

REVIEW ARTICLE

Issues for the Selection of Wheelchair-Specific Activity and Participation Outcome Measures: A Review

William B. Mortenson, MSc, William C. Miller, PhD, Claudine Auger, MSc

ABSTRACT. Mortenson WB, Miller WC, Auger C. Issues for the selection of wheelchair-specific activity and participation outcome measures: a review. *Arch Phys Med Rehabil* 2008;89:1177-86.

Objectives: To use the World Health Organization's *International Classification of Functioning, Disability and Health* as a framework to identify and to evaluate wheelchair-specific outcome instruments that are useful for measuring activity and participation.

Data Sources: CINAHL, PsychInfo, EMBASE, Google Scholar, Dissertation Abstracts Medline databases, and conference proceedings.

Study Selection: Activity and participation measures that were specifically intended for adults who use wheelchairs and that were published in English in a peer-reviewed journal were included in this review. Based on electronic database searches using a variety of search terms, articles were identified by title, and appropriate abstracts were retrieved. Articles were obtained for all relevant abstracts. For peer-reviewed measures included in the review, we obtained any instruction manuals and related publications, frequently published in conference proceedings and theses or available electronically, on the development and testing of the measure.

Data Extraction: Tools included in the review were evaluated based on their conceptual coverage, reliability, validity, responsiveness, usefulness, and wheelchair contribution, which indicated how well the tool isolated the effect of the wheelchair on activity and participation outcomes.

Data Synthesis: A number of conceptual, psychometric, and applicability issues were identified with the 11 wheelchair-specific measures included in the review. A majority of the

measures were mobility focused. No single tool received excellent ratings in all areas of the review. Some of the most frequent issues identified included a failure to account for differences attributable to different wheelchairs and wheelchair seating, limited psychometric testing, and high administrative and respondent burden.

Conclusions: Good reliability evidence was reported for most of the measures, but validity information was only available for 6 of the 11 measures, and responsiveness information for 3. This review suggests that these measures could be improved with further psychometric testing and with some modification to ensure that the contribution of the wheelchair to activity and participation outcomes is clearly identified.

Key Words: Activities of daily living; Outcome assessment; Rehabilitation; Review literature [publication type]; Wheelchairs.

© 2008 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

HISTORICALLY, WHEELCHAIR USE and participation in everyday activities were inversely related, because wheelchairs were difficult to propel and there were many problems with environmental access.¹ Although recent advances in wheelchair technology are promising,^{2,3} currently people who use wheelchairs encounter many barriers to participation including accessibility problems,⁴⁻⁷ low rates of employment,⁴ social isolation,⁵ and stigma.^{8,9} Given these issues and in light of a rapid increase in the number of wheelchair users,¹⁰ it is critical to develop better ways to measure and to understand barriers and facilitators to participation in this population.¹¹

New technology has increased wheelchair and wheelchair seating options, but it has also made the selection process more difficult for prescribers and their clients and more expensive for funding agencies. As the prescription process becomes more complicated and funding agencies begin to demand evidence to support the need for equipment, outcome measurement is becoming increasingly necessary. When attempting to decide on an appropriate measure, researchers and clinicians must choose from an increasing array of potential instruments.

A rigorous analysis of existing measurement tools should cover conceptual, reliability, validity, and practical considerations.¹²⁻¹⁵ Because no studies have critically evaluated activity and participation outcome measures for people who use wheelchairs, a comprehensive review was undertaken using the World Health Organization's (WHO) *International Classification of Functioning, Disability and Health* (ICF)^{16,17} as a framework.

