Accuracy and Precision of the Pediatric Evaluation of Disability Inventory Computer Adapted Tests (PEDI-CAT)

Stephen M. Haley, PhD, PT, FAPTA [Assoc. Director],
Health & Disability Research Inst., School of Public Health, Boston University, Boston, MA, USA

Wendy J. Coster, PhD, OTR/L, FAOTA [Chair and Professor],
Department of Occupational Therapy, Sargent College of Health & Rehabilitation Sciences, Boston University, Boston, MA, USA

Helene M. Dumas, MS, PT [Manager],
Research Center for Children with Special Health Care Needs, Franciscan Hospital for Children, Boston, MA, USA

Maria A. Fragala-Pinkham, DPT, MS, PT [Clinical Researcher],
Research Center for Children with Special Health Care Needs, Franciscan Hospital for Children, Boston, MA, USA

Jessica Kramer, PhD,OTR/L [Assistant Professor],
Occupational Therapy, Boston University Sargent College of Health & Rehabilitation Sciences

Pengsheng Ni, MD [Senior Data Analyst/Research Asst. Professor],
Health & Disability Research Inst., Boston University School of Public Health

Feng Tian, MS [Graduate Research Assistant],
Health & Disability Research Inst., Boston University School of Public Health

Ying-Chia Kao, MA, OTR [Graduate Research Assistant],
Department of Occupational Therapy, Boston University Sargent College of Health & Rehabilitation Sciences

Rich Moed, MPA [President and CEO], and
CREcare LLC, Newburyport, MA, USA

Larry H. Ludlow, Ph.D. [Chairperson and Professor]
Department of Educational Research, Measurement and Evaluation, Boston College Lynch School of Education, Chestnut Hill, MA, USA

Abstract

Aim—The aims of the study were to (1) build new item banks for a revised version of the Pediatric Evaluation of Disability Inventory (PEDI) with four content domains: Daily Activities, Mobility, Social/Cognitive, and Responsibility and 2) use post-hoc simulations based on the combined normative and disability calibration samples to assess the accuracy and precision of the PEDI computerized adaptive tests (PEDI-CAT) in comparison to the administration of all items.

Methods—Parents of typically developing children (*n*=2,205) and parents of children with disabilities (*n*=703) between ages 0 to 21 years, stratified by age and gender participated by

Corresponding Author: Stephen Haley, Boston University School of Public Health, Health and Disability Research Institute, 715 Albany St. Talbot Bldg. T5W, Boston, MA 02118, Phone: 617-638-1986, Fax: 617-638-1999, smhaley@bu.edu.

responding to PEDI-CAT surveys through an existing Internet Opt-in Survey Panel in the USA and by computer tablets in clinical sites.

Results—Confirmatory factor analyses supported four unidimensional content domains. Scores using the real data post-hoc demonstrated excellent accuracy (ICCs ≥ 0.95) with the full item banks. Simulations using item parameter estimates demonstrated relatively small bias in the 10- and 15-item CAT versions; error was generally higher at the scale extremes.

Interpretation—These results suggest the PEDI-CAT can be an accurate and precise assessment of children's daily functioning at all functional levels.

The Pediatric Evaluation of Disability Inventory (PEDI) was originally developed to provide a parent/clinician-reported functional test that was normed on children up to the age of 7 ½ years.[1] The PEDI has a long history of application in developmental medicine, serving as a functional outcome measure for clinical research and practice.[2] The original PEDI was a fixed-format test in which all items needed to be administered in order to derive a score. In addition to some finding the administration of all PEDI items potentially burdensome, the restriction of norms to only young children also limited its use.[3] The purpose of this article is to report on the revision of the PEDI into a series of computer-adaptive tests (CAT) that cover the 0 to 21 year age range.

CATs are being promoted as the next logical step to more efficient health outcome assessments.[4] To employ a CAT, Item Response Theory (IRT) modeling is used to express the association between an individual's response to an item and the outcome domain. IRT measurement models are a class of statistical procedures used to develop measurement scales. The measurement scales are comprised of items with a known relationship between item responses and positions on an underlying functional continuum. Using this approach, probabilities of children scoring a particular response on an item at various functional ability levels can be modeled. These probability estimates are used to determine the child's most likely position along the functional dimension. When assumptions of a particular IRT model are met, estimates of a child's functional ability do not strictly depend on a particular fixed set of items. This scaling feature allows one to compare persons along a functional outcome dimension even if they have not completed the identical set of functional items. CAT also optimizes the items administered, providing items most likely to yield the greatest information for score estimation. CATs will generally require fewer items than a comparable length fixed form instrument to achieve similar precision, although the gains in efficiency may not be uniform throughout the full scale.[5]

