

Validation of a Manual Ability Questionnaire in Patients With Systemic Sclerosis

MARIE VANTHUYNE,¹ VANESSA SMITH,² SEHER ARAT,³ RENÉ WESTHOVENS,³ FILIP DE KEYSER,² FRÉDÉRIC A. HOUSSIAU,¹ JEAN-LOUIS THONNARD,¹ AND LAURE VANDERVELDE¹

Objective. To adapt and validate a manual ability questionnaire, the ABILHAND, developed through the Rasch methodology in patients with systemic sclerosis (SSc).

Methods. The original version of the ABILHAND, which includes 81 manual daily activities, was presented to 156 patients with SSc. They were asked to provide their perceived difficulty in performing each manual activity on a 3-level scale: impossible, difficult, or easy. Items were selected from well-established psychometric criteria. The patients were reassessed 1 month later to test the reproducibility. Concomitantly, they were clinically evaluated for their disease activity/severity, and their functional ability was tested with the Health Assessment Questionnaire (HAQ).

Results. The 26 selected items defined a unidimensional and linear measure of manual ability and showed a continuous progression in their difficulty. The item difficulty hierarchy was invariant across 12 patient-related factors and the manual ability score was reproducible over time. Finally, the manual ability was significantly poorer in SSc patients with more severe disease, and was negatively correlated with the HAQ score ($\rho = -0.733$).

Conclusion. The SSc-adapted ABILHAND questionnaire is a reliable, valid, reproducible, linear, and unidimensional measure to assess and followup on the manual ability of patients with SSc; therefore, it could become a useful additional tool in clinical trials to assess treatment efficacy.

INTRODUCTION

Systemic sclerosis (SSc) is a life-threatening multisystemic disease purportedly resulting from the combination of 3 types of injuries: autoimmune, vascular, and fibrotic. Ac-

tivation of the immune system leads to the production of SSc-specific autoantibodies, vascular damage causes the progressive obstruction of small- and medium-sized arteries, and excessive collagen deposition is responsible for skin and internal organ fibrosis (1).

Several disease manifestations may contribute to manual disability in patients with SSc, such as sclerodactyly, sclerodema, Raynaud's phenomenon, digital ulcers, finger calcinosis, and arthritis. Manual ability may be defined as the capacity to manage daily activities requiring the use of the upper extremities, independent of the strategy involved (2). Manual ability can be inferred from the patients' perceived difficulty in performing activities, as determined by other questionnaires such as the Duruöz Hand Index (3) or the UK Scleroderma Functional Score (4). The former, although tested in patients with SSc (5), was not specifically designed, nor adapted, for SSc.

The ABILHAND questionnaire, developed using the Rasch model (2), offers the advantage of selecting and hierarchizing manual activities that patients with different diseases find specifically difficult to realize. It has been validated in patients with chronic stroke (6) or with rheumatoid arthritis (RA) (7), and in children with cerebral palsy (8). The aim of this study is to adapt and validate the ABILHAND questionnaire as a measure of manual ability in patients with SSc.

Supported by unrestricted research grants from the Société Royale Belge de Rhumatologie/Koninklijke Belgische Vereniging voor Reumatologie, the Fonds pour la Recherche Scientifique en Rhumatologie/Fonds voor Wetenschappelijk Reuma Onderzoek, and the Association des Patients Sclérodermiques de Belgique. Dr. Vanthuyne's work was supported by a research fellowship from the Fonds National de la Recherche Scientifique, Belgium. Dr. Vandervelde's work was supported by a grant from the Fondation Saint-Luc.

¹Marie Vanthuyne, MD, Frédéric A. Houssiau, MD, PhD, Jean-Louis Thonnard, PhD, Laure Vandervelde, PhD: Cliniques universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium; ²Vanessa Smith, MD, Filip De Keyser, MD, PhD: Ghent University, Ghent, Belgium; ³Seher Arat, RN, MSN, René Westhovens, MD, PhD: Katholieke Universiteit Leuven, Leuven, Belgium.

Address correspondence to Marie Vanthuyne, MD, Universit  catholique de Louvain, Rheumatology Department, Cliniques universitaires Saint-Luc, Avenue Hippocrate, 10, B-1200 Brussels, Belgium. E-mail: marie.vanthuyne@uclouvain.be.

Submitted for publication September 11, 2008; accepted in revised form January 1, 2009.

PATIENTS AND METHODS

Patients. In March 2008, 218 patients fulfilling LeRoy and Medsger's classification criteria for SSc (9) and included in the Belgian Systemic Sclerosis Cohort by 3 university hospitals were asked, in their own language, to answer a manual ability questionnaire. One hundred fifty-six patients with SSc (89 Dutch speakers and 67 French speakers) agreed to participate. According to Leroy's classification, 19 patients with limited SSc (lSSc), 104 patients with limited cutaneous SSc (lcSSc), and 33 patients with diffuse cutaneous SSc (dcSSc) composed the sample.

Patients were also asked to fill in the Health Assessment Questionnaire (HAQ; score ranging from 0 = no disability to 3 = impossible to do) (10) and 2 visual analog scales (VAS; ranging from 0 mm = no limitation to 100 mm = maximal limitation) about the interference of Raynaud's phenomenon and of digital ulcers with their daily activities. We checked for the presence of digital ulcers and digital pitting scars. We measured the Rodnan modified total skin score (TSS; range 0–51) (11) and the fingertip to palm distance in flexion (distance between the tip of the third finger and the distal palmar crease; most affected side selected). Finally, we computed the Scleroderma Disease Activity Score (DAS; range 0–10) (12) and the disease severity score (range 0–4 for each of the 9 organ systems, therefore the total disease severity score ranges from 0–36) (13).

