

Translation, Validation, and Reliability of the Dutch Late-Life Function and Disability Instrument Computer Adaptive Test

Remco M. Arensman, Martijn F. Pisters, Janneke M. de Man-van Ginkel, Marieke J. Schuurmans, Alan M. Jette, Rob A. de Bie

R.M. Arensman, PT, MSc, Physical Therapy Research, Department of Rehabilitation, Nursing Science and Sport, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands, and Center for Physical Therapy Research and Innovation in Primary Care, Julius Health Care Centers, Utrecht, the Netherlands.

M.F. Pisters, PT, PhD, Physical Therapy Research, Department of Rehabilitation, Nursing Science and Sport, Brain Center Rudolf Magnus, University Medical Center Utrecht, Room W01.121, Postbus 85500, 3508 GA Utrecht, the Netherlands, and Center for Physical Therapy Research and Innovation in Primary Care, Julius Health Care Centers, Utrecht, the Netherlands. Address all correspondence to Dr Pisters at: m.f.pisters@umcutrecht.nl.

J.M. de Man-van Ginkel, RN, PhD, Nursing Science, Department of Rehabilitation, Nursing Science and Sport, Brain Center Rudolf Magnus, University Medical Center Utrecht.

M.J. Schuurmans, RN, PhD, Nursing Science, Department of Rehabilitation, Nursing Science and Sport, Brain Center Rudolf Magnus, University Medical Center Utrecht.

Author information continues on next page.

Background. Adequate and user-friendly instruments for assessing physical function and disability in older adults are vital for estimating and predicting health care needs in clinical practice. The Late-Life Function and Disability Instrument Computer Adaptive Test (LLFDI-CAT) is a promising instrument for assessing physical function and disability in gerontology research and clinical practice.

Objective. The aims of this study were: (1) to translate the LLFDI-CAT to the Dutch language and (2) to investigate its validity and reliability in a sample of older adults who spoke Dutch and dwelled in the community.

Design. For the assessment of validity of the LLFDI-CAT, a cross-sectional design was used. To assess reliability, measurement of the LLFDI-CAT was repeated in the same sample.

Methods. The item bank of the LLFDI-CAT was translated with a forward-backward procedure. A sample of 54 older adults completed the LLFDI-CAT, World Health Organization Disability Assessment Schedule 2.0, RAND 36-Item Short-Form Health Survey physical functioning scale (10 items), and 10-Meter Walk Test. The LLFDI-CAT was repeated in 2 to 8 days (mean=4.5 days). Pearson's r and the intraclass correlation coefficient (ICC) (2,1) were calculated to assess validity, group-level reliability, and participant-level reliability.

Results. A correlation of .74 for the LLFDI-CAT function scale and the RAND 36-Item Short-Form Health Survey physical functioning scale (10 items) was found. The correlations of the LLFDI-CAT disability scale with the World Health Organization Disability Assessment Schedule 2.0 and the 10-Meter Walk Test were $-.57$ and $-.53$, respectively. The ICC (2,1) of the LLFDI-CAT function scale was .84, with a group-level reliability score of .85. The ICC (2,1) of the LLFDI-CAT disability scale was .76, with a group-level reliability score of .81.

Limitations. The high percentage of women in the study and the exclusion of older adults with recent joint replacement or hospitalization limit the generalizability of the results.

Conclusions. The Dutch LLFDI-CAT showed strong validity and high reliability when used to assess physical function and disability in older adults dwelling in the community.



Post a Rapid Response to this article at: ptjournal.apta.org

A.M. Jette, PT, PhD, FAPTA, Health and Disability Research Institute, School of Public Health, Boston University, Boston, Massachusetts.

R.A. de Bie, PT, PhD, Department of Epidemiology and CAPHRI Research School, Maastricht University, Maastricht, the Netherlands.

[Arensman RM, Pisters MF, de Man-van Ginckel JM, et al. Translation, validation, and reliability of the Dutch Late-Life Function and Disability Instrument Computer Adaptive Test. *Phys Ther*. 2016;96:1430–1437.]

© 2016 American Physical Therapy Association

Published Ahead of Print:

March 9, 2016

Accepted: March 6, 2016

Submitted: May 14, 2015

Adequate assessment of physical function and disability in older adults dwelling in the community is vital for estimating and predicting health care needs in research and clinical practice.^{1,2} As a result, physical function and disability have become part of the comprehensive geriatric assessment used in geriatric clinical care and are commonly used as outcome measures in gerontology research.^{3,4}

Not surprisingly, many measurement instruments have been developed to assess physical function or disability.⁵ Patient-reported outcome measures (PROMs) are preferred because of their low cost and convenience.⁶ However, PROMs often suffer from limitations, such as measuring only a single construct; being multidimensional, with no apparent conceptual structure; lacking sensitivity to detect important changes; being time-consuming to administer; and having floor or ceiling effects when used for evaluative purposes.⁷

