steps 2/3 and 3/4: 1.37	Reversal of step difficulties^a No, distance between Steps 1/2 and 2/3: 0.04 ; between Steps 2/3 and 3/4: 1.10 ; between Steps 3/4 and 4/5: 1.35
Outfit mean square	<2	Yes	Yes
(b) Modified four-category scale ^b			
Step difficulties	Increase monotonically	Yes	Yes
Distance between adjacent steps	1.4–5.0 logits	Yes	No, distance between Steps 1/2 and 2/3: 1.24
Outfit mean square	<2	No, Category 2: 4.94	Yes

Note. Bold fonts indicated deviation from criterion.

^aReversal of step difficulties made calculation for distance between adjacent categories misleading.

^bAfter combining original Categories 2 and 3 and Categories 4 and 5. This gave a four-category rating scale (0, 1, 2 [2 + 3], 3 [4 + 5]).

some cultural differences have been found (Schulze, Page, Lilja, & Kottorp, 2017). Although DIF does not necessarily distort individual person measures (Schulze et al., 2017), the DIF of PEDI-UG needs to be investigated more extensively in future studies.

From the cultural adaptation process, some items were questioned but kept by Kakooza-Mwesige et al. (2018), for example, the bathtub item. Studies in Sweden, Norway, Germany, and Saudi Arabia have previously reported that bathtubs are not commonly used in these

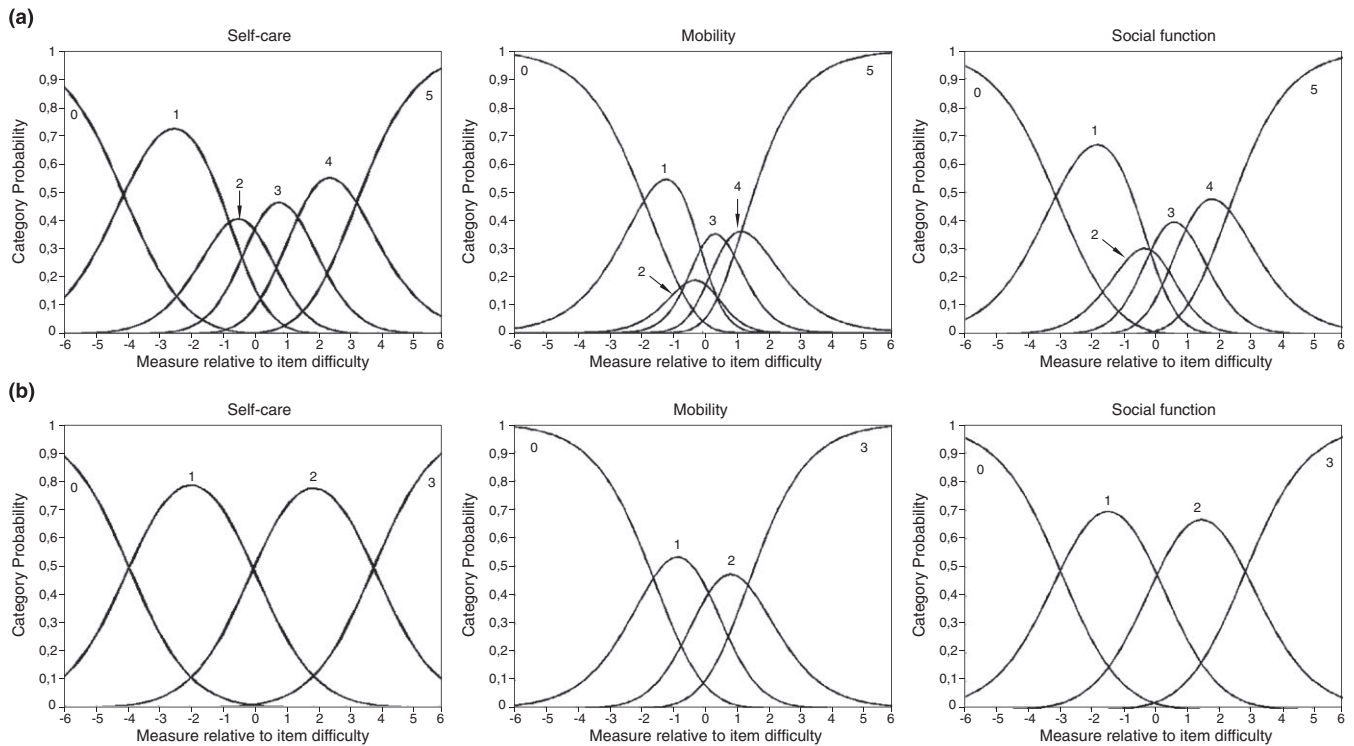


FIGURE 1 Category structure of rating scale for Caregiver Assistance scales of the Ugandan version of Pediatric Evaluation of Disability Inventory. (a) Original six-category scale. The arrows indicate the narrowness/overlap between adjacent categories. (b) Modified four-category scale after combining Categories 2 and 3 and Categories 4 and 5