The clinical and functional characteristics of the 156 patients are described in Table 1. There was no significant difference in the clinical characteristics (sex, age, disease subsets, and TSS) of the patients who did not respond to the questionnaire (62 of 218) compared with those who did respond.

Manual ability assessment. Manual ability was assessed with the ABILHAND questionnaire (2), which evaluates the difficulties a patient may experience in performing manual activities of daily living. In a preliminary study, 10 patients with lcSSc and 10 patients with dcSSc, the most regularly followed patients with that disease subset consecutively examined in the outpatient department, were asked to propose other relevant manual activities not included in the original 56-item set. Following the patients' suggestions, 25 items were added to the original set, and the experimental version of ABILHAND questionnaire consisted of 81 items. The patients were asked to provide their perceived difficulty in performing each of the 81 activities using a 3-level scale, where 0 = impossible, 1 = difficult, and 2 = easy. Each activity could be completed irrespective of the arm(s) used, and independent of the adaptive strategies. Activities not commonly performed in the previous 3 months were not scored, and were recorded as missing responses (6.4%). The items were randomly presented to avoid any item order bias.

Data analysis: measuring manual ability through the Rasch model. The data were analyzed with the Rasch Unidimensional Measurement Models computer program (RUMM2020; RUMM Laboratory, Perth, Western Australia).

Table 1. Characteristics of the patients (n = 156)*

Sex, n (%)	
Male	32 (20.5)
Female	124 (79.5)
Age, years	54.3 (21–82)
Language, n (%)	
French	67 (42.9)
Dutch	89 (57.1)
SSc subsets, n (%)	
lSSc	19 (12.2)
lcSSc	104 (66.7)
dcSSc	33 (21.1)
Disease duration, years†	9.7 (0.6–67.9)
Fingertip to palm distance in flexion, mm‡	7 (0–55)
TSS score (range 0–51)	4 (0–25)
Presence of digital pitting scars, n (%)	65 (41.7)
Presence of digital ulcers, n (%)	21 (13.5)
HAQ score (range 0–3)	0.5 (0–2.875)
VAS score for Raynaud's phenomenon, mm	26 (0–100)
VAS score for digital ulcers, mm	3 (0–100)
Presence of calcinosis on hands radiograph, n (%)	27 (17.3)
Presence of osteolysis on hands radiograph, n (%)	21 (13.5)
DAS (range 0–10)	1 (0–10)
Peripheral vascular DSS (range 0–4)	1 (0–3)
Total DSS (range 0–36)	5 (0–15)

* Values are the median (range) unless otherwise noted. SSc = systemic sclerosis; lSSc = limited SSc; lcSSc = limited cutaneous SSc; dcSSc = diffuse cutaneous SSc; TSS = Rodnan modified total skin score; HAQ = Health Assessment Questionnaire; VAS = visual analog scale; DAS = Disease Activity Score; DSS = disease severity score.

† Since onset of Raynaud's phenomenon.

‡ Distance between the tip of the third finger and the distal palmar crease, most affected side selected.

lia). The Rasch model estimates the item difficulty and the patient manual ability on a common linear scale from the responses given to each item within a probabilistic framework (14). This model is used to investigate response category functioning, item and patient fit to the model, scale unidimensionality, item local independence, patient targeting, and scale reliability (15).

The category functioning is studied by verifying that successive response categories for each item represent increasing levels of ability, and that thresholds between adjacent categories are located in the expected order. The thresholds correspond to the manual ability required to have a higher probability to select one category rather than the previous one. The Rasch model could verify, via a likelihood ratio test, whether the response categories share the same discrimination across all items. This allows the use of a rating scale model (16,17), thus making the clinical interpretation of the scores easier. The overall item and patient fit to the model is checked by testing how close the observed data are from the model expectations. Individual item fit indicates whether the item contributes to

the definition of the unidimensional manual ability construct (18).

To test item unidimensionality, the sample is divided along the variable into level groups called class intervals. For each item, the degree of similitude between the observed responses and the expected responses in each class interval are computed through a standardized residual and a chi-square fit statistic reported by the software (19). The item standardized residual corresponds to the sum of the differences between the observed and expected scores for each patient, divided by the SD of the differences, and is sensitive to item discrimination. Positive and negative values represent an under discrimination and an over discrimination of the item, respectively. The chi-square fit statistic represents the deviations from the model expectations.

The assumption of local independency and, subsequently, the scale unidimensionality is verified by a principal component analysis of the residuals (20). The correlation between items and the first residual factor could define 2 subsets of items (i.e., the positively and the negatively correlated items). Two separate patients' locations are estimated from these subsets of items, and are compared using a *t*-test (20). The targeting is checked by comparing the mean patient location with the mean item difficulty. The software also reports the reliability, indicating the level of measurement precision attained.

Item selection. Starting with the 81 experimental items, successive analyses were performed to select the items composing the final ABILHAND scale. The following criteria were used (7): 1) a response rate >80%, indicating that the activity is commonly performed by patients with SSc; 2) a difficulty targeted at the level of the sample ability, indicating that items provide enough information to estimate the patients' ability; 3) an ordered rating scale, indicating that the 3 response categories are well discriminated; 4) a discrimination of the 3 response categories similar to the other items, verified by a nonsignificant likelihood ratio test (21); 5) an adequate fit to the Rasch model confirmed by *P* values (by chi-square) greater than 0.05 and standardized residuals between -2.5 and 2.5 (21), indicating that the items contribute to the definition of the unidimensional construct of manual ability; and 6) no redundancy in item location, preferably keeping items that best fit the Rasch model.