Early work has highlighted the possible benefits of using CAT in the assessment of children's functioning, including improved efficiency with minimal loss of precision.[6–9] We have shown that the original PEDI can be successfully engineered into efficient CAT programs.[8, 10] The North American Shriners Orthopedic Hospitals have recently developed a series of parent- and child-reported CAT measures of physical functioning and participation for their network with promising results.[11–15]

The aims of this project were to build and test new PEDI-CAT item banks for an age range of 0 to 21 years and use simulations based on the combined normative and disability calibration samples to assess the accuracy and precision of the PEDI-CAT.

METHOD

Selection of Participants

The targeted normative population of interest for the PEDI-CAT was civilian households in the contiguous United States with children under 21 years of age. Normative data were

collected through the internet. An online panel from YouGovPolimetrix (www.polimetrix.com), (Palo Alto, CA) was the sample source for a nationally representative sample of 2,205 parents. YouGovPolimetrix operates a panel of over one million respondents who have provided them with their names, street addresses, email addresses, and other information, and who regularly participate in online surveys. Panel members receive modest compensation when they participate in online surveys. Panel members are restricted to respondents with fluency in English.

Quota sampling based on age was used to ensure that sufficient cases of typically developing children and youth were collected within each of the PEDI age-based strata (100 cases in each of the 21 age strata starting with 0–1 and stopping with 20–21). Within each age group, equal proportions of gender were selected and efforts were made to ensure that subjects were representative of the racial/ethnic and socio-economic distribution of the US population according to the Year 2000 Census Bureau data. (See Table 1 for sample characteristics). Eligibility for participation was determined by a series of parent-reported screening items to ensure the child was developing typically. For example any child with a disability, chronic condition or receiving specialized services was considered not eligible for the normative sample. Once eligibility was determined participation and consent were obtained.

Data for the vast majority of the disability sample ($n= 617$) were also collected through the internet by YouGovPolimetrix. Eligibility for the disability sample was determined by screening questions which identified children and youth with a physical, cognitive, or behavioral disability. We supplemented the disability sample by collecting parent-reported mobility domain data at two clinical sites that served a wide age range ($n= 86$) in order to add children with more severe physical disabilities. (See Table 2). Both clinical sites collected parent-reported data from an internet system. The study received approval from each site's governing institutional review board, and each parent respondent provided informed consent as well as Health Insurance Portability and Accountability Act authorization before participation.

Procedures

Item development—The initial item pools for the PEDI-CAT were developed through a comprehensive review of existing pediatric measures, the published literature on the functional outcomes of children and youth in hospital-based and community settings and user feedback since the original PEDI's publication in 1992.[16] Feedback regarding content coverage, content relevance, and item clarity was compiled through focus groups, expert reviews and cognitive testing. Details regarding the qualitative procedures used to develop the PEDI-CAT item pool are summarized in Dumas, et al. [17] The calibration item pool was finally condensed to a total of 298 items (76 Daily Activities, 105 Mobility, 64 Social/Cognitive and 53 Responsibility).

PEDI-CAT domains—The PEDI-CAT is comprised of a set of functional activities that are likely to be experienced by children and youth within the context of their daily lives. Functional activity is multidimensional, thus, the PEDI-CAT is comprised of four independent content domains. *Daily Activities* is the ability of a child to carry out daily living skills such as eating, dressing, and grooming activities. The Daily Activities domain also includes items related to household maintenance and the operation of electronic devices. *Mobility* is the ability of a child to move in different environments such as in the home (getting in and out of own bed) or in the community (getting on and off a public bus or school bus). Mobility items range from foundational motor skills of rolling over and sitting unsupported to more physically challenging skills of jumping, running, or carrying heavy

objects. *Social/Cognitive* is the ability to function safely and in effective social exchanges. Social/Cognitive items address communication, interaction, safety, behavior, play, attention, and problem-solving. *Responsibility* is the extent to which a young person is managing life tasks (e.g. fixing a meal, planning and following a weekly schedule) which are important for the transition to adulthood and independent living. See Appendix for sample items in each domain.