The ICF was developed to create a common language to explain function and health. The ICF model describes "[f]unctioning [. . . as] an umbrella term encompassing all body functions, activities and participation"^{17(p3)} which results from a dynamic interaction between health conditions and personal and environmental factors. In the ICF, body function is defined as the physiologic function of body systems and

From the Occupational Therapy, Long-Term Care, Vancouver Coastal Health, Vancouver, BC, Canada (Mortenson); School of Rehabilitation Sciences, University of British Columbia, Vancouver, BC, Canada (Mortenson, Miller); GF Strong Rehabilitation Research Lab, Vancouver, BC, Canada (Mortenson, Miller); International Collaboration on Repair Discoveries, Vancouver, BC, Canada (Miller); Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, BC, Canada (Miller); School of Rehabilitation, Université de Montréal, Montréal, QC, Canada (Auger); and Research Center, Institut universitaire de gériatrie de Montréal, Montréal, QC, Canada (Auger).

Presented in part to the 12th Annual North American Collaborating Centre Conference on ICF, June 6-7, 2006, Vancouver, BC, Canada, and the Festival of International Conferences on Care Giving, Disability, Aging and Technology, June 16-19, 2007, Toronto, ON, Canada.

Supported by a Quality of Life Strategic Training Fellowship in Rehabilitation Research, Canadian Institute of Health Research Musculoskeletal and Arthritis Institute; the Canadian Occupational Therapy Foundation (doctoral grant); Michael Smith Foundation for Health Research (doctoral grant); the Canadian Institute of Health Research (graduate fellowship); the Canadian Institutes of Health Institute of Aging; and l'Ordre des Ergothérapeutes du Québec (doctoral grant).

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the authors or upon any organization with which the authors are associated.

Correspondence to William B. Mortenson, MSc, G.F. Strong Rehabilitation Centre, 4255 Laurel St, Vancouver, BC V5Z 2G9, Canada, e-mail: bmortens@interchange.ubc.ca. Reprints are not available from the authors.

0003-9993/08/8906-0063\$34.00/0

doi:10.1016/j.apmr.2008.01.010

psychologic functioning of the individual. Activity is defined as the execution of a task or action, and participation is defined as involvement in a life situation.¹⁷ The term *capacity*, which is associated with activity, describes what a person can do, whereas the term *performance*, which is associated with participation, describes what a person normally does do in their natural environment. Capacity assessments often take place in controlled settings, so that the person's ability without personal assistance can be determined.¹⁶

Although a variety of generic measures of participation exist, reviews of these measures have identified a number of general limitations. First, many do not include all of the ICF activity participation domains (eg, Craig Hospital Assessment and Reporting Technique, Reintegration to Normal Living, Impact on Participation and Autonomy, World Health Organization Disability Assessment Schedule 2).¹⁸ Second, older instruments that focus on the construct of handicap (as participation restriction) may not fully capture the full scope of the concept of participation in the ICF.¹¹ Third, in their review of 100 rehabilitation outcome measures, Rust and Smith¹⁹ found that 30% ignored the impact of assistive technology, 44% reduced scores for people who used assistive technology (eg, the FIM instrument, Sickness Impact Profile), and only 22% incorporated the use of the assistive technology in the outcome measure. In light of such findings, Harris concluded that "current [generic] measures of participation are inadequate to the task of measuring activity and participation among wheeled mobility users"^{11(p137)} and advocated for the use of measures that captured device-specific information.¹¹ To advance research in this area, this review provides a conceptual, psychometric, and pragmatic critique of a variety of activity and participation outcome measures, intended exclusively for people who use wheelchairs, that are used clinically and in research.

METHODS

We searched the literature systematically and extracted relevant data about outcome instruments that are used to measure activity and participation of people who use wheelchairs. The measures included in this study came from peer-reviewed articles that reported on wheelchair-specific measures included in the ICF activity and participation domains. Measures were included only if the articles were in English and the subjects were older than 18.

