



Contents lists available at ScienceDirect

Research in Developmental Disabilities



ACTIVLIM-CP a new Rasch-built measure of global activity performance for children with cerebral palsy

Yannick Bleyenheuft (PhD)^{a,*}, Julie Paradis (OT)^{a,b}, Anne Renders (MD)^c,
Jean-Louis Thonnard (PhD)^{a,c}, Carlyne Arnould (PhD)^d

^a Institute of NeuroScience, Université catholique de Louvain, Brussels, Belgium

^b Service l'escalier, CHU Lyon, Lyon, France

^c Cliniques universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium

^d Haute Ecole Louvain en Hainaut, Physical and Occupational Therapy Department, Paramedical Category, Montignies-sur-Sambre, Belgium

ARTICLE INFO

Article history:

Received 2 April 2016

Received in revised form

13 September 2016

Accepted 13 October 2016

Available online xxx

Number of reviews completed is 2

Keywords:

Cerebral palsy

Activity

Upper extremity

Lower extremity + Rasch-model (outside

Mesh)

ABSTRACT

Objective: Children with cerebral palsy (CP) often have upper extremity (UE) and lower extremity (LE) impairments. While tools measuring separately UE and LE abilities are currently used, activities in which UE and LE are used in combination – numerous in everyday life – cannot be assessed because no instrument allows capturing global activity performance in children with CP. This study aimed to develop a clinical tool for measuring their global activity performance using the Rasch model.

Study design: The caregivers of 226 children with CP (2–18 years old) answered a 154-item experimental questionnaire. Within 4–6 weeks, 129 of them filled in the questionnaire a second time. Responses were analyzed using the Rasch RUMM2020 software.

Results: The final 43 item scale presented a high reliability ($R=0.98$) and reproducibility ($R=0.97$). The item difficulty hierarchy was consistent over time and did not vary according to age, gender, or clinical form, allowing the follow-up of children from 2 to 18 years old.

Conclusions: ACTIVLIM-CP is a unidimensional scale specifically developed to measure global activity performance in children with CP providing a reliable tool to follow children's evolution and document changes related to neurorehabilitation, especially where a combination of UE and LE is targeted. Its responsiveness is still to be tested.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

What's known on this subject

Tools for measuring performance of children with CP in activities involving both upper (UE) and lower extremities (LE) are missing. As a UE/LE combination is needed in everyday life and is increasingly stimulated in therapy, a reliable tool is required.

Abbreviations: CP, cerebral palsy; LE, lower extremity; UE, upper extremity; MACS, manual ability classification system; GMFCS, gross motor function classification system; PED1, pediatric evaluation of disability inventory.

* Corresponding author at: Institute of NeuroScience, Université catholique de Louvain, avenue Mounier 53, 1200 Brussels, Belgium.

E-mail address: yannick.bleyenheuft@uclouvain.be (Y. Bleyenheuft).

<http://dx.doi.org/10.1016/j.ridd.2016.10.005>

0891-4222/© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Please cite this article in press as: Bleyenheuft, Y., et al. ACTIVLIM-CP a new Rasch-built measure of global activity performance for children with cerebral palsy. *Research in Developmental Disabilities* (2016), <http://dx.doi.org/10.1016/j.ridd.2016.10.005>

What this study adds

- a unidimensional scale measuring global activity performance in children with CP;
- a reliable tool to document changes in upper and lower extremities;
- allowing a follow-up from 2 to 18 years old;
- validated for all clinical forms of CP.

1. Introduction

As cerebral palsy (CP) represents the most frequent cause of physical disability in children, affecting 2,9 per 1000 children (Durkin, Benedict, Christensen, Dubois, & Fitzgerald, 2016), a considerable attention has been paid over the last decade to the improvement of activity level in children with CP (Novak et al., 2013). Activity has been defined in the International Classification of Functioning, Disability and Health (ICF) model as the execution of a task or an action by an individual and includes the different sub-categories of mobility (d4, lifting & carrying objects, fine hand used, walking, moving around using equipment, using transportation, driving bi/tricycle or car), self-care (d5, washing oneself, caring for bodyparts, toileting, dressing, eating, drinking, looking after one's health) and domestic life (d6, acquisition of goods and services, preparation of meals, doing housework, assisting others) (WHO, 2016).

Measuring activity performance is thus a key point in determining the impact of functional therapies/intervention in these children. However, activity performance cannot be measured directly, but requires to be inferred either from direct observations, either from caregivers' or patients' perceptions of the difficulty in performing the activities (through questionnaires) (Penta, Tesio, Arnould, Zancan, & Thonnard, 2001). In a pediatric context, the use of caregivers as valid respondents to questionnaires is suggested as they discriminate more precisely the different difficulty levels of children's activities while children themselves tend to be more dichotomous (Arnould, Penta, Renders, & Thonnard, 2004; Vandervelde, Van den Bergh, Goemans, & Thonnard, 2007). Children with CP often have upper extremity (UE), lower extremity (LE) as well as trunk/postural disorders affecting their activity level and likely their autonomy. These deficits, depending on their location and severity are leading to different subtypes of CP (i.e. unilateral CP or bilateral CP with variable impairments).