Determining the scale invariance through differential item functioning. The invariance of the item difficulty hierarchy was checked according to 12 dichotomous patient-related factors: 1) sex (male versus female), 2) patient age (median 54 years; ≤ 54 years versus > 54 years), 3) language community (Dutch speakers versus French speakers), 4) SSc subsets (lcSSc versus dcSSc), 5) disease duration from the onset of Raynaud's phenomenon (median 6.3 years; ≤ 6.3 years versus > 6.3 years), 6) fingertip to palm distance (median 7 mm; ≤ 7 mm versus > 7 mm), 7) TSS (median 4; ≤ 4 versus > 4), 8) presence of digital pitting scars (yes versus no), 9) presence of digital ulcers (yes versus no), 10) functional ability (median HAQ score 0.5;

≤ 0.5 versus > 0.5), 11) DAS (< 3 versus ≥ 3), and 12) total disease severity score (median score 5; ≤ 5 versus > 5). When the patients' numbers were unbalanced in the dichotomous subgroups (e.g., 124 women for 32 men), a sample of the largest group equivalent to the smallest one was randomly selected to avoid any bias due to the patients' number in the item hierarchy estimate.

Scale test-retest reliability. The test-retest reliability of the patients' responses and of the item hierarchy estimate was investigated by reassessing 134 patients with SSc 1 month after the first evaluation, and determined using intraclass correlation coefficients (ICCs) (22).

Scale construct validity. The construct validity was tested by examining the degree of association between the patients' ABILHAND measures and other measures widely used in patients with SSc. Moreover, the relationships between these patients' measures and demographic (age, sex, and language community) and clinical (SSc subsets) indices were tested either by a correlation coefficient for continuous indices or by a one-way analysis of variance for groups of nominal indices. The analyses were performed with SigmaStat software (SPSS, Chicago, IL); *P* values less than 0.05 were considered statistically significant. For the multivariate analysis, the combination of a multiple logistic regression, a multiple correspondence analysis, and a discriminant analysis was used to detect the manual ability's predictors. Before performing the multivariate analysis using dichotomous variables, we checked whether they strongly correlated with each other when expressed in continuous and in dichotomous values. We found no strong relationship ($r \geq 0.8$) between any variables, thereby making colinearity unlikely.

RESULTS

Refinement of the ABILHAND questionnaire for patients with SSc. Successive analyses through the Rasch model removed 55 items from the original 81-item set. Five items were not commonly performed by patients with SSc (e.g., typewriting), 16 were too easy (e.g., turning on a television set), none presented reversed thresholds, 20 had a different discrimination of the response categories (e.g., sharpening a pencil), 5 did not contribute to the definition of the unidimensional construct (e.g., grooming), and 9 were redundant. The final ABILHAND scale contains 26 daily manual activities specifically selected for patients with SSc (Figure 1 and Table 2).

Psychometric qualities of ABILHAND in patients with SSc. The likelihood ratio test performed on the 26 selected items was not significant ($P = 0.16$), allowing the use of a rating scale model. The reported overall fit statistics for the final scale were close to a standardized normal distribution for item and person fit residuals (mean $\pm$ SD -0.301 ± 0.845 and -0.324 ± 1.105 , respectively), and the reported item-trait interaction was not significant