For overcoming these limitations, the Late-Life Function and Disability Instrument (LLFDI) was developed.^{8,9} The LLFDI is a PROM designed to assess physical function and disability in older adults living in the community.^{8,9} It consists of 2 scales: the 32-item function scale and the 16-item disability scale. The LLFDI has excellent test-retest reliability in the function component (intraclass correlation coefficient [ICC]=.87–.98) and moderate-to-good reliability in the disability component (ICC=.68–.91).^{10,11} For both components, the observation of expected differences in summary scores of known functional groups supported validity.^{8,9} Additionally, the responsiveness, construct validity, and predictive validity of the LLFDI have been shown to be comparable to those of performance-based measures.^{11,12} However, the LLFDI has 2 major limitations. Like comparable PROMs for physical function and disability, the LLFDI takes a long time to complete (>20 minutes, on average, for the combined function and disability scales), and all questions are administered to all patients regardless of applicability; these limitations make using the LLFDI in clinical care difficult.¹³

For alleviating respondent burden without sacrificing precision and sensitivity, a computer adaptive test (CAT) version of the LLFDI was developed with item response theory methods.⁴ For construction of the item bank required for the CAT version of the LLFDI, the 48 items of the fixed-item LLFDI were expanded to a 192-item database. The added items were designed to create a comprehensive pool of physical function and disability items. Item response theory-based CAT instruments have several advantages over conventional instruments.¹⁴ First, CAT instruments use existing data to individualize the measurement process and select relevant items for an individual respondent. Furthermore, CAT instruments reduce the number of questions needed, maintain measurement precision, and decrease respondent burden.⁴ To enable the use of the LLFDI-CAT in research and clinical practice in the Netherlands, we aimed to translate the LLFDI-CAT to the Dutch language and to investigate its validity and reliability in older people who spoke Dutch and dwelled in the community.

Method

Phase 1: Translation

The translation protocol was based on the guidelines proposed by Beaton et al.¹⁵ Figure 1 shows the flowchart of the translation procedure.

In stage I, all 192 items from the LLFDI-CAT function and disability item banks and all introductory texts were translated into Dutch by 2 independent translators who were bilingual, with Dutch as their mother tongue. The first translator was a clinician who was aware of the purpose and application of the questionnaire in order to create a translation of clinical equivalence. The second translator had no medical background and was unaware of the purpose and application of the LLFDI-CAT in order to create a translation reflecting the language used by the general population in the Netherlands.¹⁵ This approach ensured 2 translations with different perspectives.

In stage II, the 2 translations from stage I were incorporated into a combined translation. For this stage, the protocol of Beaton et al¹⁵ was modified for practical

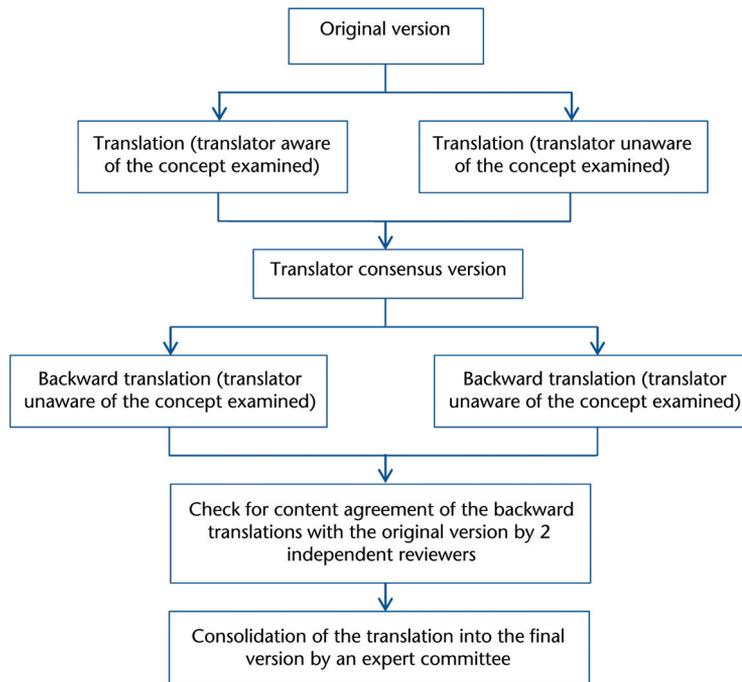


Figure 1. Translation and cross-cultural adaptation of the Dutch version of the Late-Life Function and Disability Instrument Computer Adaptive Test.

reasons. Instead of a meeting of the 2 translators to discuss the translations and resolve all discrepancies, the first translator created a first draft of the combined translation. The first draft was thoroughly checked by the second translator, who listed any inconsistencies or translations with which he disagreed. The first draft and the list from the second translator were discussed during a meeting of both translators and an independent observer until a consensus on the combined translation was reached. All changes made were registered by the observer.