Questionnaires used to measure activity performance in the field of neurorehabilitation (Chong, Mackey, Broadbent, & Stott, 2011; Gilmore, Sakzewski, & Boyd, 2009), even when providing evidences for good psychometric qualities, are always focused selectively on either UE (e.g. ABILHAND-Kids; Arnould et al., 2004, the Children's Hand-use Experience Questionnaire-CHEQ; Sköld, Hermansson, Krumlinde-Sundholm, & Eliasson, 2011) or LE ability (e.g. ABILOCO-Kids; Caty, Arnould, Thonnard, & Lejeune, 2008). However, many activities of daily life involve upper and lower extremities in a combined use (e.g. undressing), which also imply an involvement of postural/trunk control. The performance/ability of each body part (UE or LE) may be influenced by performance changes in the other, as demonstrated by the unexpected changes on LE reported during intensive UE training (Coker, Karakostas, Dodds, & Hsiang, 2010; Gillick, 2012; Gillick & Koppes, 2010; Zipp & Winning, 2012). For instance, a child could become able to walk with a walker after intensive intervention of upper extremities because his grasp, his elbow or shoulder strength or amplitude, or even his trunk, have evolved in such a way that he can use the walker in a safe or more efficient way. Opening a tap, typically considered as UE activity, might also become possible after intensive LE training because the child is able to stand, maintain his trunk in a static position, and thus reach the tap. Some global activities, like going in and out of a bath, picking something in the fridge or going in and out of a car, require a combined use of UE, LE and postural control and the children's difficulty changes in these activities cannot be captured by either separate UE or LE measurements. Therefore, in a clinical context where UE, LE and postural control are targeted together (e.g. Hand and Arm Bimanual Intensive Therapy Including Lower Extremities—HABIT-ILE (Bleyenheuft & Gordon, 2014; Bleyenheuft, Arnould, Brandao, Bleyenheuft, & Gordon, 2015) one cannot dissociate activity performance changes in UE and LE due to their global, combined, and interrelated functioning. It seems thus relevant to use a specific tool measuring the global activity performance of children with CP on a unidimensional scale. A recent review of Harvey et al. has highlighted that no instrument was specifically designed for measuring global activity performance in the daily life of children with CP (Harvey, Robin, Morris, & Baker, 2008). A few instruments, including the Children's Assessment of Participation and Enjoyment (CAPE), the Pediatric Evaluation of Disability Inventory (PEDI), the Activities Scale for Kids (ASK), and the School Function Assessment (SFA) were developed to capture more global functioning. Overall these scales were originally developed as generic measures, which is likely to induce a risk of error in the measurement tool since the difficulty of activities is diagnosis-dependent (Arnould, Vandervelde, Batcho, Penta, & Thonnard, 2012). It is thus risky to use generic scales for a given population (here children with CP) without prior investigation of the items calibration invariance between different diagnostic groups (Arnould et al., 2012). Moreover, the CAPE is clearly focused on participation and not on activity. The PEDI is focused on capacity (can or cannot do) and not on performance and lacks of precision due to its dichotomic response scale. The ASK is a well-designed tool that measures capacity and performance based on how often activities are performed. It is thus measuring another variable than the one captured by the ACTVLIM-CP (ability/facility to manage the activities). Finally though the SFA is helpful in the context of school, its subscale based on activity performance would not necessarily allow to capture this variable in another context.

A unidimensional specific tool measuring global activity performance, the ACTIVLIM, has already been developed in stroke patients (Batcho, Tennant, & Thonnard, 2012), as well as in children and adults with neuromuscular diseases (Arnould et al., 2004), for who this questionnaire is considered as one of the four best instruments to measure activity level (Seedat, James,

& Rose, 2014). However, the difficulty of the activities calibrated for children with neuromuscular diseases cannot be used in children with CP since the difficulty of an item is diagnosis-dependent (Arnould et al., 2012). Therefore a new calibration, specific to children with CP, leading to a new scale is needed.

The aim of this work was to develop the ACTIVLIM-CP, a global activity performance scale for children with CP, using the Rasch model (Rasch, 1980), in order to provide a global activity performance measure on a linear and unidimensional scale that has the potential to follow children's activities changes across time.

2. Methods

Children were recruited from Centers dedicated to CP in Belgium and in France. On the basis of centre referrals, parents were contacted by mail. Parents gave their written informed consent and filled in the questionnaires. Two hundred and twenty six children with CP (2–20 years old) were considered in this study. This study was conducted under an authorization of the ethical committees of the University clinics of the Université catholique de Louvain, Belgium. A specific agreement was introduced in France for including French participants.

2.1. Questionnaire development

The ACTIVLIM-CP questionnaire was intended to cover the widest possible range of daily activities either involving UE, either LE or combining both. The preliminary questionnaire consisted in 148 items coming from the preliminary ACTIVLIM for neuromuscular diseases (Vandervelde et al., 2007). This pool of 148 items was submitted to 10 experts in the field (1 physician, 6 physiotherapists, 3 occupational therapists) who were asked to provide an opinion on the pertinence of the items and to propose relevant activities not already included. An activity was suppressed if more than 3 experts considered it as irrelevant. Following the first expert analysis, 48 items were suppressed (not relevant or redundant), forty-nine items were proposed for targeting young children, five other items were proposed, providing the 154-item experimental version of the ACTIVLIM-CP. This version was submitted to the experts a second time for approval. No more items were added or discarded following the second round.

2.2. Instrument

The experimental version of the questionnaire explored daily activities when completed without technical or human assistance. The rationale for this choice was to measure the global activity performance in an autonomous way, avoiding a potential bias in the assessment of 2 children not having access to the same material/help. The caregivers were asked to provide their perception of the child's difficulty to achieve the activities on a three-level scale: impossible (0), difficult (1), or easy (2). Activities that were unfamiliar or not performed in the last 3 months were considered as missing responses (4,5% of the data).

2.3. Procedures

The 154 items of the experimental ACTIVLIM-CP (French version) were randomly presented (5 different orders) to the respondents, aiming to avoid a potential bias caused by the item order. Respondents were concomitantly asked to fill in 3 other questionnaires: the ABILHAND-Kids, the ABILOCO-Kids and the PEDI (self-care section), for details see Supplementary material.

Respondents were usually the parents, but in case children attended boarding schools, questionnaires were filled by caregivers (usually taking care of the child). The total time for completing the questionnaires was around 25 min. Respondents were not asked to complete them in a certain order. However, in the mail provided to parents/caregivers, the order of the 3 first questionnaires sharing the same response format were presented first in a random order, while the PEDI, with a different response structure was presented last. The questionnaires were mailed to the families/respondents.

2.4. Data analysis

2.4.1. Rasch model

Caregivers' responses to the experimental version were analyzed using the Rasch model with the Rasch Unidimensional Measurement Models computer program 2020 (RUMM2020). This model allows using the responses given to each item in a probabilistic framework to estimate on a common linear scale 1) the patient performance level, 2) the item difficulty, and 3) the item thresholds (i.e. the locations along the scale at which two successive categories are equally likely to be observed).

