

Reliability of the Dutch Pediatric Evaluation of Disability Inventory (PEDI)

Jeltje E Wassenberg-Severijnen Institute for the Study of Education and Human Development, Utrecht University, **Jan WH Custers** Department of Pediatric Physical Therapy, University Medical Centre/Children's Hospital Utrecht, **Joop J Hox, Adri Vermeer** Institute for the Study of Education and Human Development, Utrecht University and **Paul JM Helders** Department of Pediatric Physical Therapy, University Medical Centre/Children's Hospital Utrecht, The Netherlands

Received 22nd June 2001; returned for revisions 15th August 2001; revised manuscript accepted 23rd July 2002.

Objective: To evaluate the reliability of the Dutch version of the Pediatric Evaluation of Disability Inventory (PEDI), an instrument for measuring functional status (capability and performance in self-care, mobility and social function) of young children using parent interviews.

Design: Inter-interviewer reliability was studied after scoring audiotaped interviews by a second researcher. For test–retest reliability the same parent was interviewed twice within three weeks; in inter-respondent reliability both parents of a child were interviewed independently within a few days. On item level, percentage identical scores were computed, and on scale level intraclass correlation coefficients (ICC) and Cronbach's alphas were calculated.

Subjects: Parents of 63 nondisabled and 53 disabled (various diagnosis) children aged between 7 and 88 months were interviewed.

Results: On scale level, all ICCs were above 0.90 and Cronbach's alpha was 0.89 for the self-care domain, 0.74 for the mobility domain and 0.87 for the social function domain. On item level for the Functional Skills Scale, the mean percentage identical scores varied from 89 to 99, and for the Caregiver Assistance Scale from 54 to 90. Different scores between interviewers resulted partially from ambiguous interpretation of the item and/or the explanation.

Conclusions: Although small adaptations have to be made, the psychometric properties of the Dutch PEDI are found to be good.

Address for correspondence: JE Wassenberg-Severijnen, Utrecht University, Faculty of Social Sciences, M&S, PO Box 80140, 3508 TC Utrecht, The Netherlands. e-mail: J.Wassenberg@fss.uu.nl

Introduction

The Pediatric Evaluation of Disability Inventory (PEDI) was developed to measure functional status in young children.¹⁻³ The PEDI can be used both for discriminative and evaluative purposes, and meets criteria of reliability and validity.^{1,3-6} This instrument, originally developed for the North American population, has recently been translated and cross-culturally adapted for use in the Netherlands.⁷ In the adaptation process four new items were added, while many other items were adapted. Therefore, this series of investigations was designed to establish the psychometric properties of the Dutch PEDI. First, inter-interviewer reliability was studied to find out whether items and the accompanying explanation are interpreted in a uniform way. Secondly, internal consistency was established by studying the extent in which items within a scale are related. Third, test-retest reliability was examined to reach an indication of stability of measures and last, inter-respondent reliability was studied to find out if there are differences in the judgements of both parents.

PEDI

The Pediatric Evaluation of Disability Inventory measures functional status in children aged between six months and seven and a half years. Both the capability of the child and the amount of help they get from their parents as well as the equipment used in daily tasks are measured using a structured interview with the parent(s). Functional status is determined in three domains: self-care, mobility and social function. Table 1 shows an overview of topics in the PEDI.

Items of the Functional Skills Scale (FSS, 201 items), measuring capability, are dichotomous and are scored either 'capable' or 'not capable'. Summed scores can be computed in every domain; the American version also gives standardized scores. The Caregiver Assistance Scale (CAS, 20 items), measuring the amount of help, is an ordinal six-point scale, ranging from 'totally dependent' to 'totally independent'. The Modification Scale (MS, 20 items), measuring the equipment used, is also an ordinal scale.

An example of an item from the Functional Skills Scale is self-care, item 53:

<i>Item</i>	'The child takes off pants, including opening of fasteners.'
<i>Explanation</i>	'The child must be capable of opening snaps, buttons and zippers. Belt buckles are not included.'

Methods

Participants

In testing the Dutch PEDI we interviewed parents of children with disabilities ($n = 53$) and without known disabilities ($n = 63$). The children without known disabilities were between 24 and 36 months of age (mean 31 months, SD 3.6). A first group consisted of children visiting a primary health care centre for infants where growth and development of healthy children is recorded routinely. The health care centre sent a letter with an outline of the study and a request to participate to parents of all children aged two and living in a small town in the centre of the Netherlands ($n = 260$). Parents of 43 children were interviewed (a response of 17%). A second group of nondisabled children were already participating in another study when they were asked to participate in this study: 20 of them agreed.