Stage III was the backward translation of the combined translation. Two independent bilingual translators with English as their mother tongue and Dutch as their second language translated the combined translation back to English. Both translators were unaware of the LLFDI-CAT and had no medical background. When both backward translations were finished, content agreement with the original version was checked by 2 independent reviewers to ensure consistent translation. Any inconsistencies or conceptual errors in the translation were

documented, and the corresponding items were changed.

In stage IV, an expert committee consisting of a methodologist, a medical professional, one of the forward translators, one of the backward translators, and a language professional consolidated the final translation of the introductory texts and 192 items from the LLFDI-CAT item banks. The complete list of items in the LLFDI-CAT item banks is provided in the LLFDI-CAT manual of procedures.¹⁶

The expert committee ensured that the translation and adaptation were idiomatically, semantically, experientially, and conceptually equivalent. After all issues were resolved, the software for the Dutch translation of the LLFDI-CAT was produced by the original developer of the instrument.

Phase 2: Validity and Reliability Study

Study population. A convenience sample of older adults dwelling in the community was recruited from the regions of Leiden and Utrecht, the Neth-

erlands. Invitation letters were distributed to all residents (N=252) in a convenience sample of 4 senior apartment buildings and senior housing facilities. Older adults interested in participating in the study were asked to contact the researchers by phone or by email. Additionally, the older adults could contact a researcher in person at the senior apartment buildings or housing facilities 1 week after distribution of the invitation letters for more information regarding the study or to express interest in participating in the study. Information about when the researcher would be present at the senior apartment buildings or housing facilities was included in the invitation letters received by the residents.

Older adults willing to participate in the study were screened for eligibility with the following criteria: (1) age of 65 years or older, (2) independently ambulatory (with or without an assistive device), (3) community dwelling, and (4) must have provided informed consent prior to participation. Older adults were excluded if they: (1) had undergone joint replacement surgery in the lower extremities within the preceding 6 months, (2) had been hospitalized within the preceding 3 months, (3) were unable to walk 10 m without assistance from another person, or (4) were living in a nursing home or a similar facility at the time of screening. Informed consent was obtained from all participants before participation. Demographic data were collected and are shown in Table 1.

Assessment of validity. The concurrent validity of the Dutch LLFDI-CAT was determined as follows. The participants were asked to complete the Dutch LLFDI-CAT on a laptop computer. Additionally, paper forms of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) and the RAND 36-Item Short-Form Health Survey physical functioning scale (10 items) (PF-10) were completed.^{17,18} Lastly, the 10-Meter Walk Test (10MWT) was completed by all participants.¹⁹ Measurements were obtained by a researcher in the participant's own home environment or at a local physical therapy practice.

Table 1.
Characteristics of the Study Sample (n=54)^a

Characteristic	n	%	\bar{X} (SD)	Range
Age (y)			80.19 (8.0)	65–93
Sex				
Male	12	22.2	80.17 (8.0)	67–93
Female	42	77.8	80.19 (8.0)	65–93
Body mass index (kg/m ²)	53		27.3 (5.0)	20–45
Marital status				
Single	5	10		
Married	19	40		
Widow/widower	30	50		
Chronic diseases				
None	18	33.3		
1	24	44.4		
2	8	14.9		
3 or more	4	7.4		
Walking aid				
None	28	51.9		
Cane or wheeled walker	22	40.7		
Walking aid, wheelchair, or both	4	7.4		
Education				
<6 y	21	38.9		
6–12 y	32	59.3		
>12 y	1	1.9		
LLFDI-CAT function scale score			51.77 (8.6)	37.57–76.43
LLFDI-CAT disability scale score			50.99 (8.5)	34.61–64.95

^a LLFDI-CAT=Late-Life Function and Disability Instrument Computer Adaptive Test.

Assessment of reliability. The test-retest reliability of the LLFDI-CAT was assessed with a retest moment. Participants were contacted by phone within 2 to 14 days after completion of the initial testing procedure. During this contact, a researcher administered the LLFDI-CAT to obtain the retest data.

LLFDI-CAT. The LLFDI-CAT consists of a large item bank containing items for both the function scale and the disability scale. Item response theory methods were used to calibrate the items in the item bank on a scale ranging from 0 to 100, with a mean of 50 and a standard deviation of 10.⁴ A higher scale score represents better functioning or less disability. Function scale items ask, “How much difficulty do you currently have. . . ?” Response options are “none at all,” “a little,” “a lot,” “unable to do,” and “does not apply.” The disability scale

items ask, “Because of your physical or mental health, to what extent do you feel limited in. . . ?” Response options are “none at all,” “a little,” “a lot,” “completely,” and “does not apply.” After an item from a scale is completed, the software calculates a participant scale score and a participant-level standard error of measurement (SEM). When a preset SEM has been reached or when a preset number of items has been administered, the final participant scale score and participant-level SEM are calculated.