2.4.2. Item selection

From the 154 items of the experimental version, successive analyses were used to select items presenting good psychometric qualities to constitute the final ACTIVLIM-CP scale:

- 1) *Relevant items*: an item is considered as relevant when it belongs to the daily reality of most children. Items with a response rate <80% were removed, as these activities are not commonly achieved in the sample of children with CP.
- 2) *Ordered rating scale*: Perceptions of the parents were reported on a three-level scale: impossible (0), difficult (1), or easy (2). A good discrimination of these different categories is highlighted by increasing levels of performance as represented by thresholds presenting the expected order (Andrich, 1998). Items presenting threshold disorders were removed from the scale.
- 3) *Rating scale model*: The use of a rating scale model (i.e., the relative threshold locations –relative to the item location– of all items are fixed) was preferred as it allows an easy clinical interpretation of the scores. However, the use of a rating scale requires that all items share the same discrimination concerning their relative threshold locations (Linacre, 2000; Wright, 1999). Items with relative threshold locations significantly different from the average (Z-test) were removed from the scale.
- 4) *Unidimensionality*: Unidimensionality implies that no attribute of the person besides global activity performance is theorized to account for the probability of choosing a given response to a given item. The similarity between the observed and expected responses was investigated using a standardized residuals and a χ^2 fit statistic computed over 5 class intervals (CI) of increasing performance levels. Items presenting either standardized residuals values below -2.5 or above 2.5 or a p-value of the χ^2 fit statistic <0.05 were deleted from the scale (Andrich & Sheridan, 2005). However, some simulation studies have shown that good fit statistics may be reported when the scales are multidimensional. A Principal Component Analysis (PCA) was thus performed on the residuals. The PCA identifies the first residual factor that best explains observed responses variations not attributable to the Rasch global activity performance factor. Independent *t*-tests were used to compare the estimates for each subject (Smith, Conrad, Chang, & Piazza, 2002), deriving from the highest positive and negative loadings items (correlated at ≥ 0.3 with the first residual factor). The scale is considered as unidimensional when the percentage of tests outside the range ± 1.96 is less than 5%.
- 5) *No differential item functioning* (DIF): Unidimensionality also requires that patients with identical global activity performance but different demographic or clinical characteristics have the same probability of succeeding any particular item. If this is not the case, the item is biased or presents “differential functioning”. The invariance of the item difficulty hierarchy was tested for age (2–6 years old vs 7–11 years old, vs 12–20 years old), gender, and clinical form of CP (hemiplegia, diplegia, quadriplegia). DIF was measured by computing a two-way analysis of variance (ANOVA) on the standardized residuals of the different CIs; the first factor was the investigated patient characteristics (age, gender, clinical form of CP) and the second factor was the CIs (Pallant & Tennant, 2007). Items presenting DIF across age, gender, or clinical form of CP (highlighted by a significant main effect for the first patient factor) were removed from the scale.
- 6) *Local independency*: local dependency is a phenomenon due to a high correlation between some items of the same scale, affecting the measurement because it inflates the scale in one particular direction (Embretson & Reise, 2000). Residual correlations between two items higher than 0.3 led to the suppression of the item with the worse psychometric qualities.

2.5.1. Item-patient targeting

The comparison of the mean patient location with the mean item threshold difficulty (arbitrary set at 0 logit) was performed to verify that the difficulty of the scale was globally adapted to the focused sample. Item-patient targeting was also analyzed by comparing histograms frequency (top and lower panel of Fig. 1) to detect potential gaps in the scale.

2.5.2. Scale reliability

The degree of precision achieved in the measurement (i.e. the reliability of internal consistency) was reported by a Person Separation Index. This index was computed as the ratio between the true and the observed measure variances. The Person Separation Index allows the number of global activity performance levels that may be statistically distinguished in the sample to be calculated.

Moreover, the test-retest reliability was examined in a subsample of 129 caregivers who answered a second time to the experimental questionnaire within 4–6 weeks. Children's global activity performance measures and the item difficulty hierarchy obtained at the first and the second assessments were compared. To put the children's measures on the same scale for both assessments, the adjustment of the origin of both calibrations was obtained by anchoring the items/thresholds of the second calibration at the difficulty level of the first one. The test-retest reliability of the caregivers' responses was determined by an intra-class coefficient (ICC, 2 ways random model, absolute agreement, SPSS).

2.5.3. Construct validity

As criterion validity was difficult to test because no other scale can be considered as a gold-standard to measure global activity performance in children with CP, two other forms of construct validity were investigated: convergent and divergent validity. Divergent validity allows investigating whether constructs that are not supposed to relate to the variable are actually not related to the variable measured (e.g. age, gender). The relationship between ACTIVLIM-CP and patient characteristics was investigated using a Pearson correlation coefficient (age), and a *t*-test (gender). Convergent validity investigates whether constructs that are expected to relate to the variable of interest are actually related to variable of interest (e.g. topography, UE and LE scales). The relationship between ACTIVLIM-CP and topography was investigated by a one-way ANOVA on ranks (Kruskal-Wallis). Convergent validity was investigated by measuring the degree of association (Pearson correlation coefficient) between ACTIVLIM-CP and existing scales of UE activity (ABILHAND-Kids) (Arnould et al.,

