Can manual ability be measured with a generic ABILHAND scale?: a cross-sectional study conducted on six diagnostic groups

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001807
Article Type:	Research
Date Submitted by the Author:	13-Jul-2012
Complete List of Authors:	Arnould, Carlyne; Haute Ecole Louvain en Hainaut, Physical and Occupational Therapy Departments Vandervelde, Laure; Institut Parnasse Deux-Alice, Physical Therapy Department Batcho, Charles; Université catholique de Louvain, Institute of Neuroscience Penta, Massimo; Université catholique de Louvain, Institute of Neuroscience Thonnard, Jean-Louis; Université catholique de Louvain, Institute of Neuroscience
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Paediatrics, Cardiovascular medicine
Keywords:	REHABILITATION MEDICINE, Adult neurology < NEUROLOGY, Paediatric neurology < NEUROLOGY, Neuromuscular disease < NEUROLOGY, Hand & wrist < ORTHOPAEDIC & TRAUMA SURGERY

ARTICLE TITLE:

Can manual ability be measured with a generic ABILHAND scale?: a cross-sectional study conducted on six diagnostic groups

AUTHORS:

Carlyne Arnould[1] (PhD), Laure Vandervelde[2] (PhD), Charles Sèbiyo Batcho[3] (PT), Massimo Penta[3] (PhD), Jean-Louis Thonnard[3] (PhD)

1 Physical and Occupational Therapy Departments, Paramedical Category, Haute Ecole Louvain en Hainaut, Montignies-sur-Sambre, Belgium

2 Physical Therapy Department, Institut Parnasse Deux-Alice, Haute Ecole Leonard de Vinci, Brussels, Belgium

3 Institute of Neuroscience, Université catholique de Louvain, Brussels, Belgium

CORRESPONDING AUTHOR:

Dr. Carlyne Arnould

Physical and Occupational Therapy Departments

Paramedical Category

Haute Ecole Louvain en Hainaut

Rue Trieu Kaisin, 134

6061 Montignies-sur-Sambre

Belgium

Tel. + 32 498 59 31 33

E-mail. carlyne.arnould@gmail.com

KEYWORDS: motor skills disorders, upper extremity, hand, activities of daily living, questionnaires

WORD COUNTS

The title includes 105 characters.

The abstract includes 263 words.

This manuscript includes 3 tables and 3 figures, 2959 words

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Objectives: Several ABILHAND Rasch-built manual ability scales were previously developed for chronic stroke (CS), cerebral palsy (CP), rheumatoid arthritis (RA), systemic sclerosis (SSc), and neuromuscular disorders (NMD). The present study aimed to explore the applicability of a generic manual ability scale unbiased by diagnosis across various populations.

Design: cross-sectional study.

Setting: outpatient clinic homes (CS, CP, RA), specialized centers (CP), reference centers (CP, NMD), and university hospitals (SSc).

Participants: 762 patients from six diagnostic groups: 103 CS adults, 113 CP children, 112 RA adults, 156 SSc adults, 124 NMD children and 124 NMD adults.

Primary and secondary outcome measures: manual ability as measured by the ABILHAND disease-specific questionnaires, diagnosis, and nature (i.e., uni- or bi-manual involvement and proximal or distal joints involvement) of the ABILHAND items.

Results: The difficulty of most manual activities was diagnosis-dependent. A principal component analysis highlighted that 57% of the variance in item locations between diagnoses was explained by the symmetric or asymmetric nature of the disorders. A generic scale was constructed with 11 items sharing a common location among diagnoses and 41 items displaying a category-specific location (asymmetric: CS, CP; and symmetric: RA, SSc NMD). This generic scale showed that CP and NMD children had significantly less manual ability than RA patients, who had significantly less manual ability than CS, SSc, and NMD adults. However, the generic scale was less discriminative and responsive to small deficits than disease-specific instruments.

Conclusions: Our findings emphasize the importance of implementing disease-specific assessments (with ABILHAND) and highlight the risk of using generic scales without prior investigation of the diagnosis-invariance of item difficulties.

ARTICLE SUMARY

Article focus:

- To explore the applicability of a generic ABILHAND manual ability scale unbiased by diagnosis across various clinical populations.
- To analyse prior data from cross-sectional studies that developed disease-specific manual ability questionnaires in order to investigate the co-calibration of patient perceived item difficulty on a common metric.

Key messages:

- The difficulty of most manual activities was diagnosis-dependent and depends on the specificity of the underlying disease since the vast majority (85%) of the difficulty variations observed in manual activities across diagnostic groups was explained by 1) the symmetric or asymmetric nature of the disorder (57% of the variance) and 2) the proximal or distal nature of the disorder (28% of the variance).
- Our findings are consistent with several studies showing that disease-specific instruments are substantially more discriminative and responsive to small deficits than generic instruments.
- Our findings emphasize the importance of implementing disease-specific assessments (with ABILHAND) and highlight the risk of using generic scales in neurological clinical practice without prior investigation of the item difficulties invariance across diagnostic groups.

Strengths and limitations of this study:

- Our study explores a large set of data (732 patients) spread out evenly over 6 diagnostic groups (stroke adults, cerebral palsy children, adults with rheumatoid arthritis, adults with systemic sclerosis, children and adults with neuromuscular disorders).

- Our study proposes an original methodology (combining differential item functioning tests, principal component analysis, and manual activities categorization) that investigates the factors contributing to the hierarchy of manual item difficulty observed across diagnoses allowing the nature of manual ability to be better understood.

INTRODUCTION

Over the past decade, questionnaires and health status assessments have become widely used as outcome measures in clinical trials.[1] Consequently, rating scale data are becoming integral to patient care, prescribing, and policymaking. It is essential that functional rating scales provide scientifically robust and clinically meaningful results to ensure appropriate interpretations and decision-making regarding disease effects, clinical implications, treatment, health policies, and resource allocation. Unfortunately, most rating scales generate ordinal data by summating scores assigned to a set of items representing the intended variable, and metric properties of raw ordinal scores are known to have limited validity.[2, 3] In view of this limitation, the Rasch model [4] is becoming increasingly popular for health measurements because it enables the direct transformation from ordinal scores to linear measures with a constant unit.

Over the last 20 years, our research group has developed several manual ability rating scales (known under the umbrella term of ABILHAND questionnaires) by applying the Rasch model to various diagnostic groups. ABILHAND scales are self-administered questionnaires that measure “manual ability”, which is defined as, “the capacity to manage daily activities requiring the use of the upper limbs, whatever the strategies involved”.[5] Disease-specific manual ability “rulers” were previously developed for the following patient groups: chronic stroke (CS),[5] cerebral palsy (CP), [6] rheumatoid arthritis (RA),[7] systemic sclerosis (SSc),[8] and neuromuscular disorders (NMD).[9] Each ABILHAND scale has its own Rasch-derived item difficulty calibration, which defines a disease-specific manual ability measurement continuum. ABILHAND questionnaires present good psychometric qualities, including linearity, unidimensionality, construct validity, and test-retest reliability.

Disease-specific scales, which are highly sensitive and detect small, yet clinically important changes, are frequently used in research because they ensure comprehensive assessment of

health aspects directly related to the condition.[10, 11] In contrast, generic scales best meet rehabilitation requirements when disability treatment is not dependent upon a specific underlying diagnosis.[12] Generic scales enable comparisons of various diagnoses and healthcare interventions, which may provide useful data for health policies, cost-effective analyses, and resource allocation.[10, 11] From a metric point of view, it is possible to co-calibrate various disease-specific ABILHAND questionnaire items on the same scale, provided that the scales are based on an identical theoretical unidimensional construct.[13] In theory, and similar to the graduations of a metric ruler, items should have the same difficulty for all diagnostic groups, regardless of the disease being measured. In practice, item difficulty hierarchy may vary across groups, demonstrating Differential Item Functioning (DIF).[14] The Rasch model can be used to test the invariance of item difficulty hierarchy and to accommodate for DIF.[15] Nevertheless, the clinical application of generic rating scale usage across various disease populations remains questionable.

The present study explored the applicability of a generic ABILHAND scale, which is unbiased by diagnostic criteria, across various clinical populations. We analysed prior data from cross-sectional studies that developed disease-specific manual ability questionnaires in order to investigate the co-calibration of patient perceived item difficulty on a common metric.

METHODS

Subjects

Data from 732 subjects, who previously provided informed consent, were analysed. Patients with the following disorders were evaluated: 103 CS adults,[5] 113 CP children,[6] 112 RA adults,[7] 156 SSc adults,[8] 124 NMD children (NMDc) and 124 NMD adults (NMDa).[9]

Table 1 provides patient characteristics. The ethics committee of the Université catholique de Louvain, Faculty of Medicine in Brussels, Belgium, authorized and approved the study.

Table 1. Sample characteristics (n = 732)

Variables	CS	CP	RA	SSc	NMDc	NMDa
Number of subjects	103	113	112	156	124	124
Mean age (range)*	63 (24–84)	10 (6–15)	55 (25–82)	54 (21–82)	10 (6–16)	47 (16–80)
Sex						
Males	64	67	29	32	84	69
Females	39	46	83	124	40	55

* years; CS = chronic stroke; CP = cerebral palsy; RA = rheumatoid arthritis; SSc = systemic sclerosis; NMDc = neuromuscular children; NMDa = neuromuscular adults.

Manual ability measure

Original data included 83 manual ability activities, also called items, shared by at least two diagnostic groups. These items covered different domains of daily living such as feeding, grooming, or dressing. In addition, 12 items were child-specific (e.g., "Putting on a backpack/schoolbag"), 19 were adult-specific (e.g., "Hammering a nail"), and 52 were common to both groups (e.g. "Buttoning up trousers"). Adult patients and children's parents provided their perceived difficulty in performing each activity based upon a three-level scale: impossible (0), difficult (1) or easy (2). Each activity had to be completed without technical or human assistance and irrespective of the limb(s) and adaptive strategies used. Missing values were included when a given diagnostic group did not provide responses for a particular item, as the activity may not have been submitted to a group. The nature of the items was assessed by ten occupational or physical therapists according to the following criteria: uni- or bi-manual involvement required to perform the activity; and involvement of proximal or distal joints.

Data analysis

The RUMM2020® Rasch analysis computer program analysed all responses. Manual ability was the only personal attribute theorized to account for the probability of choosing a given response. This requirement, called "unidimensionality", has been tested using fit statistic indices as described elsewhere.[16, 17] Unidimensionality also requires that patients with identical ability, but different diagnoses, have the same probability of succeeding any particular item. Consequently, the invariance of item difficulty across patient diagnostic groups must be controlled. To investigate the invariance of item difficulty hierarchy, a two-way ANOVA was computed on the standardized residuals. [16, 17] Significant diagnostic main effects represented group differences in item difficulty hierarchy.

Analysis process

A co-calibration of the ABILHAND data of all diagnostic groups was performed, and subject responses (n=732) to the 83 items were analysed together. Items that presented an ordered rating scale and fit a unidimensional construct were then selected according to previously described methodology.[7-9] Items presenting a DIF among diagnoses were then split into as many disease-specific items as there were diagnostic groups responding to these items.[15] A principal component analysis (PCA) was subsequently performed to identify disease-specific patterns of item difficulty. Based upon PCA results, the DIF items were split into two main diagnostic groups: asymmetric disorders (CS and CP) and symmetric disorders (RA, SSc, NMD). Successive analyses were then performed to remove items with disordered thresholds, misfit items, and items that did not share a common location between the diagnostic groups. So, a generic co-calibrated scale was created to compare manual ability among diagnostic groups using a Kruskal-Wallis ANOVA of ranks and Dunn's method for pair-wise multiple comparisons. Finally, the metric properties of the generic scale were compared with ABILHAND disease-specific scale properties.

RESULTS

Invariance of manual ability

Thirty-two of the initial 83 items were deleted due to the unidimensionality requirement violation. Assessment of invariance from the remaining 51 unidimensional items showed that 13 items shared a common location between diagnostic groups. Thirty-eight items presented a DIF and were split into a total of 152 items with diagnosis-specific locations. Differences between item difficulty specific to each diagnostic group and the mean item difficulty for all diagnoses were computed to identify disease-specific patterns of item difficulty. Positive values indicated that the items were more difficult for a particular diagnosis than average while negative values indicated that they were easier than average (Figure 1). With respect to disease-specific item difficulties, bimanual activities, such as "spreading butter on a slice of bread," presented a greater challenge for patients with asymmetric disorders (CS and CP) than for patients with symmetric diagnoses (RA, SSc, NMDc, NMDa). Conversely, unimanual activities, such as "turning off a tap," were perceived as easier in asymmetric disordered patients. About 85% of the DIF items were related to the unimanual or bimanual nature of the activities.

In addition, we found that proximal activities, such as "ringing a door bell," were categorized as more difficult for NMD, CP, and CS patients compared to RA and SSc patients. In contrast, digital activities, such as "counting banknotes," presented the greatest challenge for SSc subjects who primary had a distal impairment. Approximately one third of the DIF items were concerned with the proximal or distal nature of the activities. It should be noted that some items fit both criteria. Moreover, DIF activities were related, to a lesser extent, to other factors such as age (about 30% of the items) or mechanical constraints induced in the upper limb joints (about 10–15% of the items).

PCA of differences between diagnosis-specific difficulty of split DIF items and average item difficulty

PCA results showed that 57% of the variance between diagnostic groups was explained by the symmetric or asymmetric nature of the disorders (Figure 2). Indeed, CS adults and CP children were located at one extremity of the first PCA component while diagnostic groups with symmetric disorders were located at the other extremity. The second PCA component explained 28% of the variance between diagnostic groups and distinguished patients expressing greater difficulties with proximal activities, such as NMD and CP, from more distal disorders such as SSc.

Based upon PCA results, the DIF items were split into two main diagnostic groups: asymmetric (CS and CP) and symmetric (RA, SSc, NMD) disorders. When the 13 items sharing a common location between diagnostic groups were co-calibrated with the 38 DIF items split into a total of 75 items (one item was responded neither by CS nor CP subjects) with locations specific to either asymmetric or symmetric disorders, 2 items with disordered thresholds, 7 misfitting items, and 27 remaining DIF items were removed. The resulting 52-item generic scale included 11 items sharing a common location between diagnostic groups and 41 items with locations specific to asymmetric (27 items) or symmetric (14 items) disorders. The 52 items are listed in Table 2 in order of decreasing difficulty (range: 3.60 to -3.93 logits).

Table 2. Final calibration of the six diagnostic groups after the splitting of the DIF items into 2 main groups: asymmetric disorders (CS, CP) and symmetric disorders (RA, SSc, NMDc, NMDa)

Item	Hands involvement	Responded by CS	CP	RA	SSc	NMDc	NMDa	Difficulty* (Logits)	SE (Logits)
a01 Hammering a nail	2C	x						3.60	0.23
a02 Cutting one's nails	2C	x	x					3.32	0.15
a03 Threading a needle	2C	x						3.31	0.23
a04 Peeling potatoes with a knife	2C	x						3.28	0.23
a05 Wrapping up gifts	2C	x						3.20	0.25
a06 Filing one's nails	2C	x						2.82	0.21
a07 Peeling onions	2C	x						2.67	0.24
s08 Shelling hazel nuts	2C			x	x			2.38	0.14
a08 Shelling hazel nuts	2C	x						2.23	0.24
a09 Winding up a wristwatch	2B	x						2.08	0.20
a10 Using a screwdriver	2B	x						2.04	0.22
s01 Hammering a nail	2C			x	x			1.88	0.15
s11 Taking the cap off a bottle	2B			x	x		x	1.84	0.12
a12 Tightening a nut	2B	x						1.81	0.25
s04 Peeling potatoes with a knife	2C			x	x		x	1.56	0.12
a13 Sharpening a pencil	2C	x	x					1.31	0.16
a11 Taking the cap off a bottle	2B	x						1.09	0.23
s07 Peeling onions	2C			x	x			0.91	0.14
a14 Spreading butter on a slice of bread	2B	x	x					0.74	0.16
b15 Fastening a snap (eg, jacket, bag)	2A	x	x	x	x	x	x	0.68	0.09
a16 Replacing a light bulb	2B	x						0.59	0.30
s06 Filing one's nails	2C			x	x			0.59	0.14
s16 Replacing a light bulb	2B			x	x			0.48	0.15
s05 Wrapping up gifts	2C			x	x			0.46	0.15
a17 Opening mail	2B	x	x					0.41	0.16
a18 Handling a stapler	2A	x						0.19	0.30
b19 Peeling a banana	2B		x			x		-0.19	0.15
b20 Filling a glass with water	2A		x			x	x	-0.20	0.12
b21 Opening a pack of biscuits	2B		x			x	x	-0.20	0.13
s22 Turning on a tap	1					x	x	-0.21	0.15
b23 Opening a car door	1		x		x			-0.48	0.15
a24 Opening a bread box	2A		x					-0.56	0.22
s25 Picking up a can	1			x	x	x	x	-0.93	0.12
b26 Throwing a ball	1		x			x		-0.99	0.16
b27 Brushing one's hair	1	x	x	x	x			-1.20	0.13
a28 Unwrapping candy	2C		x					-1.20	0.23
a22 Turning on a tap	1		x					-1.41	0.26
b29 Washing one's face	1	x	x	x	x			-1.63	0.14
a30 Placing a glass of water on a table	1	x	x					-1.70	0.22
s31 Dealing cards	2B					x		-1.80	0.25
a32 Drinking a glass of water	1	x	x					-1.87	0.25
a33 Handling a 4-colour ballpoint pen with one hand	1	x						-2.02	0.37
a34 Counting banknotes	2A	x						-2.20	0.29
b35 Turning on a radio	1	x		x	x	x		-2.28	0.18
a36 Wiping one's hands	2A		x					-2.32	0.26
s37 Using a fork	1		x			x	x	-2.51	0.21
s38 Turning on a television	1			x	x	x	x	-2.56	0.18
b39 Piling up Lego® blocks	2A		x			x		-2.74	0.21
b40 Using a spoon	1	x		x	x			-3.35	0.17
a38 Turning on a television	1	x	x					-3.43	0.38
a41 Turning off a tap	1	x						-3.54	0.56
s42 Blowing one's nose	1			x	x			-3.93	0.23

* higher logit values indicate a greater manual ability for more difficult activities.

a: asymmetric disorders (CS, CP); s: symmetric disorders (RA, SSc, NMDc, NMDa); b: both symmetric and asymmetric disorders; 1 indicates unimanual activities; 2 indicates bimanual activities manageable in several unimanual steps (2A); requiring stabilization with one hand and digital activity with the other (2B); requiring digital activity from both hands (2C). CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults; SE: standard error.