Item calibration procedures—The calibration item pool consisted of 76 Daily Activities, 105 Mobility, 64 Social/Cognitive and 53 Responsibility items. Most parents completed the PEDI-CAT items over the Internet. In some cases ($n = 86$), parents completed the assessment during their child's medical or therapy session(s) at one of the two pediatric clinical sites that serve children and youth ages birth to 21 years. IRT analysis does not require complete data on all participants as long as everyone in the sample has answered a common core set of items. Therefore a series of 12 parallel forms were developed so that no one participant responded to more than 175 items. Three quarters of the sample answered items from each domain. A unique set of cases ($n=512$, 25% of sample) completed all items from each domain. The software program provided introductory information and instructions. Filter questions were included to ensure that parents were not asked questions that did not pertain to their child. For example, if parents indicated their child used a walker, but not a wheelchair, subsequent questions pertaining to the use of a wheelchair were not administered.

Statistical analyses procedures—Item response theory methods were used to refine the item pool for each PEDI-CAT dimension. Assumptions that are checked prior to IRT modeling include unidimensionality and local independence (items measure a single trait), and stability or invariance of item parameters across groups (e.g. between normative and disability samples). Violations to these assumptions create sub-optimal modeling of the data and may restrict the accuracy of estimated scores. Our intent was to develop a combined scale for the normative and clinical samples, if possible so that both groups would be scored from the same metric.

We examined the unidimensionality of all four PEDI-CAT domains by confirmatory factor analysis (CFA) and removed items when necessary over a series of subsequent iterative analyses until satisfactory overall fit was achieved. CFA model fit was assessed by multiple fit indexes, including Comparative Fit Index (CFI), Tucker–Lewis Index (TLI) and Root Mean Square Error Approximation (RMSEA). CFI and TLI compare the model to a baseline null model; possible values range from 0 to 1; 0.95 or higher values suggest acceptable fit. RMSEA assesses misfit per degree of freedom; values less than 0.08 mean acceptable fit.[18] We evaluated local independence by inspecting the residual correlations between items using MPlus [19] software.

We developed item parameters using the two-parameter logistic Graded Response Model (GRM) with PARSCALE.[20] To test for item fit, we used likelihood ratio chi-square statistics to test each item based on the comparison of the expected and observed value across the distribution of the latent variable; a *p* value less than 0.05 indicted misfit.[20] We also used the Residual-Plot program[21] to assess item fit as it provided both item and category fit plots. We identified item misfit based on whether both the residual plots and chi-square statistics exceeded published standards. We examined Differential Item Functioning (DIF) based on the logistic regression model,[22] which determines whether there are significant differences in item calibrations between samples. We were particularly interested in DIF between the normative and disability samples.

Simulations were used to assess the accuracy and precision of the PEDI-CAT in comparison to the administration of all items in the domain banks. In simulations using real data, we estimated the subjects' scores based on the response patterns from the calibration data using the Health and Disability Research Institute (HDRI) software. [23] As items were selected for administration in the simulation, responses were taken from the actual data set. After each response, an estimated score based on all administered items to that point in the simulation and the associated standard error was calculated. The selection of the next item was based on the item that could provide the most information at the estimated score. For these simulations, we established specific stop-rules based on the number of items (5, 10, or 15-item versions for each PEDI-CAT domain). For each subject this procedure produced one simulated record of responses for each 5-, 10-, and 15-item CAT version.

In a second series of simulations to evaluate the performance of the PEDI-CAT, a series of Monte Carlo CAT simulations were conducted based only on the item parameters. For each logit point from −4 to +4 with a 0.5 interval, we simulated 100 subjects' response patterns to the item bank. We converted the IRT logit metric to the more conventional PEDI-CAT scoring metric of 20–80. As in the real data simulations, we contrasted 5-, 10 and 15-item stop rule versions of the PEDI-CAT. Using the full-bank score as the reference, we chose the following as evaluation criteria: the average standard error (level of measurement precision also defined as the reciprocal of the information function), bias (difference between the score estimated from the CAT and full item bank), absolute bias (absolute difference of the scores estimated from the CAT and full item bank, and root mean square error (RMSE) (square root of the mean square difference between the scores estimated from the CAT and the full item bank).[24–26] In addition, we provided an assessment of RMSE for the three CAT versions and the full item bank at selected intervals along each of the PEDI-CAT scales. We consider RMSE of <0.30 as the criterion for precise measurement based on a 20–80 score metric.[5]

RESULTS

Unidimensionality and Full Item Banks

We found sound evidence for the unidimensionality of each of the PEDI-CAT domains. (See Table 3). When the items of each domain were modeled as a single-factor, each PEDI-CAT domain had high fit indices (CFI all >.95; TLI all >.99), small error estimates (RMSE between .05–.08), and accounted for more than 80% of the total variance. In addition, the magnitude of the four sets of ratios of the first and second factor eigenvalues were very high, again indicating that each domain had one dominant factor.