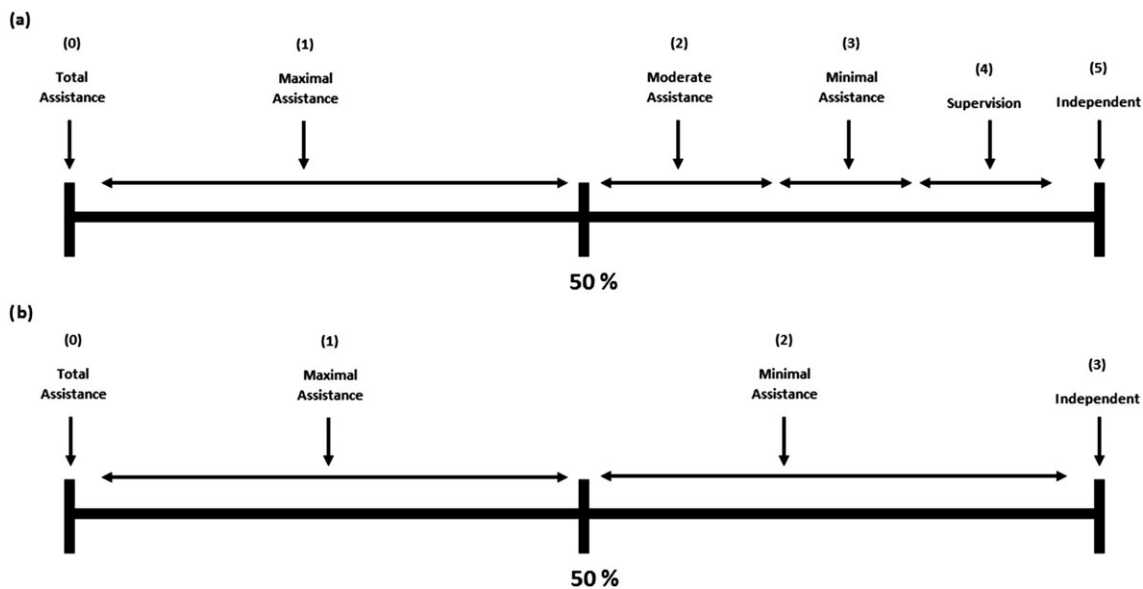


FIGURE 2 Rating scale for Caregiver Assistance scales of the Ugandan version of Pediatric Evaluation of Disability Inventory: (a) original six-category rating scale and (b) new four-category rating scale where 0 = total assistance (child is totally dependent on caregiver, or the activity is not applicable), 1 = maximal assistance (caregiver does more than half of effort of the activity; child provides some meaningful assistance), 2 = minimal assistance (child contributes more than half of the effort of the activity; caregiver provides some meaningful assistance), and 3 = independent (child performs activity independently or with very limited assistance, e.g., some verbal instruction, from caregiver)

countries (Al-Khudair & Al-Eisa, 2014; Berg, Jahnsen, Holm, & Hussain, 2003; Custers et al., 2002; Haley et al., 2010; Nordmark, Orban, Hagglund, & Jarnlo, 1999). Based on the overwhelmingly low response rate in this data collection (it was answered by only one respondent),

this item was deleted. Other items were added during the cultural adaptation process to adapt to different toilet and bed solutions in different areas of the country. The deletion of culturally inappropriate test items as well as addition of new items during the adaptation of instruments is

not unfamiliar and has been employed in other settings (Bangirana et al., 2009; Schulze et al., 2017). Likewise, it is expected that children's ability varies in different cultural contexts; for example, in Saudi Arabia, the general performance in children is lower than in American children, which is explained by differences in mother role and frequent presence of nannies (Al-Khudair & Al-Eisa, 2014).