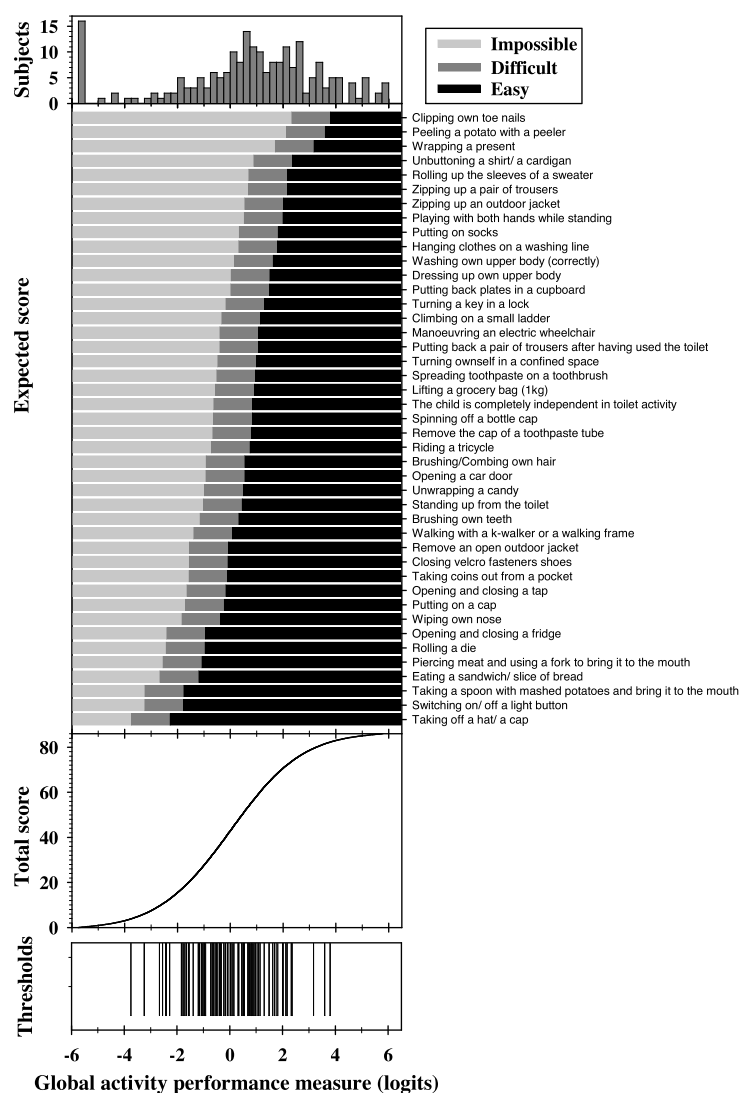


Fig. 1. Top panel: Distribution of global activity performance measures in the whole sample. Twenty-three children with extreme scores could not be assessed by the scale since all activities were either impossible ($n = 16$) or easy ($n = 7$). Middle panel/horizontal bars: children's expected response to each item related to the global activity performance measure. The average item difficulty is conventionally set at 0 logit, items being ordered from the most (top) to the less (bottom) difficult. A child with a measure of 0 logit is expected to succeed easily in the 13 first activities, to be able to perform the 17 following with difficulties and to be unable to perform the 13 most difficult. Children with a measure above 2.33 logits should be able to perform all activities easily or with some difficulties. Middle panel/ogive curve: this curve shows the relationship between the total raw score and the linear ACTIVLIM-CP measure, allowing conversion from one to the other. The range of the total score is 0–86 since the 43 items could be scored as 0,1 or 2. Lower panel: distribution of the item thresholds (86 in total – 2 by item).

2004), LE activity (ABILOCO-Kids) (Caty et al., 2008), capacity of children in self-care (PEDI- self-care section) (Feldman, Haley, & Coryell, 1990).

3. Results

3.1. Sample characteristics

From the 592 questionnaires that were sent to caregivers of children with CP, 226 (38%) were filled by the caregivers and sent back. Among them, 214 questionnaires (36%) were correctly filled and were subsequently used in the development of the scale, including 120 boys (54%). Characteristics of the sample were: mean age 10.3 YO (2–6: 25%; 7–11: 34%; 12–20: 41%), GMFCS (RUMM2020) level from I to V (I = 35.5%, II = 22%, III = 18.5%, IV = 9%, V = 15%), MACS²⁴ level from I to V (I = 31%, II = 32%, III = 15.5%, IV = 11.5%, V = 10%), topography: unilateral CP 47%, diplegia 18%, quadriplegia 30%, not available 5%.

3.2. Item selection of the ACTIVLIM-CP scale

The successive analyses of the Rasch model provided a selection of 43 items. From the 154 items presented at the start, 2 items did not fulfill the response rate criteria (<80%) but were kept due to their clinical relevance (driving electric wheelchair

Table 1
ACTIVLIM-CP calibration.

Items	Difficulty (logits)	SE (logits)	Residual (z)	Fit (χ^2)	DF	p-value
Clipping own toe nails	3.07	0.17	−0.07	6.28	4	0.18
Peeling a potato with a peeler	2.86	0.17	−0.26	4.54	4	0.34
Wrapping a present	2.44	0.15	0.12	2.22	4	0.69
Unbuttoning a shirt/a cardigan	1.63	0.14	0.06	5.40	4	0.25
Rolling up the sleeves of a sweater	1.43	0.14	−0.88	2.01	4	0.73
Zippering up a pair of trousers	1.42	0.14	−1.10	2.28	4	0.68
Zippering up an outdoor jacket	1.28	0.13	−0.20	1.89	4	0.76
Playing with both hands while standing	1.26	0.14	2.26	1.64	4	0.80
Putting on socks	1.08	0.13	−0.92	4.56	4	0.34
Hanging clothes on a washing line	1.05	0.15	−0.45	4.75	4	0.31
Washing own upper body (correctly)	0.89	0.14	0.17	3.50	4	0.48
Dressing up own upper body	0.76	0.13	−0.43	3.60	4	0.46
Putting back plates in a cupboard	0.75	0.14	−1.23	6.71	4	0.15
Turning a key in a lock	0.57	0.14	−0.26	3.92	4	0.42
Climbing on a small ladder	0.41	0.14	1.06	2.36	4	0.67
Manoeuvring an electric wheelchair	0.34	0.18	1.93	7.24	4	0.12
Putting back a pair of trousers after having used the toilet	0.34	0.14	−1.70	9.49	4	0.05
Turning oneself in a confined space	0.25	0.15	1.11	1.61	4	0.81
Spreading toothpaste on a toothbrush	0.23	0.14	−0.96	2.84	4	0.58
Lifting a grocery bag (1 kg)	0.18	0.14	−0.39	2.58	4	0.63
The child is completely independent in toilet activity (dressing up, wiping, flushing the toilet)	0.11	0.14	−0.98	4.33	4	0.36
Spinning off a bottle cap	0.10	0.14	−0.63	2.75	4	0.60
Remove the cap of a toothpaste tube	0.06	0.14	−1.65	4.30	4	0.37
Riding a tricycle	0.01	0.15	2.22	3.60	4	0.46
Brushing/Combing own hair	−0.18	0.15	−1.28	4.56	4	0.34
Opening a car door	−0.19	0.15	−0.52	3.85	4	0.43
Unwrapping a candy	−0.24	0.14	−1.22	1.65	4	0.80
Standing up from the toilet	−0.29	0.14	0.61	2.66	4	0.62
Brushing own teeth	−0.41	0.15	−0.59	1.63	4	0.80
Walking with a k-walker or a walking frame	−0.65	0.19	0.53	2.75	4	0.60
Remove an open outdoor jacket	−0.81	0.16	−1.10	1.72	4	0.79
Closing velcro fasteners shoes	−0.82	0.16	−1.37	3.84	4	0.43
Taking coins out from a pocket	−0.83	0.16	−1.49	1.36	4	0.85
Opening and closing a tap	−0.91	0.16	−1.66	8.39	4	0.08
Putting on a cap	−0.97	0.16	1.92	2.35	4	0.67
Wiping own nose	−1.10	0.16	−1.05	6.21	4	0.18
Opening and closing a fridge	−1.67	0.18	−1.32	3.79	4	0.44
Rolling a die	−1.69	0.18	0.16	3.42	4	0.49
Piercing meat and using a fork to bring it to the mouth	−1.81	0.18	−0.62	3.34	4	0.50
Eating a sandwich/slice of bread	−1.93	0.19	−0.30	1.80	4	0.77
Taking a spoon with mashed potatoes and bring it to the mouth (without dropping it)	−2.50	0.21	−0.63	1.17	4	0.88
Switching on/off a light button (general switch, power switch of computer or television)	−2.51	0.21	−0.53	3.44	4	0.49
Taking off a hat/a cap	−3.02	0.24	0.03	1.01	4	0.91