Search Strategy

We searched CINAHL, PsychInfo, EMBASE, Google Scholar, and Medline databases to identify peer-reviewed articles using keywords that included: *function, activity, assistive technology, wheelchair(s), psychometrics, responsiveness, sensitivity to change, questionnaires, participation, outcome assessment, outcomes, treatment outcomes, reproducibility of results, validity, and validation studies*. No time limits were imposed on the search, but it was completed in August 2007 and therefore does not include measures published after this time. Titles of articles were screened and, if they fit the criteria, an abstract was retrieved. If the abstract was deemed relevant, the article was obtained. When a wheelchair-specific activity or participation outcome measure was identified, all related publications on the development and testing of the measure, including grey literature such as conference proceedings, thesis publications, and the instruction manuals (if available) were acquired.

Appraisal of Measurement Tools

It is necessary to understand the conceptual basis of a tool in order to evaluate it.¹² Although this is related to the content validity of an instrument, the concept of content validity is much narrower and does not invite inter-measurement comparisons on a theoretical level. For conceptual comparisons, each instrument was classified according to the ICF based on the process proposed by Cieza et al.²⁰ All items from all measures were coded independently by 2 authors; and for items in which there was disagreement, consensus was achieved through discussion. When the potential existed for the use of multiple activity or participation item codes, a single code was selected that best represented the item. Based on this review, each measure was classified in terms of whether it measured activity or participation, or both, using the ICF approach 1,^{17(p234)} in which certain domains are designated as activity and certain domains are designated as participation. Instruments with items that assessed mobility, self-care, and domestic life domains were deemed activity measures. Measures with items that assessed interpersonal interactions and relationships, major life areas and community, and social and civic life were considered participation measures; and measures that included a balance of items from across all domains were considered to represent both. Jette et al.²¹ provide some empirical support for this approach. Their factor analysis of data from administration of the Late Life Disability and Functioning Instrument revealed a 3-factor solution: mobility activities, daily activities, and social/participation, which suggests activity and participation are distinct constructs. Finally, a determination was made about whether the tool measured capacity or performance. Instruments that measured the wheelchair user's ability (capacity) in his or her natural setting were described as "capacity in natural environment" measures.

Reliability

For the purpose of this review, we used Andresen's grading criteria¹² to compare reliability scores across instruments based on the lowest value reported. In this system, an "A" corresponds to reliability coefficients .75 or higher (intraclass correlation coefficient [ICCs] or κ value), "B" ranges between 0.4 and 0.75, and "C" was 0.4 or lower. "A+" was used for reliability coefficients greater than 0.9, because reliability coefficients of this magnitude have been recommended to allow individual comparisons.²²

Validity

Content, face, criterion, and construct validity results and responsiveness data were extracted for each measurement tool. Measures of construct validity indicate whether measurements from the instrument vary in a priori hypothesized directions with measurements of related constructs.²³ To compare construct validity scores across instruments, Andresen's grading criteria¹² were used when correlation values were available. According to this grading scheme, a grade of A corresponds to correlations with similar measures or related constructs of .60 or higher, a grade of B ranges between 0.3 and 0.6, and a grade of C is 0.3 or lower. Tools that only reported positive extreme group validity without additional validity data were given a B grade. A widely used measure of responsiveness is the standardized response mean (SRM).²⁴ To compare across instruments, SRM values 0.8 or higher were graded as A, values between 0.5 and 0.8 were graded B, and values of 0.5 or lower were graded C.²⁵ Tools that only reported significant changes over time, as hypothesized, but did not report an SRM were given a B grade.

Wheelchair Contribution

Another critical element for people who use wheelchair-specific outcome measures is whether the instrument considers the impact of assistive technology on the results obtained, because this potentially confounding variable is frequently overlooked in the field of health measurement and rehabilitation.¹⁹ It is therefore very important to detail how assistive technology is considered in the scoring. In the present paper, we refer to this characteristic as *wheelchair contribution*, which is defined as how well the tool isolates the contribution of the wheelchair to activity and participation. To compare across measures, wheelchair contribution was operationalized as follows: a grade of A was assigned when (1) the scoring system involved a dual rating to clearly differentiate the effects of wheelchair use versus no wheelchair use (or different models of wheelchair and other assistive devices) on activity and participation and (2) wheelchair specifications were clearly documented. A grade of B was given when procedural instructions documented wheelchair specifications and specified how to assess a person using different methods (such as different models of wheelchairs or components of wheelchairs) but did not support head to head comparisons. A grade of C was given when there was no documentation of wheelchair specifications.