Earlier research on the LLFDI-CAT software showed that the use of a stopping rule of at least 10 items per scale is needed to achieve precision and sensitivity levels similar to those of the original fixed-item LLFDI.⁴ Therefore, in the present study, the LLFDI-CAT software was programmed to stop administering items per scale when 10 items per scale were

completed or when the participant-level SEM reached less than 3.0.

Comparison Instruments

The 36-item PROM WHODAS 2.0 was used to assess the functioning of a participant in 6 activity domains: understanding and communicating, getting around, self-care, getting along with people, life activities, and participation in society.¹⁷ The WHODAS 2.0 items are scored on a 5-point Likert scale (none; mild; moderate; severe; and extreme, cannot do). After scoring is complete, an algorithm converts the item scores to a score ranging from 0 (no disability) to 100 (full disability). Cronbach’s α values for the subscales of the WHODAS 2.0 ranged from .7 to .97 for patients in rehabilitation settings and from .77 to .98 for patients with chronic diseases.^{20,21}

The PF-10 consists of 10 items designed to assess self-care, mobility, and other physical activities and body movements.¹⁸ Items are scored on a 3-point Likert scale (“yes, limited a lot”; “yes, limited a little”; and “no, not limited at all”). The raw scores are converted to a scale score ranging from 0 to 100, with higher scores representing better physical function. Cronbach’s α representing the internal consistency of the PF-10 in older adults is .82.²²

The 10MWT is designed to measure walking speed.¹⁹ Participants are asked to walk a distance of 10 m at a comfortable walking speed. The test-retest reliability of the 10MWT has been reported to be excellent (ICC=.96-.98), with a small SEM (0.004–0.008 m/s).²³

Data Analysis

We used IBM SPSS 20.0 (IBM Corp, Armonk, New York) to perform statistical analyses. On the basis of data from a validation study of the original LLFDI, a sample size of at least 42 participants was calculated with an α of .05, a β of .20, and an expected absolute effect size r of .50 for all correlations between the LLFDI-CAT scales and the comparison instruments.²⁴

To assess concurrent validity, we calculated Pearson’s r ; we used Spearman’s rho for data that were not normally dis-

Validity and Reliability of the Dutch LLFDI-CAT

Table 2.

Correlations of LLFDI-CAT Scales With Comparison Instruments^a

LLFDI-CAT Scale	Correlation With:		
	WHODAS 2.0	PF-10	10MWT
Function	-.65	.74 ^b	-.69
Disability	-.57 ^{b,c}	.68	-.53 ^b

^a LLFDI-CAT=Late-Life Function and Disability Instrument Computer Adaptive Test, WHODAS 2.0=World Health Organization Disability Assessment Schedule 2.0, PF-10=RAND 36-Item Short-Form Health Survey physical functioning scale (10 items), 10MWT=10-Meter Walk Test.

^b Significant correlation at *P* level of <.000.

^c Spearman's rho calculated with data from 52 participants.

tributed. The validity of the LLFDI-CAT was interpreted with Cohen's conventions for effect sizes of Pearson's *r* (0.10=small, 0.30=medium, and 0.5=large).²⁵

Item response theory models expand on measurement domain reliability by allowing the calculation of participant-level and group-level reliability.²⁶ Unlike the reliability of fixed-item measurement instruments, which is calculated with the SEM of sample scores, the CAT software provides a SEM for an individual participant. These data allow the calculation of participant-level reliability, which is the reliability of the measurement instrument for the level of functioning or disability of an individual participant.²⁶ Using the average of participant-level SEMs, we calculated group-level reliability, which more closely reflects conventional reliability statistics. To assess test-retest reliability, we calculated the ICC (2,1) absolute agreement.^{27,28} Because high between-subject variability inflates ICC scores, classifying ICC scores in categories such as low, medium, and high provides little information.²⁹ Instead, participant-level reliability was plotted against LLFDI-CAT participant scores to provide information on the reliability of the LLFDI-CAT for different participant score levels.

Results

Demographics

A total of 56 older adults expressed interest in participating in the study and were screened for eligibility. Two of the 56 potential participants were excluded because they were living in a nursing home at the time of data collection. The 54 participants in the final sample were predominantly women (77.8%), and 67% reported having one or more chronic diseases. The mean score on the LLFDI-CAT function scale was 51.8 (SD=8.6, range=37.6–76.4), and that on the LLFDI-CAT disability scale was 51.0 (SD=8.5, range=34.6–65.0). More detailed information about the participants is shown in Table 1.

The average time required to administer the complete LLFDI-CAT was 6 minutes 3 seconds (n=54). For one participant, 2 items on the PF-10 were missing. These missing values were imputed with the personal scale mean, as suggested in the RAND 36-Item Short-Form Health Survey manual.³⁰ For 2 participants, 2 items on the same subscale of the WHODAS 2.0 were missing; therefore, data from these participants had to be excluded from the correlation analysis.