The original PEDI Caregiver Assistance scales have a six-category rating scale. As has been discussed in other research from Eastern Africa (Masquillier, Wouters, Loos, & Nostlinger, 2012), the understanding and use of a grading system may vary depending on the meaning of specific terms used to define the grades in the scale. In the earlier translation process, some of the caregivers found it hard to discriminate between the six categories in PEDI-UG; the caregivers like a clear difference between the choices and are not used to fine grading systems (Kakooza-Mwesige et al., 2018). This is confirmed by the results in this study. Therefore, we combined the six categories into a new four-category rating scale, which turned out to be more discriminative for the caregivers. Thus, we recommend the four-category rating scale for future use in PEDI-UG. By applying this four-category rating scale, the Self-care, Social function, and Mobility domains seem to function well. Still, some points need to be discussed, such as the somewhat weak person reliability for the Mobility domain of the Caregiver Assistance scales. This could have been caused by a low number of items (six items only) and a high number of children who received the top total score, making it more difficult to separate children of different abilities. This situation might change if a more varied population is investigated, including children with disability as part of the general population (Haley et al., 1992). This explanation is supported by the fact that reliability and separation improved above the cut-off when children with extreme scores were excluded from the calculation. Furthermore, there was a large ceiling effect in all Caregiver Assistance scales of PEDI-UG. In the normative profile of the original PEDI Caregiver Assistance scales (Haley et al., 1992), it is clearly shown that the majority of children with typical development achieve independence in most items before 5 years of age. Hence, further investigation of the PEDI-UG in a varied population may demonstrate a sample more targeted to the item difficulty.

4.1 | Limitations

We used two language versions for data collection, and the items may have been interpreted differently in these versions. In addition, there are many local languages in Uganda and it might be possible that neither English nor Luganda was the native language for some caregivers. This is a potential limitation, but considering the administration procedure where the respondents were able to ask and get further explanation of the items, we believe that the responses are valid. Therefore, a DIF analysis of the two language versions needs to be performed in a larger and more evenly distributed sample where the native language is used; more than 100 participants for each group are recommended (Fidalgo, Ferreres, & Muñiz, 2004). As part of the translation process, some items from the original Functional Skills scales of PEDI were supplemented by culturally adapted equivalents to accommodate for the different living conditions of children in

Uganda. These worked well. Still, there might be unidentified situations where further adaptations of PEDI-UG are needed. Future use of this version and research in clinical samples may show this and suggest further adaptations. The PEDI-UG was administered in interviews conducted by different research assistants who could have interpreted questions and responses differently. However, the interviewers were trained and each contributed with a large number of data, which may limit the potential rater bias in this study. Lastly, no children with functional limitations were included in the study, which may explain some of the mistargeting of the sample.

5 | CONCLUSIONS

The culturally and contextually adapted PEDI-UG possesses good to excellent psychometric properties and is thus an acceptable and valid instrument to assess the ability of children's daily living in an environment very different from western countries. Future studies should evaluate the validity for using PEDI-UG, including misfit and deleted items, with children who do not follow the typical developmental trajectory. A valid and sensitive PEDI-UG could be invaluable for both clinicians and researchers in Uganda and other sub-Saharan countries when following the progress of children with disabilities and when planning and evaluating interventions.

ACKNOWLEDGEMENTS

We would like to thank the children and caregivers who participated in this study and the clinical staff from Mulago Hospital and Makerere University College of Health Sciences. Special thanks are expressed to Clemence Bitanihirwe, Edward Galiwango, Emmaunuel Odhiambo, Francis Ekwaru, Juliet Wamukota, Julius Sekyawa, Johnson Musinguzi, Maureen Nawenja, Precious Ahabwe, Rebecca Nabbanja, and Winnie Nandawula for collecting the data and Janet Nakigudde and Eva Brogren Carlberg for professional support.

CONFLICT OF INTEREST

The authors state that they have no competing interests.

FUNDING

This work was supported by grants from a Belgian government scholarship (L07UGA023) from the Belgian Technical Cooperation (BTC), Uganda Country Office, from a doctoral grant financed by Karolinska Institutet, Sweden, and from a grant from the Swedish International Development Agency (SIDA) awarded to A. K.-M. for PhD studies. The work was also partly supported by funds from the African Doctoral Dissertation Research Fellowship, offered by the African Population and Health Research Center (APHRC) in partnership with the International Development Research Center (IDRC) and Ford Foundation, awarded to A. K.-M, and funds from the Frimurare Barnhus Foundation and Swedish Research Council (5925), awarded to H. F. The funding agencies had no role in data collection, interpretation of results, or the decision to submit for publication.