Although the CFA results were favorable, some item-level irregularities were noted and addressed. Hence, twenty-two items across the four domains (Daily Activities-8; Mobility-8; Social/Cognitive-4; Responsibility-2) were removed due to excessive local dependence (item residual correlations >.2), poor fit (likelihood ration chi-square ration <.05), or significant DIF across samples. We did retain 16 items (Daily Activities-2; Mobility-9; Social/Cognitive-3; Responsibility-2) with either poor fit or DIF, primarily based on content expert feedback. Removal of these items would have created unacceptable floor effects and removed key content. The final PEDI-CAT item banks consisted of 68 Daily Activities items, 105 Mobility items, 64 Social/Cognitive items and 51 Responsibility items. All subsequent CAT analyses were conducted with the normative and clinical samples combined.

Accuracy and Precision of PEDI-CAT Programs

Based on the real data simulations, correlations of the domain scores from the full item banks and the three CAT versions were very high. See Table 4. All correlations were 0.95 or above, including the 5-item CAT versions.

Using simulations based only on the item parameter estimates with 100 random replications, we calculated the average standard error (SE), bias, absolute bias and RMSE across selected scoring intervals for the PEDI-CAT 20–80 scoring metric. The 15-item CAT versions all have lower SE differences from the full item banks than the 10- or 5-item CAT versions, with the 10-item CAT in the middle. We note that the Responsibility scale has generally greater standard errors than the other three PEDI-CAT scales. There are fairly large differences in bias and absolute bias between the 5-item CAT versions and the 10- and 15-item versions, with the 15-item version always the closest to the full item bank values.

Figure 1 illustrates that PEDI-CAT measurement precision is dependent upon the particular domain and score estimate. In general, measurement precision is much better in the mid-range of each scale than in the extremes. Using the RMSE value of <0.30 as a general criterion for precise measurement, all 5-item CAT versions have relatively high error across all domains at the floor and ceiling extremes. This pattern also occurs in the scoring intervals for the Responsibility scale. RSMEs for all the CAT versions and the full item banks are extremely small in the mid-ranges of the scales, indicating smaller measurement bias in the mid-ranges of the scales than the extremes.

DISCUSSION

All PEDI-CAT domains represented unidimensional constructs. Collectively they provide the foundation for a broad set of item banks. Up to now pediatric clinical programs have had to use multiple assessments that change as the child gets older. A measure such as the PEDI-CAT is needed so that one instrument can be used across all ages and levels of disability. This feature would allow assessment of outcomes for research and program evaluation purposes over the full span of childhood and adolescence as well as across populations using a common metric

Item banks should include items that fit closely with the domain and are spread evenly throughout the range of function. We did allow some items with DIF or fit problems (less than 6% of the total items) across samples to enter into the item bank. This decision was made based on content importance and was applied particularly in the Daily Routines and Mobility scales. This decision will be examined over time to see how frequently the CATs are presenting these items.

We found very high correlations between the full item bank scores and the PEDI-CAT scores for each of the 5-, 10-and 15-item versions. This is a fairly minimum requirement; these relationships are likely an over-estimate because we are using the same responses for the calibrations and the CAT scoring estimates. As in other studies,[8, 10] the results suggest that although scores from the 15-item CAT are closer to the full item bank scores in all instances, the differences in correlations between the 10-and 5-item CATs and the full item banks are relatively small. Similarly, the difference in bias (scaled score in 20–80 metric) between the scores from the 15-item CAT and the full item set are very small; bias increases slightly with the 10-item CATs, and then gets larger with the 5-item CATs.

We also examined conditional precision (Figure 1) as measured by the root mean squared error (RMSE) for the CAT strategies at selected scale locations (100 simulees per location). The 10- and 15-item CAT versions yielded excellent measurement precision throughout

most of the content range, although the 5-item CAT had difficulty at the extremes. Measurement in the mid-range of the scales, regardless of the number of CAT items, was very precise.

Findings suggest that the 15 item PEDI-CAT can be used as an accurate measure of function in clinical outcome measurement and clinical trials, reducing the burden typically placed on both parent respondents and research protocols when full item banks are administered. Other work by our group has shown that the mean time to complete 15 items in each domain (60 items total) is around 12.6 minutes.[27] Additionally, the negligible amount of bias and levels of precision in child estimates obtained from the 10 item CATs suggest that it will provide clinicians with a sound measure of function to inform intervention planning with a significant savings of time. Although we did not examine responsiveness to change in the current study, we believe that the PEDI-CAT should have sensitivity to change that is at least as good as the original PEDI[28] given the results of analyses to date.