The children with disabilities in this study were between 7 and 88 months of age (mean 42 months, SD 21.6) and are known to the children's hospital. They have neurometabolic disorders ($n = 29$), spina bifida ($n = 7$), osteogenesis imperfecta ($n = 11$) and infantile encephalopathy ($n = 6$). All children have stable or slowly progressive limitations in performing daily activities. After visiting the outpatient's clinic parents were asked to participate in this study.

Parents who agreed were asked whether they also agreed into audiotaping the interview, whether it was possible to interview also the other parent, and/or to (partially) re-do the interview within a few weeks. In sum, we audiotaped 31 interviews, we interviewed both parents of 32 children, while 20 parents were interviewed twice within a few weeks.

Analysis

Inter-respondent and inter-interviewer reliability was established using scores from both disabled and nondisabled children. In all reliability studies we first looked at the proportion of iden-

tical answers in every item. Although Cohen's kappa is the usual measure of correspondence, its size depends on variance. However, dichotomous skill questions did not always vary that much because of the rather homogeneous group of nondisabled children. Therefore, we decided to use the proportion of identical answers in every

item instead of Cohen's kappa. In addition, intra-class correlation coefficients (ICC) were calculated.

In studying test-retest reliability, the same parent of the same child was interviewed twice: mean time between the two interviews was 14.9 days (SD 3.6 days). The mean time between

Table 1 Content of the Dutch PEDI

Domains	Functional Skills Scale Subscales	Number of items	Caregiver Assistance Scale Modifications Scale
Self-care	Types of food textures	4	Eating
	Use of utensils	5 + 1 ^a	
	Use of drinking containers	5	
	Tooth brushing	5	Grooming
	Hair brushing	4	
	Nose care	5	
	Hand washing	5	Bathing
	Washing body and face	5	
	Pullover/front-opening garments	5	Dressing upper body
	Fasteners	5	
	Pants	5	Dressing lower body
	Shoes/socks	5	
	Toileting task	5	Toileting
Management of bladder	5	Bladder management	
Management of bowel	5	Bowel management	
Mobility	Toilet transfers	5	Chair/toilet transfers
	Chair/Wheelchair transfers	5	
	Car transfers	5	Car transfers
	Bed mobility/transfers	4 + 1 ^a	Bed mobility/transfers
	Tub transfers	5	Tub transfers
	Indoor locomotion methods	3	Indoor locomotion
	Indoor locomotion – distance/speed	5	
	Indoor locomotion – pulls/carries objects	5	
	Outdoor locomotion methods	2 + 1 ^a	Outdoor locomotion
	Outdoor locomotion – distance/speed	5	
	Outdoor surfaces	5	
	Up stairs	5	Stairs
	Down stairs	5	
Social function	Comprehension of word meanings	5	Functional comprehension
	Comprehension of sentence complexity	5	
	Functional use of communication	5	Functional expression
	Complexity of expressive communication	5	
	Problem – resolution	5	Joint problem-solving
	Social interactive play (adults)	5	
	Peer interactions (child of similar age)	5 + 1 ^a	Peer play
	Play with objects	5	
	Self information	5	
	Time orientation	5	
	Play with objects	5	
	Self protection	5	Safety
	Community function	5	

^aItems added in the Dutch version.

interviews of both parents was 3.9 days (SD 4.8 days) and parents were asked not to discuss the interview before they had both been interviewed.

Results

Because of the different type of scales, the results are presented in different tables. Summed scores of the FSS are presented in Table 2 and scale

results in Table 3. The scale results of the CAS and the MS are presented in Table 4.

In inter-interviewer reliability it was possible to study the differences found because the interviews were audiotaped. For the FSS, differences resulted from ambiguous interpretation of item and/or explanation in one-third of the differences, especially in items where correspondence was relatively low. For example, item 41 in the social function domain is 'Can say their own

Table 2 Summed scores: mean (standard deviation) and *t*-test

		Inter-interviewer (<i>n</i> = 31)		Test-retest (<i>n</i> = 20)		Inter-respondent (<i>n</i> = 32)	
		Researcher 1	Researcher 2	First interview	Second interview	Mother	Father
Self-care	Mean	35.4 (21.0)	35.0 (20.8)	27.3 (16.7)	26.3 (16.7)	33.3 (19.1)	32.1 (19.2)
	<i>t</i> -test	<i>t</i> = 1.20 (<i>p</i> = 0.241)		<i>t</i> = 0.80 (<i>p</i> = 0.455)		<i>t</i> = 1.78 (<i>p</i> = 0.086)	
Mobility	Mean	36.6(20.1)	36.4 (20.3)	33.8 (16.6)	33.1 (17.1)	36.3 (19.7)	35.1 (18.7)
	<i>t</i> -test	<i>t</i> = 1.10 (<i>p</i> = 0.283)		<i>t</i> = 0.73 (<i>p</i> = 0.486)		<i>t</i> = 1.51 (<i>p</i> = 0.142)	
Social function	Mean	32.2 (20.3)	31.9 (20.2)	33.4 (16.2)	35.6 (13.8)	32.6 (16.9)	30.9 (17.3)
	<i>t</i> -test	<i>t</i> = 1.17 (<i>p</i> = 0.277)		<i>t</i> = -1.83 (<i>p</i> = 0.100)		<i>t</i> = 1.79 (<i>p</i> = 0.084)	