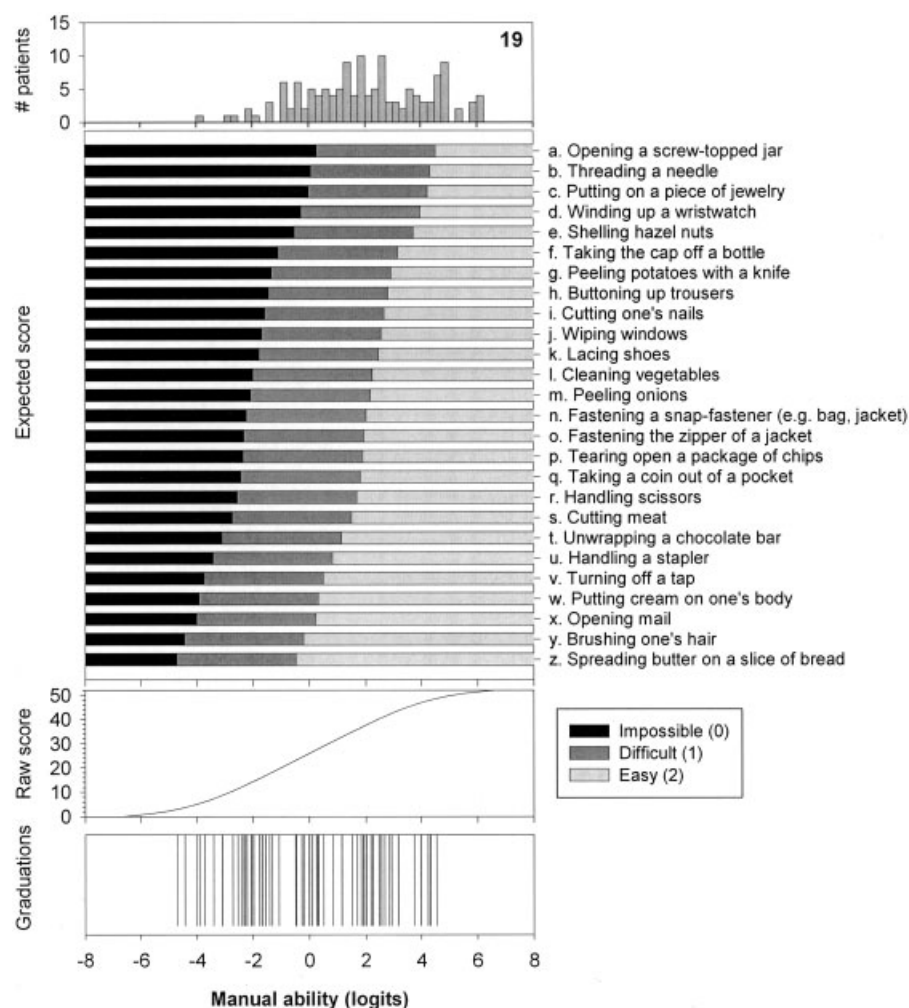


Figure 1. Structure of the ABILHAND scale in patients with systemic sclerosis (SSc). **Top:** distribution of manual ability measures of the 156 patients with SSc. Nineteen patients (12%) with extreme scores could not be assessed by the scale because they scored all of the activities as easy. **Upper middle:** a patient's expected response for each item as a function of the manual ability. Zero is, by convention, set at the average item difficulty; i.e., being able to easily lace shoes requires a measure of ≥ 2.47 logits, and a patient with a measure below -1.77 would be unable to fully accomplish this. A patient with a measure of 0 logits would be expected to perform the 2 easiest activities without difficulty, the average activities with some difficulties, and would be unable to perform the 2 most difficult activities. **Lower middle:** the Ogival relationship makes it possible to convert the ordinal raw score into a linear measure of manual ability. The raw score ranges 0–52 (from 26 items each scored 0–2). The manual ability in logits is an infinite measure. This relationship is quasi-linear in its central part, but a unitary increment in total score encompasses larger differences at both ends of the scale. **Bottom:** the scale graduations are represented by the 52 thresholds (2 per item). The first graduation corresponds to the manual ability required to choose difficult rather than impossible for spreading butter on a slice of bread, and the last one represents the manual ability required to select easy rather than difficult for opening a screw-topped jar.

($\chi^2 = 101.263$, 130 df, $P = 0.97$). The t -test between the patients' locations, estimated via the items positively correlated with the first residual factor of the principal component analysis on the residuals and those estimated by the negatively correlated items, was not significant ($t = 0.072$, 136 df, $P = 0.943$), indicating that the item local independence and the scale unidimensionality are assumed.

The calibration of the 26-item ABILHAND scale is presented in Table 2. The items are classified according to decreasing difficulty from top to bottom (range 2.41 to

-2.59 logits), with higher logit values representing more difficult items. The manual ability measure is expressed in logits (i.e., log-odds units), which is a linear unit defined as the natural logarithm of the odds of successful achievement by a patient for any item. This unit is constant along the measurement scale, and the zero is set by convention at the average difficulty of the whole selected item set. The SE of the item difficulty estimates (mean 0.21, range 0.19 to 0.26), the standardized residuals (mean -0.30 , range -1.59 to 1.95), the fit statistic computed as a chi-square, and the associated P values are also shown in

Table 2. ABILHAND calibration for patients with systemic sclerosis (n = 156)

Item	Difficulty, logits	SE, logits	Residual, z	Fit, χ^2 *	P
a. Opening a screw-top jar	2.41	0.19	-0.25	4.00	0.55
b. Threading a needle	2.20	0.19	-0.56	1.63	0.90
c. Putting on a piece of jewelry	2.12	0.19	-0.65	6.60	0.25
d. Winding up a wristwatch	1.85	0.20	-0.66	5.22	0.39
e. Shelling hazelnuts	1.62	0.20	1.95	7.61	0.18
f. Taking the cap off a bottle	1.05	0.19	-0.75	2.41	0.79
g. Peeling potatoes with a knife	0.80	0.19	0.03	2.43	0.79
h. Buttoning up trousers	0.71	0.19	-1.53	2.47	0.78
i. Cutting one's nails	0.56	0.19	-0.50	1.64	0.90
j. Wiping windows	0.44	0.20	1.16	5.74	0.33
k. Lacing shoes	0.35	0.20	-0.91	4.64	0.46
l. Cleaning vegetables	0.13	0.20	-1.60	4.59	0.47
m. Peeling onions	0.05	0.20	-0.12	6.13	0.29
n. Fastening a snap-fastener (e.g., bag, jacket)	-0.09	0.20	-1.35	3.97	0.55
o. Fastening the zipper of a jacket	-0.19	0.20	-1.59	1.48	0.92
p. Tearing open a package of chips	-0.22	0.21	-0.12	3.96	0.56
q. Taking a coin out of a pocket	-0.30	0.20	0.35	4.19	0.52
r. Handling scissors	-0.43	0.20	0.00	1.46	0.92
s. Cutting meat	-0.61	0.20	-0.73	1.52	0.91
t. Unwrapping a chocolate bar	-0.97	0.22	-0.17	4.93	0.42
u. Handling a stapler	-1.29	0.22	-0.52	4.13	0.53
v. Turning off a tap	-1.61	0.22	-0.42	0.45	0.99
w. Putting cream on one's body	-1.79	0.23	0.72	4.49	0.48
x. Opening mail	-1.88	0.24	0.67	8.88	0.11
y. Brushing one's hair	-2.31	0.26	0.49	2.82	0.73
z. Spreading butter on a slice of bread	-2.59	0.26	-0.78	3.88	0.57

* Degrees of freedom = 5.