Summary

This report describes the development of a CAT approach to measuring functional performance with the new PEDI-CAT. The four domains of the PEDI-CAT demonstrated good unidimensionality and IRT fit. All CATs were accurate, showed small bias except for the 5-item PEDI-CAT, and provided extremely good measurement in the middle of the range. The PEDI-CAT is a broad instrument that covers a wide-age range and meets precision requirements for research and clinical practice uses for children with functional difficulties affecting daily routines, mobility and social/cognitive skills. Future work is needed to examine additional validity aspects of the PEDI-CAT, such as responsiveness and feasibility of use by parents with more limited English skills.

Acknowledgments

References

1. Haley, SM.; Coster, WJ.; Ludlow, LH.; Haltiwanger, JT.; Andrellos, PA. Pediatric Evaluation of Disability Inventory: Development, Standardization and Administration Manual. Boston, MA: Trustees of Boston University; 1992.

2. Haley SMCW, Kao YC, Dumas HM, Fragala-Pinkham MA, Kramer JM, Ludlow LH, Moed R. Lessons from Use of the Pediatric Evaluation of Disability Inventory: Where Do We Go from Here? Pediatr Phys Ther. 2010; 22(1):69–75. [PubMed: 20142708]

3. Ziviani J, Ottenbacher K, Shephard K, Foreman S, Astbury W, Ireland P. Concurrent validity of the Functional Independence Measure for children (WeeFIM) and the Pediatric Evaluation of Disabilitites Inventory in children with developmental disabilities and acquired brain injuries. Phys Occup Ther Pediatr. 2001; 21(2/3):91–101. [PubMed: 12029858]

4. Cella D, Gershon R, Lai J, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. Qual Life Res. 2007; 16(1):133–41. [PubMed: 17401637]

5. Choi SRS, Pilkonis PA, Hays RD, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Qual Life Res. 2009; (19):125–36. [PubMed: 19941077]

6. Gardner W, Shear K, Kelleher K, Pajer K, Mammen O, Buysse D, et al. Computerized adaptive measurement of depression: a simulation study. BMC Psychiatry. 2004; 4:13. [PubMed: 15132755]

7. Haley S, Ni P, Fragala-Pinkham M, Skrinar A, Corzo D. A computer adaptive testing approach for assessing physical function in children and adolescents. Dev Med Child Neurol. 2005; 47(2):113–20. [PubMed: 15707234]

8. Haley SM, Raczek AE, Coster WJ, Dumas HM, Fragala-Pinkham MA. Assessing mobility in children using a computer adaptive testing version of the Pediatric Evaluation of Disability Inventory (PEDI). Arch Phys Med Rehabil. 2005; 86:932–9. [PubMed: 15895339]

9. Jacobusse G, van Buuren S. Computerized adaptive testing for measuring development of young children. Statistics in Medicine. 2007; 26(13):2629–38. [PubMed: 17133649]

10. Coster W, Haley S, Ni P, Dumas H, Fragala Pinkham MA. Assessing self-care and social function using a computer adaptive testing version of the pediatric evaluation of disability inventory. Arch Phys Med Rehabil. 2008; 89(4):622–9. [PubMed: 18373991]

11. Haley S, Chafetz R, Tian F, Montpetit K, Watson K, Gorton G, et al. Validity and reliability of physical functioning computer adaptive tests for children with cerebral palsy. J Pediatr Orthop. 2010; 30:71–5. [PubMed: 20032746]

12. Tucker CA, Montpetit K, Bilodeau N, Gorton GE, Watson K, Fragala-Pinkham MA, et al. Assessment of children with cerebral palsy using a parent–report computer adaptive test: II. Upper extremity skills. Dev Med Child Neurol. 2009:725–31. [PubMed: 19416341]

13. Tucker CA, Gorton GE, Watson K, Fragala-Pinkham MA, Dumas HD, Montpetit K, et al. Assessment of children with cerebral palsy using a parent–report computer adaptive test: I. Lower extremity and mobility skills. Dev Med Child Neurol. 2009:717–24. [PubMed: 19486108]

14. Haley SM, Ni P, Dumas HD, Fragala Pinkham MA, Tucker CA. Measuring global physical health in children with cerebral palsy: illustration of a bi-factor model and computerized adaptive testing. Qual Life Res. 2009; 18:359–65. [PubMed: 19221892]

15. Haley S, Fragala-Pinkham M, Dumas H, Ni P, Gorton G, Watson K, et al. Evaluation of an Item Bank for a Computerized Adaptive Test of Activity in Children With Cerebral Palsy. Phys Ther. 2009; 89(6):589–600. [PubMed: 19423642]

16. Haley, S.; Coster, W. PEDI-CAT: Development, Standardization and Administration Manual. Boston, MA: CRECare, LLC; 2010.