Validity

The absolute correlations (*r*) of the LLFDI-CAT function scale with all comparison instruments exceeded .65; the correlation with the PF-10 was .74 (Tab. 2). Additionally, the absolute correlations (*r*) of the LLFDI-CAT disability scale with all comparison instruments exceeded .50; the correlations with the WHODAS 2.0 and the 10MWT were -.57 and -.53, respectively (Tab. 2).

Reliability

All participants were available for the retest of the LLFDI-CAT, and the number of days between the test and the retest ranged from 2 to 8 (median=5). The ICC (2,1) scores and the group-level reliability of the LLFDI-CAT function scale and the LLFDI-CAT disability scale are shown in Table 3. Figures 2 and 3 show the participant-level reliability plots of the relationship between participant-level reliability and the LLFDI-CAT scale scores. For both LLFDI-CAT scales, participant-level reliability scores were between .8 and .9, until a participant ability score of 60 was reached. Participants with scale scores between 60 and 70 had participant-level reliability scores ranging from .6 to .8. Scale scores over 70 resulted in participant-level reliability scores of less than .6. Eleven and 12 participants scored higher than 60 on the function scale and on the disability scale, respectively.

Discussion

The results found in the validity study confirmed the concurrent validity of the Dutch language version of both the function scale and the disability scale of the LLFDI-CAT. The magnitude and direction of the correlations of the LLFDI-CAT scales and the comparison instruments were as expected. The group-level reliability and test-retest reliability of the LLFDI-CAT were found to be good, with all reliability scores exceeding .80.

The results of the present study were compared with the results of studies investigating the validity of the original LLFDI. The correlation of the LLFDI-CAT function scale with the PF-10 in the present study was slightly lower than the values found by Dubuc et al (*r*=.85),³¹ Lapier (*r*=.83),³² and Hand et al

Table 3.

Reliability of LLFDI-CAT Scales^a

LLFDI-CAT Scale	Test X̄ (SD)	Retest X̄ (SD)	ICC (2,1) (95% CI)	Group-Level Reliability (SEM)
Function	51.77 (8.6)	51.21 (8.7)	.84 (.74, .90)	.85 (3.36)
Disability	50.99 (8.5)	49.25 (8.9)	.76 (.62, .86)	.81 (3.69)

^a LLFDI-CAT=Late-Life Function and Disability Instrument Computer Adaptive Test, ICC=intraclass correlation coefficient, CI=confidence interval, SEM=average participant standard error of measurement.

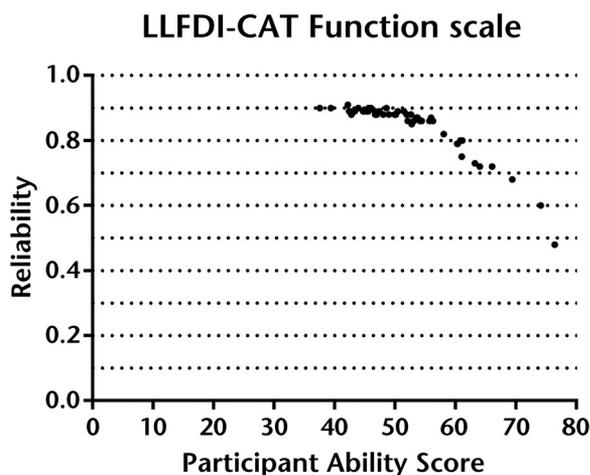


Figure 2.

Participant-level reliability plotted against participant scores for the Late-Life Function and Disability Instrument Computer Adaptive Test (LLFDI-CAT) function scale.

($r=.88$)³³ for the original LLFDI. However, Roaldsen et al¹⁰ reported a lower correlation ($r=.52$). A possible explanation for the lower correlation found by Roaldsen et al¹⁰ is that their sample consisted of older adults who dwelled in the community and had self-reported balance deficits and fear of falling. Because the correlations found in the present study are comparable to or higher than those of the original LLFDI, the validity of the LLFDI-CAT function scale is supported.¹⁰

The correlations of the LLFDI-CAT disability scale with the WHODAS 2.0 and

the 10MWT in the present study were strong, according to Cohen's conventions, confirming the concurrent validity of the LLFDI-CAT.²⁵ A possible explanation for the lower correlations of the LLFDI-CAT disability scale compared with the LLFDI-CAT function scale is the broad construct of disability. As a result, fixed-item instruments lack the large number of items required to measure the entirety of the construct.³⁴ Similarly, a performance test, such as the 10MWT, does not capture the parts of disability caused by mental health, social, or environmental factors. Therefore, it was expected that the construct "disability"

would be only partially measured by the comparison instruments, resulting in lower correlations.