Applicability of Measures

Applicability, as a complement to psychometric properties of a measurement instrument, also requires careful appraisal.^{12,26,27} Applicability represents the pragmatic qualities of a tool that enable its use with a given population or in a specific context.²⁸ When the applicability of a measurement tool is appraised, many criteria are addressed: respondent burden, administrative burden, and format compatibility. Respondent burden indicates the level of effort required, acceptance of the items, or discomfort caused by the assessment procedure.²⁹⁻³¹ Administrative burden comprises practical elements that the clinician or researcher considers as advantages or limitations when the measurement tool is applied in a specific context, such as training required, and time to administer and score the measure.^{14,27} Finally, format compatibility examines the match between the target population characteristics and the assessment format to avoid age, sex, cultural, or disability biases.^{12,32} For respondent burden, tools were given an A grade if they took 15 minutes or less and had high acceptability to respondents, a B grade if they were either longer (but appropriately so) or had some reported problems with acceptability, or a C grade if both length and acceptability were problematic.¹² For the administrative burden and format compatibility aspects, a description was provided for each tool.

RESULTS

Fifty-eight measures were identified for consideration, and 11 measures were included in the final review. Two wheelchair-specific measures were excluded because they focused on physical activity as exercise. The Physical Activity Scale for Individuals with Physical Disabilities³³ was excluded because it was designed to identify metabolic equivalence for physical activity. The Wheelchair Physical Functional Performance test³⁴ was excluded because it includes a number of functional activities that intend to measure the domains of upper-body strength, upper-body flexibility, lower-body strength, endurance, and balance and coordination, which are more closely related to body structure and function than to activity and participation in the ICF. The other measures (n=45) were excluded because they were not wheelchair user specific. For example, the Participation Survey/Mobility (PARTS/M)³⁵ was

not included in the review because it was developed for people with mobility impairments living in the community rather than for people who use wheelchairs exclusively. Similarly, the FIM instrument was excluded, because it is not exclusively for people who use wheelchairs and, in fact, penalizes subjects for wheelchair use.³⁶

Table 1 lists the 11 tools included in the review and includes a description of their intended population(s), number of items, response scales, and conceptual appropriateness.

Conceptual Appropriateness

The fourth column of table 1 indicates how items from each measure were classified into ICF chapters using the criteria of Cieza et al.²⁰ All but 3 measures cover the domain of mobility exclusively. The Wheelchair Users Functional Assessment (WUFA)³⁷ includes 2 items from the self-care domain (dressing, bathing). The Wheelchair Outcome Measure (WhOM)³⁸ is the only instrument that uses items nominated and weighted by the client. As such, it could potentially, but not necessarily, include items from all participation domains, because the rater solicits information that is participation focused. The Functional Evaluation in a Wheelchair Questionnaire version (FEW-Q)³⁹ includes a mix of body function, activity and participation, and environmental elements. The FEW-Q was the most difficult measure to classify, because all items are multi-barreled. For instance, question 5 asks respondents to rate their agreement with the following statement (with ICF codes added): "The size, fit, postural support and functional features of my wheelchair/scooter [e1201: Assistive products and technology for personal indoor and outdoor mobility and transportation] allow me to operate it [d465: Mobility chapter-moving around using equipment] as independently [not classified in the ICF], safely [not classified in the ICF] and efficiently [not classified in the ICF] as possible."^{40(p182)} In this case, as in most FEW-Q items, the respondent needs to consider the impact of 4 features of the assistive technology on 3 characteristics of the quality of 1 facet of activity or participation by providing only 1 response on a 7-point scale.