17. Dumas H, Fragala-Pinkham M, Haley S, Coster W, Ying-Chia K, Kramer J, et al. Item bank development for a revised Pediatric Evaluation of Disability Inventory (PEDI). Phys Occup Ther Pediatr. 2010; 30:168–84. [PubMed: 20608855]

18. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care. 2007; 45(5 Suppl 1):S22–31. [PubMed: 17443115]

19. Muthen, B.; Muthen, L. Mplus User's Guide. Los Angeles: Muthen & Muthen; 2001.

20. Muraki, E.; Bock, RD. PARSCALE: IRT Item Analysis and Test Scoring for Rating--Scale Data. Chicago: Scientific Software International; 1997.

21. Liang, T.; Han, K.; Hambleton, R. User's guide for ResidPlots-2: Computer Software for IRT Graphical Residual Analyses, Version 2.0.; 2008. Amherst, MA: University of Massachusetts, Center for Educational Assessment; 2008.

22. Swanson DB, Clauser BE, Case SM, Nungester RJ, Featherman C. Analysis of differential item functioning (DIF) using hierarchical logistic regression models. J Educ Behav Stat. 2002; 27(1): 53–75.

23. Haley SM, Ni P, Hambleton RK, Slavin MD, Jette AM. Computer adaptive testing improves accuracy and precision of scores over random item selection in a physical functioning item bank. J Clin Epidemiol. 2006; 59:1174–82. [PubMed: 17027428]

24. van der Linden, W.; Hambleton, R. Handbook of Modern Item Response Theory. Berlin: Springer; 1997.

25. Wang SD, Wang TY. Precision of Warm's Weighted likelihood estimates for a polytomous model in computerized adaptive testing. Appl Psychol Meas. 2001; 25(4):317–31.

26. Warm TA. Weighted likelihood estimation of ability in item response theory. Psychometrika. 1989; 54(3):427–50.

27. Dumas HM, Fragala-Pinkham MA, Haley SM, Ni P, Coster WJ, Kramer J, Kao Y-C, Moed R, Ludlow LH. Computer adaptive test performance in children with and without disabilities: A prospective field study for the PEDI-CAT. Disabil Rehabil. In press.

28. Iyer LV, Haley SM, Watkins MP, Dumas HM. Establishing minimal clinically important differences for scores on the Pediatric Evaluation of Disability Inventory for Inpatient Rehabilitation. Phys Ther. 2003; 83:888–98. [PubMed: 14519060]

Appendix- Sample PEDI-CAT Items

Domain	Content Areas	Sample Item	Response Scale
Daily Activities (68 items)	Eating & Mealtime	Inserts a straw into a juice box	Please choose which response below best describes your child's ability in the following: • Unable = <u>Can't do</u>, doesn't know how or is too young. • Hard = Does with <u>a lot</u> of help, extra time, or effort. • A little hard = Does with <u>a little</u> help, extra time or effort. • Easy = Does with <u>no</u> help, extra time or effort, or child's skills are past this level. • I don't know.
	Getting Dressed	Puts on winter, sport, or work gloves	
	Keeping Clean	Puts toothpaste on brush and brushes teeth thoroughly	
	Home Tasks	Opens door lock using key	
Mobility (97 items)	Basic Movement & Transfers	When lying on back, turns head to both sides	
	Standing & Walking	Walks while wearing a light backpack	
	Steps & Inclines	Goes up and down an escalator	
	Running & Playing	Pulls self out of swimming pool not using ladder	
	Wheelchair	Goes up and down ramp with wheelchair	
Social/Cognitive (60 items)	Interaction	Greets new people appropriately when introduced	
	Communication	Writes short notes or sends text messages or email	
	Everyday Cognition	Recognizes his/her printed name	
	Self Management	When upset, responds without punching, hitting, or biting	
Responsibility (51 items)	Organization & Planning	Keeping personal electronic devices in working order (e.g., cell phone, computer)	How much responsibility does your child take for the following activities? • Adult/caregiver has <u>full</u> responsibility; the child does not take any responsibility.