Hand involvement, whether particular groups responded to each item, and item difficulty with standard errors (mean: 0.20 logits; range: 0.09 to 0.56 logits) are also reported. It should be noted that only one diagnostic group responded to 21 items (40%) of the generic scale, while two or three diagnostic groups responded to as many as 12 items (23%). All diagnostic groups responded to the item, "Fastening a snap (e.g., jacket, bag)." The person separation reliability of the generic scale was 0.93, indicating that 5.19 strata of manual ability can be distinguished in our sample. The average measure of the entire sample was 2.34 logits indicating that the patients' ability level exceeded the scale average difficulty.

Manual ability across diagnostic groups

Figure 3 shows the distribution of manual ability across the six diagnostic groups. Significant differences in manual ability measures were observed among diagnoses ($p < 0.001$). The CP and NMDc groups had significantly less manual ability than the RA group, who in turn had less manual ability than CS, SSc and NMDa patients ($p < 0.05$, Dunn's pair-wise comparisons).

Comparison of the generic and disease-specific scales

As reported in Table 3, standard errors of patient locations were greater for the generic scale than for disease-specific scales, and a smaller range of patient measures was observed in the generic scale. The generic scale is globally less accurate than the disease-specific scales leading to a higher number of extreme persons. In addition, Table 3 shows that manual ability measures of generic and disease-specific scales were highly correlated (range: 0.94–0.97).

Table 3. Comparison of generic and disease-specific scales

Diagnostic group	Generic scale			Disease-specific scales			Relationship between the measures of generic and disease-specific scales
	Median SE of measures (logits)	Range of measures (logits)	Extreme subjects (%)	Median SE of measures (logits)	Range of measures (logits)	Extreme subjects (%)	R (p-value)
CS	0.50	7.06	11	0.46	8.56	08	0.95 (< 0.001)
CP	0.53	10.87	09	0.49	11.98	11	0.95 (< 0.001)
RA	0.61	8.27	09	0.44	8.68	09	0.96 (< 0.001)
SSc	0.66	8.45	21	0.46	10.64	12	0.95 (< 0.001)
NMDc	0.70	6.78	23	0.53	8.12	15	0.94 (< 0.001)
NMDa	1.08	10.48	40	0.70	11.14	31	0.97 (< 0.001)

CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults; SE: standard error associated to subjects' measure; R: Pearson correlation.

DISCUSSION

The present study investigated the applicability of a generic manual ability scale unbiased by diagnosis across six populations. We analysed previous subject responses gathered during calibrations of disease-specific ABILHAND questionnaires, and we examined similarities and differences in manual ability among diagnostic groups. A unidimensional scale was constructed with 11 items sharing a common location between diagnostic groups and 41 items having a location specific to asymmetric (CS and CP) or symmetric (NMD, RA, and SSc) disorders. The resulting generic scale revealed that CP and NMD children had significantly less manual ability than RA patients, who in turn had significantly less manual ability than CS, SSc, and NMD adults.

A generic manual ability scale should best meet the requirements of upper limb rehabilitation, insofar as a common instrument with a diagnosis-independent calibration can be used across clinical settings. Of course, use of a generic scale assumes that individuals achieving identical activities have the same manual ability level regardless of their diagnosis. However, this assumption may not hold true in clinical practice. For example, we found that only 11 out of 52 items had difficulties unbiased by diagnosis. Our results differ from those of Simone et al.,[18] who found that the 23 CS-specific ABILHAND item scale "can be routinely applied to a variety of motor impairments." These authors argue that the item hierarchy can be

successfully preserved across diagnoses. Using our patient responses, we conducted a comparable analysis on the same 23 items from the CS-specific ABILHAND scale as Simone et al.[18] Our findings showed that 21 items (91%) presented a significant DIF, which contrasts with the apparent invariance reported by Simone et al.[18] Two factors may contribute to the observed differences in results: sample size and case mix. Our sample included 732 patients, which is significantly more than the 150 subjects in the Simone et al. study.[18] At least 200 subjects are required to detect uniform DIF with adequate power (> 80%).[19] In addition, the unbalanced case mix in the Simone et al.[18] project (83 CS, 17 multiple sclerosis, 13 ataxia, 10 tetraplegics, 3 Parkinson's disease, and 24 healthy controls) may have concealed possible disease influences on difficulty ratings. Taken together, these findings indicate that individuals' underlying diseases or disorders may differentially influence the difficulty of manual activities.

An explicit construct theory initiated the development of disease-specific ABILHAND scales. For each diagnosis, the scale content was selected to delineate a single unidimensional construct, correlated to the patients' functional, clinical, and demographic characteristics.[5-9] The nature of the measured variable, namely manual ability, can be determined by investigating the factors contributing to the hierarchy of manual item difficulty that is observed across diagnoses. To address this issue, we developed an original methodology that combines DIF tests, PCA, and manual activities categorization about their nature. Although an activity is expressed in the same way for all patients, its perceived difficulty may vary according to one's disease or disorder and the specificity of underlying motor impairments. Indeed, various studies show that manual activity limitations are partially related to underlying hand impairments.[5, 20]

The current PCA results suggest that the vast majority (85%) of the difficulty variations observed in manual activities across diagnostic groups was explained by two characteristics:

1) the symmetric or asymmetric nature of the disorder (57% of the variance), and 2) the proximal or distal nature of the disorder (28% of the variance). For example, activities requiring greater bimanual involvement (e.g., "peeling potatoes with a knife") tended to be rated as more difficult by patients with asymmetric disorders (CP children and CS adults) than by patients with more symmetric disorders (RA, SSc, NMDc, NMDa). On the other hand, unimanual activities (e.g., "turning on a television") or bimanual activities manageable in several unimanual steps (e.g., "handling a stapler") were rated as less difficult for patients with asymmetric disorders, likely because these activities can be achieved by exclusively using the unaffected or less affected hand.[6, 21] Activities involving the shoulder (e.g., "drinking a glass of water") were generally more difficult for NMD and CP patients. Indeed, the NMD groups included several diseases in which proximal segments were more likely to be affected than distal ones (e.g., Duchenne/limb girdle muscular dystrophy, facio-scapulo-humeral dystrophy, spinal muscular atrophy).[9] Moreover, and contrary to other diagnoses, NMD and CP groups included subjects in a wheelchair, which may prevent the achievement of activities such as, "ringing a door bell", or "replacing a light bulb". In contrast, digital activities (e.g. "winding up a wristwatch") were particularly difficult for SSc subjects, who have reduced digital dexterity.[8]

The strong correlations ($R \geq 0.94$) observed between the generic scale and each of the disease-specific ABILHAND scales supported the assertion that these instruments successfully measure manual ability. Overall, children had less manual ability than adults. This finding is consistent with previous results[22, 23] showing that children have relatively greater difficulty with manipulation activities than adults.

In our study, the generic scale was globally less accurate than the disease-specific scales, which often included a greater number of disease-relevant activities. Overall, our findings are consistent with several studies showing that disease-specific instruments are substantially

more discriminative and responsive to small deficits than generic instruments.[24, 25] Consequently, this increased sensitivity allows for the detection and quantification of small, yet clinically significant health changes.[10, 11] For example, ABILHAND disease-specific scales should be used to determine pathology impacts on manual ability, to measure clinical changes consecutive to specific treatments, and to tailor interventions to the specific needs of individuals with a particular diagnosis. All of these concerns are important for patients and clinicians in their daily practice. In contrast, generic scales may identify the relative burden of diagnoses, compare various health-care programs, and demonstrate evidence of cost-effectiveness of different healthcare interventions[10, 11, 26] Thus, each type of scale has clinical value and provides unique information corresponding to specific objectives.[27]

In conclusion, the present study proposed an original methodology combining DIF tests with a splitting procedure and PCA to better understand the nature of manual ability. Our results showed that scale content and item type remained the same among diagnostic groups; however, item difficulty changed depending on patient diagnosis. The ABILHAND scale centres on the individual and his/her impairments, adaptive strategies, environmental, and personal factors. Thus, it is not surprising that disease characteristics contribute to the difficulties experienced in performing manual activities, such that manual ability level change will differ between patients with different diagnoses even if exactly the same activities show similar improvement. Our finding that item difficulties were disease-dependent emphasizes the danger of using generic scales without prior investigation of item invariance across diagnostic groups.

Nevertheless, using 11 linked items unbiased by diagnosis, we successfully constructed a unidimensional scale common to six diagnostic groups by separating asymmetric from symmetric disorders. This new generic scale that allows the manual ability of patients with different diagnoses to be compared may be used to examine cost-effectiveness of health

interventions across a variety of diseases and disorders. However, ABILHAND disease-specific scales should be preferred in daily clinical settings for treatment, planning and follow-up of patients.

COMPETING INTERESTS

None declared.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

CONTRIBUTORSHIP STATEMENT

CA performed the statistical analyses, conducted the literature search, and drafted the manuscript. CA, LV, and MP participated in the data collection. CA, LV, MP, and JLT contributed to the study design and the data analysis. All authors participated in the data interpretation, critically revised the draft of the manuscript for important intellectual content, and contributed to the writing. All authors have read and approved the final manuscript.

DATA SHARING STATEMENT

Dataset is available from the corresponding author at carlyne.arnould@gmail.com

REFERENCES

1. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med* 2012;**44**:97–8.
2. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil* 1989;**70**:308–12.
3. Wright BD, Linacre JM. Observations are always ordinal; measurement, however, must be interval. *Arch Phys Med Rehabil* 1989;**70**:857–60.
4. Rasch G. *Probabilistic models for some intelligence and attainment tests.* Chicago: Mesa Press, 1980.
5. Penta M, Tesio L, Arnould C, *et al.* The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke* 2001;**32**:1627–34.
6. Arnould C, Penta M, Renders A, *et al.* ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. *Neurology* 2004;**63**:1045–52.
7. Durez P, Fraselle V, Houssiau F, *et al.* Validation of the ABILHAND questionnaire as a measure of manual ability in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**:1098–105.
8. Vanthuyne M, Smith V, Arat S, *et al.* Validation of a manual ability questionnaire in patients with systemic sclerosis. *Arthritis Rheum* 2009;**61**:695–703.
9. Vandervelde L, Van den Bergh PY, Penta M, *et al.* Validation of the ABILHAND questionnaire to measure manual ability in children and adults with neuromuscular disorders. *J Neurol Neurosurg Psychiatry* 2010;**81**:506–12.
10. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;**27**(3 Suppl):217–32S.

11. Marra CA, Woolcott JC, Kopec JA, *et al.* A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQol and the HAQ) in rheumatoid arthritis. *Soc Sci Med* 2005;**60**:1571–82.
12. Haigh R, Tennant A, Biering-Sorensen F, *et al.* The use of outcome measures in physical medicine and rehabilitation within Europe. *J Rehabil Med* 2001;**33**:273–8.
13. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 2004; **7**(Suppl 1):22–6.
14. Holland PW, Wainer H. *Differential Item Functioning.* Hillsdale, New Jersey: Lawrence Erlbaum, 1993.
15. Tennant A, Penta M, Tesio L, *et al.* Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care* 2004;**42**(1 Suppl): 37–48S.
16. RUMM Laboratory. *Interpreting RUMM2020: Part 1, dichotomous data.* Perth, Western Australia: RUMM Laboratory, 2004.
17. RUMM Laboratory. *Interpreting RUMM2020: Part 2, polytomous data.* Perth, Western Australia: RUMM Laboratory, 2004.
18. Simone A, Rota V, Tesio L, *et al.* Generic ABILHAND questionnaire can measure manual ability across a variety of motor impairments. *Int J Rehabil Res* 2011;**34**:131–40.
19. Scott NW, Fayers PM, Aaronson NK, *et al.* Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health Qual Life Outcomes* 2010;**8**:81.
20. Arnould C, Penta M, Thonnard J-L. Hand impairments and their relationship with manual ability in children with cerebral palsy. *J Rehabil Med* 2007;**39**:708–14.

21. Sakzewski L, Ziviani J, Boyd R. The relationship between unimanual capacity and bimanual performance in children with congenital hemiplegia. *Dev Med Child Neurol* 2010;**52**:811–6.

22. Haley SM, Ludlow LH. Applicability of the hierarchical scales of the Tufts Assessment of Motor Performance for school-aged children and adults with disabilities. *Phys Ther* 1992;**72**:191–202.

23. van Eck M, Dallmeijer AJ, van Lith IS, *et al.* Manual ability and its relationship with daily activities in adolescents with cerebral palsy. *J Rehabil Med* 2010;**42**:493–8.

24. Murawski MM, Miederhoff PA. On the generalizability of statistical expressions of health related quality of life instrument responsiveness: a data synthesis. *Qual Life Res* 1998;**7**:11–22.

25. Wiebe S, Guyatt G, Weaver B, *et al.* Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol* 2003;**56**:52–60.

26. Mazur W, Kupiainen H, Pitkäniemi J, *et al.* Comparison between the disease-specific Airways Questionnaire 20 and the generic 15D instruments in COPD. *Health Qual Life Outcomes* 2011;**9**:4.

27. Le Pen C, Lévy E, Loos F, *et al.* "Specific" scale compared with "generic" scale: a double measurement of the quality of life in a French community sample of obese subjects. *J Epidemiol Community Health* 1998;**52**:445–50.

FIGURE LEGENDS

Figure 1. Disease-specific patterns of item difficulty according to the bimanual or proximal nature of the activities showing a Differential Item Functioning. Differences between item difficulty ratings specific to each diagnostic group ($\delta_{specific}$) and the average item difficulty for all diagnoses (δ_{mean}) are shown for each disorder (CS = chronic stroke; CP = cerebral palsy; RA = rheumatoid arthritis; SSc = systemic sclerosis; NMDc = neuromuscular children; and NMDa = neuromuscular adults). Boxes indicate the 25% and 75% limits (the interquartile range); the vertical line inside each box indicates the median; vertical bars outside each box indicate the 10% and 90% limits and dots indicate the 5% and 95% outliers.

Figure 2. PCA results based upon differences between disease-specific difficulty of the split DIF items and the average item difficulty across all diagnoses. CP = cerebral palsy; NMDc = neuromuscular children; RA = rheumatoid arthritis; SSc = systemic sclerosis; and NMDa = neuromuscular adults. CS = chronic stroke;

Figure 3. Box plots showing the distribution of manual ability measures for each diagnosis. Boxes indicate the 25% and 75% limits (the interquartile range); the vertical line inside each box indicates the median; vertical bars outside each box indicate the 10% and 90% limits and dots indicate the 5% and 95% outliers. CP = cerebral palsy; NMDc = neuromuscular children; RA = rheumatoid arthritis; NMDa = neuromuscular adults; SSc = systemic sclerosis; and CS = chronic stroke.