ORCID

A. Amer  <http://orcid.org/0000-0002-9222-8066>

A. Kakooza-Mwesige  <http://orcid.org/0000-0002-1189-3671>

G. Jarl  <http://orcid.org/0000-0002-6410-2474>

REFERENCES

- Al-Khudair, R. I., & Al-Eisa, E. (2014). Applicability of the pediatric evaluation of disability inventory among Saudi children. *Disability and Rehabilitation, 36*, 658–664.
- Bangirana, P., Seggane, M., Allebeck, P., Giordani, B., John, C. C., Opoka, O. R., ... Boivin, M. J. (2009). A preliminary examination of the construct validity of the KABC-II in Ugandan children with a history of cerebral malaria. *African Health Sciences, 9*, 186–192.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976), 25*, 3186–3191.
- Berg, M., Aamodt, G., Stanghelle, J., Krumlinde-Sundholm, L., & Hussain, A. (2008). Cross-cultural validation of the Pediatric Evaluation of Disability Inventory (PEDI) norms in a randomized Norwegian population. *Scandinavian Journal of Occupational Therapy, 15*, 143–152.
- Berg, M., Frøslie, K. F., & Hussain, A. (2003). Applicability of Pediatric Evaluation of Disability Inventory in Norway. *Scandinavian Journal of Occupational Therapy, 10*, 118–126.
- Berg, M., Jahnsen, R., Holm, I., & Hussain, A. (2003). Translation of a multi-disciplinary assessment—Procedures to achieve functional equivalence. *Advances in Physiotherapy, 5*, 57–66.
- Berg, M. M., Dolva, A. S., Kleiven, J., & Krumlinde-Sundholm, L. (2016). Normative scores for the Pediatric Evaluation of Disability Inventory in Norway. *Physical & Occupational Therapy in Pediatrics, 36*, 131–143.
- Bond, T., & Fox, C. (2001). *Applying the Rasch model*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Chen, K. L., Hsieh, C. L., Sheu, C. F., Hu, F. C., & Tseng, M. H. (2009). Reliability and validity of a Chinese version of the Pediatric Evaluation of Disability Inventory in children with cerebral palsy. *Journal of Rehabilitation Medicine, 41*, 273–278.
- Custers, J. W., Wassenberg-Severijnen, J. E., Van Der Net, J., Vermeer, A., Hart, H. T., & Helders, P. J. (2002). Dutch adaptation and content validity of the 'Pediatric Evaluation of Disability Inventory (PEDI)'. *Disability and Rehabilitation, 24*, 250–258.
- Durkin, M., Zaman, S., Thorburn, M., Hasan, M., & Davidson, L. (1991). Population-based studies of childhood disability in developing countries: Rationale and study design. *International Journal of Mental Health, 20*, 47–60.
- Erkin, G., Elhan, A. H., Aybay, C., Sirzai, H., & Ozel, S. (2007). Validity and reliability of the Turkish translation of the Pediatric Evaluation of Disability Inventory (PEDI). *Disability and Rehabilitation, 29*, 1271–1279.
- Fidalgo, A. M., Ferreres, D., & Muñiz, J. (2004). Utility of the Mantel-Haenszel procedure for detecting differential item functioning in small samples. *Educational and Psychological Measurement, 64*, 925–936.
- Fisher, W. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions, 21*, 1.
- Fleiss, J., Levin, B., & Cho, P. (2003). *Statistical methods for rates and proportions*. (pp. 598–626). New York, NY: Wiley.
- Gannotti, M. E., & Cruz, C. (2001). Content and construct validity of a Spanish translation of the Pediatric Evaluation of Disability Inventory for children living in Puerto Rico. *Physical & Occupational Therapy in Pediatrics, 20*, 7–24.
- Haley, S. M., Coster, W. I., Kao, Y. C., Dumas, H. M., Fragala-Pinkham, M. A., Kramer, J. M., ... Moed, R. (2010). Lessons from use of the Pediatric Evaluation of Disability Inventory: Where do we go from here? *Pediatric Physical Therapy, 22*, 69–75.
- Haley, S. M., Coster, W. J., Ludlow, L. H., Haltiwanger, J. T. & Andrellos, P. J. (1992) *Pediatric Evaluation of Disability Inventory (PEDI): Development, standardization and administration manual*. Boston university, Boston, USA: PEDI Resarch Group.