Domain	Content Areas	Sample Item	Response Scale
		Includes: Having devices charged and available when needed; Updating software	• Adult/caregiver has most responsibility and child takes a little responsibility. • Adult/caregiver and child share responsibility about equally. • Child has most responsibility with a little direction, supervision or guidance from an adult/caregiver. • Child takes full responsibility without any direction, supervision or guidance from an adult/caregiver. • I don't know.
	Taking Care of Daily Needs	Buying clothing at a store, from a catalog or online Includes: Purchasing clothing, including outerwear and undergarments	
	Health Management	Following health and medical treatment requirements Includes: Taking prescribed medication as directed; Following dietary restrictions; Adhering to exercise or other treatment routines	
	Staying Safe	Using the internet safely Includes: Recognizing scams and inappropriate approaches from strangers; Avoiding posting inappropriate images; Evaluating safety of files before downloading	

What this paper adds

- PEDI-CAT measures four domains (daily activities, mobility, cognitive/social and responsibility in children 0–21 years.

- Psychometric properties (accuracy and precision) of the 10- and 15-item CAT are very good.

- Precision varies with position on the functional continuum; extreme scores are less precise than scores in the mid-range.
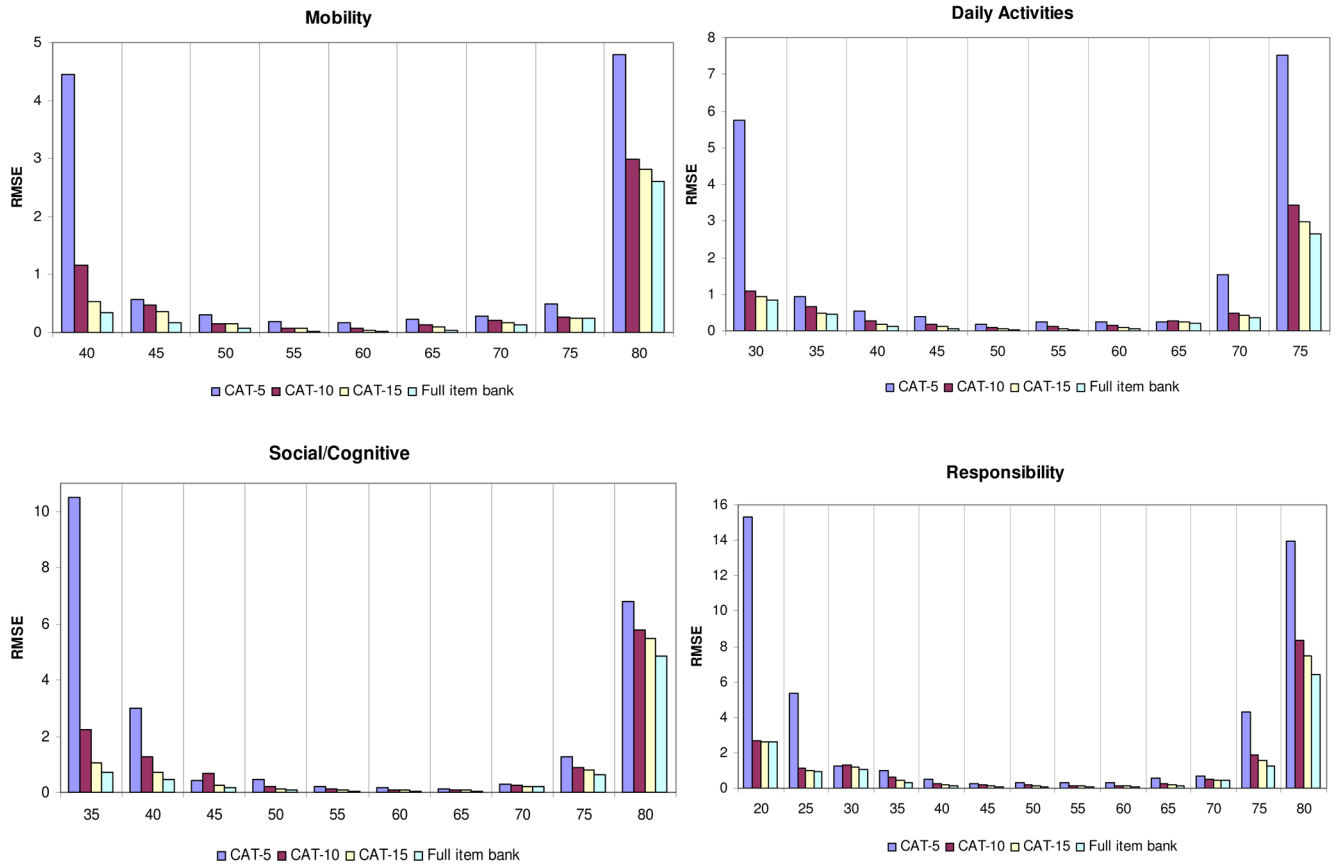
Figure 1.
Comparison of Full Item Bank and RMSE Conditioned on PEDI-CAT Score