Figure 1

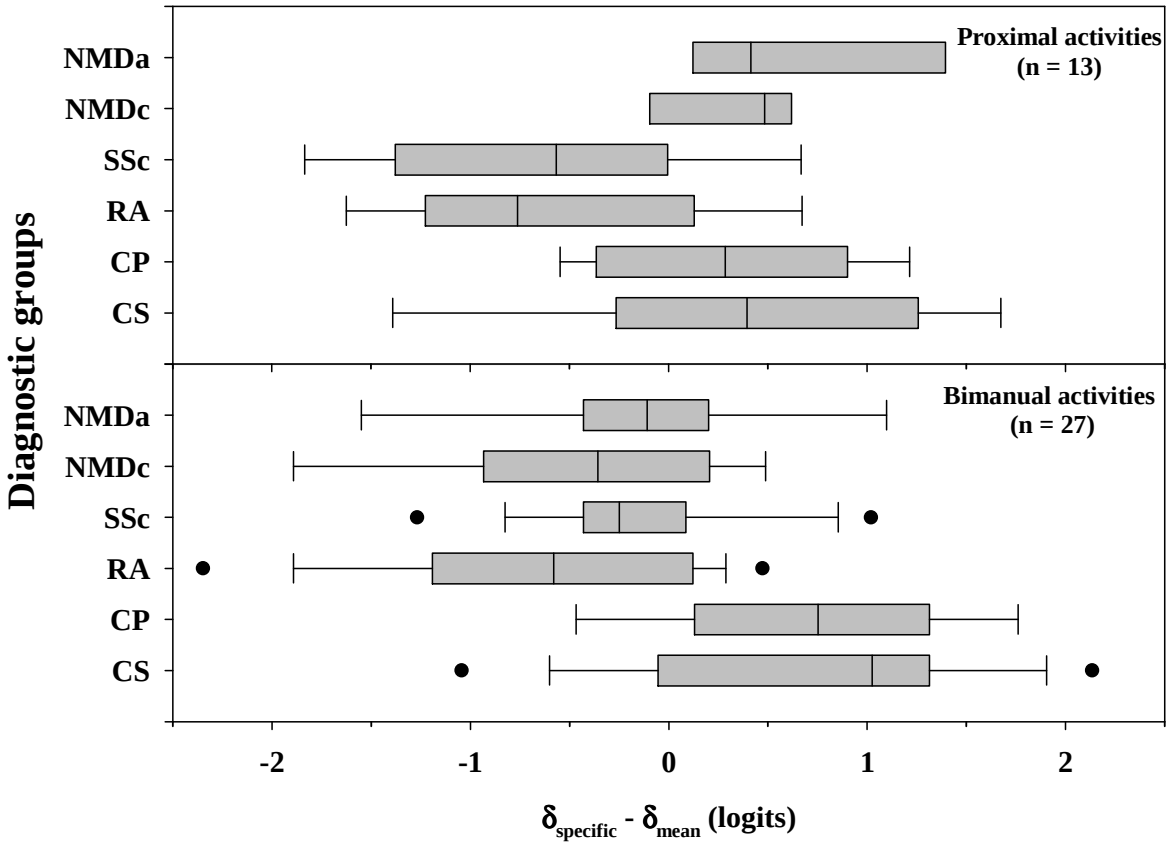


Figure 2

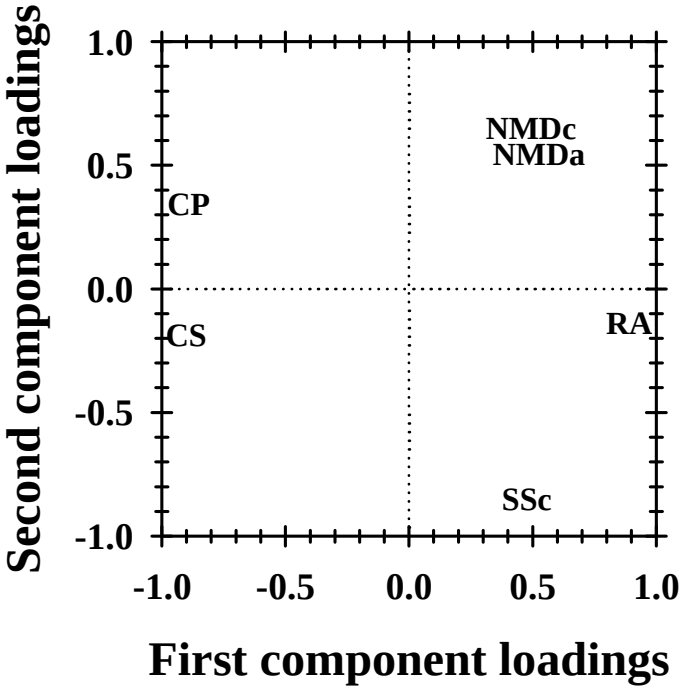
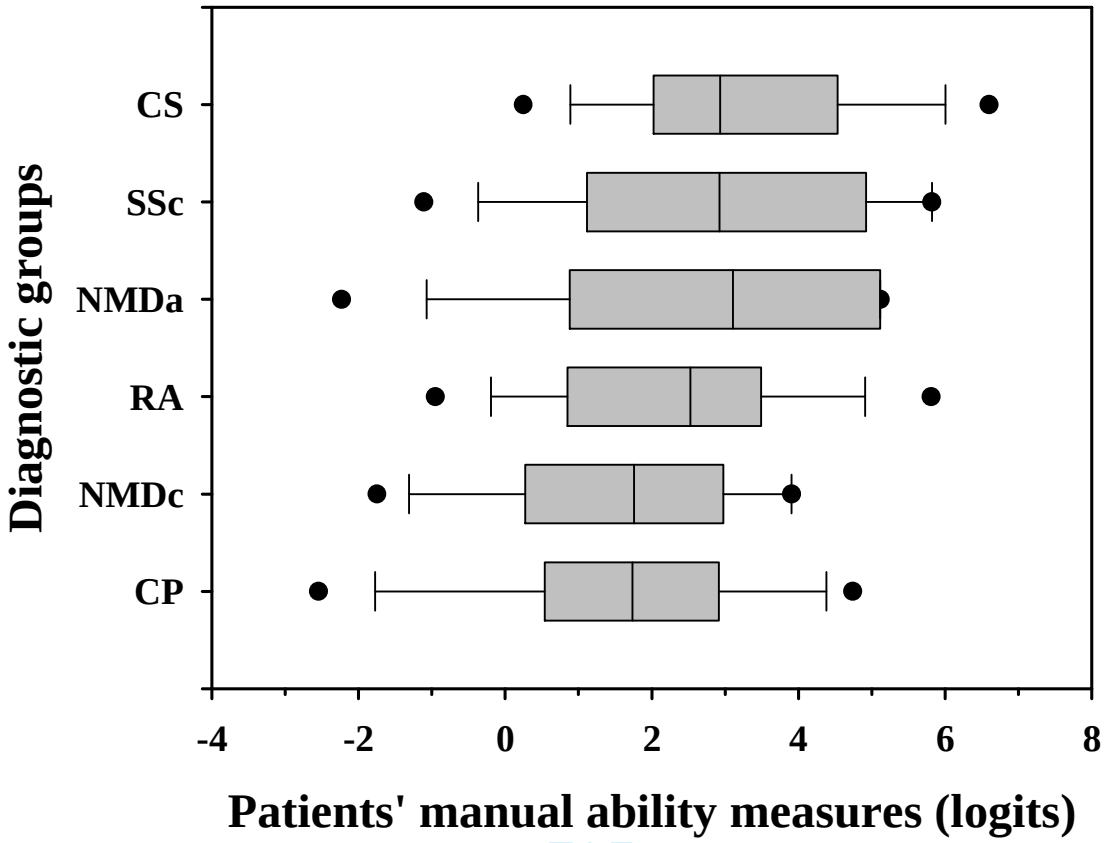


Figure 3



Can manual ability be measured with a generic ABILHAND scale? A cross-sectional study conducted on six diagnostic groups

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001807.R1
Article Type:	Research
Date Submitted by the Author:	19-Sep-2012
Complete List of Authors:	Arnould, Carlyne; Haute Ecole Louvain en Hainaut, Physical and Occupational Therapy Departments Vandervelde, Laure; Institut Parnasse Deux-Alice, Physical Therapy Department Batcho, Charles; Université catholique de Louvain, Institute of Neuroscience Penta, Massimo; Université catholique de Louvain, Institute of Neuroscience Thonnard, Jean-Louis; Université catholique de Louvain, Institute of Neuroscience
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Paediatrics, Cardiovascular medicine
Keywords:	REHABILITATION MEDICINE, Adult neurology < NEUROLOGY, Paediatric neurology < NEUROLOGY, Neuromuscular disease < NEUROLOGY, Hand & wrist < ORTHOPAEDIC & TRAUMA SURGERY

ARTICLE TITLE:

Can manual ability be measured with a generic ABILHAND scale? A cross-sectional study conducted on six diagnostic groups

AUTHORS:

Carlyne Arnould[1] (PhD), Laure Vandervelde[2] (PhD), Charles Sèbiyo Batcho[3] (PT), Massimo Penta[3] (PhD), Jean-Louis Thonnard[3] (PhD)

1 Physical and Occupational Therapy Departments, Paramedical Category, Haute Ecole Louvain en Hainaut, Montignies-sur-Sambre, Belgium

2 Physical Therapy Department, Institut Parnasse Deux-Alice, Haute Ecole Leonard de Vinci, Brussels, Belgium

3 Institute of Neuroscience, Université catholique de Louvain, Brussels, Belgium

CORRESPONDING AUTHOR:

Dr. Carlyne Arnould

Physical and Occupational Therapy Departments

Paramedical Category

Haute Ecole Louvain en Hainaut

Rue Trieu Kaisin, 134

6061 Montignies-sur-Sambre

Belgium

Tel. + 32 498 59 31 33

E-mail. carlyne.arnould@gmail.com

KEYWORDS: motor skills disorders, upper extremity, hand, activities of daily living, questionnaires

WORD COUNTS

The title includes 103 characters.

The abstract includes 297 words.

This manuscript includes 3 tables and 4 figures, 3994 words

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Objectives: Several ABILHAND Rasch-built manual ability scales were previously developed for chronic stroke (CS), cerebral palsy (CP), rheumatoid arthritis (RA), systemic sclerosis (SSc), and neuromuscular disorders (NMD). The present study aimed to explore the applicability of a generic manual ability scale unbiased by diagnosis and to study the nature of manual ability across diagnoses.

Design: cross-sectional study.

Setting: outpatient clinic homes (CS, CP, RA), specialized centers (CP), reference centers (CP, NMD), and university hospitals (SSc).

Participants: 762 patients from six diagnostic groups: 103 CS adults, 113 CP children, 112 RA adults, 156 SSc adults, 124 NMD children and 124 NMD adults.

Primary and secondary outcome measures: manual ability as measured by the ABILHAND disease-specific questionnaires, diagnosis, and nature (i.e., uni- or bi-manual involvement and proximal or distal joints involvement) of the ABILHAND manual activities.

Results: The difficulty of most manual activities was diagnosis-dependent. A principal component analysis highlighted that 57% of the variance in the item difficulty between diagnoses was explained by the symmetric or asymmetric nature of the disorders. A generic scale was constructed, from a metric point of view, with 11 items sharing a common difficulty among diagnoses and 41 items displaying a category-specific location (asymmetric: CS, CP; and symmetric: RA, SSc, NMD). This generic scale showed that CP and NMD children had significantly less manual ability than RA patients, who had significantly less manual ability than CS, SSc, and NMD adults. However, the generic scale was less discriminative and responsive to small deficits than disease-specific instruments.

Conclusions: Our finding that most of the manual item difficulties were disease-dependent emphasizes the danger of using generic scales without prior investigation of item invariance

across diagnostic groups. Nevertheless, a generic manual ability scale could be developed by adjusting and accounting for activities perceived differently in various disorders.

ARTICLE SUMMARY

Article focus:

- To explore the applicability of a generic ABILHAND manual ability scale unbiased by diagnosis across various clinical populations.
- To analyse prior data from cross-sectional studies that developed disease-specific manual ability questionnaires in order to investigate the co-calibration of patient perceived item difficulty on a common metric.
- To better understand the nature of the measured variable, namely manual ability.

Key messages:

- The difficulty of most manual activities was diagnosis-dependent, emphasizing the danger of using generic scales without prior investigation of item invariance across diagnostic groups.
- The vast majority (85%) of the difficulty variations observed in manual activities across diagnostic groups was explained by 1) the symmetric or asymmetric nature of the disorder (57% of the variance) and 2) the proximal or distal nature of the disorder (28% of the variance).
- Although less sensitive than diagnosis-specific scales, a generic manual ability scale could be developed by adjusting and accounting for activities perceived differently in various disorders, which allows quantitative comparisons of manual ability between diagnostic groups.

Strengths and limitations of this study:

- Our study explores a large set of data (732 patients) spread out evenly over 6 diagnostic groups (stroke adults, cerebral palsy children, adults with rheumatoid arthritis, adults with systemic sclerosis, children and adults with neuromuscular disorders).

- Our study proposes an original methodology (combining differential item functioning tests, principal component analysis, and manual activities categorization) that investigates the factors contributing to the hierarchy of manual item difficulty observed across diagnoses allowing the nature of manual ability to be better understood.

INTRODUCTION

One fundamental goal of rehabilitation is to improve the subjects' ability to manage the daily activities necessary for autonomous living.[1] Such an ability belongs to the domain of latent variables concealed within the person, such as pain or intelligence. It cannot be observed directly, but it can be inferred from subject's perceived difficulty in performing activities, also called items, using patient self-reported questionnaires. Over the past decade, questionnaires have therefore become widely used as outcome measures in clinical trials[2] and rating scale data are becoming integral to patient care, prescribing, and policymaking. It is essential that functional rating scales provide scientifically robust and clinically meaningful results to ensure appropriate interpretations and decision-making regarding disease effects, clinical implications, treatment, health policies, and resource allocation. Unfortunately, most rating scales generate ordinal data by summating scores assigned to a set of items representing the intended variable, and metric properties of raw ordinal scores are known to have limited validity.[3, 4] In view of this limitation, the Rasch model[5] is becoming increasingly popular for health measurements because it enables the direct transformation from ordinal scores to linear measures with a constant unit.

Over the last 20 years, our research group has developed several manual ability rating scales (known under the umbrella term of ABILHAND questionnaires) by applying the Rasch model to various diagnostic groups. ABILHAND scales are self-administered questionnaires that measure "manual ability", which is defined as, "the capacity to manage daily activities requiring the use of the upper limbs, whatever the strategies involved".[6] Disease-specific manual ability "rulers" were previously developed for the following patient groups: chronic stroke (CS),[6] cerebral palsy (CP),[7] rheumatoid arthritis (RA),[8] systemic sclerosis (SSc),[9] and neuromuscular disorders (NMD).[10] Each ABILHAND scale has its own Rasch-derived item difficulty calibration, which defines a disease-specific manual ability

measurement continuum. ABILHAND questionnaires present good psychometric qualities, including linearity, unidimensionality, construct validity, and test-retest reliability.

Disease-specific scales, which are highly sensitive and detect small, yet clinically important changes, are frequently used in research because they ensure comprehensive assessment of health aspects directly related to the condition.[11, 12] In contrast, generic scales enable comparisons of various diagnoses and healthcare interventions, which may provide useful data for health policies, cost-effective analyses, and resource allocation.[11, 12] They best meet rehabilitation requirements when disability treatment is not dependent upon a specific underlying diagnosis. [13] For instance, as a single bathroom scale can be used to weigh all patients, a generic manual ability scale would enable quantitative comparisons of the ability to use the upper limbs in daily activities across patients of various diagnoses (and also with healthy subjects).

From a metric point of view, it is possible to co-calibrate various disease-specific ABILHAND questionnaire items on the same scale, provided that the scales are based on an identical unidimensional construct.[14] In theory, and similar to the graduations of a metric ruler, items should have the same difficulty for all diagnostic groups, regardless of the disease being measured. Nevertheless, the main implicit assumption made by the users of generic scales is that the difficulties of daily activities are invariant across diagnoses. However, in practice, item difficulty hierarchy may vary across groups, demonstrating Differential Item Functioning (DIF).[15] The Rasch model can be used to test the invariance of item difficulty hierarchy and to accommodate for DIF.[16] When the items of a generic scale are unstable across diagnoses, the measurements generated by them cannot be used to make meaningful comparisons.

The present study explored the applicability of a generic ABILHAND manual ability scale, which is unbiased by diagnoses, across various clinical populations. Setting out this objective,

we also intended to improve the current understanding of the nature of manual ability and especially its interaction with diagnosis. We analysed prior data from cross-sectional studies that developed disease-specific manual ability questionnaires in order to investigate the co-calibration of patient perceived item difficulty on a common metric.