Table 3 Results of the Functional Skills Scale: mean (standard deviation)

		Self-care 74 items)	Mobility (61 items)	Social function (66 items)
Inter-interviewer	% identical score	97.7 (3.1)	98.8 (2.5)	96.7 (3.6)
	ICC	0.99	0.99	0.99
Test-retest	% identical scores	92.1 (11.6)	95.0 (6.8)	91.5 (10.3)
	ICC	0.98	0.98	0.98
Inter-respondent	% identical scores	89.9 (6.9)	91.4 (6.5)	89.2 (7.0)
	ICC	0.99	0.99	0.97

ICC, intraclass correlation coefficient.

Table 4 Results of the Caregiver Assistance Scale (CAS) and the Modifications Scale (MS): mean (standard deviation)

		Self-care (8 items)	Mobility (7 items)	Social function (5 items)
Inter-interviewer	% identical scores	CAS 85.6 (8.3) MS 84.1 (8.4)	CAS 90.1 (6.0) MS 87.7 (5.6)	CAS 80.6 (9.4) MS 99.2 (1.8)
	ICC	CAS 0.99	CAS 0.99	CAS 0.99
Test-retest	% identical scores	CAS 81.3 (18.7) MS 92.7 (9.1)	CAS 71.9 (11.9) MS 91.1 (11.9)	CAS 54.0 (11.4) MS 94.0 (13.4)
	ICC	CAS 0.97	CAS 0.94	CAS 0.91
Inter-respondent	% identical scores	CAS 66.8 (11.9) MS 87.6 (7.0)	CAS 66.9 (13.5) MS 82.9 (13.2)	CAS 61.4 (6.6) MS 95.0 (11.2)
	ICC	CAS 0.91	CAS 0.97	CAS 0.93

ICC, intraclass correlation coefficient.

name'. From the explanation it was not clear whether a child is capable only if they pronounce their name correctly, or also when they call themselves consequently in the same manner but not correct (which is often the case when a child has a name that is hard to pronounce). In the other two-thirds of the differences, one of the researchers had scored inaccurately. Sometimes it was very obvious, in other cases parents gave an explanation after their judgement 'capable' or 'not capable', by which the initial answer turned out to be incorrect. Both the researcher who took the interviews and the researcher who scored the audiotaped interviews sometimes gave inaccurate scores. For the CAS and the MS, it was not always possible to determine the score that best fitted the parent's answer. In the CAS, especially in items concerning different activities, it was not always clear how to relate those activities into one score. Because the CAS is a six-point ordinal scale, it was also possible to look at the size of the differences found. In our study, the size of the differences found in the CAS was 1 point in more than 80% of the differences. This means that the score in the first interview was just above or below the score in the second interview, the smallest possible difference. In the MS it was sometimes difficult to determine how to score the equipment used. For example, an electric toothbrush is not a modification as meant, but it is when a parent chooses to use an electric toothbrush for their child because of its functional limitations.

Differences found in retesting after a short period of time showed a decrease in independence in the self-care domain, while in the social function domain most differences showed an increase in independence and in the mobility domain both were found. In case of differences between two parents of the same child, mothers judged their child as more independent than fathers did, especially mothers of disabled children.

Cronbach's alpha was computed on the sample of 63 nondisabled children; for the self-care domain $\alpha = 0.89$, for the mobility domain $\alpha = 0.74$ and for the social function domain $\alpha = 0.87$.

Discussion

The purpose of this study was to examine the reliability of the Dutch adaptation of the PEDI. Intraclass correlation coefficients and Cronbach's alphas were high; therefore, reliability was found to be good, and by considering not only scales but also individual items we could improve some items.

For the American PEDI Cronbach's alpha was computed on the normative sample ($n = 410$) that was made up of children aged between six months and seven years and six months.¹ For the Functional Skills Scale alpha was 0.99, 0.97 and 0.98 for the self-care, mobility and social function domains, respectively. Our results are not completely comparable: we controlled for development by computing alpha on a sample that was homogeneous for age, and we found lower alphas (0.89, 0.74 and 0.87, respectively). Dutch research with children of other ages will have to confirm the high alphas found in the studies of the American PEDI.

