Research Report

Improved Scaling of the Gross Motor Function Measure for Children With Cerebral Palsy: Evidence of Reliability and Validity

Background and Purpose. This study examined the reliability, validity, and responsiveness to change of measurements obtained with a 66-item version of the Gross Motor Function Measure (GMFM-66) developed using Rasch analysis. Subjects and Methods. The validity of measurements obtained with the GMFM-66 was assessed by examining the hierarchy of items and the GMFM-66 scores for different groups of children from a stratified random community-based sample of 537 children with cerebral palsy (CP). A subset of 228 children who had been reassessed at 12 months was used to test the hypothesis that children who are young (<5 years of age) and have "mild" CP will demonstrate greater change in GMFM-66 scores than children who are older (≥5 years of age) and whose CP is more severe. Data from an additional 19 children with CP who were assessed twice, one week apart, were used to examine test-retest reliability. Results. The overall changes in GMFM-66 scores over 12 months and a time × severity × age interaction supported our hypotheses. Test-retest reliability was high (intraclass correlation coefficient=.99). Conclusion and Discussion. This study demonstrated that the GMFM-66 has good psychometric properties. By providing a hierarchical structure and interval scaling, the GMFM-66 can provide a better understanding of motor development for children with CP than the 88 item GMFM and can improve the scoring and interpretation of data obtained with the GMFM. [Russell DJ, Avery LM, Rosenbaum PL, et al. Improved scaling of the Gross Motor Function Measure for children with cerebral palsy: evidence of reliability and validity. *Phys Ther.* 2000;80:873–885.]

Key Words: *Cerebral palsy, Gross Motor Function Measure, Motor function, Rasch analysis, Validation.*

Dianne J Russell

Lisa M Avery

Peter L Rosenbaum

Parminder S Raina

Stephen D Walter

Robert J Palisano

The Gross Motor Function Measure (GMFM) is a criterion-referenced observational measure that was developed and validated to assess children with cerebral palsy (CP).[1] The original GMFM was modified in 1990 based on feedback from the clinicians involved in the validation study. Three items were added to the original 85-item measure in an effort to allow the skills tested by those items to be assessed bilaterally. Prior to re-establishing the reliability of the GMFM measurements with the 3 items added, administration and scoring guidelines were developed. The reliability of scores obtained with the 88-item GMFM was established with the revised guidelines using videotaped examples, and this reliability was sufficiently high to permit the revised guidelines to replace the original guidelines (intraclass correlation coefficient [ICC]=.90).[2] Further evidence of the reliability of measurements obtained with the 88-item GMFM has been established by several investigators for its use with children with CP[3–5] and for children with Down syndrome.[6]

The 88 items of the GMFM are measured by observation of the child and scored on a 4-point ordinal scale (0=does not initiate, 1=initiates <10% of activity, 2=partially completes 10% to <100% of activity, 3=completes activity). The items are weighted equally and grouped into 5 dimensions: (1) lying and rolling (17 items), (2) sitting (20 items), (3) crawling and kneeling (14 items), (4) standing (13 items), and (5) walking, running, jumping (24 items). By the age of 5 years, children without motor delays can generally accomplish all of the items of the GMFM.[1] Scores for each dimension are expressed as a percentage of the maximum score for that dimension. The total score is obtained by averaging the percentage scores across the 5 dimensions. The original intent of the GMFM developers was to have one measure that could be used for children across a spectrum of ability levels in order to make it possible for children with different gross motor abilities to enter clinical trials and be assessed with the same measurement tool.[2] In addition, the measure needed to be useful for tracking individual children over time. Although items tend to increase in difficulty within a dimension, their ordering was based on clinical judgment and the literature, and the order had not been substantiated with data-based evidence. Items were grouped into dimensions primarily for ease of administration.

Evidence of construct validity of the measure's capacity to detect change in motor function over time was supported by several analyses of the scores of children who were administered the GMFM twice by the same therapist over a 5- to 7-month interval.[1] For children with CP, change scores on the GMFM were correlated with parents' judgments of change (r=.54), the child's treating therapist's judgment of change (r=.65), and ratings of change made by therapists who were familiar with the GMFM but unfamiliar with the children by viewing pairs of videotapes, which were in random order (r=.82). As hypothesized in the original validation study,[1] the

DJ Russell, MSc, is Assistant Professor, School of Rehabilitation Science, Faculty of Health Sciences, McMaster University, and Research Coordinator, *CanChild*, Centre for Childhood Disability Research, Room 408, Institute for Applied Health Sciences, McMaster University, 1400 Main St W, Hamilton, Ontario, Canada L8S 1C7 (russelld@fhs.mcmaster.ca). Address all correspondence to Ms Russell.

LM Avery, BEng, is Research Assistant, *CanChild*, Centre for Childhood Disability Research, McMaster University.

PL Rosenbaum, MD (FRCP), is Professor, Department of Pediatrics, McMaster University, Co-Director, *CanChild*, Centre for Childhood Disability Research, McMaster University, and Chief of Medical Staff, Bloorview MacMillan Centre, Toronto, Ontario, Canada.

PS Raina, PhD, is Assistant Professor, Department of Health Care and Epidemiology, Centre for Community Health and Health Evaluation Research, BC Research Institute for Children's and Women's Health, University of British Columbia, Vancouver, British Columbia, Canada.

SD Walter, PhD, is Professor, Clinical Epidemiology and Biostatistics, McMaster University, and Investigator, *CanChild*, Centre for Childhood Disability Research, McMaster University.

RJ Palisano, PT, ScD, is Professor, Department of Rehabilitation Sciences, MCP Hahnemann University, Philadelphia, Pa, and Co-investigator, *CanChild*, Centre for Childhood Disability Research, McMaster University.

All authors provided concept/research design, writing, and consultation (including review of manuscript before submission). Ms Avery, Dr Walter, and Dr Raina provided data analysis. Ms Russell and Ms Avery provided project management, and Ms Russell and Dr Rosenbaum provided fund procurement.

The study was approved by the McMaster University Research Ethics Board.

This research was supported by grant R01 HD34947 from the National Institute for Child Health and Human Development, National Institutes of Health.

This article was submitted June 22, 1999, and was accepted May 10, 2000.

change in GMFM scores was greatest in children recovering from acute head injury, less in preschool-aged children without motor delays, and least in children with CP. For the children with CP, the amount of change in GMFM scores was related to their age and severity of motor disability. The GMFM scores of children who were young (<3 years of age) and had "mild" CP changed more than the scores of children who were older (>6 years of age) and had "severe" CP. Furthermore, children who were judged by both their parents and therapists not to have changed did not demonstrate a change in GMFM scores, whereas children who were judged by both their parents and therapists to have changed demonstrated an increase in GMFM scores.

Bjornson and colleagues[7] replicated the original GMFM validation study with children with spastic diplegia and quadriplegia and provided support for the construct validity of measurements of change in motor function obtained with the GMFM over 12- and 24-month periods. The responsiveness of the GMFM for infants under the age of 24 months with CP and motor delay has also been established.[8] Researchers have used the GMFM with children with CP to assess the effectiveness of rhizotomy,[9–11] intrathecal baclofen,[12,13] physical therapy,[14,15] horseback riding,[16] therapeutic electrical stimulation,[17] orthoses,[18,19] strength training,[20,21] and muscle tendon surgery,[22] as well as to determine the correlation between these measurements and measurements of gait and physical fitness.[23–26]

As the GMFM has been used in a variety of clinical and research situations, its limitations have become more apparent. Some users of the GMFM have chosen to administer only those dimensions that are most relevant to their clients' current level of functional needs. This selective use of the GMFM allows fewer items to be administered and increases the measure's responsiveness to change by eliminating items that are not relevant to the therapeutic intervention or are unlikely to change as a result of intervention. However, the evidence for the reliability and the validity of the dimension scores is generally not as strong as for the measure as a whole.[1]

Another limitation has been the interpretation of the GMFM total score. Children with different skills and abilities within and between dimensions can receive the same total score. A further limitation is that scores of children functioning in the middle of the scale have greater potential to change than scores of children whose initial assessment is either very low or very high because more items are in the middle of the scale than at the extremes.