METHODS

Subjects

Data from 732 subjects, who previously provided informed consent, were analysed. Patients with the following disorders were evaluated: 103 CS adults,[6] 113 CP children,[7] 112 RA adults,[8] 156 SSc adults,[9] 124 NMD children (NMDc) and 124 NMD adults (NMDa).[10] Table 1 provides patient characteristics. The ethics committee of the Université catholique de Louvain, Faculty of Medicine in Brussels, Belgium, authorized and approved the study.

Table 1. Sample characteristics (n = 732)

Variables	CS	CP	RA	SSc	NMDc	NMDa
Number of subjects	103	113	112	156	124	124
Age, y, mean (range)	63 (24-84)	10 (6-15)	55 (25-82)	54 (21-82)	10 (6-16)	47 (16-80)
Sex						
Males	64	67	29	32	84	69
Females	39	46	83	124	40	55
Diagnosis	R hemi: 55	tetra: 35	No UL disorder: 20	lcSSc: 104	DMD/BMD or LGMD: 47	MD: 24
	L hemi: 48	di: 24	DH disorder: 9	dcSSc: 33	HN: 35	HN: 24
		R hemi: 26	NDH disorder: 4	lSSc: 19	SMA: 3	DMD/BMD or LGMD: 19
		L hemi: 28	2 UL disorder: 79		Others (CM, CMD, PPS, …): 29	SMA: 7
						FSHD: 7
						Others (CM, CMD, PPS, …): 43

CS = chronic stroke; CP = cerebral palsy; RA: rheumatoid arthritis; NMDc = neuromuscular children; NMDa = neuromuscular adults; SSc = systemic sclerosis; R hemi = right hemiplegia; L hemi = left hemiplegia; tetra = tetraplegia; di = diplegia; UL = upper limb; DH = dominant hand; NDH = non-dominant hand; DMD = Duchenne muscular dystrophy; BMD = Becker muscular dystrophy; LGMD = limb girdle muscular dystrophy; HN = hereditary neuropathy; SMA = spinal muscular atrophy; CM = congenital myopathy; CMD = congenital muscular dystrophy; PPS = post-polio syndrome; MD = myotonic dystrophy; FSHD = facio-scapulo-humeral dystrophy; lcSSc = limited cutaneous SSc; dcSSc = diffuse cutaneous SSc; lSSc = limited SSc.

Manual ability measure

Original data included 83 manual activities shared by at least 2 diagnostic groups (the 83 items are provided in the supplementary table). Original items covered different domains of

daily living such as feeding, grooming, or dressing and were selected in previous studies based on literature review and patient and experts interviews. Twelve items were child-specific (e.g., ""throwing a ball"), 19 were adult-specific (e.g., "hammering a nail"), and 52 were common to both groups (e.g. "buttoning up trousers"). Adult patients and children's parents provided their perceived difficulty in performing each activity based upon a three-level response scale: impossible (0), difficult (1) or easy (2). Each activity had to be completed without technical or human assistance and irrespective of the limb(s) and adaptive strategies used. Missing values were included when a given diagnostic group did not provide responses for a particular item, as the activity may not have been submitted to a group. The nature of the items was assessed by ten occupational or physical therapists according to the following criteria: uni- or bi-manual involvement required to perform the activity; and involvement of proximal or distal joints.

Data analysis

All responses were analysed with RUMM2020®, a Rasch analysis computer program. The Rasch model[5] can be used to estimate, on a single manual ability construct, the location of each patient , i.e. their manual ability, the location of each item, i.e. the difficulty of the manual activities, and the location of each threshold between successive categories of the response scale, i.e. the locations along the latent construct at which two successive categories are equally likely to be observed. The model can be used to verify that successive response categories for each item represent increasing levels of ability and that thresholds between successive response categories are located in the anticipated order.[17]

The model also requires that the probability of endorsing any response category to an item depends solely on the subject's ability, the item difficulty and the location of the threshold between adjacent response categories. In the case of manual ability measurement, no attribute of the person - such as diagnosis - besides manual ability is theorized to account for the

probability of choosing a given response to a given item. The similarity between the observed and expected responses can be investigated using a χ^2 fit statistic computed over 5 class intervals (CI) of patients with increasing ability.[18] Items with a p-value lower than 0.05 indicate a threat to the fit requirement.

Invariance of the item difficulty hierarchy

Unidimensionality also requires that patients with identical ability, but different diagnoses, have the same probability of succeeding any particular item. Consequently, the invariance of item difficulties across patient diagnostic groups must be controlled using Differential Item Functioning (DIF) tests.[15] To investigate the invariance of item difficulty hierarchy, a two-way ANOVA was computed on the standardized residuals of the different CIs [19, 20]; the first factor was the diagnostic group and the second factor was the CI of increasing manual ability. Significant diagnostic main effects represented group differences in item difficulty hierarchy. A solution to the presence of DIF by diagnosis is the removal of items showing difficulty variations. Another solution is to allow for the variations that exist across DIF items by splitting them into disease-specific items, one for each diagnostic, with a difficulty peculiar to the corresponding diagnosis.[16] In this case, the different diagnostic groups can be compared on the same continuum even if they have specific items provided that there are common linking items unbiased by DIF.

Analysis process

Two different approaches can be used to combine data from different scales responded by different samples. The ‘co-calibration’, also called ‘concurrent equating’, merges all items together as one scale with empty spaces for missing values. The ‘anchoring’ approach anchors items that are common to all diagnoses and then includes diagnosis-specific items in the same frame of reference. The anchoring approach requires that the common linking items be free of DIF,[21–23] which was not the case in our dataset. Therefore, the co-calibration approach,

also applied in previous rehabilitation studies,[24–27] was followed and the analysis process is illustrated in Figure 1. The first step in the data analysis was to co-calibrate the ABILHAND data of all diagnostic groups by analysing all responses (n=732) to the 83 items. The second step was to remove items with disordered thresholds and items that misfit a unidimensional variable (i.e., presenting a χ^2 p-value < 0.05). In the third step, the invariance of item difficulty hierarchy was detected across diagnostic groups through DIF tests. The fourth step consisted in splitting the items presenting a DIF by diagnosis providing one specific item for each diagnostic group who answered the item.[16] In the fifth step, a principal component analysis (PCA) was performed to identify the potential factors explaining item difficulty hierarchy variations observed across the diagnostic groups. The PCA was performed on the differences between item difficulty specific to each diagnostic group and the average item difficulty for all diagnoses as these differences reflect disease-specific patterns of item difficulty. In the sixth step, the items presenting a DIF among diagnoses (detected in the third step) were split into two main groups: asymmetric disorders (CS and CP) and symmetric disorders (RA, SSc, NMD). Finally, the seventh step included successive analyses performed to remove items with disordered thresholds, misfitting items, and items presenting a DIF by diagnosis for another reason than the symmetric/asymmetric nature of the disorders. So, a generic co-calibrated scale was created and manual ability was compared among diagnostic groups using a Kruskal-Wallis ANOVA of ranks and Dunn's method for pairwise multiple comparisons. Finally, the metric properties of the generic scale were compared with ABILHAND disease-specific scale properties.

RESULTS

Invariance of the item difficulty hierarchy

Thirty-two of the initial 83 items were deleted due to the unidimensionality requirement violation. Assessment of invariance from the remaining 51 unidimensional items showed that

13 items shared a common location between diagnostic groups. Thirty-eight items presented a DIF and were split into a total of 152 items with diagnosis-specific locations. Differences between item difficulty specific to each diagnostic group and the mean item difficulty for all diagnoses were computed to identify disease-specific patterns of item difficulty. Positive values indicated that the items were more difficult for a particular diagnosis than average while negative values indicated that they were easier than average (Figure 2). With respect to disease-specific item difficulties, bimanual activities, such as "spreading butter on a slice of bread," presented a greater challenge for patients with asymmetric disorders (CS and CP) than for patients with symmetric diagnoses (RA, SSc, NMDc, NMDa). Conversely, unimanual activities, such as "turning off a tap," were perceived as easier in asymmetric disordered patients. About 85% of the DIF items were related to the unimanual or bimanual nature of the activities.

In addition, we found that proximal activities, such as "ringing a door bell," were categorized as more difficult for NMD, CP, and CS patients compared to RA and SSc patients. In contrast, digital activities, such as "counting banknotes," presented the greatest challenge for SSc subjects who primary had a distal impairment. Approximately one third of the DIF items were concerned with the proximal or distal nature of the activities. It should be noted that some items fit both criteria. Moreover, DIF activities were related, to a lesser extent, to other factors such as age (about 30% of the items) or mechanical constraints induced in the upper limb joints (about 10–15% of the items).

PCA on diagnosis-specific-to-average item difficulty differences

PCA results showed that 57% of the variation of item difficulty hierarchy between diagnostic groups was explained by the symmetric or asymmetric nature of the disorders (Figure 3). Indeed, CS adults and CP children were located at one extremity of the first PCA component while symmetric disorders were located at the other extremity. The second PCA component

explained 28% of the variation of item difficulty hierarchy between diagnostic groups and distinguished patients expressing greater difficulties with proximal activities, such as NMD and CP, from more distal disorders such as SSc.

A "generic" ABILHAND manual ability scale

Based upon PCA results, the DIF items were split into two main groups: asymmetric (CS and CP) and symmetric (RA, SSc, NMD) disorders. When the 13 items sharing a common location between diagnostic groups were co-calibrated with the 38 DIF items split into a total of 75 items (one item was responded neither by CS nor CP subjects) with locations specific to either asymmetric or symmetric disorders, 2 items with disordered thresholds, 7 misfitting items, and 27 remaining DIF items were removed. The resulting 52-item generic scale included 11 items sharing a common location between diagnostic groups and 41 items with locations specific to asymmetric (27 items) or symmetric (14 items) disorders. The 52 items are listed in Table 2 in order of decreasing difficulty (range: 3.60 to -3.93 logits).

Hand involvement, whether particular groups responded to each item, and item difficulty with standard errors (mean: 0.20 logits; range: 0.09 to 0.56 logits) are also reported. It should be noted that only one diagnostic group responded to 21 items (40%) of the generic scale, while two or three diagnostic groups responded to as many as 12 items (23%). All diagnostic groups responded to the item "fastening a snap (e.g., jacket, bag)." The person separation reliability of the generic scale was 0.93, indicating that 5.19 strata of manual ability can be distinguished in our sample. The average measure of the entire sample was 2.34 logits indicating that the patients' ability level exceeded the scale average difficulty.

Table 2. Final calibration of the six diagnostic groups after the splitting of the DIF items into 2 main groups: asymmetric disorders (CS, CP) and symmetric disorders (RA, SSc, NMDc, NMDa)

Item	Hands involvement	Responded by CS	CP	RA	SSc	NMDc	NMDa	Difficulty* (Logits)	SE (Logits)
a01 Hammering a nail	2C	x						3.60	0.23
a02 Cutting one's nails	2C	x	x					3.32	0.15
a03 Threading a needle	2C	x						3.31	0.23
a04 Peeling potatoes with a knife	2C	x						3.28	0.23
a05 Wrapping up gifts	2C	x						3.20	0.25
a06 Filing one's nails	2C	x						2.82	0.21
a07 Peeling onions	2C	x						2.67	0.24
s08 Shelling hazel nuts	2C			x	x			2.38	0.14
a08 Shelling hazel nuts	2C	x						2.23	0.24
a09 Winding up a wristwatch	2B	x						2.08	0.20
a10 Using a screwdriver	2B	x						2.04	0.22
s01 Hammering a nail	2C			x	x			1.88	0.15
s11 Taking the cap off a bottle	2B			x	x		x	1.84	0.12
a12 Screwing on a nut	2B	x						1.81	0.25
s04 Peeling potatoes with a knife	2C			x	x		x	1.56	0.12
a13 Sharpening a pencil	2C	x	x					1.31	0.16
a11 Taking the cap off a bottle	2B	x						1.09	0.23
s07 Peeling onions	2C			x	x			0.91	0.14
a14 Spreading butter on a slice of bread	2B	x	x					0.74	0.16
b15 Fastening a snap (eg, jacket, bag)	2A	x	x	x	x	x	x	0.68	0.09
a16 Replacing a light bulb	2B	x						0.59	0.30
s06 Filing one's nails	2C			x	x			0.59	0.14
s16 Replacing a light bulb	2B			x	x			0.48	0.15
s05 Wrapping up gifts	2C			x	x			0.46	0.15
a17 Opening mail	2B	x	x					0.41	0.16
a18 Handling a stapler	2A	x						0.19	0.30
b19 Peeling a banana	2B		x			x		-0.19	0.15
b20 Filling a glass with water	2A		x			x	x	-0.20	0.12
b21 Opening a pack of biscuits	2B		x			x	x	-0.20	0.13
s22 Turning on a tap	1					x	x	-0.21	0.15
b23 Opening a car door	1		x		x			-0.48	0.15
a24 Opening a bread box	2A		x					-0.56	0.22
s25 Picking up a can	1			x	x	x	x	-0.93	0.12
b26 Throwing a ball	1		x			x		-0.99	0.16
b27 Brushing one's hair	1	x	x	x	x			-1.20	0.13
a28 Unwrapping candy	2C		x					-1.20	0.23
a22 Turning on a tap	1		x					-1.41	0.26
b29 Washing one's face	1	x	x	x	x			-1.63	0.14
a30 Placing a glass of water on a table	1	x	x					-1.70	0.22
s31 Dealing cards	2B					x		-1.80	0.25
a32 Drinking a glass of water	1	x	x					-1.87	0.25
a33 Handling a 4-colour ballpoint pen with one hand	1	x						-2.02	0.37
a34 Counting banknotes	2A	x						-2.20	0.29
b35 Turning on a radio	1	x		x	x	x		-2.28	0.18
a36 Wiping one's hands	2A		x					-2.32	0.26
s37 Using a fork	1		x			x	x	-2.51	0.21
s38 Turning on a television	1			x	x	x	x	-2.56	0.18
b39 Piling up Lego® blocks	2A		x			x		-2.74	0.21
b40 Using a spoon	1	x		x	x			-3.35	0.17
a38 Turning on a television	1	x	x					-3.43	0.38
a41 Turning off a tap	1	x						-3.54	0.56
s42 Blowing one's nose	1			x	x			-3.93	0.23

* higher logit values indicate a greater manual ability for more difficult activities.

a: asymmetric disorders (CS, CP); s: symmetric disorders (RA, SSc, NMDc, NMDa); b: both symmetric and asymmetric disorders; 1 indicates unimanual activities; 2 indicates bimanual activities manageable in several unimanual steps (2A); requiring stabilization with one hand and digital activity with the other (2B); requiring digital activity from both hands (2C). CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults; SE: standard error.

Manual ability across diagnostic groups

Figure 4 shows the distribution of manual ability across the six diagnostic groups. Significant differences in manual ability measures were observed among diagnoses ($p < 0.001$). The CP and NMDc groups had significantly less manual ability than the RA group, who in turn had less manual ability than CS, SSc and NMDa patients ($p < 0.05$, Dunn's pairwise comparisons).

Comparison of the generic and disease-specific scales

As reported in Table 3, standard errors of patient locations were greater for the generic scale than for disease-specific scales, and a smaller range of patient measures was observed in the generic scale. The generic scale is globally less accurate than the disease-specific scales leading to a higher number of extreme persons. In addition, Table 3 shows that manual ability measures of generic and disease-specific scales were highly correlated (range: 0.94–0.97).

Table 3. Comparison of generic and disease-specific scales

Diagnostic group	Generic scale			Disease-specific scales			Relationship between the measures of generic and disease-specific scales
	Median SE of measures (logits)	Range of measures (logits)	Extreme subjects (%)	Median SE of measures (logits)	Range of measures (logits)	Extreme subjects (%)	R (p-value)
CS	0.50	7.06	11	0.46	8.56	08	0.95 (< 0.001)
CP	0.53	10.87	09	0.49	11.98	11	0.95 (< 0.001)
RA	0.61	8.27	09	0.44	8.68	09	0.96 (< 0.001)
SSc	0.66	8.45	21	0.46	10.64	12	0.95 (< 0.001)
NMDc	0.70	6.78	23	0.53	8.12	15	0.94 (< 0.001)
NMDa	1.08	10.48	40	0.70	11.14	31	0.97 (< 0.001)

CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults; SE: standard error associated to subjects' measure; R: Pearson correlation.