Table 2. The 26 items define a unidimensional scale of manual ability in patients with SSc, because all the *P* values are not significant. The reliability index of the final scale is 0.97.

Description of the ABILHAND scale. The structure and the targeting of the linear ABILHAND scale in patients with SSc are shown in Figure 1. The measurement range represented by the thresholds distribution (-4.71 to 4.54 logits) is well targeted to the range of the patients' abilities (-4.02 to 5.75 logits). Nevertheless, the overall mean patient location is equal to 1.79 logits, indicating that patients report moderate difficulties in performing manual activities. Indeed, 19 patients (11 with lcSSc, 1 with dcSSc, and 7 with lSSc) with a manual ability >5.75 logits reported they could perform all activities easily, and 96 patients with a manual ability >0.29 logits reported they could achieve all activities easily or with some difficulty. None reported they could not perform any of the 26 manual activities.

Differential item functioning. The item hierarchy invariance of the 26 selected items was tested across 12 dichotomous subgroups of patients using the differential item functioning plots represented in Figure 2. As most of the manual activities lie within the 95% confidence interval (95% CI) of the identity line, the item diffi-

culty tends to be consistently estimated across demographic and clinical subgroups of patients. There are some exceptions identified by their label, as shown in Table 2.

Test-retest reliability. The test-retest reliability of the patients' measures is shown in Figure 3 (right panel). Most of the measures lie within the 95% CI of the identity line, indicating that patients consistently estimate their manual ability. Moreover, the ICC = 0.96 for the patient measures. The left panel shows the item difficulty hierarchy between the first and the second assessment. No item lies outside the 95% CI of the identity line, and the item estimates ICC = 0.98, indicating an excellent reproducibility of the item hierarchy.

Construct validity. The manual ability measures were significantly different across SSc subsets. A post hoc analysis showed 1) that patients with lSSc had a higher manual ability (mean $\pm$ SD 4.10 $\pm$ 2.5 logits) than patients with lcSSc (mean $\pm$ SD 2.44 $\pm$ 2.44 logits) and patients with dcSSc (mean $\pm$ SD 1.17 $\pm$ 2.47 logits), and 2) that patients with lcSSc have a significantly higher manual ability than patients with dcSSc. The ABILHAND measures were negatively correlated with the following clinical and para-clinical indices: disease duration ($r = -0.28$; $P < 0.001$), fingertip to palm distance ($r = -0.39$; $P < 0.001$), TSS

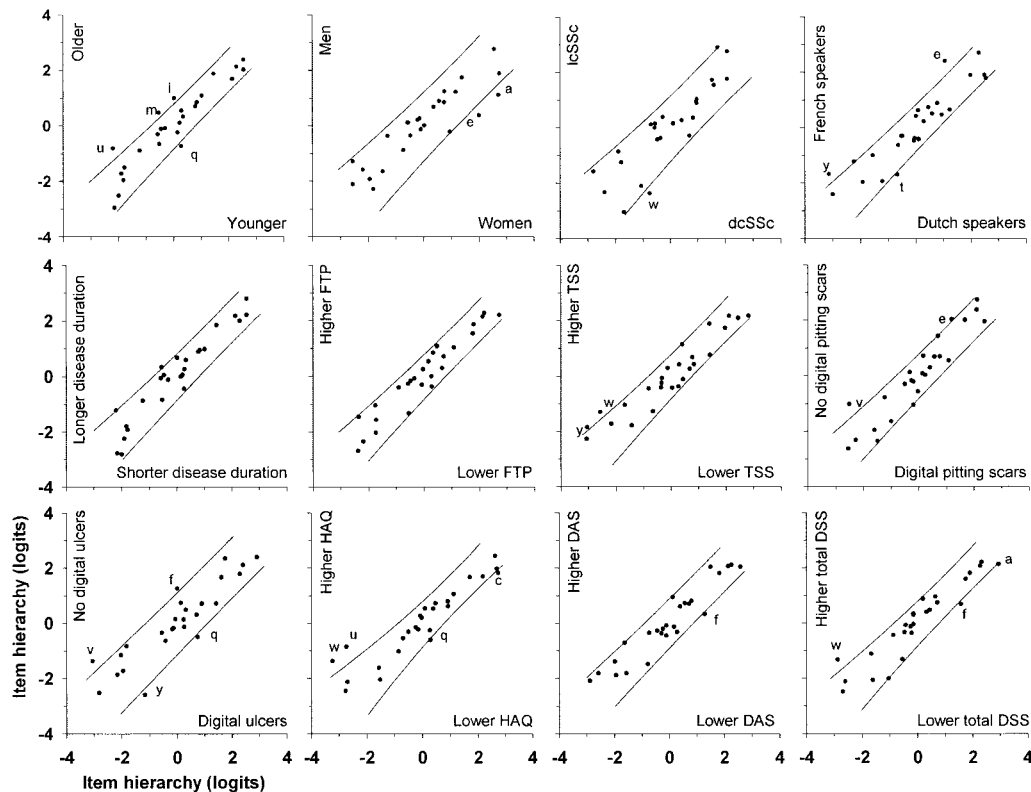


Figure 2. Invariance of the ABILHAND scale in subgroups of patients with systemic sclerosis. Each panel shows the differential item functioning plot for dichotomous subgroups of patients, divided according to 12 splitting criteria. The most difficult items are plotted in the top/right part of each. The difficulties of the items lying outside the **solid lines** (the 95% confidence interval of the ideal invariance) and identified by their labels (as shown in Table 2) were differently estimated in both subgroups of patients. lcSSc = limited cutaneous systemic sclerosis; dcSSc = diffuse cutaneous systemic sclerosis; disease duration = disease duration since onset of Raynaud's phenomenon; FTP = fingertip to palm distance in flexion (distance between the tip of the third finger and the distal palmar crease; most affected side selected); TSS = Rodnan modified total skin score; HAQ = Health Assessment Questionnaire; DAS = Disease Activity Score; Total DSS = total disease severity score.