Another explanation for the lower correlations is the theoretical basis on which the instruments are based. The LLFDI-CAT disability scale aims to measure disability in older adults, where the WHODAS 2.0 was designed for adults in general.^{8,9,17} The concurrent validity of the LLFDI-CAT disability scale was comparable to that of the original LLFDI. The LLFDI has been compared with self-report questionnaires, such as the London Handicap Scale ($r=.47-.66$), the Western Ontario and McMaster Universities Osteoarthritis Index ($r=-.47$), and the 20-Meter Walk Test (a performance test) ($r=.37$).^{31,32,35} The correlations found in those studies are similar to or lower than the correlations found in the present study, further confirming the concurrent validity of the LLFDI-CAT disability scale.

The test-retest reliability scores of the function and disability scales were sufficiently high to indicate that the instrument is stable over repeated measurements when no change is expected. These reliability findings are consistent with those of reliability studies of the original LLFDI (ICC range=.44-.98), the Hebrew translation of the conventional LLFDI (ICC range=.46-.90), and the Swedish translation of the conventional LLFDI (ICC range=.82-.91).^{8-10,36}

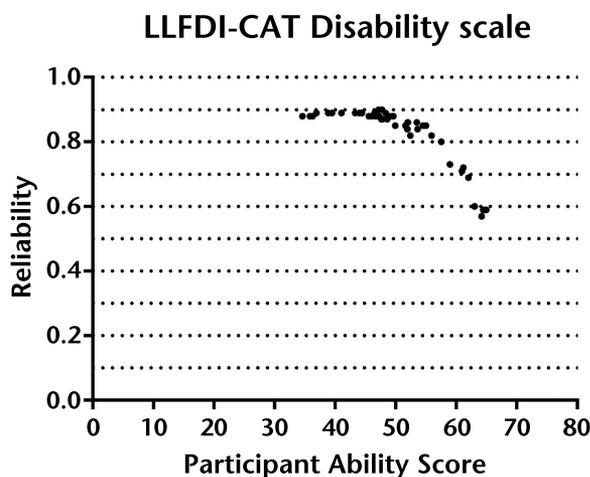


Figure 3.

Participant-level reliability plotted against participant scores for the Late-Life Function and Disability Instrument Computer Adaptive Test (LLFDI-CAT) disability scale.

Additionally, the reliability of item response theory-based instruments expands on the concept of reliability in classical test theory. Because the CAT software provides a participant-level SEM tied to the participant's ability score, it is possible to tailor the reliability of the instrument to the specific needs of the application. In the present study, the stopping rule for the software was a participant-level SEM of less than 3.0 or the administration of 10 questions. The resulting group-level reliability scores of .85 for the function scale and .81 for the disability scale are high and, combined with the observed excellent test-retest reliability, indicate that the LLFDI-CAT is excellent for use in research pertaining

Validity and Reliability of the Dutch LLFDI-CAT

to older adults dwelling in the community.

The participant-level reliability scores shown in Figures 2 and 3 lie between .8 and .9 until an ability score of approximately 60 on both LLFDI-CAT scales. Additionally, the plotted lines show a limitation of CAT. At extreme ability scores, participant-level reliability decreases and, as a result, more items are required to achieve higher reliability scores. Therefore, we advise changing the stopping rules of the LLFDI-CAT to incorporate only the SEM and not a number of questions in order to maintain high reliability when targeting populations of older adults with more extreme expected ability scores.

In addition to high precision, the other advantage of a CAT instrument over a fixed-item instrument is the short time required for completion of the instrument. Completing both domains of the LLFDI-CAT took less than 9 minutes, thereby reducing respondent burden by as much as 50% compared with that associated with the fixed-item LLFDI.⁴

The present study has some limitations. First, to our knowledge, the present study is the first study in which translation of the CAT version of the LLFDI was attempted and is one of the first studies in which an existing CAT instrument was translated and validated. Consequently, there are no standardized guidelines or protocols for the translation and validation of existing CAT instruments. To overcome this problem, we adapted an existing protocol originally designed for use in the translation of fixed-item instruments and applied it to the translation of the LLFDI-CAT.¹⁵

Second, the high percentage of women in the sample can make generalization of the results to the general population of the same age difficult. However, as the age of the population advances, the percentage of women increases, to up to 72.1% of people who are 90 years old.³⁷ Given that the mean age of the participants in the present study was 80 years and that 78% were women, the number of women in the sample was only slightly

higher than that in the general population of the same age.

Finally, the exclusion criteria used in the present study prevented the participation of older adults with recent joint replacement surgery or hospitalization. Although both joint replacement surgery and hospitalization are not uncommon in older adults, the results of the present study cannot be generalized to older adults with recent joint replacement surgery or hospitalization.

The thoroughness with which the validity and reliability of the LLFDI-CAT were tested provides a clear understanding of the psychometric properties of the instrument and reveals that the Dutch language version of the LLFDI-CAT has acceptable levels of validity and reliability for the assessment of physical function and disability in older adults dwelling in the community.