In an effort to improve the interpretability and the clinical usefulness of the GMFM, we applied the Rasch model of item analysis to the GMFM.[27] Rasch analysis uses a one-parameter logistic model to derive an equal interval measure from the raw score.[28] Rasch analysis originated in the areas of education and psychology, and during the last 5 to 10 years has been used to construct and validate measures used in rehabilitation.[29–33] There are several reasons to apply the Rasch scaling to the GMFM. First, the items can be arranged in order of relative difficulty (hierarchical structure). Second, it makes it possible to create interval scales from what are believed to be ordinal scales, because scores take into account how much more difficult one item is to accomplish than the previous item. Third, Rasch analysis allows the elimination of items that do not fit the unidimensional construct (ie, misfitting items). The unidimensionality assumption is met when the items of the test demonstrate that only one ability is being measured (in our case, gross motor ability). Fourth, Rasch analysis allows calculation of a total score when not all items are administered. All of these factors would lead to improved scoring and interpretation of the GMFM.

Although Rasch scaling has potential advantages, potential challenges also needed to be considered. These challenges include the fact that the GMFM is already in widespread use and a revised measure would require modification of the manual and training materials. In addition, because Rasch analysis is used primarily to create unidimensional measures, it could identify items for removal that do not fit the model but are considered clinically relevant. Of major concern was whether the removal of items would affect the GMFM's responsiveness to change and require new validation of the modified scale. A computer program would also be necessary to analyze and interpret scores based on Rasch scaling.

With these considerations in mind, our group applied the Rasch partial credit model to the GMFM.[27] The Rasch partial credit model does not make any assumptions about the difficulty of each response option, either within an item or between items. For example, it does not assume that the difficulty of going from a score of 1 to a score of 2 is the same difficulty as going from a score of 2 to a score of 3 or that the difficulty of going from a score of 1 to a score of 2 is the same for different items. Rasch modeling helped us to identify 66 items from the original 88-item GMFM that form a unidimensional hierarchical scale, the GMFM-66. Table 1 lists the original 88 GMFM items and indicates which items we removed for the GMFM-66.

Following modification of the scale, we believed that it was important to assess the psychometric properties of the GMFM-66. One of the key concerns in reducing the number of items in the GMFM was not to lose the measure's responsiveness to change over time. We also

Table 1.
Listing of Gross Motor Function Measure (GMFM) Items Indicating Items Removed, Mean Difficulty Estimates, and Standard Errors[a]

Dimension A: Lying & Rolling	Difficulty	SE
1—SUP, HEAD IN MIDLINE: TURNS HEAD WITH EXTREMITIES SYMMETRICAL	Removed	
2—SUP: BRINGS HANDS TO MIDLINE, FINGERS ONE WITH THE OTHER	22.90	0.59
3—SUP: LIFTS HEAD 45°	Removed	
4—SUP: FLEXES R HIP & KNEE THROUGH FULL RANGE	Removed	
5—SUP: FLEXES L HIP AND KNEE THROUGH FULL RANGE	Removed	
6—SUP: REACHES OUT WITH R ARM, HAND CROSSES MIDLINE TOWARD TOY	24.66	0.59
7—SUP: REACHES OUT WITH L ARM, HAND CROSSES MIDLINE TOWARD TOY	24.54	0.59
8—SUP: ROLLS TO PR OVER R SIDE	Removed	
9—SUP: ROLLS TO PR OVER L SIDE	Removed	
10—PR: LIFTS HEAD UPRIGHT	17.25	0.71
11—PR ON FOREARMS: LIFTS HEAD UPRIGHT, ELBOWS EXT, CHEST RAISED	Removed	
12—PR ON FOREARMS: WEIGHT ON R FOREARM, FULLY EXTENDS OPPOSITE ARM FORWARD	Removed	
13—PR ON FOREARMS: WEIGHT ON L FOREARM, FULLY EXTENDS OPPOSITE ARM FORWARD	Removed	
14—PR: ROLLS TO SUP OVER R SIDE	Removed	
15—PR: ROLLS TO SUP OVER L SIDE	Removed	
16—PR: PIVOTS TO R 90° USING EXTREMITIES	Removed	
17—PR: PIVOTS TO L 90° USING EXTREMITIES	Removed	
Dimension B: Sitting	**Difficulty**	**SE**
18—SUP, HANDS GRASPED BY EXAMINER: PULLS SELF TO SITTING WITH HEAD CONTROL	24.31	0.59
19—SUP: ROLLS TO R SIDE, ATTAINS SITTING	Removed	
20—SUP: ROLLS TO L SIDE, ATTAINS SITTING	Removed	
21—SIT ON MAT, SUPPORTED AT THORAX BY THERAPIST: LIFTS HEAD UPRIGHT, MAINTAINS 3 SEC	13.07	0.88
22—SIT ON MAT, SUPPORTED AT THORAX BY THERAPIST: LIFTS HEAD TO MIDLINE, MAINTAINS 10 SEC	18.13	0.71
23—SIT ON MAT, ARM(S) PROPPING: MAINTAINS, 5 SEC	23.07	0.59
24—SIT ON MAT: MAINTAINS, ARMS FREE, 3 SEC	30.08	0.53
25—SIT ON MAT WITH SMALL TOY IN FRONT: LEANS FORWARD, TOUCHES TOY, RE-ERECTS WITHOUT ARM PROPPING	33.84	0.53
26—SIT ON MAT: TOUCHES TOY PLACED 45° BEHIND CHILD'S R SIDE, RETURNS TO START	37.67	0.53
27—SIT ON MAT: TOUCHES TOY PLACED 45° BEHIND CHILD'S L SIDE, RETURNS TO START	37.08	0.53
28—R SIDE SIT: MAINTAINS, ARMS FREE, 5 SEC	Removed	
29—L SIDE SIT: MAINTAINS, ARMS FREE, 5 SEC	Removed	
30—SIT ON MAT: LOWERS TO PR WITH CONTROL	38.02	0.53
31—SIT ON MAT WITH FEET IN FRONT: ATTAINS 4 POINT OVER R SIDE	44.20	0.47
32—SIT ON MAT WITH FEET IN FRONT: ATTAINS 4 POINT OVER L SIDE	44.97	0.47
33—SIT ON MAT: PIVOTS 90°, WITHOUT ARMS ASSISTING	Removed	
34—SIT ON BENCH: MAINTAINS, ARMS AND FEET FREE, 10 SEC	36.55	0.53
35—STD: ATTAINS SIT ON SMALL BENCH	47.62	0.47
36—ON THE FLOOR: ATTAINS SIT ON SMALL BENCH	45.03	0.47
37—ON THE FLOOR: ATTAINS SIT ON LARGE BENCH	47.85	0.47
Dimension C: Crawling & Kneeling	**Difficulty**	**SE**
38—PR: CREEPS FORWARD 6'	Removed	
39—4 POINT: MAINTAINS WEIGHT ON HANDS AND KNEES, 10 SEC	38.79	0.53
40—4 POINT: ATTAINS SIT ARMS FREE	43.20	0.47
41—PR: ATTAINS 4 POINT, WEIGHT ON HANDS AND KNEES	39.43	0.53
42—4 POINT: REACHED FORWARD WITH R ARM, HAND ABOVE SHOULDER LEVEL	44.32	0.47
43—4 POINT: REACHED FORWARD WITH L ARM, HAND ABOVE SHOULDER LEVEL	44.67	0.47
44—4 POINT: CRAWLS OR HITCHES FORWARD 6'	42.44	0.47
45—4 POINT: CRAWLS RECIPROCALLY FORWARD 6'	46.56	0.47
46—4 POINT: CRAWLS UP 4 STEPS ON HANDS AND KNEES/FEET	47.32	0.47
47—4 POINT: CRAWLS BACKWARDS DOWN 4 STEPS ON HANDS AND KNEES/FEET	Removed	
48—SIT ON MAT: ATTAINS HIGH KN USING ARMS, MAINTAINS, ARMS FREE, 10 SEC	45.44	0.47
49—HIGH KN: ATTAINS HALF KN ON R KNEE USING ARMS, MAINTAINS, ARMS FREE, 10 SEC	Removed	
50—HIGH KN: ATTAINS HALF KN ON L KNEE USING ARMS, MAINTAINS, ARMS FREE, 10 SEC	Removed	
51—HIGH KN: KN WALKS FORWARD 10 STEPS, ARMS FREE	53.03	0.47
Dimension D: Standing	**Difficulty**	**SE**
52—ON THE FLOOR: PULLS TO STD AT LARGE BENCH	43.14	0.47
53—STD: MAINTAINS, ARMS FREE, 3 SEC	46.97	0.47
54—STD: HOLDING ON TO LARGE BENCH WITH ONE HAND, LIFTS R FOOT, 3 SEC	50.68	0.47
55—STD: HOLDING ON TO LARGE BENCH WITH ONE HAND, LIFTS L FOOT, 3 SEC	50.97	0.47
56—STD: MAINTAINS, ARMS FREE, 20 SEC	54.56	0.47
57—STD: LIFTS L FOOT, ARMS FREE, 10 SEC	74.81	0.59
58—STD: LIFTS R FOOT, ARMS FREE, 10 SEC	74.63	0.59
59—SIT ON SMALL BENCH: ATTAINS STD WITHOUT USING ARMS	52.09	0.47
60—HIGH KN: ATTAINS STD THROUGH HALF KN ON R KNEE, WITHOUT USING ARMS	61.04	0.53
61—HIGH KN: ATTAINS STD THROUGH HALF KN ON L KNEE, WITHOUT USING ARMS	61.57	0.53
62—STD: LOWERS TO SIT ON FLOOR WITH CONTROL, ARMS FREE	57.39	0.53
63—STD: ATTAINS SQUAT, ARMS FREE	58.15	0.53
64—STD: PICKS UP OBJECT FROM FLOOR, ARMS FREE, RETURNS TO STAND	55.03	0.47
Dimension E: Walking, Running & Jumping	**Difficulty**	**SE**
65—STD, 2 HANDS ON LARGE BENCH: CRUISES 5 STEPS TO R	45.50	0.47
66—STD, 2 HANDS ON LARGE BENCH: CRUISES 5 STEPS TO L	45.50	0.47
67—STD, 2 HANDS HELD: WALKS FORWARD 10 STEPS	40.67	0.53
68—STD, 1 HAND HELD: WALKS FORWARD 10 STEPS	49.15	0.47
69—STD: WALKS FORWARD 10 STEPS	55.44	0.53
70—STD: WALKS FORWARD 10 STEPS, STOPS, TURNS 180°, RETURNS	57.39	0.53
71—STD: WALKS BACKWARD 10 STEPS	61.27	0.59
72—STD: WALKS FORWARD 10 STEPS CARRYING A LARGE OBJECT WITH 2 HANDS	57.68	0.59
73—STD: WALKS FORWARD 10 CONSECUTIVE STEPS BETWEEN PARALLEL LINES, 8" APART	66.16	0.59
74—STD: WALKS FORWARD 10 CONSECUTIVE STEPS ON A STRAIGHT LINE ¾" WIDE	73.04	0.59
75—STD: STEPS OVER STICK AT KNEE LEVEL, R FOOT LEADING	67.27	0.59
76—STD: STEPS OVER STICK AT KNEE LEVEL, L FOOT LEADING	67.16	0.59
77—STD: RUNS 15 FEET, STOPS AND RETURNS	65.10	0.59
78—STD: KICKS BALL WITH R FOOT	59.68	0.59
79—STD: KICKS BALL WITH L FOOT	60.15	0.59
80—STD: JUMPS 12" HIGH, BOTH FEET SIMULTANEOUSLY	74.75	0.59
81—STD: JUMPS FORWARD 12", BOTH FEET SIMULTANEOUSLY	69.45	0.59
82—STD ON R FOOT: HOPS ON R FOOT 10 TIMES WITHIN A 24" CIRCLE	83.93	0.65
83—STD ON L FOOT: HOPS ON L FOOT 10 TIMES WITHIN A 24" CIRCLE	83.76	0.65
84—STD, HOLDING 1 RAIL: WALKS UP 4 STEPS, HOLDING 1 RAIL, ALTERNATING FEET	62.74	0.59
85—STD, HOLDING 1 RAIL: WALKS DOWN 4 STEPS, HOLDING 1 RAIL, ALTERNATING FEET	66.57	0.59
86—STD: WALKS UP 4 STEPS, ALTERNATING FEET	72.40	0.59
87—STD: WALKS DOWN 4 STEPS, ALTERNATING FEET	77.28	0.59
88—STD ON 6" STEP: JUMPS OFF, BOTH FEET SIMULTANEOUSLY	70.04	0.59