($\rho = -0.39$; $P < 0.001$), VAS dealing with the interference of Raynaud's phenomenon ($\rho = -0.32$; $P < 0.001$), DAS ($\rho = -0.4$; $P < 0.001$), total disease severity score ($\rho = -0.53$; $P < 0.001$), and HAQ score ($\rho = -0.73$; $P < 0.001$). No relationship was found between manual ability and sex, language community, presence of digital pitting scars, or the presence of calcinosis on hand radiographs. The multivariate analysis indicates that a HAQ score ≥ 0.5 and a total disease severity score ≥ 5 were the only independent predictors of a poor manual ability (Table 3).

DISCUSSION

This manuscript describes the validation of the ABILHAND questionnaire in a large population of patients with SSc. We applied the Rasch model, which allows one to hierarchize items to create a linear scale, provided that they fit the unidimensionality imposed by the model (14). In this case, manual ability tasks have been classified according to their difficulty after eliminating those influenced by factors other than the manual ability itself. The resulting 26 selected items define a unidimensional scale (all $P > 0.05$; Table 2) and allow the patient's manual ability to be estimated and followed up on a linear scale through a

self-administered, patient-friendly, non-time-consuming questionnaire.

To identify the manual disabilities associated with SSc, we expanded the original 56-item set with 25 additional manual tasks considered to be difficult by 20 patients with SSc in a preliminary survey. Of these 81 items, 26 were retained by the Rasch analysis. These include unimanual and bimanual activities, as well as tasks requiring strength and dexterity. They encompass a large range of activities, such as dressing, cooking, hygiene, working in an office, and cleaning. Interestingly, and not surprisingly, 6 of the 26 final items were among the 25 suggested by patients themselves. Thus, putting on a piece of jewelry, handling scissors, or lacing shoes indeed requires exquisite digital ability that can be reduced by the presence of digital ulcer(s). Wiping windows and putting cream on one's body might be affected by defective hand stretching due to skin fibrosis. Finally, Raynaud's phenomenon can compromise one's ability to clean vegetables in cold water.

The hierarchy of the 26 items is consistent with the type of hand involvement observed in patients with SSc. More difficult items require either manual strength (e.g., opening a screw-top jar) or finger precision (e.g., threading a needle), whereas the easiest involve rather gross manual

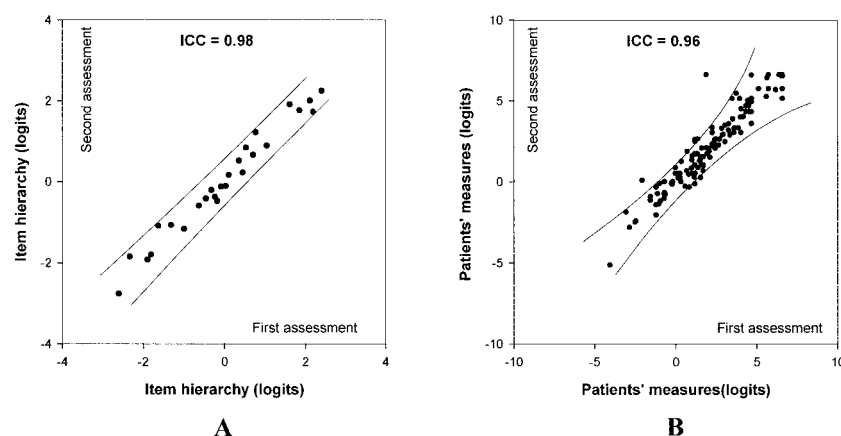


Figure 3. Test-retest reliability of the ABILHAND scale. **A**, item difficulty perceived by patients with systemic sclerosis at the first and second assessment, and the 95% confidence interval (95% CI) (solid line) of the ideal invariance. The most difficult items are plotted in the top/right part of the figure. Items within the solid lines were ranked with the same hierarchy between both assessments. **B**, manual ability self-perceived by patients at the first and second assessment, and the 95% CI (solid lines) of the ideal invariance. More able patients are plotted in the top/right part of the figure. Patients' measures (dots) lying within the solid lines have the same estimated manual ability at the first and the second assessment. ICC = intraclass correlation coefficient.

activities (e.g., putting cream on one's body). When we compare the retained items with those retained in patients with RA (7), some activities were perceived differently in terms of difficulty. For instance, threading a needle is more difficult for patients with SSc than for patients with RA because it involves greater fine digital dexterity (i.e., requiring intact fingertips). On the other hand, activities such as handling a stapler are more difficult for patients with RA due to wrist and metacarpophalangeal joint involvement. These observations indicate that the 26 items correspond to appropriate activities that specifically assess manual ability in patients with SSc.

The ABILHAND scale for patients with SSc presented a good reliability index, indicating that more than 7 groups of significantly different manual abilities can be distin-

guished in our sample (23). The 26 items are well targeted to the patients' abilities, covering a wide range of functional abilities, despite the fact that 12% of the patients were able to easily complete all the activities. Moreover, the item hierarchy and the self-perceived manual ability were consistently reassessed after 1 month (Figure 3), confirming the excellent test-retest reliability of the ABILHAND questionnaire.