Problems in physical function and disability in older adults dwelling in the community are often treated by physical therapists. For assessment of the effectiveness of treatment strategies used by physical therapists, reliable and valid measurement instruments are required. However, the fact that many older adults have multiple morbidities complicates the choice of a disease-specific instrument. Furthermore, the effects of treatment strategies are difficult to compare in individual adults with different morbidities when different measurement instruments are used. Finally, because most PROMs have a relatively large measurement error, the precision of these measurement instruments is often too low to reveal treatment effects in individual adults.

The LLFDI-CAT can overcome these difficulties because it was designed to function as a generic instrument, independent of underlying morbidities. Also, it can reduce administrative burden and time investment for both physical therapists and patients. Furthermore, an advantage of CAT is that it allows the user to specify the stopping rule for a particular application. For individual assessment, in which high precision is desirable, a 15-item stopping rule or a

criterion reflecting a smaller degree of measurement error may be more desirable. In contrast, for research in large-scale studies, in which efficiency of administration is essential and less precision is required, a 5-item CAT may be acceptable.

Furthermore, the use of PROMs is becoming increasingly important as a health system performance indicator in the measurement of health care quality. Generic and low-burden measures (such as the LLFDI-CAT) that can be used independently of morbidities and with high precision—enabling individual decision making—are needed and preferable to disease-specific measures.

Future research should focus on responsiveness and the ability to detect long-term change. When the responsiveness of the LLFDI-CAT has been confirmed, it can be confidently used to assess the effectiveness of treatment strategies aimed at improving physical function and disability in older adults dwelling in the community. Furthermore, the validity and reliability of the LLFDI-CAT in specific groups of older adults, such as patients after stroke, patients after recent hospitalization, or older adults with cognitive impairments, should be examined further.

In conclusion, the Dutch language version of the LLFDI-CAT has good concurrent validity and high reliability for the assessment of physical function and disability in older adults dwelling in the community and can be used for evaluative purposes in research and clinical practice. Furthermore, the advantages of the LLFDI-CAT over traditional, fixed-item instruments make it preferable over those instruments.

Mr Arensman, Dr Pisters, Dr de Man-van Ginkel, Professor Schuurmans, and Professor de Bie provided concept/idea/research design. Mr Arensman and Dr Pisters provided writing, data analysis, and project management. Mr Arensman provided data collection and participants. Dr Pisters provided facilities/equipment. Dr Pisters, Dr de Man-van Ginkel, Professor Schuurmans, Dr Jette, and Professor de Bie provided consultation (including review of manuscript before submission).

The authors thank Mark Bakker for his assistance during data collection and Rachel Grubb, Pieter Miltenburg, Reyn Wagenaar, and Els Niele for their assistance with the translation and finalization of the Dutch Late-Life Function and Disability Instrument Computer Adaptive Test.

The study was approved by the Medical Ethics Committee of the University Medical Center Utrecht, Utrecht, the Netherlands (14-1111/C).

Dr Jette holds stock in CreCare LLC, a small business that distributes and licenses patient-reported outcome measures.