Looking more closely at the mobility item codes assigned to each measure, a majority of items were classified as d465 (moving around using equipment). For some measures, this code covers up to 30 different items. Other measures reflect a wider spectrum of mobility tasks, including transfers, reaching, manipulating wheelchair accessories, and using transportation.

The fifth column of table 1 indicates whether measures were deemed capacity or performance assessments. Most measures involve standardized tasks in a standardized environment, suggesting a capacity assessment. Other measures are based on what the person can do in his/her natural environment. Two involve the observation of customized tasks, selected based on the individual needs of the client, which are accomplished in his/her natural environment (Powered Indoor Driving Assessment [PIDA],⁴¹ Powered Community Driving Assessment [PCDA]⁴²). As such, these are considered to be capacity assessments in the natural environment. As a self-report measure, the FEW-Q also involves a capacity assessment in the natural environment, because the questions focus on what the person can do with his/her wheelchair. The WhOM belongs to a unique category, because it captures the satisfaction with performance of activities or participation when using a wheelchair.

Reliability

The reliability coefficients for each measure are presented in table 2. Reliability evidence for wheelchair user specific mea-

Table 1: Instrument Characteristics of Wheelchair-Specific Activity and Participation Measures

Instrument	Population	Items	ICF Constructs (items per ICF chapter)	Performance or Capacity	Response Scale (Range)
Wheelchair Users Functional Assessment (WUFA)	MW	13	Mobility (9) Self-care (2) Not classified (2)	C	7 point (1–7)
Functional Evaluation in a Wheelchair Questionnaire (FEW-Q)	PM/MW	10	Mobility (6) Self-care (3) General tasks and demands (1)	C/P	6 point (1–6)
4 functional tasks	MW	4	Mobility (4)	C	Time distance
Wheelchair Skills Test, version 1.0 (WST 1.0)	MW	33	Mobility (33)	C	3-point scale (0–2)
Wheelchair Skills Test, version 2.4 (WST 2.4)	MW	50	Mobility (50)	C	Pass/fail
Additional Mobility and Locomotor Items (AML) assessment	MW*	6 ^s	Mobility (6)	C	6 point (1–6)
Obstacle Course Assessment of Wheelchair User Performance (OCAWUP)	PM/MW	10	Mobility (10)	C	Time 4 point (0–3)
Wheelchair Outcome Measure (WhOM)	PM/MW	Varies	Client identified participation issues	P [†]	11 point (0–10)
Wheelchair Circuit	MW [†]	8	Mobility (8) and exercise capacity (body function)	C	2 (0–1) and 3 point (0, 0.5, 1) Time Maximal heart rate
Power-mobility Indoor Driving Assessment (PIDA)	PM	30	Mobility (30)	C/P	4 point (1–4)
Power-mobility Community Driving Assessment (PCDA)	PM	Varies	Mobility (14–36)	C/P	4 point (0–3)

Abbreviations: C, capacity in standardized environment; C/P, capacity in natural environment; MW, manual wheelchair; P, performance; PM, power mobility.

*People with paraplegia.

[†]People with spinal cord injury.

[‡]Measures satisfaction with performance.

^sA 5-item version of this tool (5-AML) has been suggested to augment the FIM.

asures is limited. One measure has been subjected to no reliability study (PCDA); and, for most other studies, reliability estimates are established based on only 1 study. The published reliability coefficients are generally good ($>.75$), except for the intrarater reliability coefficient of the PIDA, which was .67,⁴¹ however, only 4 measures have coefficients recommended for

individual comparisons: 4 functional tasks,⁴³ Wheelchair Skills Test (WST), version 2.4,⁴⁴ WhOM,⁴⁵ and Wheelchair Circuit.^{46,47} The type of ICC is not reported for 3 tools (WST 2.4,⁴⁴ WUFA,³⁷ PIDA⁴¹), and the confidence intervals (CIs) are missing for 3 of them (WUFA, WST 2.4, WST 1.0⁴⁸). For the WST 1.0, a Spearman correlational coefficient for intrarater reliability is