DF = degrees of freedom.

and walking with a walker); 13 items presented a threshold disorder; 2 items did not share the same discrimination of the response categories; 54 items did not fit the model; and 38 items presented a DIF. Finally, 4 items were deleted because they presented a local dependency with other items (correlation > 0.3).

3.3. Metric properties

Table 1 displays the calibration of the final 43-item ACTIVLIM-CP scale. From top to bottom, items are ordered in a decreasing difficulty level, with higher values representing more difficult items. Values are expressed in logit (natural logarithm of the odds of successful achievement by a patient for any item). Table 1 presents also the standard error (SE) associated with each item, the standard residual and the fit statistic (χ^2 and associated probability). The scale presented a mean Chi-square probability of 0.84 indicating that all 43 items contribute to the definition of a unidimensional measure of global activity performance. This is confirmed by the PCA results showing that the percentage of individual *t*-tests outside the range ± 1.96

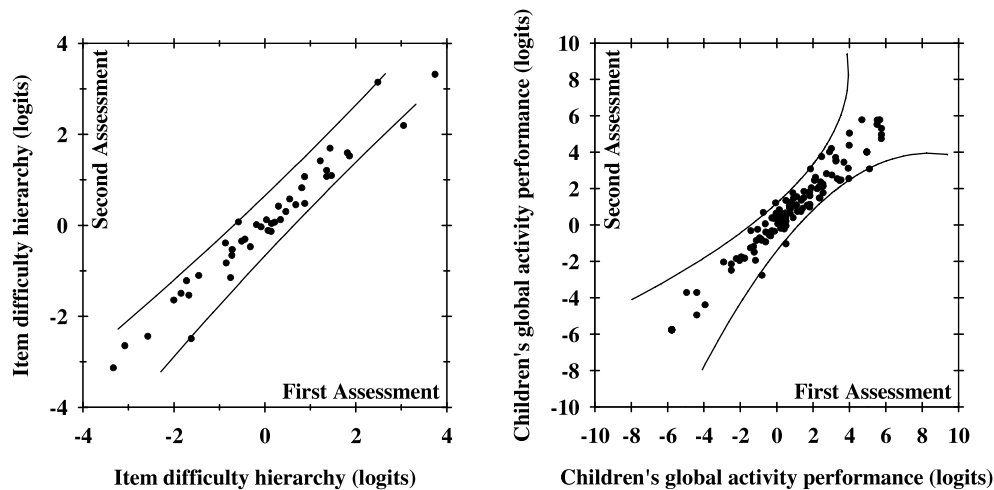


Fig. 2. Left Panel: Item difficulty perceived by the caregivers at first and second assessments in the subsample ($n = 129$) that answered twice the questionnaires within 4–6 weeks, and the 95% of confidence interval (solid lines) of the ideal invariance. Items requiring the highest performance levels to be performed are plotted in the top right part of the panel. All items (dots) lying within the control lines have the same estimated difficulties at the first and second assessments. Right Panel: relationship between children's locations at first and second assessments and the 95% confidence interval (solid lines) of the ideal invariance. More performant children are plotted in the top right part of the panel. Children's measures (dots) included in the confidence interval have the same estimated performance levels at the first and second assessments.

(95% confidence interval (CI)) was 3.96%, showing that the responses variations not attributable to the children's global activity performance were not sufficient to threat the unidimensionality of the scale. The final scale includes 7 items that are common with the ABILHAND-Kids scale and 1 item common with the ABILOCO-Kids scale.

3.4. Description of ACTIVLIM-CP

The final ACTIVLIM-CP scale includes 43 items. In order to characterize the involvement of upper extremities, lower extremities and trunk in the different items included in the scale, a group of 13 experts (PTs and OTs) was solicited. They defined that 5% of the items were focused on LE only (not UE), 16% of the items required standing and using stand & use the UE, 9% of the items required the combined use of UE & LE but not in a standing position, 28% of the items were usually performed with UE & LE standing but could be managed in a wheelchair. This leads to a total of 58% involving LE, while 42% of the items were focused solely on the UE. All items required trunk use: 67% of them in static conditions; 33% of them in dynamic conditions. The 43 items were mostly related to self-care and mobility, while a few of them were related to domestic life. It is also of interest to note that though the majority of the activities could be considered as activities of daily living (ADL), a few of them could be considered as instrumental activities of daily living (i.e. activities necessary for adaptation to the environment like cooking, shopping, housekeeping and transportation; Roehrig, Hoeffken, Pientka, & Wedding, 2007).