[a] SUP=supine, PR=prone, R=right, L=left, SIT=sitting, STD=standing, KN=kneeling, EXT=extended. An item's difficulty corresponds to the ability required to receive a score of 3 on that item. "Removed" refers to items that were removed from the original 88-item GMFM by the Rasch analysis and are no longer part of the GMFM-66 scoring.

Table 2.
Sample Characteristics by Mean 66-Item Gross Motor Function Measure (GMFM-66) Score and Age (N=537)

	Baseline GMFM-66 Score				Age (y) at Baseline				
	$\bar{X}$	SD	Minimum	Maximum	$\bar{X}$	SD	Minimum	Maximum	N
Type of cerebral palsy									
Spastic	53.93	23.46	0.00	100.00	6.48	2.79	0.95	12.71	411
Dystonic/athetotic	37.35	14.96	17.01	67.75	5.11	2.15	1.67	9.23	32
Ataxic	60.84	11.75	41.79	74.75	6.42	2.88	2.13	10.02	14
Low tone-hypotonic	47.87	23.02	0.00	100.00	6.67	2.52	1.71	10.66	27
Mixed	40.90	21.19	0.00	100.00	6.61	2.73	1.62	11.06	53
Distribution									
Leg dominant	62.21	16.04	14.83	100.00	6.41	2.69	0.95	11.74	183
Three-limb dominant	47.80	14.67	22.66	89.70	6.63	2.45	1.81	11.06	53
Four-limb dominant	33.57	17.00	0.00	100.00	6.45	2.80	1.52	12.71	215
Right hemiplegic	75.08	15.81	43.26	100.00	6.09	3.17	1.72	11.57	43
Left hemiplegic	78.62	16.71	34.84	100.00	6.52	2.80	1.59	11.82	42
Missing									1
GMFCS[a] level									
I	78.06	13.29	45.91	100.00	6.52	2.83	1.59	11.82	155
II	60.92	11.16	34.84	89.70	5.89	2.77	1.69	10.78	70
III	49.98	7.07	29.31	67.04	6.68	2.83	0.95	11.73	104
IV	37.94	7.77	19.72	52.85	6.38	2.63	1.62	11.74	105
V	20.63	8.66	0.00	46.67	6.45	2.68	1.68	12.71	103
Sex									
Male	52.04	23.85	0.00	100.00	6.38	2.84	0.95	12.71	299
Female	51.18	22.38	0.00	100.00	6.49	2.65	1.52	11.82	238

[a] GMFCS=Gross Motor Function Classification System.[34]

wanted to determine whether the GMFM-66 is stable over a short period of time, when true change in gross motor function is not expected to occur. We hypothesized that (1) test-retest reliability for the GMFM-66 would not differ from the test-retest reliability of data obtained with the original 88-item GMFM, (2) children with CP would demonstrate an increase in GMFM-66 scores over 12 months, and (3) children classified as younger and having "mild" CP would demonstrate a greater change score on the GMFM-66 over 12 months compared with children classified as older and having "severe" CP. The purposes of this article are (1) to report the psychometric properties of reliability, validity, and responsiveness to change of data obtained with the GMFM-66 and (2) to discuss the research and clinical implications of the GMFM-66 for users of the measure.