DISCUSSION

The present study investigated the applicability of a generic manual ability scale unbiased by diagnosis across six populations. We analysed previous subject responses gathered during calibrations of disease-specific ABILHAND questionnaires, and we examined similarities and differences in manual ability among diagnostic groups. A unidimensional scale was constructed with 11 items sharing a common location between diagnostic groups and 41 items

having a location specific to asymmetric (CS and CP) or symmetric (NMD, RA, and SSc) disorders. The resulting generic scale revealed that CP and NMD children had significantly less manual ability than RA patients, who in turn had significantly less manual ability than CS, SSc, and NMD adults.

A generic manual ability scale should best meet the requirements of upper limb rehabilitation, insofar as a common instrument with a diagnosis-independent calibration can be used across clinical settings. Of course, the use of a generic scale assumes that individuals achieving identical activities have the same manual ability level regardless of their diagnosis. However, this assumption may not hold true in clinical practice. In our study, we found that only 11 out of 52 items had difficulties unbiased by diagnosis indicating that individuals' underlying diseases may bias the perceived difficulty of manual activities. Using a sample size of 100 patients per diagnostic group, a DIF of 1 logit, namely the approximate amplitude of DIF observed for the items split between symmetric and asymmetric disorders (see Figure 2), in a test containing 10 items or more answered by at least 100 subjects can be detected at a significance level of 0.05 with a power of 95% or more.[28] This indicates that the power of the DIF observed in our study is more than adequate considering the study setup (i.e., test length, sample size, and significance level). Our results differ from those of Simone et al.[27] who found that the 23 CS-specific ABILHAND item scale "can be routinely applied to a variety of motor impairments." These authors argue that the item hierarchy can be successfully preserved across diagnoses. Using our patient responses, we conducted a comparable analysis on the same 23 items from the CS-specific ABILHAND scale as Simone et al.[27] Our findings showed that 21 items (91%) presented a significant DIF, which contrasts with the apparent invariance reported by Simone et al.[27] Two factors may contribute to the observed differences in results: sample size and case mix. Our sample included 732 patients which is significantly more than the 150 subjects in the Simone et al.

study.[27] In addition, the unbalanced case mix in the Simone et al.[27] project (83 CS, 17 multiple sclerosis, 13 ataxia, 10 tetraplegics, 3 Parkinson's disease, and 24 healthy controls) may have concealed possible disease influences on difficulty ratings.

An explicit construct theory initiated the development of disease-specific ABILHAND scales. For each diagnosis, the scale content was selected to delineate a single unidimensional construct, correlated to the patients' functional, clinical, and demographic characteristics.[6-10] The nature of the measured variable, namely manual ability, can be determined by investigating the factors contributing to the hierarchy of manual item difficulty that is observed across diagnoses. To address this issue, we developed an original methodology that combines DIF tests, PCA, and manual activities categorization about their nature. Although an activity is expressed in the same way for all patients, its perceived difficulty may vary according to one's disease or disorder and the specificity of underlying motor impairments. Several studies have also shown that manual ability limitations are, at least partially, related to underlying upper limb impairments.[6, 29] Hence, it is not surprising that disease characteristics contribute to the difficulties experienced in performing manual activities.

The PCA results suggest that the vast majority (85%) of the difficulty variations observed in manual activities across diagnostic groups was explained by two characteristics: 1) the symmetric or asymmetric nature of the disorder (57% of the item difficulty hierarchy variations observed across disorders), and 2) the proximal or distal nature of the disorder (28% of the item difficulty variations). For example, activities requiring greater bimanual involvement (e.g., "peeling potatoes with a knife") tended to be rated as more difficult by patients with asymmetric disorders (CP children and CS adults) than by patients with more symmetric disorders (RA, SSc, NMDc, NMDa). On the other hand, unimanual activities (e.g., "turning on a television") or bimanual activities manageable in several unimanual steps (e.g., "handling a stapler") were rated as less difficult for patients with asymmetric disorders, likely

because these activities can be achieved by exclusively using the unaffected or less affected hand.[7, 30] Activities involving the shoulder (e.g., “drinking a glass of water”) were generally more difficult for NMD and CP patients. Indeed, the NMD groups included several diseases in which proximal segments were more likely to be affected than distal ones (e.g., Duchenne/limb girdle muscular dystrophy, facio-scapulo-humeral dystrophy, spinal muscular atrophy).[10] Moreover, and contrary to other diagnoses, NMD and CP groups included subjects in a wheelchair, which may prevent the achievement of activities such as, “ringing a door bell”, or “replacing a light bulb”. In contrast, digital activities (e.g. “winding up a wristwatch”) were particularly difficult for SSc subjects, who have reduced digital dexterity.[9] Other characteristics of the diseases than their symmetric/asymmetric or proximal/digital nature may explain, even though to a lesser extent, the variations of item difficulty hierarchy between disorders. Activities inducing high mechanical constraints on the upper limb joints (e.g. “screwing on a nut”) presented the highest challenge for RA patients due to wrist and metacarpophalangeal joint involvement.[8] Similar to a previous study,[31] activities related to dressing (e.g., “fastening the zipper of a jacket”) and self-care (e.g., “cutting one’s nails”) were more challenging for children than for adults as well as activities requiring turning something (e.g., “turning on/off a tap”). Parents of unhealthy children may inhibit some activities to prevent risk (e.g., “cutting one’s nails”) or save time (e.g., dressing items).[32] Activities related to eating (e.g., “unwrapping a chocolate bar”) were easier for children than for adults. It can be hypothesized that children are more motivated to compensate their hand impairments by learning adapted strategies (such as breaking down a bimanual activity into several unimanual sequences) for eating activities than for dressing or self-care tasks.[29] It is also important to note that several activities presented a DIF for more than one reason.

Nevertheless, using 11 linked items unbiased by diagnoses, we successfully constructed, from a metric point of view, a unidimensional scale common to six diagnostic groups by separating items with difficulties specific to asymmetric and to symmetric disorders. In our study, the obtained standard errors on items estimates on the generic ABILHAND scale range from 0.09 to 0.56 logits, average 0.20 logits, and correspond to the expected values regarding sample size and targeting.[33] The strong correlations ($R \geq 0.94$) observed between the generic scale and each of the disease-specific ABILHAND scales point out that they measure the same construct, namely, manual ability. However, disease-specific scales which often included a greater number of disease-relevant activities enable more accurate measures (i.e., patient estimates have lower standard errors) than the generic scale. This is most likely due to the fact that disease-specific scales have been constructed to maximize their person separation reliability and therefore also their accuracy. Overall, our findings are consistent with several studies showing that disease-specific instruments are substantially more discriminative and responsive to small deficits than generic instruments.[34, 35] Consequently, this increased sensitivity allows for the detection and quantification of small, yet clinically significant health changes.[11, 12] For example, ABILHAND disease-specific scales should be used to determine pathology impacts on manual ability, to measure clinical changes consecutive to specific treatments, and to tailor interventions to the specific needs of individuals with a particular diagnosis. All of these concerns are important for patients and clinicians in their daily practice. In contrast, the generic ABILHAND scale allows the manual ability of patients with different diagnoses to be compared and can be used, for example, to identify the relative burden of diagnoses, compare various health-care programs, and demonstrate evidence of cost-effectiveness of different healthcare interventions.[11, 12, 36] According to this generic scale, children had, on the whole, less manual ability than adults. This finding is consistent

with previous results[31, 37] showing that children have relatively greater difficulty with manipulation activities than adults.

The generic ABILHAND scale includes 52 items: 11 items sharing a common location between diagnostic groups and 41 items having a location specific to asymmetric or symmetric disorders. The 11 common items were used to establish links that connect the 41 items specific to the symmetry of the disorders to place all measures in the same frame of reference (i.e., on the same "ruler"). From a metric point of view, the common-item linking has enabled the development of a generic scale that can be used to compare subjects with various diagnoses since they are located on one single continuum. However, only one fifth of the items of the "generic" scale are common to several diagnoses. From a clinical point of view, this means that most manual activities present a difficulty that varies according to the underlying diagnosis and that various pathologies may affect differently the achievement of daily activities. The finding that the difficulties of most manual activities were disease-dependent emphasizes the danger of using generic scales without prior investigation of item invariance across diagnostic groups.

COMPETING INTERESTS

None declared.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

CONTRIBUTORSHIP STATEMENT

CA performed the statistical analyses, conducted the literature search, and drafted the manuscript. CA, LV, and MP participated in the data collection. CA, LV, MP, and JLT contributed to the study design and the data analysis. All authors participated in the data interpretation, critically revised the draft of the manuscript for important intellectual content, and contributed to the writing. All authors have read and approved the final manuscript.

DATA SHARING STATEMENT

Dataset is available from the corresponding author at carlyne.arnould@gmail.com

REFERENCES

1. Cardol M, de Jong BA, van den Bos GA, *et al.* Beyond disability: perceived participation in people with a chronic disabling condition. Clin Rehabil 2002;**16**:27–35.
2. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med* 2012;**44**:97–8.
3. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil* 1989;**70**:308–12.
4. Wright BD, Linacre JM. Observations are always ordinal; measurement, however, must be interval. *Arch Phys Med Rehabil* 1989;**70**:857–60.
5. Rasch G. *Probabilistic models for some intelligence and attainment tests.* Chicago: Mesa Press, 1980.
6. Penta M, Tesio L, Arnould C, *et al.* The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke* 2001;**32**:1627–34.
7. Arnould C, Penta M, Renders A, *et al.* ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. *Neurology* 2004;**63**:1045–52.
8. Durez P, Fraselle V, Houssiau F, *et al.* Validation of the ABILHAND questionnaire as a measure of manual ability in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**:1098–105.
9. Vanthuyne M, Smith V, Arat S, *et al.* Validation of a manual ability questionnaire in patients with systemic sclerosis. *Arthritis Rheum* 2009;**61**:695–703.
10. Vandervelde L, Van den Bergh PY, Penta M, *et al.* Validation of the ABILHAND questionnaire to measure manual ability in children and adults with neuromuscular disorders. *J Neurol Neurosurg Psychiatry* 2010;**81**:506–12.

11. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;**27**(3 Suppl):217–32S.

12. Marra CA, Woolcott JC, Kopec JA, *et al.* A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQol and the HAQ) in rheumatoid arthritis. *Soc Sci Med* 2005;**60**:1571–82.

13. Haigh R, Tennant A, Biering-Sorensen F, *et al.* The use of outcome measures in physical medicine and rehabilitation within Europe. *J Rehabil Med* 2001;**33**:273–8.

14. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 2004; **7**(Suppl 1):22–6.

15. Holland PW, Wainer H. *Differential Item Functioning.* Hillsdale, New Jersey: Lawrence Erlbaum, 1993.

16. Tennant A, Penta M, Tesio L, *et al.* Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care* 2004;**42**(1 Suppl): 37–48S.

17. Andrich D. Category ordering and their utility. *Rasch Measurement Transactions* 1996;**9**:464–5.

18. Andrich D. *Rasch analysis for measurement.* London: Sage Publications Ltd; 1988.

19. RUMM Laboratory. *Interpreting RUMM2020: Part 1, dichotomous data.* Perth, Western Australia: RUMM Laboratory, 2004.

20. RUMM Laboratory. *Interpreting RUMM2020: Part 2, polytomous data.* Perth, Western Australia: RUMM Laboratory, 2004.

21. Equating/linking with anchors. *Rasch MeasurementTransactions* 2004;**18**:993.

22. Wright B. Anchoring & standard-errors. *Rasch Measurement Transactions* 1993;**6**:259.

23. Ingebo G. Linking tests with the Rasch model. *Rasch Measurement Transactions* 1997;**11**:549.

24. Lundgren-Nilsson A, Tennant A, Grimby G, *et al.* Cross-diagnostic validity in a generic instrument: an example from the Functional Independence Measure in Scandinavia. *Health Qual Life Outcomes* 2006;**4**:55.

25. Dallmeijer AJ, de Groot V, Roorda LD, *et al.* Cross-diagnostic validity of the SF-36 physical functioning scale in patients with stroke, multiple sclerosis and amyotrophic lateral sclerosis: a study using Rasch analysis. *J Rehabil Med* 2007;**39**:163–9.

26. Wann-Hansson C, Klevsgard R, Hagell P. Cross-diagnostic validity of the Nottingham health profile index of distress (NHPD). *Health Qual Life Outcomes* 2008;**6**:47.

27. Simone A, Rota V, Tesio L, *et al.* Generic ABILHAND questionnaire can measure manual ability across a variety of motor impairments. *Int J Rehabil Res* 2011;**34**:131–40.

28. Scott NW, Fayers PM, Aaronson NK, *et al*. A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *J Clin Epidemiol* 2009;**62:**288–95.

29. Arnould C, Penta M, Thonnard J-L. Hand impairments and their relationship with manual ability in children with cerebral palsy. *J Rehabil Med* 2007;**39**:708–14.

30. Sakzewski L, Ziviani J, Boyd R. The relationship between unimanual capacity and bimanual performance in children with congenital hemiplegia. *Dev Med Child Neurol* 2010;**52**:811–6.

31. Haley SM, Ludlow LH. Applicability of the hierarchical scales of the Tufts Assessment of Motor Performance for school-aged children and adults with disabilities. *Phys Ther* 1992;**72**:191–202.

32. Sperle PA, Ottenbacher KJ, Braun SL, *et al.* Equivalence reliability of the Functional Independence Measure for Children (WeeFIM®) administration methods. *Am J Occup Ther* 1997;**51**:35–41.

33. Linacre JM. Sample Size and Item Calibration Stability. *Rasch Measurement Transactions* 1994;**7**:328.

34. Murawski MM, Miederhoff PA. On the generalizability of statistical expressions of health related quality of life instrument responsiveness: a data synthesis. *Qual Life Res* 1998;**7**:11–22.

35. Wiebe S, Guyatt G, Weaver B, *et al.* Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol* 2003;**56**:52–60.23.

36. Mazur W, Kupiainen H, Pitkäniemi J, *et al.* Comparison between the disease-specific Airways Questionnaire 20 and the generic 15D instruments in COPD. *Health Qual Life Outcomes* 2011;**9**:4.

37. van Eck M, Dallmeijer AJ, van Lith IS, *et al.* Manual ability and its relationship with daily activities in adolescents with cerebral palsy. *J Rehabil Med* 2010;**42**:493–8.

FIGURE LEGENDS

Figure 1. Flow diagram illustrating the analysis process steps. DIF = differential item functioning; PCA = principal component analysis; CS = chronic stroke; CP = cerebral palsy; NMDc = neuromuscular children; NMDa = neuromuscular adults; RA = rheumatoid arthritis; and SSc = systemic sclerosis.

Figure 2. Disease-specific patterns of item difficulty according to the bimanual or proximal nature of the activities showing a Differential Item Functioning. Differences between item difficulty ratings specific to each diagnostic group ($\delta_{specific}$) and the average item difficulty for all diagnoses (δ_{mean}) are shown for each disorder (CS = chronic stroke; CP = cerebral palsy; RA = rheumatoid arthritis; SSc = systemic sclerosis; NMDc = neuromuscular children; and NMDa = neuromuscular adults). Boxes indicate the 25% and 75% limits (the interquartile range); the vertical line inside each box indicates the median; vertical bars outside each box indicate the 10% and 90% limits and dots indicate the 5% and 95% outliers.

Figure 3. PCA results based upon differences between disease-specific difficulty of the split DIF items and the average item difficulty across all diagnoses. CP = cerebral palsy; NMDc = neuromuscular children; RA = rheumatoid arthritis; SSc = systemic sclerosis; and NMDa = neuromuscular adults. CS = chronic stroke;

Figure 4. Box plots showing the distribution of manual ability measures for each diagnosis. Boxes indicate the 25% and 75% limits (the interquartile range); the vertical line inside each box indicates the median; vertical bars outside each box indicate the 10% and 90% limits and dots indicate the 5% and 95% outliers. CP = cerebral palsy; NMDc = neuromuscular children; RA = rheumatoid arthritis; NMDa = neuromuscular adults; SSc = systemic sclerosis; and CS = chronic stroke.

ARTICLE TITLE:

Can manual ability be measured with a generic ABILHAND scale? A cross-sectional study conducted on six diagnostic groups

AUTHORS:

Carlyne Arnould[1] (PhD), Laure Vandervelde[2] (PhD), Charles Sèbiyo Batcho[3] (PT), Massimo Penta[3] (PhD), Jean-Louis Thonnard[3] (PhD)

1 Physical and Occupational Therapy Departments, Paramedical Category, Haute Ecole Louvain en Hainaut, Montignies-sur-Sambre, Belgium

2 Physical Therapy Department, Institut Parnasse Deux-Alice, Haute Ecole Leonard de Vinci, Brussels, Belgium

3 Institute of Neuroscience, Université catholique de Louvain, Brussels, Belgium

CORRESPONDING AUTHOR:

Dr. Carlyne Arnould

Physical and Occupational Therapy Departments

Paramedical Category

Haute Ecole Louvain en Hainaut

Rue Trieu Kaisin, 134

6061 Montignies-sur-Sambre

Belgium

Tel. + 32 498 59 31 33

E-mail. carlyne.arnould@gmail.com

KEYWORDS: motor skills disorders, upper extremity, hand, activities of daily living, questionnaires

WORD COUNTS

The title includes 103 characters.