Table 1

Demographics

	Normative	Disability
Mean Age (SD)	10.12(6.1)	11.67(4.7)
Gender (Female%)	1079(49)	268(38)
Race and Ethnicity		
White	1438 (65.2)	491 (69.8)
Black	241 (10.9)	63 (9.0)
Hispanic	207 (9.4)	63 (9.0)
Asian	30 (1.4)	13 (1.8)
Native American	13 (0.6)	7 (1.0)
Mixed	222 (10.1)	57 (8.1)
Other	49 (2.2)	7 (1.0)
Middle Eastern	4 (0.2)	2 (0.3)
School Placement		
Preschool/Early Childhood Program/ Kindergarten	294 (13.3)	132(13.8)
Elementary/middle/high school	1236 (56.1)	475 (67.6)
Ungraded	20 (0.9)	20 (2.8)
Undergraduate/College	196 (8.9)	20 (2.8)
Not in school	459 (20.8)	55 (7.8)
Parent/Respondent Education Level* n (%)		
No High School	47 (2.1)	58 (8.3)
High school Graduate	392 (17.8)	141 (20.1)
Some College	846 (38.4)	260 (37.0)
College Graduate	573 (26)	142 (20.2)
Post-Graduate	346 (15.7)	97 (13.8)

Table 2

Conditions Reported by Disability Sample

Developmental delay n (% Yes)	192 (31.1)
Intellectual disability n (% Yes)	50 (8.1)
Hearing impairment n (% Yes)	30 (4.9)
Speech/Language Impairment n (% Yes)	171 (27.7)
Vision impairment n (% Yes)	46 (7.5)
Serious Emotional Disturbance n (% Yes)	50 (8.1)
Orthopedic/Movement Impairment n (% Yes)	23 (3.7)
Autism Spectrum disorder n (% Yes)	108 (17.5)
Attention Deficit Disorder (ADD) n (% Yes)	232 (37.6)
Traumatic Brain Injury n (% Yes)	9 (1.5)
Specific Learning Disability n (% Yes)	78 (12.6)
Health impairment n (% Yes)	58 (9.4)
Multiple Disabilities n (% yes)	28 (4.5)
Other Impairments/Problem n (% Yes)	96 (15.6)

Note: More than one choice was possible

Table 3

Results of Confirmatory Factor Analyses

Scale	# items	% variance	Ratio of 1st and 2nd Eigen values	CFI	TLI	RMSEA
Daily Activities	68	83	37.6	.999	.999	.059
Mobility	75	87	26.7	.988	.992	.088
Social/Cognitive	60	82	20.1	.999	.999	.078
Responsibility	48	81	19.9	.995	.995	.057

Table 4

Accuracy and Scale Coverage- (20–80 scale)

Real Data Correlations	Simulations on 100 random replications				
	Correlation	Average DIFF Standard Error	Average DIFF Bias	Average DIFF Absolute Bias	Average DIFF RMSE
Daily Routines					
5-CAT	0.97	2.44	0.88	1.46	0.88
10-CAT	0.99	2.15	−0.1	0.39	0.1
15-CAT	0.99	1.95	0	0.2	0
Mobility					
5-CAT	0.95	2.03	0.68	1.35	0.87
10-CAT	0.98	1.84	0.08	0.39	0.19
15-CAT	0.99	1.64	−0.01	0.19	0.1
Social/Cognitive					
5-CAT	0.96	2.3	1.72	1.72	1.53
10-CAT	0.98	2.11	−0.1	0.57	0.38
15-CAT	0.99	1.82	0	0.29	0.1
Responsiveness					
5-CAT	0.98	3.48	0.67	2.13	2.36
10-CAT	0.99	3.14	−0.11	0.45	0.34
15-CAT	0.99	2.92	−0.11	0.34	0.11