Method

Data used for the Rasch analyses came from the cross-sectional time 1 GMFM assessments of 537 children stratified by age and severity of motor disability using the Gross Motor Function Classification System (GMFCS).[34] The children were participating in a large longitudinal study of motor development in children with cerebral palsy. This group included a subset of 228 children who had had a second GMFM assessment 12 months later. Children (N=2,108) were eligible for the longitudinal study if they had a diagnosis of CP and were on the case list of 1 of 18 publicly funded children's treatment centers in the province of Ontario, Canada, as of June 1, 1996. Children were not included in the study if they had other neuromotor disorders such as spina bifida, neuromuscular disease, or musculoskeletal disease or if they had had a selective dorsal rhizotomy or intrathecal baclofen prior to recruitment for the study. Information on the sample characteristics, including sex, age, type and distribution of CP, and severity of motor disability (GMFCS level), are listed in Table 2. The characteristics of the subset of 228 children whose scores were used to examine the responsiveness of the GMFM-66 are presented in Table 3.

Reliability

To determine the test-retest reliability of the GMFM-66 scores, data were used from 19 children with CP who were assessed twice, 1 week apart, by the same therapist. These data were from a previous study of the reliability of data obtained with the GMFM.[3] Each child's GMFM-88 total scores were computed according to the GMFM guidelines[2] and were also computed using the Gross Motor Ability Estimator (GMAE),[35] which is software that analyzes the interval level scale of the GMFM-66.

Validity and Responsiveness

Face validity was assessed by examining the hierarchy of items (Tab. 4) and the GMFM-66 total scores for different groups of children to determine whether they made

Table 3.
Sample Characteristics by Mean Baseline 66-Item Gross Motor Function Measure (GMFM-66) Score and Age of Children Used in Responsiveness Analysis (n=228)

	Baseline GMFM-66 score				Age at Baseline				
	$\bar{X}$	SD	Minimum	Maximum	$\bar{X}$	SD	Minimum	Maximum	N
Type of cerebral palsy									
Spastic	54.79	21.74	0.00	100.00	6.45	2.84	1.57	11.82	173
Dystonic/athetotic	39.47	13.78	17.01	60.39	6.04	2.41	2.31	11.02	17
Ataxic	66.99	8.49	54.15	74.75	9.48	0.67	8.71	10.02	5
Low tone-hypotonic	47.43	18.82	4.12	73.63	7.14	2.33	2.38	10.26	15
Mixed	46.27	22.15	21.25	100.00	7.18	2.76	2.18	10.92	18
Distribution									
Leg dominant	61.07	15.42	31.19	100.00	6.84	2.79	1.59	11.12	74
Three-limb dominant	45.93	13.73	22.66	82.99	7.11	2.49	2.62	10.90	30
Four-limb dominant	36.35	15.75	0.00	74.75	6.44	2.76	1.57	11.70	82
Right hemiplegic	74.09	16.54	43.26	96.00	5.64	3.01	1.97	10.52	24
Left hemiplegic	76.33	17.32	35.26	100.00	6.62	2.85	1.75	11.82	18
GMFCS[a] level									
I	76.69	14.05	46.32	100.00	6.17	2.82	1.59	11.82	61
II	63.49	12.03	37.14	89.70	6.60	2.90	1.69	10.72	35
III	50.31	7.23	31.78	67.04	6.58	2.86	1.57	11.12	49
IV	38.70	6.53	27.31	52.85	6.73	2.74	2.17	11.74	48
V	23.06	9.28	0.00	46.67	7.12	2.53	2.68	11.02	35
Gender									
Male	52.34	22.03	0.00	100.00	6.50	2.78	1.57	11.74	142
Female	53.46	20.38	13.54	100.00	6.73	2.79	1.75	11.82	86

[a] GMFCS=Gross Motor Function Classification System.[34]

clinical sense (Tab. 2) based on what is known about CP. We expected that items from the lying and rolling and the sitting dimensions would be easier for children with CP to accomplish than items in standing and the walking, running, and jumping dimensions and would, therefore, have lower difficulty estimates. We also expected that children with hemiplegia would have higher scores, on average, than children with more limbs involved (eg, children with diplegia, triplegia, or quadriplegia) and that mean GMFM-66 scores would vary systematically by GMFCS level,[34] with children in level I (mild disability) having the highest GMFM-66 scores.

Construct validity of the responsiveness of GMFM-66 scores was also assessed by testing an *a priori* hypothesis, similar to the method used in the construct validation of the original GMFM,[1] and centered on the measure's ability to respond to change using information about the natural history of CP.[36] The criterion for inclusion in the responsiveness analyses was that children had 2 GMFM assessments 12 months apart (±1 month). These 228 children were assigned to an expected change category based on their age (<5 years or ≥5 years) and severity of motor disability, as assessed using the GMFCS ("mild"=levels I and II, "moderate"=level III, "severe"=levels IV and V).

Data Analysis

Test-retest reliability data were analyzed using an ICC based on the 1-way analysis of variance (ANOVA) model 1,1[37] for both the new and the original scoring methods. To test the difference in reliability between the 2 scoring methods, 95% confidence limits for the difference in ICCs were computed.[38] To determine whether an effect of time existed and whether there were interactions of age and severity, a 3-way repeated measures ANOVA was calculated with time as the within-subject factor and age and severity as between-subject factors.

Results

Reliability

Test-retest reliability data using the new GMAE scoring method[35] showed that the GMFM-66 has a high level of stability over time (ICC=.9932) and is only slightly different from that of the original scoring system using the 88-item GMFM (ICC=.9944, mean difference=.0013 [95% confidence interval of −.082 to .0109]).

Validity

The items of the GMFM-66 in order of difficulty ascertained by Rasch analysis (ie, with an item score of 3) are presented in Table 4. To remove negative values and be more clinically interpretable, the linear estimates of difficulty expressed in logits (log-odds ratios) were trans-

Table 4.
Hierarchy of Items (From Easy to Difficult) in the 66-Item Gross Motor Function Measure (GMFM-66) for Children with Cerebral Palsy[a]