- Kakooza-Mwesige, A., Tumwine, J. K., Forsberg, H., & Eliasson, A.-C. (2018). The Uganda version of the Pediatric Evaluation of Disability Inventory (PEDI). Part I: Cross-cultural adaptation. *Child Care Health Dev, 1–10*. <https://doi.org/10.1111/cch.12563>
- Law, M. (2003). Outcome measurement in pediatric rehabilitation. *Physical & Occupational Therapy in Pediatrics, 23*, 1–4.
- Linacre, J. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16*, 1.
- Linacre, J. (2003). Rasch power analysis: Size vs. significance: Standardized chi-square fit statistic. *Rasch Measurement Transactions, 17*.
- Linacre, J. (2004). Optimizing rating scale category effectiveness. In E. Smith, & R. Smith (Eds.), *Introduction to Rasch measurement 11 11* (pp. 258–278). Maple Grove, Minnesota, USA: JAM Press.
- Linacre, J. (2014a) Dimensionality: Contrasts & variances. Available at: <http://www.winsteps.com/winman/principalcomponents.htm> (24 March 2015)
- Linacre, J. (2014b) Winsteps® Rasch measurement computer program. Winsteps.com, Beaverton, Oregon.
- Linacre, J. (2014c). *Winsteps® Rasch measurement computer program user's guide*. Beaverton, Oregon: Winsteps.com.
- Masquillier, C., Wouters, E., Loos, J., & Nostlinger, C. (2012). Measuring health-related quality of life of HIV-positive adolescents in resource-constrained settings. *PLoS One, 7*, e40628.
- Miller, L. A., Mcintire, S. A., & Lovler, R. L. (2013). *Foundations of psychological testing: A practical approach* (pp. 157–189). California, USA: SAGE Publications.
- Nordmark, E., Orban, K., Hagglund, G., & Jarnlo, G. B. (1999). The American Paediatric Evaluation of Disability Inventory (PEDI). Applicability of PEDI in Sweden for children aged 2.0-6.9 years. *Scandinavian Journal of Rehabilitation Medicine, 31*, 95–100.
- Schulze, C., Kottorp, A., Meichtry, A., Lilja, M., & Page, J. (2014). Inter-rater and test-retest reliability of the German Pediatric Evaluation of Disability Inventory (PEDI-G). *Physical & Occupational Therapy in Pediatrics*.
- Schulze, C., Page, J., Kottorp, A., & Lilja, M. (2013). Adapting functional assessments for use in a new context: A balancing act. *Scandinavian Journal of Occupational Therapy, 20*, 336–342.
- Schulze, C., Page, J., Lilja, M., & Kottorp, A. (2017). Cross-cultural validity of the German version of the Pediatric Evaluation of Disability Inventory (PEDI-G)—A Rasch model application. *Child: Care, Health and Development, 43*, 48–58.
- Schumacker, R. (2004). Rasch measurement: The dichotomous model. In E. Smith, & R. Smith (Eds.), *Introduction to Rasch measurement* (p. 236). Maple Grove, Minnesota, USA: JAM Press.
- Srsen, K. G., Vidmar, G., & Zupan, A. (2005). Applicability of the Pediatric Evaluation of Disability Inventory in Slovenia. *Journal of Child Neurology, 20*, 411–416.
- Stahlhut, M., Gard, G., Aadahl, M., & Christensen, J. (2011). Discriminative validity of the Danish version of the Pediatric Evaluation of Disability Inventory (PEDI). *Physical & Occupational Therapy in Pediatrics, 31*, 78–89.
- Wild, D., Grove, A., Martin, M., Eremenco, S., Mcelroy, S., Verjee-Lorenz, A., ... Cultural, A. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health, 8*, 94–104.
- World Health Organization. (2015) Process of translation and adaptation of instruments. Available at: http://www.who.int/substance_abuse/research_tools/translation/en/ (2015)
- Wright, B., & Linacre, J. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*.

Wright, B., & Mok, M. (2004). An overview of the family of Rasch measurement models. In E. Smith, & R. Smith (Eds.), *Introduction to Rasch measurement*. Maple Grove, Minnesota, USA: JAM Press.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Amer A, Kakooza-Mwesige A, Jarl G, et al. The Ugandan version of the Pediatric Evaluation of Disability Inventory (PEDI-UG). Part II: Psychometric properties. *Child Care Health Dev.* 2018;1-10. <https://doi.org/10.1111/cch.12562>