The abstract includes 297 words.

This manuscript includes 3 tables and 4 figures, 3994 words

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Objectives: Several ABILHAND Rasch-built manual ability scales were previously developed for chronic stroke (CS), cerebral palsy (CP), rheumatoid arthritis (RA), systemic sclerosis (SSc), and neuromuscular disorders (NMD). The present study aimed to explore the applicability of a generic manual ability scale unbiased by diagnosis and to study the nature of manual ability across diagnoses.

Design: cross-sectional study.

Setting: outpatient clinic homes (CS, CP, RA), specialized centers (CP), reference centers (CP, NMD), and university hospitals (SSc).

Participants: 762 patients from six diagnostic groups: 103 CS adults, 113 CP children, 112 RA adults, 156 SSc adults, 124 NMD children and 124 NMD adults.

Primary and secondary outcome measures: manual ability as measured by the ABILHAND disease-specific questionnaires, diagnosis, and nature (i.e., uni- or bi-manual involvement and proximal or distal joints involvement) of the ABILHAND manual activities.

Results: The difficulty of most manual activities was diagnosis-dependent. A principal component analysis highlighted that 57% of the variance in the item difficulty between diagnoses was explained by the symmetric or asymmetric nature of the disorders. A generic scale was constructed, from a metric point of view, with 11 items sharing a common difficulty among diagnoses and 41 items displaying a category-specific location (asymmetric: CS, CP; and symmetric: RA, SSc, NMD). This generic scale showed that CP and NMD children had significantly less manual ability than RA patients, who had significantly less manual ability than CS, SSc, and NMD adults. However, the generic scale was less discriminative and responsive to small deficits than disease-specific instruments.

Conclusions: Our finding that most of the manual item difficulties were disease-dependent emphasizes the danger of using generic scales without prior investigation of item invariance

across diagnostic groups. Nevertheless, a generic manual ability scale could be developed by adjusting and accounting for activities perceived differently in various disorders.

ARTICLE SUMMARY

Article focus:

- To explore the applicability of a generic ABILHAND manual ability scale unbiased by diagnosis across various clinical populations.
- To analyse prior data from cross-sectional studies that developed disease-specific manual ability questionnaires in order to investigate the co-calibration of patient perceived item difficulty on a common metric.
- To better understand the nature of the measured variable, namely manual ability.

Key messages:

- The difficulty of most manual activities was diagnosis-dependent, emphasizing the danger of using generic scales without prior investigation of item invariance across diagnostic groups.
- The vast majority (85%) of the difficulty variations observed in manual activities across diagnostic groups was explained by 1) the symmetric or asymmetric nature of the disorder (57% of the variance) and 2) the proximal or distal nature of the disorder (28% of the variance).
- Although less sensitive than diagnosis-specific scales, a generic manual ability scale could be developed by adjusting and accounting for activities perceived differently in various disorders, which allows quantitative comparisons of manual ability between diagnostic groups.

Strengths and limitations of this study:

- Our study explores a large set of data (732 patients) spread out evenly over 6 diagnostic groups (stroke adults, cerebral palsy children, adults with rheumatoid arthritis, adults with systemic sclerosis, children and adults with neuromuscular disorders).

- Our study proposes an original methodology (combining differential item functioning tests, principal component analysis, and manual activities categorization) that investigates the factors contributing to the hierarchy of manual item difficulty observed across diagnoses allowing the nature of manual ability to be better understood.

INTRODUCTION

One fundamental goal of rehabilitation is to improve the subjects' ability to manage the daily activities necessary for autonomous living.[1] Such an ability belongs to the domain of latent variables concealed within the person, such as pain or intelligence. It cannot be observed directly, but it can be inferred from subject's perceived difficulty in performing activities, also called items, using patient self-reported questionnaires. Over the past decade, questionnaires have therefore become widely used as outcome measures in clinical trials[2] and rating scale data are becoming integral to patient care, prescribing, and policymaking. It is essential that functional rating scales provide scientifically robust and clinically meaningful results to ensure appropriate interpretations and decision-making regarding disease effects, clinical implications, treatment, health policies, and resource allocation. Unfortunately, most rating scales generate ordinal data by summating scores assigned to a set of items representing the intended variable, and metric properties of raw ordinal scores are known to have limited validity.[3, 4] In view of this limitation, the Rasch model[5] is becoming increasingly popular for health measurements because it enables the direct transformation from ordinal scores to linear measures with a constant unit.

Over the last 20 years, our research group has developed several manual ability rating scales (known under the umbrella term of ABILHAND questionnaires) by applying the Rasch model to various diagnostic groups. ABILHAND scales are self-administered questionnaires that measure "manual ability", which is defined as, "the capacity to manage daily activities requiring the use of the upper limbs, whatever the strategies involved".[6] Disease-specific manual ability "rulers" were previously developed for the following patient groups: chronic stroke (CS),[6] cerebral palsy (CP),[7] rheumatoid arthritis (RA),[8] systemic sclerosis (SSc),[9] and neuromuscular disorders (NMD).[10] Each ABILHAND scale has its own Rasch-derived item difficulty calibration, which defines a disease-specific manual ability

measurement continuum. ABILHAND questionnaires present good psychometric qualities, including linearity, unidimensionality, construct validity, and test-retest reliability.

Disease-specific scales, which are highly sensitive and detect small, yet clinically important changes, are frequently used in research because they ensure comprehensive assessment of health aspects directly related to the condition.[11, 12] In contrast, generic scales enable comparisons of various diagnoses and healthcare interventions, which may provide useful data for health policies, cost-effective analyses, and resource allocation.[11, 12] They best meet rehabilitation requirements when disability treatment is not dependent upon a specific underlying diagnosis. [13] For instance, as a single bathroom scale can be used to weigh all patients, a generic manual ability scale would enable quantitative comparisons of the ability to use the upper limbs in daily activities across patients of various diagnoses (and also with healthy subjects).

From a metric point of view, it is possible to co-calibrate various disease-specific ABILHAND questionnaire items on the same scale, provided that the scales are based on an identical unidimensional construct.[14] In theory, and similar to the graduations of a metric ruler, items should have the same difficulty for all diagnostic groups, regardless of the disease being measured. Nevertheless, the main implicit assumption made by the users of generic scales is that the difficulties of daily activities are invariant across diagnoses. However, in practice, item difficulty hierarchy may vary across groups, demonstrating Differential Item Functioning (DIF).[15] The Rasch model can be used to test the invariance of item difficulty hierarchy and to accommodate for DIF.[16] When the items of a generic scale are unstable across diagnoses, the measurements generated by them cannot be used to make meaningful comparisons.

The present study explored the applicability of a generic ABILHAND manual ability scale, which is unbiased by diagnoses, across various clinical populations. Setting out this objective,

we also intended to improve the current understanding of the nature of manual ability and especially its interaction with diagnosis. We analysed prior data from cross-sectional studies that developed disease-specific manual ability questionnaires in order to investigate the co-calibration of patient perceived item difficulty on a common metric.

METHODS

Subjects

Data from 732 subjects, who previously provided informed consent, were analysed. Patients with the following disorders were evaluated: 103 CS adults,[6] 113 CP children,[7] 112 RA adults,[8] 156 SSc adults,[9] 124 NMD children (NMDc) and 124 NMD adults (NMDa).[10] Table 1 provides patient characteristics. The ethics committee of the Université catholique de Louvain, Faculty of Medicine in Brussels, Belgium, authorized and approved the study.

Table 1. Sample characteristics (n = 732)

Variables	CS	CP	RA	SSc	NMDc	NMDa
Number of subjects	103	113	112	156	124	124
Age, y, mean (range)	63 (24-84)	10 (6-15)	55 (25-82)	54 (21-82)	10 (6-16)	47 (16-80)
Sex						
Males	64	67	29	32	84	69
Females	39	46	83	124	40	55
Diagnosis	R hemi: 55	tetra: 35	No UL disorder: 20	lcSSc: 104	DMD/BMD or LGMD: 47	MD: 24
	L hemi: 48	di: 24	DH disorder: 9	dcSSc: 33	HN: 35	HN: 24
		R hemi: 26	NDH disorder: 4	lSSc: 19	SMA: 3	DMD/BMD or LGMD: 19
		L hemi: 28	2 UL disorder: 79		Others (CM, CMD, PPS, …): 29	SMA: 7
						FSHD: 7
						Others (CM, CMD, PPS, …): 43

CS = chronic stroke; CP = cerebral palsy; RA: rheumatoid arthritis; NMDc = neuromuscular children; NMDa = neuromuscular adults; SSc = systemic sclerosis; R hemi = right hemiplegia; L hemi = left hemiplegia; tetra = tetraplegia; di = diplegia; UL = upper limb; DH = dominant hand; NDH = non-dominant hand; DMD = Duchenne muscular dystrophy; BMD = Becker muscular dystrophy; LGMD = limb girdle muscular dystrophy; HN = hereditary neuropathy; SMA = spinal muscular atrophy; CM = congenital myopathy; CMD = congenital muscular dystrophy; PPS = post-polio syndrome; MD = myotonic dystrophy; FSHD = facio-scapulo-humeral dystrophy; lcSSc = limited cutaneous SSc; dcSSc = diffuse cutaneous SSc; lSSc = limited SSc.

Manual ability measure

Original data included 83 manual activities shared by at least 2 diagnostic groups (the 83 items are provided in the supplementary table). Original items covered different domains of

daily living such as feeding, grooming, or dressing and were selected in previous studies based on literature review and patient and experts interviews. Twelve items were child-specific (e.g., ""throwing a ball"), 19 were adult-specific (e.g., "hammering a nail"), and 52 were common to both groups (e.g. "buttoning up trousers"). Adult patients and children's parents provided their perceived difficulty in performing each activity based upon a three-level response scale: impossible (0), difficult (1) or easy (2). Each activity had to be completed without technical or human assistance and irrespective of the limb(s) and adaptive strategies used. Missing values were included when a given diagnostic group did not provide responses for a particular item, as the activity may not have been submitted to a group. The nature of the items was assessed by ten occupational or physical therapists according to the following criteria: uni- or bi-manual involvement required to perform the activity; and involvement of proximal or distal joints.

Data analysis

All responses were analysed with RUMM2020®, a Rasch analysis computer program. The Rasch model[5] can be used to estimate, on a single manual ability construct, the location of each patient , i.e. their manual ability, the location of each item, i.e. the difficulty of the manual activities, and the location of each threshold between successive categories of the response scale, i.e. the locations along the latent construct at which two successive categories are equally likely to be observed. The model can be used to verify that successive response categories for each item represent increasing levels of ability and that thresholds between successive response categories are located in the anticipated order.[17]

The model also requires that the probability of endorsing any response category to an item depends solely on the subject's ability, the item difficulty and the location of the threshold between adjacent response categories. In the case of manual ability measurement, no attribute of the person - such as diagnosis - besides manual ability is theorized to account for the

probability of choosing a given response to a given item. The similarity between the observed and expected responses can be investigated using a χ^2 fit statistic computed over 5 class intervals (CI) of patients with increasing ability.[18] Items with a p-value lower than 0.05 indicate a threat to the fit requirement.

Invariance of the item difficulty hierarchy

Unidimensionality also requires that patients with identical ability, but different diagnoses, have the same probability of succeeding any particular item. Consequently, the invariance of item difficulties across patient diagnostic groups must be controlled using Differential Item Functioning (DIF) tests.[15] To investigate the invariance of item difficulty hierarchy, a two-way ANOVA was computed on the standardized residuals of the different CIs [19, 20]; the first factor was the diagnostic group and the second factor was the CI of increasing manual ability. Significant diagnostic main effects represented group differences in item difficulty hierarchy. A solution to the presence of DIF by diagnosis is the removal of items showing difficulty variations. Another solution is to allow for the variations that exist across DIF items by splitting them into disease-specific items, one for each diagnostic, with a difficulty peculiar to the corresponding diagnosis.[16] In this case, the different diagnostic groups can be compared on the same continuum even if they have specific items provided that there are common linking items unbiased by DIF.

Analysis process

Two different approaches can be used to combine data from different scales responded by different samples. The 'co-calibration', also called 'concurrent equating', merges all items together as one scale with empty spaces for missing values. The 'anchoring' approach anchors items that are common to all diagnoses and then includes diagnosis-specific items in the same frame of reference. The anchoring approach requires that the common linking items be free of DIF,[21–23] which was not the case in our dataset. Therefore, the co-calibration approach,

also applied in previous rehabilitation studies,[24–27] was followed and the analysis process is illustrated in Figure 1. The first step in the data analysis was to co-calibrate the ABILHAND data of all diagnostic groups by analysing all responses (n=732) to the 83 items. The second step was to remove items with disordered thresholds and items that misfit a unidimensional variable (i.e., presenting a χ^2 p-value < 0.05). In the third step, the invariance of item difficulty hierarchy was detected across diagnostic groups through DIF tests. The fourth step consisted in splitting the items presenting a DIF by diagnosis providing one specific item for each diagnostic group who answered the item.[16] In the fifth step, a principal component analysis (PCA) was performed to identify the potential factors explaining item difficulty hierarchy variations observed across the diagnostic groups. The PCA was performed on the differences between item difficulty specific to each diagnostic group and the average item difficulty for all diagnoses as these differences reflect disease-specific patterns of item difficulty. In the sixth step, the items presenting a DIF among diagnoses (detected in the third step) were split into two main groups: asymmetric disorders (CS and CP) and symmetric disorders (RA, SSc, NMD). Finally, the seventh step included successive analyses performed to remove items with disordered thresholds, misfitting items, and items presenting a DIF by diagnosis for another reason than the symmetric/asymmetric nature of the disorders. So, a generic co-calibrated scale was created and manual ability was compared among diagnostic groups using a Kruskal-Wallis ANOVA of ranks and Dunn's method for pairwise multiple comparisons. Finally, the metric properties of the generic scale were compared with ABILHAND disease-specific scale properties.

RESULTS

Invariance of the item difficulty hierarchy

Thirty-two of the initial 83 items were deleted due to the unidimensionality requirement violation. Assessment of invariance from the remaining 51 unidimensional items showed that

13 items shared a common location between diagnostic groups. Thirty-eight items presented a DIF and were split into a total of 152 items with diagnosis-specific locations. Differences between item difficulty specific to each diagnostic group and the mean item difficulty for all diagnoses were computed to identify disease-specific patterns of item difficulty. Positive values indicated that the items were more difficult for a particular diagnosis than average while negative values indicated that they were easier than average (Figure 2). With respect to disease-specific item difficulties, bimanual activities, such as "spreading butter on a slice of bread," presented a greater challenge for patients with asymmetric disorders (CS and CP) than for patients with symmetric diagnoses (RA, SSc, NMDc, NMDa). Conversely, unimanual activities, such as "turning off a tap," were perceived as easier in asymmetric disordered patients. About 85% of the DIF items were related to the unimanual or bimanual nature of the activities.

In addition, we found that proximal activities, such as "ringing a door bell," were categorized as more difficult for NMD, CP, and CS patients compared to RA and SSc patients. In contrast, digital activities, such as "counting banknotes," presented the greatest challenge for SSc subjects who primary had a distal impairment. Approximately one third of the DIF items were concerned with the proximal or distal nature of the activities. It should be noted that some items fit both criteria. Moreover, DIF activities were related, to a lesser extent, to other factors such as age (about 30% of the items) or mechanical constraints induced in the upper limb joints (about 10–15% of the items).

PCA on diagnosis-specific-to-average item difficulty differences

PCA results showed that 57% of the variation of item difficulty hierarchy between diagnostic groups was explained by the symmetric or asymmetric nature of the disorders (Figure 3). Indeed, CS adults and CP children were located at one extremity of the first PCA component while symmetric disorders were located at the other extremity. The second PCA component

explained 28% of the variation of item difficulty hierarchy between diagnostic groups and distinguished patients expressing greater difficulties with proximal activities, such as NMD and CP, from more distal disorders such as SSc.