Item	Difficulty[b]
21—Sit on mat, supported at thorax by therapist: lifts head upright, maintains 3 sec	15.72
10—PR: lifts head upright	23.25
23—Sit on mat, arm(s) propping: maintains, 5 sec	25.13
22—Sit on mat, supported at thorax by therapist: lifts head to midline, maintains 10 sec	25.60
2—SUP: brings hands to midline, fingers one with the other	26.84
7—SUP: reaches out with L arm, hand crosses midline towards toy	29.72
18—SUP, hands grasped by examiner: pulls self to sitting with head control	30.84
6—SUP: reaches out with R arm, hand crosses midline toward toy	31.43
24—Sit on mat: maintains, arms free, 3 sec	32.14
27—Sit on mat: touches toy placed 45° behind child's L side, returns to start	39.49
25—Sit on mat with small toy in front: leans forward, touches toy, re-erects without arm propping	39.67
34—Sit on bench: maintains, arms and feet free, 10 sec	39.91
26—Sit on mat: touches toy placed 45° behind child's R side, returns to start	40.44
39—4 point: maintains, weight on hands and knees, 10 sec	41.20
30—Sit on mat: lowers to PR with control	41.38
41—PR: attains 4 point, weight on hands and knees	41.49
67—STD, 2 hands held: wallks forward 10 steps	42.26
44—4 point: crawls or hitches forward 6'	43.03
40—4 point: attains sit arms free	44.91
52—On the floor: pulls to STD at large bench	45.09
31—Sit on mat with feet in front: attains 4 point over R side	45.73
42—4 point: reaches forward with R arm, hand above shoulder level	46.44
32—Sit on mat with feet in front: attains 4 point over L side	46.73
43—4 point: reaches forward with L arm, hand above shoulder level	46.91
66—STD, 2 hands on large bench: cruises 5 steps to L	47.62
65—STD, 2 hands on large bench: cruises 5 steps to R	47.68
45—4 point: crawls reciprocally forward 6'	48.03
46—4 point: crawls up four steps on hands and knees/feet	48.38
36—On the floor: attains sit on small bench	48.44
35—STD: attains sit on small bench	49.03
48—Sit on mat: attains high KN using arms, maintains, arms free, 10 sec	49.79
68—STD, 1 hand held: walks forward 10 steps	51.27
37—On the floor: attains sit on large bench	52.80
53—STD: maintains, arms free, 3 sec	53.97
55—STD: holding on to large bench with one hand, lifts L foot, 3 sec	55.74
54—STD: holding on to large bench with one hand, lifts R foot, 3 sec	55.92
69—STD: walks forward 10 steps	55.92
51—High KN: KN walks forward 10 steps, arms free	56.74
56—STD: maintains, arms free, 20 sec	57.21
59—Sit on small bench: attains STD without using arms	57.56
64—STD: picks up object from floor, arms free, returns to stand	57.62
70—STD: walks forward 10 steps, stops, turns 180°, returns	57.74
72—STD: walks forward 10 steps carrying a large object with 2 hands	57.98
78—STD: kicks ball with R foot	60.04
79—STD: kicks ball with L foot	60.62
71—STD: walks backward 10 steps	62.45
63—STD: attains squat, arms free	65.27
62—STD: lowers to sit on floor with control, arms free	65.57
84—STD, holding 1 rail: walks up 4 steps, holding 1 rail, alternating feet	65.98
77—STD: runs 15 feet, stops & returns	67.22
76—STD: steps over stick at knee level, L foot leading	69.75
60—High KN: attains STD through half KN on R knee, without using arms	70.22
75—STD: steps over stick at knee level, R foot leading	70.22
88—STD on 6" step: jumps off, both feet simultaneously	70.81
73—STD: walks forward 10 consecutive steps between parallel lines 8" apart	71.45
61—High KN: attains STD through half KN on L knee, without using arms	71.92
85—STD, holding 1 rail: walks down 4 steps, holding 1 rail, alternating feet	72.16
81—STD: jumps forward 12", both feet simultaneously	73.34
86—STD: walks up 4 steps, alternating feet	74.28
74—STD: walks forward 10 consecutive steps on a straight line ¾" wide	80.22
87—STD: walks down 4 steps, alternating feet	81.22
58—STD: lifts R foot, arms free, 10 sec	85.23
57—STD: lifts L foot, arms free, 10 sec	86.23
83—STD on L foot: hops on L foot 10 times within a 24" circle	88.05
80—STD: jumps 12" high, both feet simultaneously	88.11
82—STD on R foot: hops on R foot 10 times within a 24" circle	88.52

[a] SUP=Supine, PR=prone, R=right, L=left, SIT=sitting, STD=standing, KN=kneeling.
[b] Difficulty values correspond to the GMFM-66 score at which a child is likely to have a score of 3 on the item.

formed to a scale with values ranging from 0 to 100. The GMFM-66 has difficulty estimates ranging from 15.72 for the easiest item (item 21: "sits on mat, supported at thorax by therapist, lifts head upright, maintains 3 seconds") to 88.52 for the most difficult item (item 82: "stands on right foot, hops on right foot 10 times within a 24-inch circle"). The Rasch scale is in logits (log-odds units), which runs from negative infinity to positive infinity. A linear transformation was applied to the ability scale so that the minimum ability was 0 and the maximum ability was 100.[28] In terms of the items on the GMFM-66, a score of 0 on the ability scale means that a child has a 100% probability of having a score of 0 on every item. A score of 100 means that a child has a 100% probability of having a score of 3 on every item. The item difficulty scale was transformed using the same linear transformation. This transformation placed the easiest item at a difficulty of 15.7 and the most difficult item at a difficulty of 88.52. An item's difficulty corresponds to the ability required for a score of 3 on that item. Therefore, for a child to be likely (P=.5) to have a score of 3 on the easiest item, he or she needs to have an ability estimate of 15.7.

Table 2 shows the mean GMFM-66 scores or "ability estimates" for the 537 children included in the Rasch analysis by age, sex, type and distribution of CP, and GMFCS level. The mean ability estimates are similar for males and females. Children with hemiplegia have the highest mean ability estimates, and children with 4-limb involvement have the lowest mean ability estimates. Children who were classified as having "mild" (ie, level I) motor impairment using the GMFCS had a mean ability estimate of 78.06, followed by mean ability estimates of 60.92 for children with level II motor impairment, 49.98 for children with level III motor impairment, 37.94 for children with level IV motor impairment, and 20.63 for children with level V motor impairment.

As hypothesized, there was an overall change in mean GMFM-66 scores for the sample of children assessed at baseline and at 12 months, from 52.76 to 54.61 (F=116.3; df=1,222; P<.0001). There was also a time × severity × age interaction (F=12.6; df=2,222; P<.0001). Figure 1 shows the 3-way interaction plotted as mean change scores by severity and age category. The means and standard deviations by group are reported in Table 5. Figure 1 shows that children under 5 years of age changed more than children aged 5 years and over. This change was greater for children whose GMFCS level was I or II than for children with other GMFCS levels. The mean change for children aged 5 years and over was approximately zero regardless of GMFCS level.

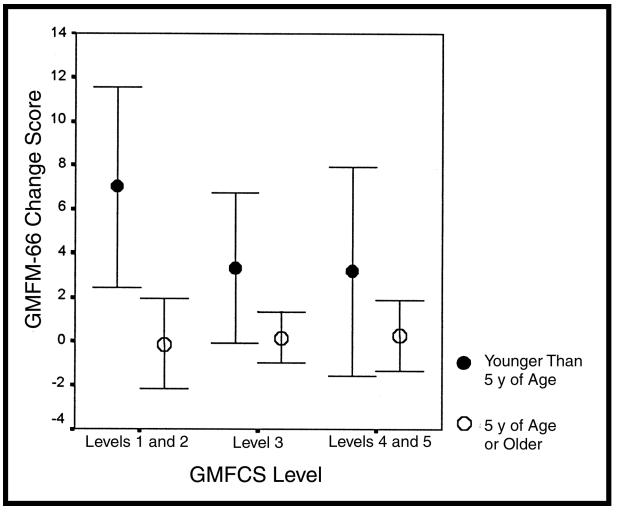

Figure 1.
Mean change (±1 standard deviation) in the 66-item Gross Motor Function Measure (GMFM-66) scores by age and severity of motor disability, as assessed using the Gross Motor Function Classification System (GMFCS).

Discussion

Our study provides evidence of test-retest reliability and construct validity of the GMFM-66 scores. Evidence of the responsiveness of the GMFM-66 includes the findings that the mean GMFM-66 scores of the children with CP changed over 12 months and that the mean change scores were related to age and severity of motor disability.

Test-retest reliability of the GMFM-66 scores was examined using data collected on the 88-item GMFM. The magnitude of the ICCs for the GMFM-88 and for the GMFM-66 was high. The ICCs may have been different had only the items of the GMFM-66 been administered; however, there is no reason to expect that the reliability estimates would have decreased.

Although we believe that the results support the construct validity of the GMFM-66 scores, the measure detected less change in the older children regardless of severity of CP. This finding suggests either that the GMFM-66 is not as sensitive to changes in motor function made by children with CP aged 5 years and over as it is to changes in motor function made by children under 5 years of age or that a smaller mean change is happening in children over age 5 years. This pattern of varying responsiveness for children over 5 years of age is similar to current findings from the models of motor growth using the 88-item GMFM.[36] The emphasis of the GMFM is on motor abilities associated with gross motor development that typically are achieved by age 5 years in children without motor impairments. For children with CP over the age of 5 years, especially those who are unable to walk without assistive mobility devices, change may be associated with performance of motor functions

Table 5.
Mean 66-Item Gross Motor Function Measure (GMFM-66) Baseline and Follow-up Scores by Severity of Motor Disability (Gross Motor Function Classification System[34] [GMFCS] Level) and Age (n=228)

	Younger Than 5 Years of Age					5 Years of Age or Older				
	Baseline		12-mo Follow-up			Baseline		12-mo Follow-up		
GMFCS	$\bar{X}$	SD	$\bar{X}$	SD	N	$\bar{X}$	SD	$\bar{X}$	SD	N
Levels 1 and 2	63.16	13.48	70.16	13.58	39	77.84	12.50	77.72	12.56	57
Level 3	47.59	6.31	50.94	6.50	16	51.62	7.36	51.82	6.88	33
Levels 4 and 5	31.12	9.64	34.31	10.11	24	32.50	11.54	32.81	11.50	59

at home, at school, in the community, and during social participation rather than acquisition of basic gross motor skills. In our view, therefore, measures of disability such as the Pediatric Evaluation Disability Inventory[29] may be more appropriate for evaluating change in older children with CP.