The overall fit statistics of the 26 items showed that the observed data fit the expectations of the Rasch model because the item and person fit residuals were close to a normal distribution. The 26 items were also locally independent because the *t*-test between patients' locations was not significant (20). Moreover, the overall chi-square statistic was not significant, reflecting invariance

Table 3. Multiple logistic regression: predictors of poor manual ability (logit <2.14)*

Variable	P	OR (95% CI)
Age ≥ 54 years	0.21	1.9 (0.7–5.5)
Disease duration ≥ 9.7 years	0.41	1.5 (0.6–4.1)
dSSc rather than lcSSc or lSSc	0.86	0.9 (0.3–3.1)
TSS ≥ 15	0.80	1.2 (0.2–6.7)
Fingertip to palm distance ≥ 7 mm	0.37	2.5 (0.3–18.0)
Presence of digital pitting scars	0.25	0.5 (0.1–1.7)
Presence of digital ulcers	0.60	1.6 (0.3–9.5)
HAQ ≥ 0.5	0.0006	8.3 (2.5–27.7)
Digital ulcers VAS score ≥ 3 mm	0.92	2.5 (0.8–7.4)
Raynaud's phenomenon VAS score ≥ 26 mm	0.95	1.0 (0.3–3.0)
DAS ≥ 3	0.41	2.0 (0.4–9.8)
Vascular DSS ≥ 1	0.56	1.4 (0.4–4.5)
Total DSS ≥ 5	0.04	3.7 (1.1–13.0)
Calcinosis on hand radiograph	0.22	0.4 (0.1–1.8)
Osteolysis on hand radiograph	0.90	1.1 (0.2–6.4)

* Except for the TSS and DAS, cut offs were based on median values. OR = odds ratio; 95% CI = 95% confidence interval. See Table 1 for additional definitions.

of the item hierarchy across the variable (15). The differential item functioning plots confirmed the invariance of the item hierarchy across 12 dichotomous subgroups of patients. First, this hierarchy did not vary across different disease subsets, suggesting that the ABILHAND questionnaire can be used to assess manual ability across the spectrum of patients with SSc. Nevertheless, before using ABILHAND in other SSc populations, it would be important to verify that patients' responses to the 26 items fit to the Rasch model. Second, the item difficulty hierarchy was not influenced by disease activity, disease severity, global physical ability, or digital joints' limitation. There are, however, some exceptions; for instance, 2 activities requiring more manual strength (i.e., opening a screw-top jar and shelling hazelnuts) were considered to be more difficult for women than for men (Figure 2).

Our data demonstrate a stringent correlation between manual disability and disease severity; the latter is not only assessed by musculoskeletal and skin involvements (e.g., fingertip to palm distance, TSS, distal phalangeal osteolysis) but also by global disease scores (e.g., DAS and disease severity score). In this respect, it is worth stating that the skin score has been correlated with major organ involvement and survival (24). In a multivariate analysis, of 15 covariates tested, only the combination of a total disease severity score ≥ 5 and a HAQ score ≥ 0.5 was predictive of a low manual ability.

The HAQ is a self-report measure that quantifies functional ability in daily activities (10). It is commonly used in SSc patient evaluation, and was previously shown to correlate with disease subsets, TSS, the presence of tendon friction rubs, joint pain, muscle weakness, and the fingertip to palm distance (25). In our study, however, even if the univariate analysis indicated that ABILHAND measures were highly correlated with HAQ, the HAQ variance explains only 50% of the manual ability variance. Those results are not surprising because the HAQ was not designed for SSc and only tests 7 activities dealing specifically with manual ability.

In 1997, Steen and Medsger partially bridged the HAQ's lack of disease specificity by adding 6 patient-assessed VAS evaluating SSc organ system symptoms, namely digital ulcers, Raynaud's phenomenon, pain, overall disease severity, and gastrointestinal and pulmonary symptoms. It was shown that the vascular VAS (ulcers and Raynaud's) was strongly correlated with the presence of digital pitting scars, tip ulcers, and gangrene (26). Nevertheless, these VAS do not specifically address manual ability.

Two other functional questionnaires have been used in SSc. The UK Scleroderma Functional Score is a self-completed questionnaire, developed to assess upper extremity function and muscle weakness caused by skin tightness in patients with SSc (4). It comprises 11 items, not specifically dealing with manual ability, each of which are scored from 0 (able to perform) to 3 (impossible to achieve). The items were, however, selected by physicians and patients with SSc based on a consensus and not by a stringent statistical approach as in ABILHAND. Nevertheless, Smyth et al (27) observed a strong correlation between the UK Scleroderma Functional Score and the HAQ in a group of 115 patients with SSc, a rather expected

finding because 5 of the 11 items of the UK Scleroderma Functional Score are part of the HAQ.

The Duruöz Hand Index (also called Cochin scale and Hand Functional Disability Scale) is an 18-item questionnaire dealing with manual ability, initially designed for patients with RA (28), then applied in osteoarthritis (29) and 2 small groups of patients with SSc (3,5). The mean Duruöz Hand Index score was significantly higher in patients with dcSSc compared with patients with lcSSc, and was found to correlate with the HAQ. The major difficulty with this questionnaire stems from its 6 response levels, ranging from possible without difficulty to impossible to do. In previous studies, patients were unable to discriminate between more than 4 response levels (2), a finding further confirmed by a statistical analysis demonstrating that a 3 response level questionnaire was ideal (6). In this respect, the 3 levels of answers (impossible, difficult, or easy) used in the ABILHAND questionnaire are easily distinguished from each other by patients, a result confirmed here because no item presented reversed thresholds. Unlike ABILHAND, neither the UK Scleroderma Functional Score nor the Duruöz Hand Index display the fundamental qualities of measurement, such as linearity and unidimensionality.