The item-patient targeting of the ACTIVLIM-CP is depicted in Fig. 1. In the top panel, the distribution of global activity performance measures in the whole sample is displayed. The overall mean children location is 0.62 logit (± 2.76), indicating that the selected items are correctly targeted for this sample covering a range from -3.75 to 3.80 logits. No major gaps in the measurement scale were observed showing that the range of measurement fits the distribution of children's performance measures. Twenty-three children with extreme scores could not be assessed by the scale since all activities were either impossible ($n = 16$, i.e. 7% of the children) or easy ($n = 7$, i.e. 3% of the children).

3.5. Scale reliability

A Person Separation Index of 0.98 was found for ACTIVLIM-CP indicating that the scale has a good precision and allows more than 9 global activity performance levels to be statistically distinguished in our sample.

Fig. 2 represents test-retest of items' and persons' locations based on the 129 parents' responses. The items location at first assessment presented an excellent agreement with location at second assessment (ICC: 0.99; 95% CI: 0.97–0.99; $p < 0.001$). The person location at first assessment was also highly correlated with the person location at second assessment (ICC: 0.96; 95% CI: 0.94–0.97; $p < 0.001$).

3.6. Construct validity

3.6.1. Divergent validity

No significant effects of age (Pearson correlation $r = 0.07$; $p = 0.326$) and gender (t -test; $p = 0.762$) were observed in the ACTIVLIM-CP measures.

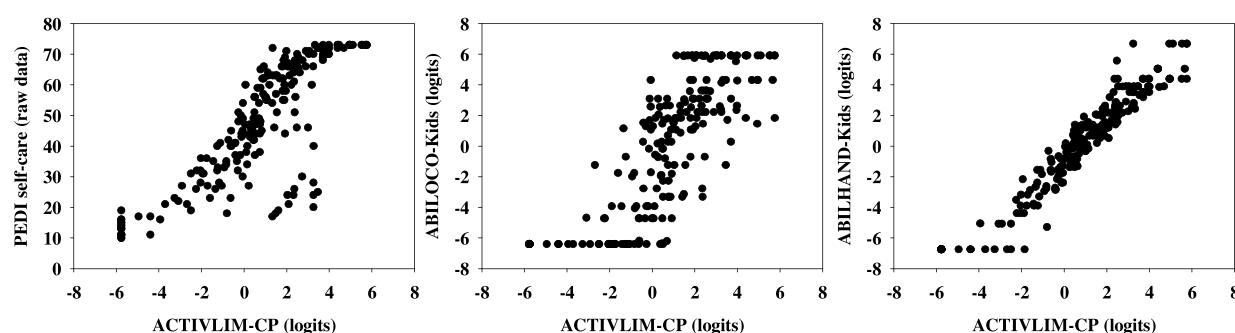


Fig. 3. Correlations between children's ACTIVLIM-CP measures and the PEDI raw scores (self-care section, on the left), the ABILOCO-Kids measures (in the middle) and the ABILHAND-Kids measures (on the right).

3.6.2. Convergent validity

A significant difference was observed when comparing groups of topography (Kruskall-Wallis, $p < 0.001$). Dunn's post-hoc analysis showed that children with quadriplegia presented significantly lower global activity performance measures (median[Q1,Q3]; $-0.82[-3.27, -0.51]$ logits) than children with diplegia ($1.04[0.47, 2.48]$ logits) or hemiplegia ($1.96[0.48, 3.26]$ logits). Significant correlations were observed between the ACTIVLIM-CP measures and the PEDI ($r = 0.87$, $p < 0.001$); the ABILOCO-Kids ($r = 0.80$, $p < 0.001$) and the ABILHAND-Kids ($r = 0.95$, $p < 0.001$) (Fig. 3). As highlighted in Fig. 3, while highly correlated, ACTIVLIM-CP demonstrated less ceiling effects than the scales concurrently used, especially the PEDI (self-care) and the ABILOCO-Kids. It has also less floor effect than ABILHAND-Kids and ABILOCO-Kids, but is less able to discriminate between children with a very low level of motor performance than the PEDI.

4. Discussion

This study allowed the building of a new scale, the ACTIVLIM-CP, measuring global activity performance in children with CP. The parents' response rate was of 38%, which closely matches the acceptable response rate estimated around 40% (Cohen, Manion, & Morrison, 2000). The scale is based on caregivers' perception of children's performance in 43 items related to ICF activity domains of self-care, mobility and domestic life, including mostly ADL, providing a unidimensional measure of global activity expressed on an interval scaling. The scale is invariant between groups (age, gender, and topography), allowing the measurement of global activity in children with CP from 2 to 18 years old whatever the topography of their pathology.

The ACTIVLIM-CP also demonstrated a good reliability (Pearson Separation Index = 0.98) indicating that an excellent degree of precision is obtained when using the questionnaire to assess children (either a single subject or a group of patients; Tennant & Conaghan, 2007). Excellent test-retest reliability is also observed for ACTIVLIM-CP since ICC of 0.99 and 0.96 were respectively obtained for item difficulty hierarchy and children's measure. The item difficulty hierarchy of ACTIVLIM-CP is congruent with clinical observations. Activities that require together a high bimanual precision and an involvement of LE and/or trunk (cutting one's foot nail) were rated as the most difficult. Unimanual activities of low precision and less trunk/LE involvement (put off a hat) were rated as the easiest. Previous development of a global activity performance scale in children with neuromuscular diseases showed that the demand in bimanual coordination and the physical load of the activities were the ingredients of increased difficulty in the items (Vandervelde et al., 2007). ACTIVLIM-CP is consistent regarding bimanual precision tasks, but adds the idea that difficulty might be linked to the combination with LE. Though starting with a very close set of items compared to Vandervelde et al. (2007) in the experimental questionnaire, the item hierarchy in ACTIVLIM-CP is not related to physical load. Most activities with a high physical load (mostly locomotor items) were removed in the course of defining a unidimensional variable. This shows the crucial importance of developing specific tools for different pathologies, i.e. using specific and not generic tools (Arnould et al., 2012).