Table 2: Instrument Reliability

Instrument	n	Test-Retest Coefficient	n	Intrarater	n	Interrater
WUFA		NR	5	ICC=.78 ^{†‡}	5	ICC=.96 ^{†‡}
FEW-Q*	30	ICC=.86		NR		NR
FFT	10	r=.99 [†]			10	ICC=.99 ^{‡§}
WST 1.0	21	r=.65 [†]	21	r=.96 [†]	24	r=.95 [†]
WST 2.4	20	ICC=.90 ^{†‡}	20	ICC=.96 ^{†‡}	20	ICC=.97 ^{†‡}
AML					20	κ range, .82–.96
OCAWUP	17	ICC range, .74–.99 (time) κ range, 0.09–1.00 (EU) ICC \geq .96 (OS)		NR	17	ICC \geq .98 (time) κ range, 0.29–1.00 (EU) ICC \geq .97 (OS)
WhOM		NR	50	ICC=.93	50	ICC=.91
Wheelchair Circuit		NR	27	ICC=.98	27	ICC=.97
PIDA		NR	15	ICC=.67 [†]	15	ICC=.87 [†]
PCDA		NR		NR		NR

Abbreviations: EU, ease of use; FFT, 4 functional tasks; NR, not reported; OS, overall score.

*Reliability data from FEW 2.0.

[†]CI not reported.

[‡]Type of ICC not reported.

[§]Excluded 1 item.

Table 3: Validity and Wheelchair Contribution

Instrument	Face/Content	Construct/Concurrent	Wheelchair Contribution
WUFA	1,2	NR	NI
FEW-Q	2	NR	NI
FFT	1	NR	NI
WST 1.0	1	TGR $r=.45$	WSR
WST 2.4	2	Age (all) $r=-.43$ Age (wheelchair users) $r=-.27$ TGR $r=.40$ Admission FIM $r=.38$	WSR
AML	1	Extreme group	NI
OCAWUP	1,2,3	FIM mobility and global score of ease $r=.84$	WSR
WhOM	1,2	QUEST (total score) $r=.58$ QUEST (assistive device scale) $r=.66$ LIFE-H (entering and exiting home) $r=.33$ LIFE-H (moving around outside home) $r=.35$	WSR
Wheelchair Circuit	3	FIM mobility and wheelchair ability $r=.52$	NI
PIDA	1,2	NR	WSR
PCDA	1,2	NR	WSR

Abbreviations: LIFE-H, Assessment of Life Habits; NI, no instruction about subject's use of wheelchairs or other assistive technology; NR, not reported; QUEST, Quebec User Evaluation With Assistive Technology; TGR, therapist global rating; WSR, wheelchair specifications recorded. Legend: 1, experts/health professional involved in development; 2, wheelchair users involved in development; 3, based on other related instruments.

reported, but this is corrected in the next version (WST 2.4). For the FEW-Q data for the FEW 2.0 are presented.⁴⁰

Validity

Content and face validity, construct validity, and wheelchair contribution are described in table 3. With the exception of the WST, Obstacle Course Assessment of Wheelchair User Performance (OCAWUP),^{49,50} Wheelchair Circuit, Additional Mobility and Locomotor Items (AML) assessment, and WhOM, validity information is very limited. External comparisons using multitrait multimethod techniques were a common approach to evaluating the construct validity of measures. For example, the FIM instrument³⁶ was used to establish the construct validity for the WST 2.4, Wheelchair Circuit, and OCAWUP, with correlations of .38 to .31, .52, and .84 respectively. Often, however, the direction and magnitude of the a priori hypotheses was not reported, which makes the results equivocal.