The improvements of the GMFM-66 over the GMFM include (1) the ordering of items according to difficulty, (2) the interval properties of the scale allowing for improved interpretability of a total score and of change scores, and (3) the decrease in administration time. The difficulty estimates for each item on the GMFM-66 (Tab. 4) provide information that is unique to gross motor function of children with CP. Items for lying, rolling, and head control in supported sitting are the least difficult, whereas items for standing on one foot, jumping, and hopping are the most difficult. The difficulty estimates for items of sitting, crawling, kneeling, standing, and walking tend to overlap in the middle of the scale, indicating that motor abilities in these different areas may be developing simultaneously for children with CP. These findings have implications for determining outcomes and planning intervention.

Interpretation of a child's GMFM-66 score should be enhanced by using the item difficulty map (Fig. 2). The item map provides a graphic presentation of items ordered along a continuum of difficulty. As each step represents a single response option (eg, a score of 1, 2, or 3 on a GMFM item), there are 198 steps (66 items $\times$ 3 response options). Because items are now ordered in terms of their difficulty, the GMFM-66 clarifies the next task that a child is likely to accomplish and how difficult it might be to accomplish this task. For example, if a child has a GMFM-66 score of 80 and we look at the item map (Fig. 2), we would be fairly confident that he or she could accomplish all the items to the left of 80 and indicated by a black box (as completing the response option of 3). This includes all items in the lying and rolling, sitting, and crawling and kneeling categories and most items in the standing category (with the exception of items 57 and 58). The items to the right of a GMFM-66 score of 80 are those that the child is unlikely to have accomplished fully. The child would likely be able to accomplish items 74 and 87 (if not, the child would likely have a score of 2 on these 2 items, and these would be the next items that the child will likely complete). Item maps have been used with the Pediatric Evaluation of Disability Inventory[29] and the School Function Assessment[39] to facilitate clinical interpretation of summary scores. A further discussion of how item maps are useful for clinical interpretation of assessment scores can be found in the article by Coster et al.[40] Transferring the GMFM-66 score onto the item map can help in the interpretation of assessment scores by presenting both the child's gross motor abilities and the difficulty of items that have not been achieved. The GMAE program[35] provides a measure of error that is useful in assessing whether an individual's change score is significant.

The interval scale for items of the GMFM-66 also has advantages for interpretation of scores and decision making. The magnitude of change in GMFM-66 scores for individual children or groups of children can be directly compared upon retesting. For example, a child whose GMFM-66 scores increase over 2 consecutive 6-month intervals from 24 to 29 and from 29 to 39 demonstrates 5-point and 10-point changes, respectively. The child's change in GMFM-66 score for the second 6-month interval is twice as much as the change made for the first interval. The interval scale also enables a direct comparison of change among children with different functional abilities. This comparison is particularly important for program evaluation and clinical research.

Profiles of 2 children with CP illustrate differences between the original method of scoring the GMFM and the Rasched interval scale for scoring the GMFM-66 (Fig. 3). John is a 3-year 8-month-old boy with a diagnosis of CP with dystonic movements and primary involvement of the legs. He is classified at level I on the GMFCS. When initially tested, John achieved a GMFM score of 59 and a GMFM-66 score of 54. Upon retesting 6 months later, John achieved a GMFM score of 77 and a GMFM-66 score of 65.

Sara is a 3-year 8-month-old child with a diagnosis of CP (spastic diplegia). She is also classified at level I on the

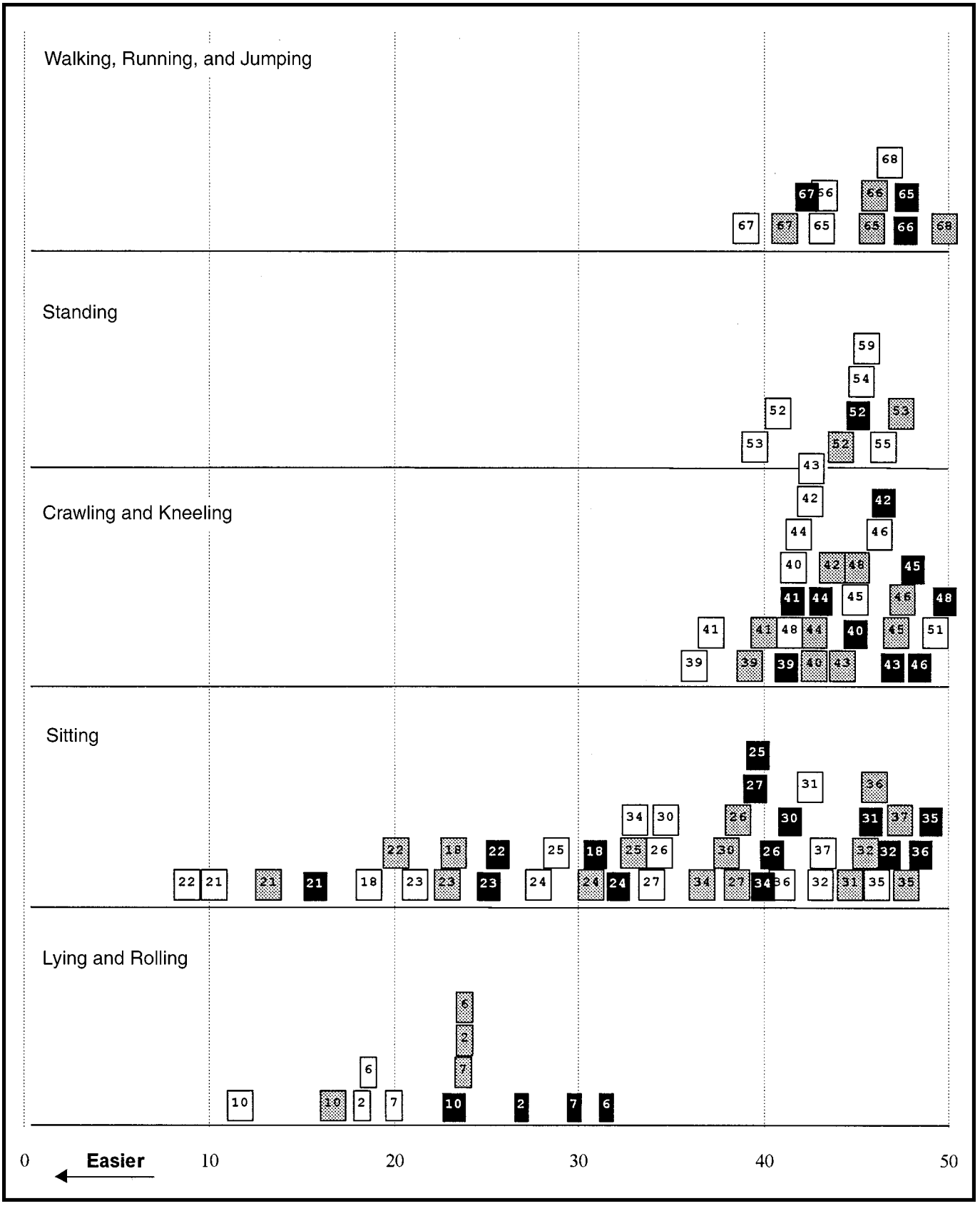

Figure 2.
Item map representing the step difficulty estimates for the 66-item Gross Motor Function Measure (GMFM-66).