Taken together, we propose the use of a new questionnaire to assess manual ability in patients with SSc. The questionnaire has been developed and validated using the gold standard Rasch methodology in a large and clinically well-studied group of patients with SSc. The linearity of the scale, together with the invariance of the item hierarchy, allows reliable assessment and reassessment of a SSc patient's manual ability. ABILHAND could therefore contribute to the evaluation of the effects of new therapies.

Scoring sheets and instructions for ABILHAND are available online at www.rehab-scales.org. Moreover, online analyses taking into account the missing values will directly convert raw scores into linear measures of manual ability.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Vanthuyne had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Vanthuyne, Smith, Arat, Westhovens, De Keyser, Houssiau, Thonnard, Vandervelde.

Acquisition of data. Vanthuyne, Smith, Arat, Westhovens, De Keyser, Houssiau, Thonnard, Vandervelde.

Analysis and interpretation of data. Vanthuyne, Smith, Arat, Westhovens, De Keyser, Houssiau, Thonnard, Vandervelde.

REFERENCES

1. Varga J, Abraham D. Systemic sclerosis: a prototypic multi-system fibrotic disorder. *J Clin Invest* 2007;117:557–67.
2. Penta M, Thonnard JL, Tesio L. ABILHAND: a Rasch-built measure of manual ability. *Arch Phys Med Rehabil* 1998;79:1038–42.
3. Brower LM, Poole JL. Reliability and validity of the Duruöz Hand Index in persons with systemic sclerosis (scleroderma). *Arthritis Rheum* 2004;51:805–9.

4. Silman A, Akesson A, Newman J, Henriksson H, Sandquist G, Nihill M, et al. Assessment of functional ability in patients with scleroderma: a proposed new disability assessment instrument. *J Rheumatol* 1998;25:79–83.
5. Rannou F, Poiraudau S, Berezne A, Baubet T, Le-Guern V, Cabane J, et al. Assessing disability and quality of life in systemic sclerosis: construct validities of the Cochin Hand Function Scale, Health Assessment Questionnaire (HAQ), Systemic Sclerosis HAQ, and Medical Outcomes Study 36-Item Short Form Health Survey. *Arthritis Rheum* 2007;57:94–102.
6. Penta M, Tesio L, Arnould C, Zancan A, Thonnard JL. 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. Durez P, Fraselle V, Houssiau F, Thonnard JL, Nielens H, Penta M. Validation of the ABILHAND questionnaire as a measure of manual ability in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;66:1098–105.
8. Arnould C, Penta M, Renders A, Thonnard JL. ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. *Neurology* 2004;63:1045–52.
9. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001;28:1573–6.
10. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26:1346–53.
11. Clements PJ, Lachenbruch PA, Seibold JR, Zee B, Steen VD, Brennan P, et al. Skin thickness score in systemic sclerosis: an assessment of interobserver variability in 3 independent studies. *J Rheumatol* 1993;20:1892–6.
12. Valentini G, Della Rossa A, Bombardieri S, Bencivelli W, Silman AJ, D'Angelo S, et al. European multicentre study to define disease activity criteria for systemic sclerosis. II. Identification of disease activity variables and development of preliminary activity indexes. *Ann Rheum Dis* 2001;60:592–8.
13. Medsger TA Jr. Natural history of systemic sclerosis and the assessment of disease activity, severity, functional status, and psychologic well-being. *Rheum Dis Clin North Am* 2003;29:255–73.vi.
14. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: MESA; 1980.
15. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 2007;57:1358–62.
16. Wright BD. Model selection: rating scale or partial credit? [letter]. *Rasch Meas Trans* 1999;12:641–2.
17. Linacre JM. Comparing “partial credit” and “rating scale” models [letter]. *Rasch Meas Trans* 2000;14:768.
18. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Mahwah (NJ): Lawrence Erlbaum; 2001.
19. Wright BD, Masters GN. Rating scale analysis. Chicago: Mesa Press; 1982.
20. Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 2002;3:205–31.
21. Andrich D, Sheridan BE. Interpreting RUMM2020: part II. Polytomous data. 3rd ed. Perth (WA): RUMM Laboratory; 2005.
22. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
23. Fisher W. Reliability statistics [letter]. *Rasch Meas Trans* 1992;6:238.
24. Steen VD, Medsger TA Jr. Improvement in skin thickening in systemic sclerosis associated with improved survival. *Arthritis Rheum* 2001;44:2828–35.
25. Poole JL, Steen VD. The use of the Health Assessment Questionnaire (HAQ) to determine physical disability in systemic sclerosis. *Arthritis Care Res* 1991;4:27–31.
26. Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. *Arthritis Rheum* 1997;40:1984–91.
27. Smyth AE, MacGregor AJ, Mukerjee D, Brough GM, Black CM, Denton CP. A cross-sectional comparison of three self-reported functional indices in scleroderma. *Rheumatology* (Oxford) 2003;42:732–8.
28. Duruoz MT, Poiraudau S, Fermanian J, Menkes CJ, Amor B, Dougados M, et al. Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. *J Rheumatol* 1996;23:1167–72.
29. Poiraudau S, Chevalier X, Conrozier T, Flippo RM, Liote F, Noel E, et al. Reliability, validity, and sensitivity to change of the Cochin hand functional disability scale in hand osteoarthritis. *Osteoarthritis Cartilage* 2001;9:570–7.