DOI: 10.2522/ptj.20150265

References

- Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*. 1994;49:M85-M94.
- Guralnik JM, Ferrucci L, Simonsick EM, et al. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med*. 1995;332:556-561.
- Brown AS, Brummel-Smith K, Burgess L, et al. National Institutes of Health Consensus Development Conference statement: geriatric assessment methods for clinical decision-making. *J Am Geriatr Soc*. 1988;36:342-347.
- Jette AM, Haley SM, Ni P, et al. Creating a computer adaptive test version of the Late-Life Function and Disability Instrument. *J Gerontol A Biol Sci Med Sci*. 2008;63:1246-1256.
- McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 2nd ed. New York, NY: Oxford University Press; 1996.
- Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med*. 1994;38:1-14.
- McHorney CA. Methodological and psychometric issues in health status assessment across populations and applications. In: Albrecht GL, Fitzpatrick R, eds. *Advances in Medical Sociology*. Vol 5. Greenwich, CT: JAI Press Ltd; 1994:281-304.
- Jette AM, Haley SM, Coster WJ, et al. Late Life Function and Disability Instrument, I: development and evaluation of the disability component. *J Gerontol A Biol Sci Med Sci*. 2002;57:M209-M216.
- Haley SM, Jette AM, Coster WJ, et al. Late Life Function and Disability Instrument, II: development and evaluation of the function component. *J Gerontol A Biol Sci Med Sci*. 2002;57:M217-M222.
- Roaldsen KS, Halvarsson A, Sarlija B, et al. Self-reported function and disability in late life: cross-cultural adaptation and validation of the Swedish version of the Late-Life Function and Disability Instrument. *Disabil Rehabil*. 2014;36:813-817.
- Beauchamp MK, Schmidt CT, Pedersen MM, et al. Psychometric properties of the Late-Life Function and Disability Instrument: a systematic review. *BMC Geriatr*. 2014;14:12.
- Beauchamp MK, Jette AM, Ward RE, et al. Predictive validity and responsiveness of patient-reported and performance-based measures of function in the Boston RISE study. *J Gerontol A Biol Sci Med Sci*. 2015;70:616-622.
- Gignac MA, Cao X, McAlpine J, Badley EM. Measures of disability: Arthritis Impact Measurement Scales 2 (AIMS2), Arthritis Impact Measurement Scales 2-Short Form (AIMS2-SF), The Organization for Economic Cooperation and Development (OECD) Long-Term Disability (LTD) Questionnaire, EQ-5D, World Health Organization Disability Assessment Schedule II (WHODASII), Late-Life Function and Disability Instrument (LLFDI), and Late-Life Function and Disability Instrument-Abbreviated Version (LLFDI-Abbreviated). *Arthritis Care Res (Hoboken)*. 2011;63(suppl 11):S308-S324.
- Bjorner JB, Chang CH, Thissen D, Reeve BB. Developing tailored instruments: item banking and computerized adaptive assessment. *Qual Life Res*. 2007;16(suppl 1):95-108.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25:3186-3191.
- Jette AM, Kopits IM, McDonough CM, et al. Boston University Late-Life Function and Disability Instrument: manual of procedures. Version 1.04. Available at: <http://www.bu.edu/bostonroc/files/2012/04/LLFDI-CAT-Manual-Version-1.04.pdf>. 2013. Accessed September 28, 2015.
- The World Health Organization Disability Assessment Schedule Phase II Field Trial Instrument*. Geneva, Switzerland: World Health Organization; 1999.
- Stewart A, Ware J. *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach*. Durham, NC: Duke University Press; 1991.
- Collen FM, Wade DT, Bradshaw CM. Mobility after stroke: reliability of measures of impairment and disability. *Int Disabil Stud*. 1990;12:6-9.
- Pösl M, Cieza A, Stucki G. Psychometric properties of the WHODASII in rehabilitation patients. *Qual Life Res*. 2007;16:1521-1531.
- Garin O, Ayuso-Mateos JL, Almansa J, et al. Validation of the "World Health Organization Disability Assessment Schedule, WHODAS-2" in patients with chronic diseases. *Health Qual Life Outcomes*. 2010;8:51.
- Bohannon RW, DePasquale L. Physical Functioning Scale of the Short-Form (SF) 36: internal consistency and validity with older adults. *J Geriatr Phys Ther*. 2010;33:16-18.
- Peters DM, Fritz SL, Krotish DE. Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults. *J Geriatr Phys Ther*. 2013;36:24-30.
- Sayers SP, Jette AM, Haley SM, et al. Validation of the Late-Life Function and Disability Instrument. *J Am Geriatr Soc*. 2004;52:1554-1559.
- Cohen J. A power primer. *Psychol Bull*. 1992;112:155-159.
- Raju NS, Price LR, Oshima TC, Nering ML. Standardized conditional SEM: a case for conditional reliability. *Appl Psychol Meas*. 2007;31:169-180.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420-428.
- McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1:30-46.
- Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19:231-240.
- van der Zee K, Sanderman R. *Het Meten van de Algemene Gezondheidstoestand met de RAND-36: Een Handleiding*. 2nd ed. Groningen, the Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken; 2012.
- Dubuc N, Haley S, Ni P, et al. Function and disability in late life: comparison of the Late-Life Function and Disability Instrument to the Short-Form-36 and the London Handicap Scale. *Disabil Rehabil*. 2004;26:362-370.
- Lapier TK. Utility of the Late Life Function and Disability Instrument as an outcome measure in patients participating in outpatient cardiac rehabilitation: a preliminary study. *Physiother Can*. 2012;64:53-62.
- Hand C, Richardson J, Letts L, Stratford P. Construct validity of the Late Life Function and Disability Instrument for adults with chronic conditions. *Disabil Rehabil*. 2010;32:50-56.
- Jette AM. Assessing disability in studies on physical activity. *Am J Prev Med*. 2003;25:122-128.
- Riddle DL, Jensen MP. Construct and criterion-based validity of brief pain coping scales in persons with chronic knee osteoarthritis pain. *Pain Med*. 2013;14:265-275.
- Melzer I, Kurz I, Sarid O, Jette AM. Relationship between self-reported function and disability and balance performance measures in the elderly. *J Rehabil Res Dev*. 2007;44:685-691.
- Giesbers H, Verweij A, de Beer J. Bevolking: wat is de huidige situatie? In: *Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid*. Bilthoven, the Netherlands: Nationaal Kompas Volksgezondheid/Bevolking. Available at: <http://www.nationaalkompas.nl>. Accessed March 17, 2014.

Copyright of Physical Therapy is the property of American Physical Therapy Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.