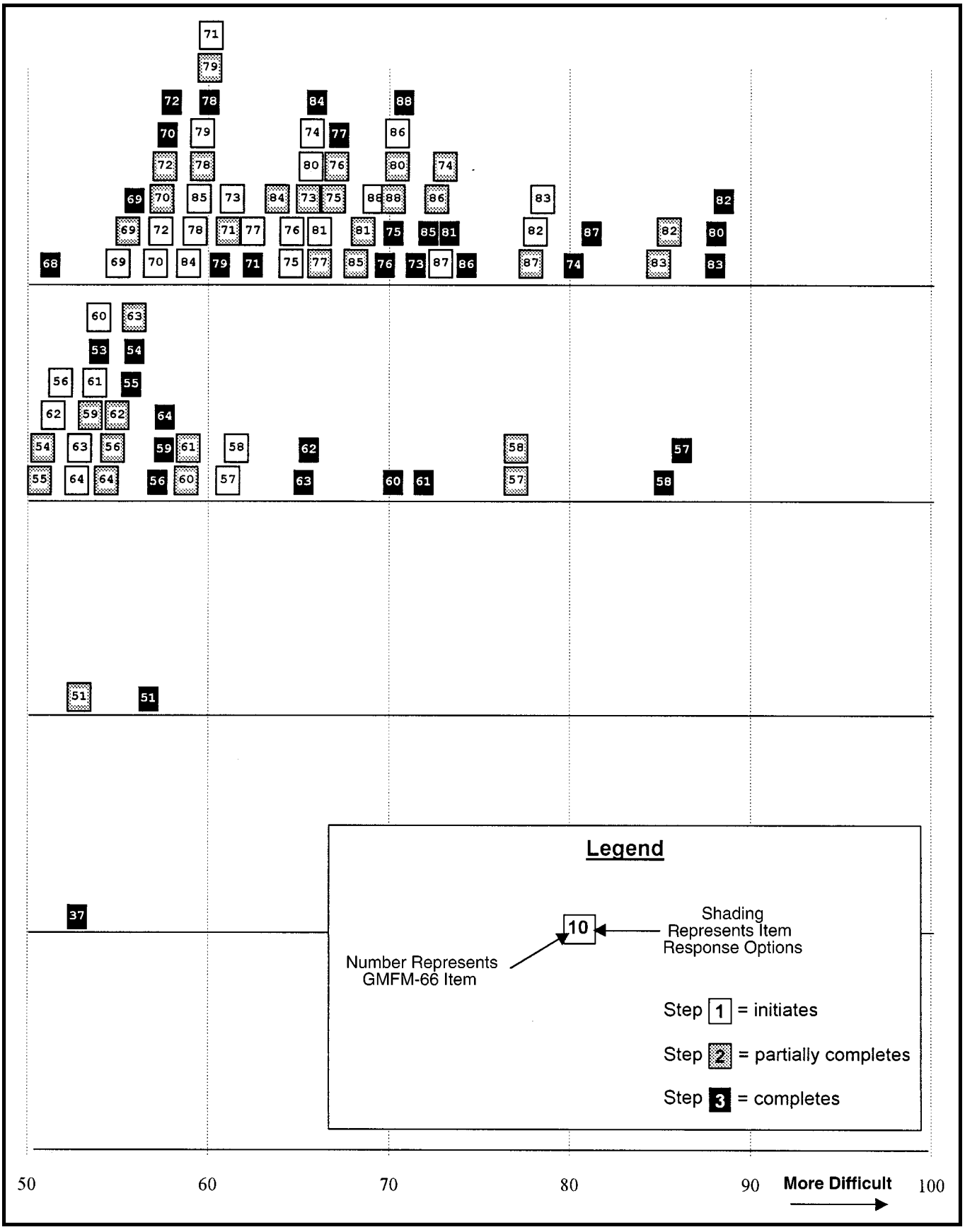

Figure 2.
Continued

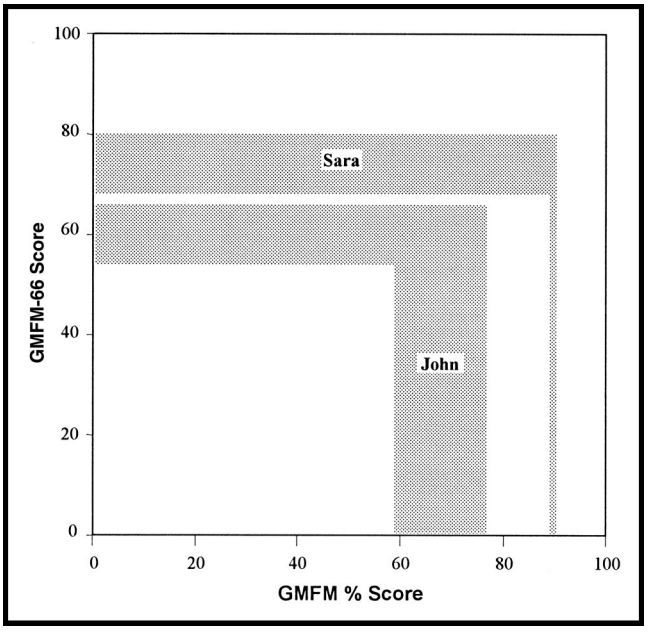

Figure 3.
Plot of 2 children displaying the same amount of change on the 66-item Gross Motor Function Measure (GMFM-66) and different amounts of change on the 88-item Gross Motor Function Measure (GMFM).

GMFCS. When initially tested, Sara achieved a GMFM score of 89 and a GMFM-66 score of 68. Six months later, upon retesting, Sara achieved a GMFM score of 90 and a GMFM-66 score of 80.

Although John demonstrated a change score of 12% and Sara demonstrated a change score of 1% on the GMFM, they had similar change scores on the GMFM-66. Sara's larger change score on the GMFM-66 is attributable to the large spacing between certain item difficulties (Fig. 2). Although Sara did not improve on as many items as John, the items she accomplished were more spread out along the difficulty continuum.

The GMFM-66 should take less time to administer with 22 fewer items and the ability of the GMAE program[35] to estimate a child's ability score even when all items have not been administered. A potential limitation of the GMFM-66 is the need for a computer program to score it. A computer program has been designed to convert the raw scores from either the GMFM-88 or the GMFM-66 into difficulty estimates.[35] It will calculate the GMAE score based on the 66 items, with a standard error. There are no longer separate scores for each dimension. One major difference in the administration and scoring of the GMFM-66 than the scoring of the 88-item GMFM is the importance of differentiating a true score of 0 (child is unable to perform) from an item that was not tested. In the original administration and scoring guidelines, a score of 0 was assigned for items that were not administered or that the child did not perform during the assessment. The new GMFM score form has been modified to include the response "not tested." Although not all 66 items have to be administered to calculate a GMAE score, the more items that are administered, the more accurate the estimate of a child's gross motor ability.[27]

We believe that the GMFM-66 has applications for evaluating the effectiveness of interventions. Three studies,[9–11] for example, have been published recently with conflicting results on the effectiveness of dorsal rhizotomy surgery for children with CP. All 3 studies used common outcome measures, including the GMFM. From discussions among the 3 groups of investigators, several hypotheses arose as to why there may have been differences in the study outcomes, including the possibility that the relative difficulty of items at different intervals of the GMFM is not reflected in the total score.[9] It will be important to determine whether the results of these studies would be different using the GMFM-66, in which the change in score is interval and reflects the difficulty of items, regardless of a child's initial motor ability.

Summary and Conclusions

The advantages of the GMFM-66 include: (1) items are arranged in order of difficulty, (2) the interval properties of the scale allow for improved interpretability of a total score and change scores, (3) a decrease in administration time with 22 fewer items to administer and score, (4) a computer scoring system that allows calculation of a child's total score and the standard error around an individual's score and that can estimate a child's score if some items are missing, and (5) psychometric properties of reliability, validity, and responsiveness. The potential disadvantages of the GMFM-66 include: (1) many items in the lower dimensions have been removed and the GMFM-66, therefore, may be less descriptive for children functioning at low ability levels, (2) the need for the GMAE software in order to score the GMFM-66, and (3) the need to learn to interpret item maps. Recognizing that not all service providers have access to computers for scoring and that service providers use the GMFM for purposes other than measuring change (eg, for descriptive purposes), we have maintained the original 88 items on the GMFM score sheet. In this way, service providers have the option of using the version that best suits their purpose. Because the item difficulties are calibrated for use with children with CP, availability of the 88-item GMFM will allow the measure to continue to be used for clients with diagnoses other than CP.