While physical load – mainly represented by LE walking activities – does not seem to be a key element in the selection of items representing global activity performance in children with CP, the high correlation with the ABILOCO-Kids questionnaire, a tool solely focused on walking abilities in children with CP (Caty et al., 2008), ensures that LE is part of the unidimensional ACTIVLIM-CP scale. However, the correlation scatterplot between ACTIVLIM-CP and ABILOCO-Kids (see Fig. 3B) illustrates that ACTIVLIM-CP presents less ceiling and floor effects than ABILOCO-Kids. Actually only 7% of the children were unable to perform at least one activity and only 3% were easily able to complete all activities. This low percentage of children at the extremes of the scale shows that ACTIVLIM-CP has not significant ceiling or floor effect (McHorney, Lu, & Sherbourne, 1994).

Similarly to the correlation with the ABILOCO-Kids, a good relationship was observed with the measure of capacity in self-care as assessed by the self-care part of the PEDI ($r = 0.85$). Despite the good correlation, the PEDI, as previously noticed (van Empelen et al., 2005) did present an important ceiling effect (see Fig. 3A). Conversely, this scale allowed more children with a very low level of motor performance to be measured.

The last scale used for attesting convergent validity was the ABILHAND-Kids, a Rasch-based measure of manual ability. While at the start the original set of items coming from ACTIVLIM for neuromuscular diseases included upper extremity

activities items also used in the ABILHAND-Kids, only 7 of them remained in the final ACTIVLIM-CP scale (16% of common items). However a high correlation was observed between the ACTIVLIM-CP and ABILHAND-Kids showing a good congruence between the measures of the two scales. This high correlation might be explained as a general congruence between upper extremities deficits, lower extremities deficits and global motor activity performance. This idea of a congruence between gross motor deficits and manual abilities in children with CP has previously been highlighted when observing the relationship between upper extremities and gross motor classifications. In a sample of 120 adolescents with cerebral palsy, Majnemer et al. (2013), observed a high correlation (up to 0.89) between scores of the Gross Motor Function Classification System (GMFCS) (Palisano and Rosenbaum et al., 1997), a classification mainly based on locomotor abilities; and the Manual Ability Classification System (MACS), a classification characterizing the manual ability of children with CP (Eliasson et al., 2006). The authors suggested this correlation was representative of the relationship between locomotor and manual abilities. However, the degree of association between the 2 classifications varied following the topography of the pathology, showing lower correlation coefficient for children with diplegia or hemiplegia than quadriplegia.

As high correlations were highlighted between ACTIVLIM-CP and existing UE and LE questionnaires, a supplementary clinical investigation was performed to ascertain this questionnaire is not capturing in one scale the two different variables included in ABILHAND-Kids and ABILOCO-Kids (see Supplementary material).

The interest of targeting both upper and lower extremities, as well as combined activities in ACTIVLIM-CP is further enhanced when observing the differences between the children with various topographies: children with quadriplegia presented significantly lower global activity measures than children with diplegia or hemiplegia, as highlighted by the DIF analysis (see Results, construct validity). This is in agreement with clinical observations, highlights the good fitting of the scale for measuring combined upper and lower extremities deficits and reinforces its construct validity. In addition, this global activity performance measure, including UE, LE and combined activities, is closer to the concept of children's autonomy in daily life activities than separate questionnaires that, however, can be useful to provide more details on UE or LE solely.

ACTIVLIM-CP responds to the key criteria for scale assessment defined in the guidelines for psychometric standards for rating scales developed by the Scientific Advisory Committee of the Medical Outcomes Trust (2002) and the US Food and Drug Administration (2009). A Rasch online analysis will be freely available (www.rehab-scales.org).

4.1. Limitations

Though unidimensionality is ensured through Rasch analysis and reliability has been demonstrated for ACTIVLIM-CP, the responsiveness of the scale is not addressed in this study. This psychometric quality should be tested to ensure the ability of the scale to detect changes and to confirm the absence of floor or ceiling effect before recommending the use of ACTIVLIM-CP for clinical and research purposes.

Though the item difficulty hierarchy of ACTIVLIM-CP is invariant across age and gender, normative data for typically developing children should be provided for the questionnaire, allowing to locate each child with CP, during development along the continuum of typical development.

Funding source

All authors confirm they have no external funding was secured for this study.

Financial disclosure

All authors confirm they have no financial disclosure.

Conflicts of interest

All authors confirm they have no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ridd.2016.10.005>.

References

- Andrich, D. (1998). *Rasch models for measurement*. Newbury Park, CA: Sage Publications.
- Andrich, D., & Sheridan, B. E. (2005). *Interpreting RUMM2020: part II, polytomous data* (3rd ed.). Perth: RUMM Laboratory Pty Ltd.
- Arnould, C., Penta, M., Renders, A., & Thonnard, J. L. (2004). ABILHAND-kids: A measure of manual ability in children with cerebral palsy. *Neurology*, 63(6), 1045–1052.
- Arnould, C., Vandervelde, L., Batcho, C. S., Penta, M., & Thonnard, J. L. (2012). Can manual ability be measured with a generic ABILHAND scale? A cross-sectional study conducted on six diagnostic groups. *BMJ Open*, 2(6) <http://dx.doi.org/10.1136/bmjopen-2012-001807>, pii: e001807
- Batcho, C. S., Tennant, A., & Thonnard, J. L. (2012). ACTIVLIM-Stroke: a crosscultural Rasch-built scale of activity limitations in patients with stroke. *Stroke*, 43(3), 815–823.