In our study inter-interviewer reliability was greater than test-retest reliability, and both were greater than inter-respondent reliability. Other researchers^{4,5} found the same results for test-retest and inter-respondent reliability, although their methods were slightly different from ours. We feel that differences between first and second measurement are due to a test effect: parents mentioned that as a result of the first measurement they watched the performances of the child and their assistance more consciously. In a Swedish study,⁸ the Functional Skills Scale was administered as a written questionnaire and the other scales were administered as an interview: sending a list with the subjects to be dis-

Clinical messages

- The Dutch PEDI is a reliable instrument to measure functional status in Dutch children.
- In using the PEDI as an evaluative instrument, it is important to always interview the same parent to be sure that differences found are not due to differences in judgements between parents.

cussed could be a way to increase reliability by focusing attention on these subjects before the parents are interviewed.

In all parts of this study, except for internal consistency, the mobility domain scored better than the self-care and social function domain. An explanation for this finding might be that items of the mobility domain, more than in other domains, are less subjected to choices parents make when raising their children. For example, 'carrying small objects' (mobility, item 35) is different from 'using a knife' (self-care, item 9) where the choice of parents whether or not to let their child practise with a knife is, in some age groups, the main explanation for the score 'capable' or 'not capable'.

Both researchers scored inaccurately while they were interviewing and scoring; we do not feel this to be an interviewer effect. Test length might be an important factor because the PEDI consists of 241 items. Despite the fact that only a part of the 201 Functional Skills Scale items are offered, the PEDI interview lasts about 45–60 minutes and it must be difficult to concentrate fully all that time. Furthermore, parents do not always answer in accordance with the scoring criteria and more information is required than a simple 'capable' or 'not capable' to be sure that the parent has understood the question well and the child meets the scoring criteria. Sometimes, this leads to answers other than those initially given. This is also why administration of the FSS as a written questionnaire did not reduce the interview time⁸. Respondents in the content validity study⁷ also mentioned the length of the test as a possible disadvantage of the PEDI, but did not want to remove items.

In the inter-respondent study it became clear that parents differ in their judgement of the functional status of their child. Mothers judged their children more capable and less dependent on their assistance than fathers did, especially when the child was disabled. This may be because mothers spend more time taking care of their children. We cannot conclude that the judgement of the mother is better than the judgement of the father, just because the mother spends more time with her child. But when the PEDI is used for

evaluative purposes, it seems important to interview the same parent to ensure that differences between two measures are the result of real changes in functional status rather than the result of differences in judgements between parents.

In conclusion, the Dutch PEDI can be used in further research to establish validity, in computing standardized scores for the Dutch population and finally in paediatric rehabilitation practice.

Acknowledgements

We gratefully thank Stichting Nationale Fonds Het Gehandicapte Kind and Zorg Onderzoek Nederland (ZON, 1435.0011) for granting this study.

References

- 1 Haley SM, Coster WJ, Ludlow LH, Haltiwanger JT, Andrellos PJ. *Pediatric Evaluation of Disability Inventory: Development, standardization, and administration manual*. Boston, MA: New England Medical Centre Inc. and PEDI Research Group, 1992.
- 2 Coster WJ, Haley SM. Conceptualization and measurement of disablement in infants and young children. *Infants Young Children* 1992; **4**: 11–22.
- 3 Ketelaar M, Vermeer A, Helders PJM. Functional motor abilities of children with cerebral palsy: A systematic literature review of assessment measures. *Clin Rehabil* 1998; **12**: 369–80.
- 4 Nichols DS, Case-Smith J. Reliability and validity of the Pediatric Evaluation of Disability Inventory. *Pediatr Phys Ther* 1996; **8**:15–24.
- 5 Wright FV, Boschen KA. The Pediatric Evaluation of Disability Inventory (PEDI): Validation of a new functional assessment outcome instrument. *Can J Rehabil* 1993; **7**: 41–42.
- 6 Feldman AB, Haley SM, Coryell J. Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. *Phys Ther* 1990; **10**: 602–10.
- 7 Custers JWH, Wassenberg-Severijnen JE, Net Jvd, Vermeer A, Hart H't, Helders PJM. Dutch adaptation and content validity of the 'Pediatric Evaluation of Disability Inventory (PEDI)'. *Disabil Rehabil* 2002; **24**: 250–58
- 8 Nordmark E, Orban K, Häggglund G, Jarnlo GB. The American Paediatric Evaluation of Disability Inventory (PEDI). Applicability of PEDI in Sweden for children aged 2.0–6.9 years. *Scand J Rehabil Med* 1999; **31**: 95–100.