References

1 Russell DJ, Rosenbaum PL, Cadman DT, et al. The Gross Motor Function Measure: a means to evaluate the effects of physical therapy. *Dev Med Child Neurol.* 1989;31:341–352.

2 Russell DJ, Rosenbaum PL, Gowland C, et al. *Manual for the Gross Motor Function Measure*. 2nd ed. 1993. (Available from *CanChild* Centre for Childhood Disability Research, Room 408, Institute for Applied Health Sciences, McMaster University, 1400 Main St W, Hamilton, Ontario, Canada L8S 1C7.)

3 Bjornson KF, Graubert C, McLaughlin JF, et al. Test-retest reliability of the Gross Motor Function Measure in children with cerebral palsy. *Physical and Occupational Therapy in Pediatrics*. 1998;18(2):51–61.

4 Nordmark E, Hägglund G, Jarnlo GB. Reliability of the Gross Motor Function Measure in cerebral palsy. *Scand J Rehabil Med*. 1997;29:25–28.

5 Bjornson KF, Graubert C, McLaughlin JF, Astley SJ. Inter-rater reliability of the Gross Motor Function Measure. *Dev Med Child Neurol*. 1994;36(SUPpl 70):27–28.

6 Russell DJ, Palisano RJ, Walter S, et al. Evaluating motor function in children with Down syndrome: validity of the GMFM. *Dev Med Child Neurol*. 1998;40:693–701.

7 Bjornson KF, Graubert C, Buford V, McLaughlin JF. Validity of the Gross Motor Function Measure. *Pediatric Physical Therapy*. 1998; 10(2):43–47.

8 Kolobe TH, Palisano RJ, Stratford PW. Comparison of two outcome measures for infants with cerebral palsy and infants with motor delay. *Phys Ther*. 1998;78:1062–1072.

9 McLaughlin JF, Bjornson KF, Astley SJ, et al. Selective dorsal rhizotomy: efficacy and safety in an investigator-masked randomized clinical trial. *Dev Med Child Neurol*. 1998;40:220–232.

10 Steinbok P, Reiner AM, Beauchamp R, et al. A randomized clinical trial to compare selective dorsal rhizotomy plus physiotherapy with physiotherapy alone in children with spastic diplegic cerebral palsy. *Dev Med Child Neurol*. 1997;39:178–184.

11 Wright V, Sheil E, Drake J, Wedge J. Evaluation of selective dorsal rhizotomy for the reduction of spasticity in cerebral palsy: a randomized controlled trial. *Dev Med Child Neurol*. 1998;40:239–247.

12 Krach L, Gilmartin R, Bruce D, et al. Functional changes noted following treatment of individuals with cerebral palsy with intrathecal baclofen. *Dev Med Child Neurol*. 1997;39(suppl 75):12–13.

13 Almeida GL, Campbell SK, Girolami GL, et al. Multidimensional assessment of motor function in a child with cerebral palsy following intrathecal administration of baclofen. *Phys Ther*. 1997;77:751–764.

14 Bower E, McLellan DL, Arney J, Campbell MJ. A randomised controlled trial of different intensities of physiotherapy and different goal-setting procedures in 44 children with cerebral palsy. *Dev Med Child Neurol*. 1996;38:226–237.

15 Bower E, McLellan DL. Effect of increased exposure to physiotherapy on skill acquisition of children with cerebral palsy. *Dev Med Child Neurol*. 1992;34:25–39.

16 MacKinnon J, Noh S, Lariviere J, et al. A study of therapeutic effects of horseback riding for children with cerebral palsy. *Physical and Occupational Therapy in Pediatrics*. 1995;15(1):17–31.

17 Steinbok P, Reiner A, Kestle JR. Therapeutic electrical stimulation following selective dorsal rhizotomy in children with spastic diplegic cerebral palsy: a randomized clinical trial. *Dev Med Child Neurol*. 1997;39:515–520.

18 Wright V, Belbin G, Slac M, et al. A pilot evaluation of the David Hart Walker Orthosis (DHWO): a new assistive device for children with cerebral palsy (CP). *Dev Med Child Neurol*. 1997;39(suppl 75):35.

19 Evans C, Gowland C, Rosenbaum PL, et al. The effectiveness of orthoses for children with cerebral palsy. *Dev Med Child Neurol*. 1994; 36(suppl 70):26.

20 MacPhail A, Kramer JF. Effect of isokinetic strength training on functional ability and walking efficiency in adolescents with cerebral palsy. *Dev Med Child Neurol*. 1995;37:763–775.

21 Kramer JF, MacPhail A. Relationships among measures of walking efficiency, gross motor ability, and isokinetic strength in adolescents with cerebral palsy. *Pediatric Physical Therapy*. 1994;6:3–8.

22 Abel MF, Damiano DL, Pannunzio M, Bush J. Role of multiple muscle-tendon recessions and releases to improve motor function in diplegic cerebral palsy. *Dev Med Child Neurol*. 1997;39(suppl 75):16–17.

23 Damiano DL, Abel MF. Relation of gait analysis to gross motor function in cerebral palsy. *Dev Med Child Neurol*. 1996;38:389–396.

24 Harris T, Damiano DL, Abel M. Gait efficiency in diplegic cerebral palsy. *Dev Med Child Neurol*. 1997;39(suppl 75):28–29.

25 Drouin LM, Malouin F, Richards CL, Marcoux S. Correlation between the GMFM scores and gait spatiotemporal measures in children with neurological impairments. *Dev Med Child Neurol*. 1996;38:1007–1019.

26 Parker DF, Carriere L, Hebestreit H, et al. Muscle performance and gross motor function of children with spastic cerebral palsy. *Dev Med Child Neurol*. 1993;35:17–23.

27 Avery LM, Russell DJ, Raina PS, et al. Rasch analysis of an established health outcome measure: a case study of the Gross Motor Function Measure (GMFM). Manuscript submitted for publication.

28 Wright B, Masters G. *Rating Scale Analysis*. Chicago, Ill: MESA Press; 1980.

29 Haley SM, Coster W, Ludlow L, et al. *Pediatric Evaluation of Disability Inventory (PEDI): Version 1.0*. Boston, Mass: New England Medical Centre Hospitals Inc; 1992.

30 Campbell SK, Osten E, Kolobe TH, Fisher A. Development of the Test of Infant Motor Performance. *Phys Med Rehabil Clin N Am*. 1993;4:541–550.

31 Fisher W, Fisher A. Application of Rasch analysis to studies in occupational therapy. *Phys Med Rehabil Clin N Am*. 1993;4:551–569.

32 Haley SM, McHorney CA, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10), I: unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol*. 1997;47:671–684.

33 McHorney CA, Haley SM, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10), II: comparison of relative precision using Likert and Rasch scoring methods. *J Clin Epidemiol*. 1997; 50:451–461.

34 Palisano RJ, Rosenbaum PL, Walter S, et al. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997;39:214–223.

35 Burrows L. *Gross Motor Ability Estimator* [software]. Hamilton, Ontario, Canada: *CanChild* Centre for Childhood Disability Research, McMaster University; 1999. (Software available from the *CanChild* website at: http://www.fhs.mcmaster.ca/canchild).

36 Palisano RJ, Hanna SE, Rosenbaum PL, et al. Validation of a model of gross motor function for children with cerebral palsy. *Phys Ther*. In press.

37 Shrout PE, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428.

38 Olkin E. Correlations revisited. In: Stanley JC, ed. *Improving Experimental Design and Statistical Analysis*. Chicago, Ill: Rand McNally; 1967.

39 Coster W, Beeney T, Haltiwinger J, Haley SM. *School Function Assessment*. San Antonio, Tex: The Psychological Corporation/Therapy Skill Builders; 1998.

40 Coster W, Ludlow L, Mancini M. Using IRT variable maps to enrich understanding of rehabilitation data. *J Outcome Meas*. 1999;3:23–133.