Please cite this article in press as: Bleyenheuft, Y., et al. ACTIVLIM-CP a new Rasch-built measure of global activity performance for children with cerebral palsy. *Research in Developmental Disabilities* (2016), <http://dx.doi.org/10.1016/j.ridd.2016.10.005>

- Bleyenheuft, Y., Arnould, C., Brandao, M., Bleyenheuft, C., & Gordon, A. M. (2015). Hand and arm bimanual intensive therapy including lower extremity (HABIT-ILE) in children with unilateral spastic cerebral palsy: A randomized trial. *Neurorehabilitation and Neural Repair*, 29, 645–657.
- Bleyenheuft, Y., & Gordon, A. M. (2014). Hand-arm bilateral intensive therapy including lower extremities (HABIT-ILE) for children with cerebral palsy, physiotherapy and occupational therapy in pediatrics. *Physical and Occupational Therapy in Pediatrics*, 34, 390–403.
- Caty, G. D., Arnould, C., Thonnard, J. L., & Lejeune, T. M. (2008). ABILOCO-kids: A Rasch-built 10-item questionnaire for assessing locomotion ability in children with cerebral palsy. *Journal of Rehabilitation Medicine*, 40(10), 823–830. <http://dx.doi.org/10.2340/16501977-0267>
- Chong, J., Mackey, A. H., Broadbent, E., & Stott, S. (2011). Relationship between walk test and parental reports of walking abilities in children with cerebral palsy. *Archives of Physical Medicine and Rehabilitation*, 92, 265–270.
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education*. London: Routledge.
- Coker, P., Karakostas, T., Dodds, C., & Hsiang, S. (2010). Gait characteristics of children with hemiplegic cerebral palsy before and after modified constraint-induced movement therapy. *Disability & Rehabilitation*, 32(5), 402–408.
- Durkin, M. S., Benedict, R. E., Christensen, D., Dubois, L. A., Fitzgerald, R. T., et al. (2016). Prevalence of cerebral palsy among 8-year-old children in 2010 and preliminary evidence of trends in its relationship to low birthweight. *Paediatric and Perinatal Epidemiology*, 30, 496–510.
- Eliasson, A. C., Krumlinde-Sundholm, L., Rösblad, B., Beckung, E., Arner, M., Ohrvall, A. M., et al. (2006). The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. *Developmental Medicine & Child Neurology*, 48(7), 549–554.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.
- Feldman, A. B., Haley, S. M., & Coryell, J. (1990). Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. *Physical Therapy*, 70, 602–610.
- Gillick, B. (2012). Effects of constraint-induced movement therapy (CIMT) on gross motor outcomes. *Pediatric Physical Therapy*, 24(3), 300. <http://dx.doi.org/10.1097/PEP.0b013e31825f7e45>
- Gillick, B. T., & Koppes, A. (2010). Gross motor outcomes in children with hemiparesis involved in a modified constraint-induced therapy program. *Journal of Pediatric Rehabilitation Medicine*, 3(3), 171–175.
- Gilmore, R., Sakzewski, L., & Boyd, R. (2009). Upper-limb assessment activity measures for 5- to 16-year-old children with congenital hemiplegia: A systematic review. *Developmental Medicine & Child Neurology*, 52, 14–21.
- Harvey, A., Robin, J., Morris, M., & Baker, R. (2008). A systematic review of measures of activity limitation for children with cerebral palsy. *Developmental Medicine & Child Neurology*, 50, 190–198.
- Linacre, J. M. (2000). New approaches to determining reliability and validity. *Research Quarterly for Exercise and Sport*, 71(2 Suppl), S129–36.
- Majnemer, A., Shikako-Thomas, K., Shevell, M., Poulin, C., Lach, L., Law, M., et al. (2013). The relationship between manual ability and ambulation in adolescents with cerebral palsy. *Physical and Occupational Therapy in Pediatrics*, 33(2), 243–252. <http://dx.doi.org/10.3109/01942638.2012.754394>
- McHorney, C. A., Ware, J. E., Jr., Lu, J. F., & Sherbourne, C. D. (1994). The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, 32(1), 40–66.
- Novak, I., McIntyre, S., Morgan, C., Campbell, L., Dark, L., Morton, N., et al. (2013). A systematic review of interventions for children with cerebral palsy: State of the evidence. *Developmental Medicine & Child Neurology*, 55(10), 885–910. <http://dx.doi.org/10.1111/dmcn.12246>
- Palisano, R. J., Rosenbaum, P. L., et al. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine & Child Neurology*, 39(4), 214–223.
- Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, 46, 1–18.
- Penta, M., Tesio, L., Arnould, C., Zancan, A., & Thonnard, J. L. (2001). The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke*, 32(7), 1627–1634.
- RUMM2020, RUMM Laboratory Pty. Ltd., Perth, Western Australia.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: MESA Press.
- Roehrig, B., Hoeffken, K., Pientka, L., & Wedding, U. (2007). How many and which items of activities of daily living (ADL) and instrumental activities of daily living (IADL) are necessary for screening. *Critical Reviews in Oncology/Hematology*, 62, 164–171.
- Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality of life instruments: Attributes and review criteria. *Quality of Life Research*, 11, 193–205.
- Seedat, F., James, M. K., & Rose, M. R. (2014). Activity rating scales in adult muscle disease: How well do they actually measure? *Muscle and Nerve*, 50(July (1)), 24–33. <http://dx.doi.org/10.1002/mus.24090>. Epub 2014 Jan 11
- Sköld, A., Hermansson, L. N., Krumlinde-Sundholm, L., & Eliasson, A. C. (2011). Development and evidence of validity for the children's hand-use experience questionnaire (CHEQ). *Developmental Medicine & Child Neurology*, 53(May (5)), 436–442.
- Smith, E. V., Jr., Conrad, K. M., Chang, K., & Piazza, J. (2002). An introduction to Rasch measurement for scale development and person assessment. *Journal of Nursing Measurement*, 10(Winter (3)), 189–206.
- Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatology*, 57, 1358–1362.
- US Food and Drug Administration. (2009). *Patient reported outcome measures: use in medical product development to support labeling claims*. Available at www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf Accessed 24.05.11
- van Empelen, R., Jennekens-Schinkel, A., Gorter, J. W., Volman, M. J., van Nieuwenhuizen, O., Helders, P. J., et al. (2005). Epilepsy surgery does not harm motor performance of children and adolescents. *Brain*, 128(Pt. 7), 1536–1545.
- Vandervelde, L., Van den Bergh, P. Y., Goemans, N., & Thonnard, J. L. (2007). ACTIVLIM: A Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscular Disorders*, 17(6), 459–469.
- WHO. (2016). *International classification of functioning, disability and health*. <http://www.who.int/classifications/icf/en/>
- Wright, B. D. (1999). Model selection: Rating scale or partial credit? *Rasch Measurement Transactions*, 12, 641–642.
- Zipp, G. P., & Winning, S. (2012). Effects of constraint-induced movement therapy on gait, balance, and functional locomotor mobility. *Pediatric Physical Therapy*, 24(1), 64–68.