A "generic" ABILHAND manual ability scale

Based upon PCA results, the DIF items were split into two main groups: asymmetric (CS and CP) and symmetric (RA, SSc, NMD) disorders. When the 13 items sharing a common location between diagnostic groups were co-calibrated with the 38 DIF items split into a total of 75 items (one item was responded neither by CS nor CP subjects) with locations specific to either asymmetric or symmetric disorders, 2 items with disordered thresholds, 7 misfitting items, and 27 remaining DIF items were removed. The resulting 52-item generic scale included 11 items sharing a common location between diagnostic groups and 41 items with locations specific to asymmetric (27 items) or symmetric (14 items) disorders. The 52 items are listed in Table 2 in order of decreasing difficulty (range: 3.60 to -3.93 logits).

Hand involvement, whether particular groups responded to each item, and item difficulty with standard errors (mean: 0.20 logits; range: 0.09 to 0.56 logits) are also reported. It should be noted that only one diagnostic group responded to 21 items (40%) of the generic scale, while two or three diagnostic groups responded to as many as 12 items (23%). All diagnostic groups responded to the item "fastening a snap (e.g., jacket, bag)." The person separation reliability of the generic scale was 0.93, indicating that 5.19 strata of manual ability can be distinguished in our sample. The average measure of the entire sample was 2.34 logits indicating that the patients' ability level exceeded the scale average difficulty.

Table 2. Final calibration of the six diagnostic groups after the splitting of the DIF items into 2 main groups: asymmetric disorders (CS, CP) and symmetric disorders (RA, SSc, NMDc, NMDa)

Item	Hands involvement	Responded by CS	CP	RA	SSc	NMDc	NMDa	Difficulty* (Logits)	SE (Logits)
a01 Hammering a nail	2C	x						3.60	0.23
a02 Cutting one's nails	2C	x	x					3.32	0.15
a03 Threading a needle	2C	x						3.31	0.23
a04 Peeling potatoes with a knife	2C	x						3.28	0.23
a05 Wrapping up gifts	2C	x						3.20	0.25
a06 Filing one's nails	2C	x						2.82	0.21
a07 Peeling onions	2C	x						2.67	0.24
s08 Shelling hazel nuts	2C			x	x			2.38	0.14
a08 Shelling hazel nuts	2C	x						2.23	0.24
a09 Winding up a wristwatch	2B	x						2.08	0.20
a10 Using a screwdriver	2B	x						2.04	0.22
s01 Hammering a nail	2C			x	x			1.88	0.15
s11 Taking the cap off a bottle	2B			x	x		x	1.84	0.12
a12 Screwing on a nut	2B	x						1.81	0.25
s04 Peeling potatoes with a knife	2C			x	x		x	1.56	0.12
a13 Sharpening a pencil	2C	x	x					1.31	0.16
a11 Taking the cap off a bottle	2B	x						1.09	0.23
s07 Peeling onions	2C			x	x			0.91	0.14
a14 Spreading butter on a slice of bread	2B	x	x					0.74	0.16
b15 Fastening a snap (eg, jacket, bag)	2A	x	x	x	x	x	x	0.68	0.09
a16 Replacing a light bulb	2B	x						0.59	0.30
s06 Filing one's nails	2C			x	x			0.59	0.14
s16 Replacing a light bulb	2B			x	x			0.48	0.15
s05 Wrapping up gifts	2C			x	x			0.46	0.15
a17 Opening mail	2B	x	x					0.41	0.16
a18 Handling a stapler	2A	x						0.19	0.30
b19 Peeling a banana	2B		x			x		-0.19	0.15
b20 Filling a glass with water	2A		x			x	x	-0.20	0.12
b21 Opening a pack of biscuits	2B		x			x	x	-0.20	0.13
s22 Turning on a tap	1					x	x	-0.21	0.15
b23 Opening a car door	1		x		x			-0.48	0.15
a24 Opening a bread box	2A		x					-0.56	0.22
s25 Picking up a can	1			x	x	x	x	-0.93	0.12
b26 Throwing a ball	1		x			x		-0.99	0.16
b27 Brushing one's hair	1	x	x	x	x			-1.20	0.13
a28 Unwrapping candy	2C		x					-1.20	0.23
a22 Turning on a tap	1		x					-1.41	0.26
b29 Washing one's face	1	x	x	x	x			-1.63	0.14
a30 Placing a glass of water on a table	1	x	x					-1.70	0.22
s31 Dealing cards	2B					x		-1.80	0.25
a32 Drinking a glass of water	1	x	x					-1.87	0.25
a33 Handling a 4-colour ballpoint pen with one hand	1	x						-2.02	0.37
a34 Counting banknotes	2A	x						-2.20	0.29
b35 Turning on a radio	1	x		x	x	x		-2.28	0.18
a36 Wiping one's hands	2A		x					-2.32	0.26
s37 Using a fork	1		x			x	x	-2.51	0.21
s38 Turning on a television	1			x	x	x	x	-2.56	0.18
b39 Piling up Lego® blocks	2A		x			x		-2.74	0.21
b40 Using a spoon	1	x		x	x			-3.35	0.17
a38 Turning on a television	1	x	x					-3.43	0.38
a41 Turning off a tap	1	x						-3.54	0.56
s42 Blowing one's nose	1			x	x			-3.93	0.23

* higher logit values indicate a greater manual ability for more difficult activities.

a: asymmetric disorders (CS, CP); s: symmetric disorders (RA, SSc, NMDc, NMDa); b: both symmetric and asymmetric disorders; 1 indicates unimanual activities; 2 indicates bimanual activities manageable in several unimanual steps (2A); requiring stabilization with one hand and digital activity with the other (2B); requiring digital activity from both hands (2C). CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults; SE: standard error.

Manual ability across diagnostic groups

Figure 4 shows the distribution of manual ability across the six diagnostic groups. Significant differences in manual ability measures were observed among diagnoses ($p < 0.001$). The CP and NMDc groups had significantly less manual ability than the RA group, who in turn had less manual ability than CS, SSc and NMDa patients ($p < 0.05$, Dunn's pairwise comparisons).

Comparison of the generic and disease-specific scales

As reported in Table 3, standard errors of patient locations were greater for the generic scale than for disease-specific scales, and a smaller range of patient measures was observed in the generic scale. The generic scale is globally less accurate than the disease-specific scales leading to a higher number of extreme persons. In addition, Table 3 shows that manual ability measures of generic and disease-specific scales were highly correlated (range: 0.94–0.97).

Table 3. Comparison of generic and disease-specific scales

Diagnostic group	Generic scale			Disease-specific scales			Relationship between the measures of generic and disease-specific scales
	Median SE of measures (logits)	Range of measures (logits)	Extreme subjects (%)	Median SE of measures (logits)	Range of measures (logits)	Extreme subjects (%)	R (p-value)
CS	0.50	7.06	11	0.46	8.56	08	0.95 (< 0.001)
CP	0.53	10.87	09	0.49	11.98	11	0.95 (< 0.001)
RA	0.61	8.27	09	0.44	8.68	09	0.96 (< 0.001)
SSc	0.66	8.45	21	0.46	10.64	12	0.95 (< 0.001)
NMDc	0.70	6.78	23	0.53	8.12	15	0.94 (< 0.001)
NMDa	1.08	10.48	40	0.70	11.14	31	0.97 (< 0.001)

CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults; SE: standard error associated to subjects' measure; R: Pearson correlation.

DISCUSSION

The present study investigated the applicability of a generic manual ability scale unbiased by diagnosis across six populations. We analysed previous subject responses gathered during calibrations of disease-specific ABILHAND questionnaires, and we examined similarities and differences in manual ability among diagnostic groups. A unidimensional scale was constructed with 11 items sharing a common location between diagnostic groups and 41 items

having a location specific to asymmetric (CS and CP) or symmetric (NMD, RA, and SSc) disorders. The resulting generic scale revealed that CP and NMD children had significantly less manual ability than RA patients, who in turn had significantly less manual ability than CS, SSc, and NMD adults.

A generic manual ability scale should best meet the requirements of upper limb rehabilitation, insofar as a common instrument with a diagnosis-independent calibration can be used across clinical settings. Of course, the use of a generic scale assumes that individuals achieving identical activities have the same manual ability level regardless of their diagnosis. However, this assumption may not hold true in clinical practice. In our study, we found that only 11 out of 52 items had difficulties unbiased by diagnosis indicating that individuals' underlying diseases may bias the perceived difficulty of manual activities. Using a sample size of 100 patients per diagnostic group, a DIF of 1 logit, namely the approximate amplitude of DIF observed for the items split between symmetric and asymmetric disorders (see Figure 2), in a test containing 10 items or more answered by at least 100 subjects can be detected at a significance level of 0.05 with a power of 95% or more.[28] This indicates that the power of the DIF observed in our study is more than adequate considering the study setup (i.e., test length, sample size, and significance level). Our results differ from those of Simone et al.[27] who found that the 23 CS-specific ABILHAND item scale "can be routinely applied to a variety of motor impairments." These authors argue that the item hierarchy can be successfully preserved across diagnoses. Using our patient responses, we conducted a comparable analysis on the same 23 items from the CS-specific ABILHAND scale as Simone et al.[27] Our findings showed that 21 items (91%) presented a significant DIF, which contrasts with the apparent invariance reported by Simone et al.[27] Two factors may contribute to the observed differences in results: sample size and case mix. Our sample included 732 patients which is significantly more than the 150 subjects in the Simone et al.

study.[27] In addition, the unbalanced case mix in the Simone et al.[27] project (83 CS, 17 multiple sclerosis, 13 ataxia, 10 tetraplegics, 3 Parkinson's disease, and 24 healthy controls) may have concealed possible disease influences on difficulty ratings.

An explicit construct theory initiated the development of disease-specific ABILHAND scales. For each diagnosis, the scale content was selected to delineate a single unidimensional construct, correlated to the patients' functional, clinical, and demographic characteristics.[6-10] The nature of the measured variable, namely manual ability, can be determined by investigating the factors contributing to the hierarchy of manual item difficulty that is observed across diagnoses. To address this issue, we developed an original methodology that combines DIF tests, PCA, and manual activities categorization about their nature. Although an activity is expressed in the same way for all patients, its perceived difficulty may vary according to one's disease or disorder and the specificity of underlying motor impairments. Several studies have also shown that manual ability limitations are, at least partially, related to underlying upper limb impairments.[6, 29] Hence, it is not surprising that disease characteristics contribute to the difficulties experienced in performing manual activities.

The PCA results suggest that the vast majority (85%) of the difficulty variations observed in manual activities across diagnostic groups was explained by two characteristics: 1) the symmetric or asymmetric nature of the disorder (57% of the item difficulty hierarchy variations observed across disorders), and 2) the proximal or distal nature of the disorder (28% of the item difficulty variations). For example, activities requiring greater bimanual involvement (e.g., "peeling potatoes with a knife") tended to be rated as more difficult by patients with asymmetric disorders (CP children and CS adults) than by patients with more symmetric disorders (RA, SSc, NMDc, NMDa). On the other hand, unimanual activities (e.g., "turning on a television") or bimanual activities manageable in several unimanual steps (e.g., "handling a stapler") were rated as less difficult for patients with asymmetric disorders, likely

because these activities can be achieved by exclusively using the unaffected or less affected hand.[7, 30] Activities involving the shoulder (e.g., “drinking a glass of water”) were generally more difficult for NMD and CP patients. Indeed, the NMD groups included several diseases in which proximal segments were more likely to be affected than distal ones (e.g., Duchenne/limb girdle muscular dystrophy, facio-scapulo-humeral dystrophy, spinal muscular atrophy).[10] Moreover, and contrary to other diagnoses, NMD and CP groups included subjects in a wheelchair, which may prevent the achievement of activities such as, “ringing a door bell”, or “replacing a light bulb”. In contrast, digital activities (e.g. “winding up a wristwatch”) were particularly difficult for SSc subjects, who have reduced digital dexterity.[9] Other characteristics of the diseases than their symmetric/asymmetric or proximal/digital nature may explain, even though to a lesser extent, the variations of item difficulty hierarchy between disorders. Activities inducing high mechanical constraints on the upper limb joints (e.g. “screwing on a nut”) presented the highest challenge for RA patients due to wrist and metacarpophalangeal joint involvement.[8] Similar to a previous study,[31] activities related to dressing (e.g., “fastening the zipper of a jacket”) and self-care (e.g., “cutting one’s nails”) were more challenging for children than for adults as well as activities requiring turning something (e.g., “turning on/off a tap”). Parents of unhealthy children may inhibit some activities to prevent risk (e.g., “cutting one’s nails”) or save time (e.g., dressing items).[32] Activities related to eating (e.g., “unwrapping a chocolate bar”) were easier for children than for adults. It can be hypothesized that children are more motivated to compensate their hand impairments by learning adapted strategies (such as breaking down a bimanual activity into several unimanual sequences) for eating activities than for dressing or self-care tasks.[29] It is also important to note that several activities presented a DIF for more than one reason.

Nevertheless, using 11 linked items unbiased by diagnoses, we successfully constructed, from a metric point of view, a unidimensional scale common to six diagnostic groups by separating items with difficulties specific to asymmetric and to symmetric disorders. In our study, the obtained standard errors on items estimates on the generic ABILHAND scale range from 0.09 to 0.56 logits, average 0.20 logits, and correspond to the expected values regarding sample size and targeting.[33] The strong correlations ($R \geq 0.94$) observed between the generic scale and each of the disease-specific ABILHAND scales point out that they measure the same construct, namely, manual ability. However, disease-specific scales which often included a greater number of disease-relevant activities enable more accurate measures (i.e., patient estimates have lower standard errors) than the generic scale. This is most likely due to the fact that disease-specific scales have been constructed to maximize their person separation reliability and therefore also their accuracy. Overall, our findings are consistent with several studies showing that disease-specific instruments are substantially more discriminative and responsive to small deficits than generic instruments.[34, 35] Consequently, this increased sensitivity allows for the detection and quantification of small, yet clinically significant health changes.[11, 12] For example, ABILHAND disease-specific scales should be used to determine pathology impacts on manual ability, to measure clinical changes consecutive to specific treatments, and to tailor interventions to the specific needs of individuals with a particular diagnosis. All of these concerns are important for patients and clinicians in their daily practice. In contrast, the generic ABILHAND scale allows the manual ability of patients with different diagnoses to be compared and can be used, for example, to identify the relative burden of diagnoses, compare various health-care programs, and demonstrate evidence of cost-effectiveness of different healthcare interventions.[11, 12, 36] According to this generic scale, children had, on the whole, less manual ability than adults. This finding is consistent

with previous results[31, 37] showing that children have relatively greater difficulty with manipulation activities than adults.

The generic ABILHAND scale includes 52 items: 11 items sharing a common location between diagnostic groups and 41 items having a location specific to asymmetric or symmetric disorders. The 11 common items were used to establish links that connect the 41 items specific to the symmetry of the disorders to place all measures in the same frame of reference (i.e., on the same "ruler"). From a metric point of view, the common-item linking has enabled the development of a generic scale that can be used to compare subjects with various diagnoses since they are located on one single continuum. However, only one fifth of the items of the "generic" scale are common to several diagnoses. From a clinical point of view, this means that most manual activities present a difficulty that varies according to the underlying diagnosis and that various pathologies may affect differently the achievement of daily activities. The finding that the difficulties of most manual activities were disease-dependent emphasizes the danger of using generic scales without prior investigation of item invariance across diagnostic groups.

COMPETING INTERESTS

None declared.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

CONTRIBUTORSHIP STATEMENT

CA performed the statistical analyses, conducted the literature search, and drafted the manuscript. CA, LV, and MP participated in the data collection. CA, LV, MP, and JLT contributed to the study design and the data analysis. All authors participated in the data interpretation, critically revised the draft of the manuscript for important intellectual content, and contributed to the writing. All authors have read and approved the final manuscript.

DATA SHARING STATEMENT

Dataset is available from the corresponding author at carlyne.arnould@gmail.com

REFERENCES

1. Cardol M, de Jong BA, van den Bos GA, *et al.* Beyond disability: perceived participation in people with a chronic disabling condition. Clin Rehabil 2002;**16**:27–35.
2. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med* 2012;**44**:97–8.
3. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil* 1989;**70**:308–12.
4. Wright BD, Linacre JM. Observations are always ordinal; measurement, however, must be interval. *Arch Phys Med Rehabil* 1989;**70**:857–60.
5. Rasch G. *Probabilistic models for some intelligence and attainment tests.* Chicago: Mesa Press, 1980.
6. Penta M, Tesio L, Arnould C, *et al.* The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke* 2001;**32**:1627–34.
7. Arnould C, Penta M, Renders A, *et al.* ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. *Neurology* 2004;**63**:1045–52.
8. Durez P, Fraselle V, Houssiau F, *et al.* Validation of the ABILHAND questionnaire as a measure of manual ability in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**:1098–105.
9. Vanthuyne M, Smith V, Arat S, *et al.* Validation of a manual ability questionnaire in patients with systemic sclerosis. *Arthritis Rheum* 2009;**61**:695–703.
10. Vandervelde L, Van den Bergh PY, Penta M, *et al.* Validation of the ABILHAND questionnaire to measure manual ability in children and adults with neuromuscular disorders. *J Neurol Neurosurg Psychiatry* 2010;**81**:506–12.

11. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;**27**(3 Suppl):217–32S.

12. Marra CA, Woolcott JC, Kopec JA, *et al.* A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQol and the HAQ) in rheumatoid arthritis. *Soc Sci Med* 2005;**60**:1571–82.

13. Haigh R, Tennant A, Biering-Sorensen F, *et al.* The use of outcome measures in physical medicine and rehabilitation within Europe. *J Rehabil Med* 2001;**33**:273–8.

14. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 2004; **7**(Suppl 1):22–6.

15. Holland PW, Wainer H. *Differential Item Functioning.* Hillsdale, New Jersey: Lawrence Erlbaum, 1993.

16. Tennant A, Penta M, Tesio L, *et al.* Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care* 2004;**42**(1 Suppl): 37–48S.

17. Andrich D. Category ordering and their utility. *Rasch Measurement Transactions* 1996;**9**:464–5.

18. Andrich D. *Rasch analysis for measurement.* London: Sage Publications Ltd; 1988.

19. RUMM Laboratory. *Interpreting RUMM2020: Part 1, dichotomous data.* Perth, Western Australia: RUMM Laboratory, 2004.

20. RUMM Laboratory. *Interpreting RUMM2020: Part 2, polytomous data.* Perth, Western Australia: RUMM Laboratory, 2004.

21. Equating/linking with anchors. *Rasch MeasurementTransactions* 2004;**18**:993.

22. Wright B. Anchoring & standard-errors. *Rasch Measurement Transactions* 1993;**6**:259.

23. Ingebo G. Linking tests with the Rasch model. *Rasch Measurement Transactions* 1997;**11**:549.

24. Lundgren-Nilsson A, Tennant A, Grimby G, *et al.* Cross-diagnostic validity in a generic instrument: an example from the Functional Independence Measure in Scandinavia. *Health Qual Life Outcomes* 2006;**4**:55.

25. Dallmeijer AJ, de Groot V, Roorda LD, *et al.* Cross-diagnostic validity of the SF-36 physical functioning scale in patients with stroke, multiple sclerosis and amyotrophic lateral sclerosis: a study using Rasch analysis. *J Rehabil Med* 2007;**39**:163–9.

26. Wann-Hansson C, Klevsgard R, Hagell P. Cross-diagnostic validity of the Nottingham health profile index of distress (NHPD). *Health Qual Life Outcomes* 2008;**6**:47.

27. Simone A, Rota V, Tesio L, *et al.* Generic ABILHAND questionnaire can measure manual ability across a variety of motor impairments. *Int J Rehabil Res* 2011;**34**:131–40.

28. Scott NW, Fayers PM, Aaronson NK, *et al*. A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *J Clin Epidemiol* 2009;**62:**288–95.

29. Arnould C, Penta M, Thonnard J-L. Hand impairments and their relationship with manual ability in children with cerebral palsy. *J Rehabil Med* 2007;**39**:708–14.

30. Sakzewski L, Ziviani J, Boyd R. The relationship between unimanual capacity and bimanual performance in children with congenital hemiplegia. *Dev Med Child Neurol* 2010;**52**:811–6.

31. Haley SM, Ludlow LH. Applicability of the hierarchical scales of the Tufts Assessment of Motor Performance for school-aged children and adults with disabilities. *Phys Ther* 1992;**72**:191–202.

32. Sperle PA, Ottenbacher KJ, Braun SL, *et al.* Equivalence reliability of the Functional Independence Measure for Children (WeeFIM®) administration methods. *Am J Occup Ther* 1997;**51**:35–41.

33. Linacre JM. Sample Size and Item Calibration Stability. *Rasch Measurement Transactions* 1994;**7**:328.

34. Murawski MM, Miederhoff PA. On the generalizability of statistical expressions of health related quality of life instrument responsiveness: a data synthesis. *Qual Life Res* 1998;**7**:11–22.

35. Wiebe S, Guyatt G, Weaver B, *et al.* Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol* 2003;**56**:52–60.23.

36. Mazur W, Kupiainen H, Pitkäniemi J, *et al.* Comparison between the disease-specific Airways Questionnaire 20 and the generic 15D instruments in COPD. *Health Qual Life Outcomes* 2011;**9**:4.

37. van Eck M, Dallmeijer AJ, van Lith IS, *et al.* Manual ability and its relationship with daily activities in adolescents with cerebral palsy. *J Rehabil Med* 2010;**42**:493–8.

FIGURE LEGENDS

Figure 1. Flow diagram illustrating the analysis process steps. DIF = differential item functioning; PCA = principal component analysis; CS = chronic stroke; CP = cerebral palsy; NMDc = neuromuscular children; NMDa = neuromuscular adults; RA = rheumatoid arthritis; and SSc = systemic sclerosis.

Figure 2. Disease-specific patterns of item difficulty according to the bimanual or proximal nature of the activities showing a Differential Item Functioning. Differences between item difficulty ratings specific to each diagnostic group ($\delta_{specific}$) and the average item difficulty for all diagnoses (δ_{mean}) are shown for each disorder (CS = chronic stroke; CP = cerebral palsy; RA = rheumatoid arthritis; SSc = systemic sclerosis; NMDc = neuromuscular children; and NMDa = neuromuscular adults). Boxes indicate the 25% and 75% limits (the interquartile range); the vertical line inside each box indicates the median; vertical bars outside each box indicate the 10% and 90% limits and dots indicate the 5% and 95% outliers.

Figure 3. PCA results based upon differences between disease-specific difficulty of the split DIF items and the average item difficulty across all diagnoses. CP = cerebral palsy; NMDc = neuromuscular children; RA = rheumatoid arthritis; SSc = systemic sclerosis; and NMDa = neuromuscular adults. CS = chronic stroke;

Figure 4. Box plots showing the distribution of manual ability measures for each diagnosis. Boxes indicate the 25% and 75% limits (the interquartile range); the vertical line inside each box indicates the median; vertical bars outside each box indicate the 10% and 90% limits and dots indicate the 5% and 95% outliers. CP = cerebral palsy; NMDc = neuromuscular children; RA = rheumatoid arthritis; NMDa = neuromuscular adults; SSc = systemic sclerosis; and CS = chronic stroke.

Supplementary table. Original set of items answered by at least two diagnostic groups

Item	Submitted to*					
	CS	CP	RA	SSc	NMDc	NMDa
Eating						
Using a spoon	x		x	x		
Using a fork		x			x	x
Cutting meat	**X**	x	**X**	**X**	x	x
Eating a sandwich	x	x	x	x		
Picking up a can	x	x	**X**	x	x	x
Placing a glass of water on a table	x	x	x	x	x	x
Drinking a glass of water	x	x	x	x	x	x
Filling a glass with water		**X**			**X**	**X**
Unscrewing a bottle cap		**X**			x	x
Taking the cap off a bottle	**X**		**X**	**X**		**X**
Spreading butter on a slice of bread	**X**	x	x	**X**	**X**	**X**
Opening a bread box		**X**			**X**	**X**
Tearing open a pack of chips	**X**	**X**	**X**	**X**	**X**	
Unwrapping a chocolate bar	**X**	**X**	x	**X**	**X**	
Opening a pack of biscuits		x			**X**	**X**
Unwrapping candy		x			x	x
Peeling a banana		x			x	
Shelling hazel nuts	**X**		x	**X**		
Opening a screw-topped jar	**X**	**X**	**X**	**X**	x	x
Peeling onions	**X**		**X**	**X**		
Peeling potatoes with a knife	**X**		**X**	**X**		x
Making pancake batter	x		x	x		
Grooming						
Opening the cap of a toothpaste tube		**X**			**X**	**X**
Squeezing toothpaste onto a toothbrush	**X**	**X**	x	x	**X**	**X**
Brushing one's teeth	x	x	x	x	x	x
Brushing one's hair	x	x	**X**	**X**		
Combing one's hair	x		**X**	x	x	x
Washing one's face	x	x	x	x		
Washing one's hands	**X**	x	x	x	**X**	**X**
Wiping one's hands		x			**X**	**X**
Cutting one's nails	**X**	x	**X**	**X**		**X**
Filing one's nails	**X**		**X**	x		
Blowing one's nose	x	x	x	x		
Dressing						
Fastening the zipper of a jacket	**X**	**X**	**X**	**X**	**X**	**X**
Fastening a snap (eg, jacket, bag)	**X**	**X**	**X**	**X**	**X**	**X**
Buttoning up a shirt	**X**	**X**	x	x	**X**	**X**
Pilling up the zipper of trousers	**X**	**X**	x	x	x	x
Buttoning up trousers	**X**	**X**	x	**X**	x	x
Lacing shoes		x		**X**	x	x
Rolling up a sleeve of a sweater		**X**			x	
Putting on gloves					x	x

* Items included in disease-specific ABILHAND scales are marked by a large bold cross.
CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults.

Supplementary table. Original set of items answered by at least two diagnostic groups

Item	Submitted to					
	CS	CP	RA	SSc	NMDc	NMDa
Environment						
Turning on a radio	x		x	x	x	
Turning on a television	x	x	x	x	x	x
Switching on a bedside lamp	x	**X**	x	x		x
Turning off a tap	x		**X**	**X**	**X**	**X**
Turning on a tap		x			**X**	**X**
Closing a door	x		x	x		
Inserting a key in a keyhole		x				**X**
Turning a key in a keyhole	x	x	**X**	x	**X**	**X**
Ringing a door bell	x	x	x	x		x
Communication						
Handling a 4-colour ballpoint pen with one hand	x		**X**	x		
Writing a sentence	x		**X**	x	x	x
Opening mail	**X**	x	x	**X**		x
Dialling on a keypad phone	x	x	x	x	x	x
Inserting a diskette in a disk drive	x	x	x	x	x	x
Using a computer keyboard					x	x
Typewriting	x		x	x		
Turning over the pages of a book	x	x	x	x	x	x
Do-it-yourserlf						
Sharpening a pencil	**X**	**X**	**X**	x	**X**	
Drawing	x		x	x	x	
Drawing a line with a ruler		x			x	
Colouring		x			x	
Painting		x			x	
Using an eraser		x			x	
Handling scissors		x		**X**	x	
Handling a stapler	x		**X**	**X**		
Threading a needle	**X**		**X**	**X**		
Screwing on a nut	x		**X**	x		
Using a screwdriver	x		**X**	x		
Hammering a nail	**X**		**X**	x		
Replacing a light bulb	x		**X**	x		
Leisure and play						
Throwing a ball		x			x	
Catching a ball		x			x	
Piling up Lego® blocks		x			x	
Dealing cards		x			**X**	
Using a joystick		x			x	
Miscellaneous						
Taking a coin out of the pocket	x	**X**	**X**	**X**		x
Grasping a coin on a table	x	x	**X**	x	x	x
Putting a coin in a piggy bank		x			x	
Counting banknotes	x		x	x		**X**
Winding up a wristwatch	x		x	**X**		
Wrapping up gifts	**X**		**X**	x		
Opening a car door		x		x		

* Items included in disease-specific ABILHAND scales are marked by a large bold cross.

CS: chronic stroke; CP: cerebral palsy; RA: rheumatoid arthritis; SSc: systemic sclerosis; NMDc: neuromuscular children; NMDa: neuromuscular adults.

Figure 1

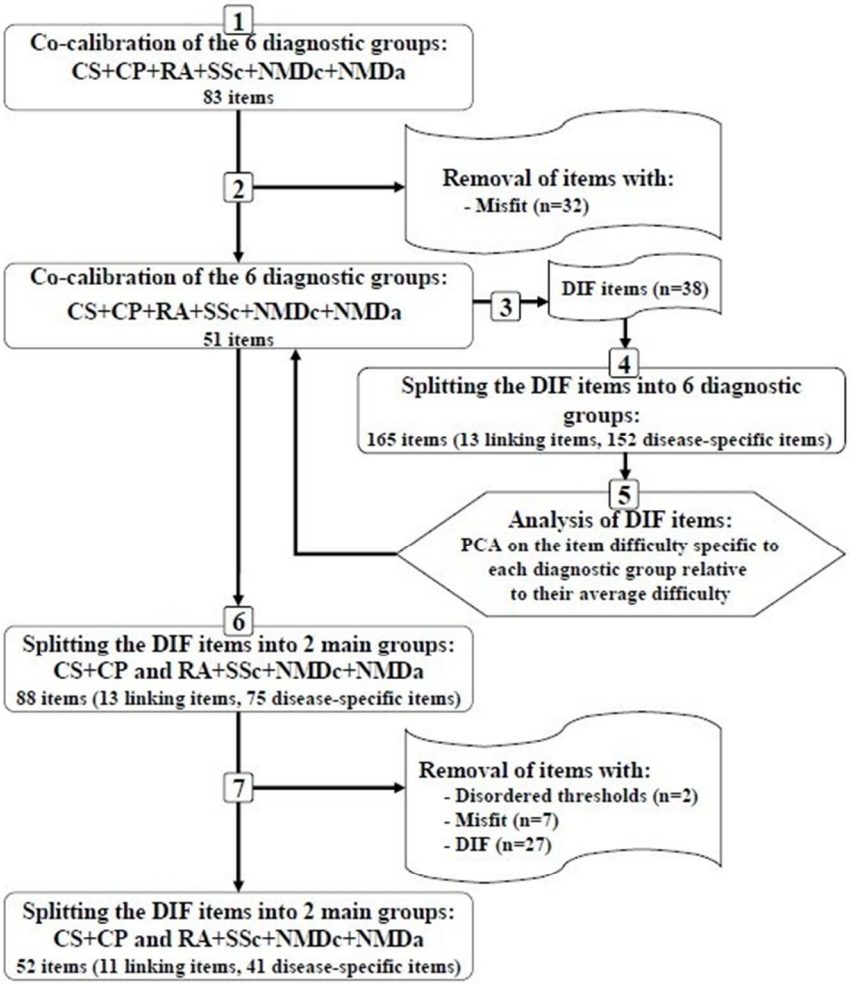


Figure 2

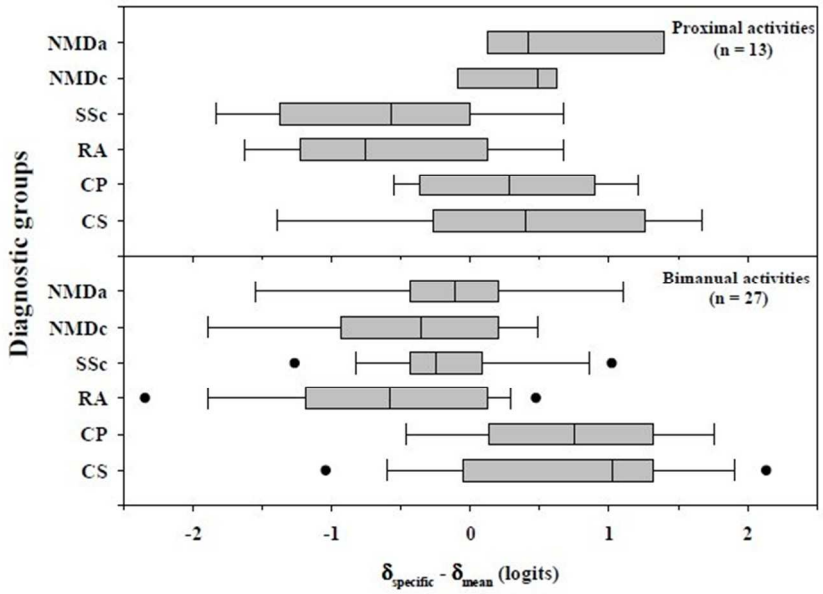


Figure 3

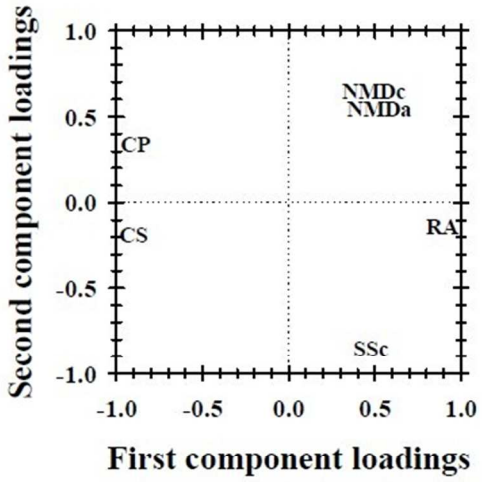


Figure 4