Only the Wheelchair Circuit, AML,^{51,52} and WST 1.0 have reported responsiveness. Significant improvement in wheelchair skills postrehabilitation were found with the Wheelchair Circuit. SRMs were 0.6 for ability, 0.9 for performance time, and 0.8 for physical strain.⁴⁶ In contrast, correlation between WST 1.0 change scores between time 1 and time 2 and therapist rated change scores were not statistically significant.⁴⁸ Research on the AML (5-item version) found significant changes in scores over time post spinal cord injury (SCI).⁵²

Extreme group validity was reported for 4 measures: the WST 2.4, OCAWUP, AML, and Wheelchair Circuit. For the WST 2.4, people who used wheelchair longer than 21 days scored significantly better; and scores varied significantly between diagnostic groups.⁴⁴ For the OCAWUP, there was a significant difference in global score of ease between groups using different propulsion methods.⁵⁰ For the Wheelchair Circuit ability score differed significantly based on level of lesion.⁴⁴ For the AML (5-item version), scores on the tool differed significantly between people with tetraplegia and those with paraplegia.⁵²

Wheelchair Contribution

The wheelchair contribution for each measure is included in the last column of table 3. The FEW-Q, 4 functional tasks, Wheelchair Circuit, WUFA, and AML did not document or otherwise consider the type of wheelchair system used by the person.

Applicability

The applicability of each measure is described in table 4. The respondent burden is high for most measures, especially for those based on task observation. The administration time exceeds 15 minutes for all measures except the 4 functional tasks and FEW-Q. Despite its brevity, however, the respondent burden of the FEW-Q is considered fair, because of the double-barreled nature of its questions. The PCDA is the most demanding measure, because 3 sessions may be needed to complete the assessment. All of the measures are available for free and the scoring is done by hand. None require raters to attend special training sessions, and free manuals are available for the WST, FEW-Q, WhOM, PCDA, PIDA, and OCAWUP (in French). Most measures are only applicable to 1 category of wheelchair, specifically to manual wheelchairs (n=6) or to powered wheelchairs (n=2). Three measures in the review are intended for people who use manual and power wheelchairs: the FEW-Q, OCAWUP, and WhOM. A power mobility version of the WST is available, but no validity or reliability results are available to date.

With respect to format compatibility, the populations studied were generally heterogeneous in terms of diagnostic categories and sex. However, 4 measures were tested exclusively with people with a diagnosis of SCI (4 functional tasks, AML, WhOM, Wheelchair Circuit); and 2 were only tested with men (4 functional tasks, PIDA). The age span covers adults and older adults, except for the PCDA, which was validated for older adults only (66–93y) and for the Wheelchair Circuit, WhOM, OCAWUP and 4 functional tasks that included adults under the age of 65 years. Only the Wheelchair Circuit, WhOM, OCAWUP, and WST have alternative forms avail-

Table 4: Applicability of Measures

Instrument	Alternative Forms	Administration Time (min)	Respondent Demands	Rater Training	Expense (\$)	Availability	No. of Studies
WUFA	No	45	TP	0	0 [#]	Authors	1
FEW-Q	Yes*	15	P&P	0	0	Online	1
FFT	No	15	TP	0	0 [#]	Authors	1
WST 1.0	Yes	45	TP	¶	0	Online	1
WST 2.4	Yes [†]	45	TP	¶	0 [#]	Online	1
AML	No	30	TP	0	0 [#]	Authors	2
OCAWUP	Yes [‡]	45	TP	¶	0 [#]	Authors	2
WhOM	Yes [‡]	30	P&P	¶	0	Authors	1
WC	Yes [§]	45	TP	0	0 [#]	Authors	2
PIDA	No	45	TP	¶	0 [#]	Online	1
PCDA	No	60	TP	¶	0	Online	1

Abbreviations: P&P, pencil and paper; TP, task performance; WC, Wheelchair Circuit.

*An observational capacity and performance based FEW-Q has been developed.

[†]Questionnaire format available.

[‡]French and English forms available.

[§]Dutch and English forms available.

^{||}Potential risk for injury requires use of spotter strap.

[¶]Raters are instructed to review the manual.

[#]Zero dollars for measure, but other items may need to be acquired.

able. A Dutch version of the Wheelchair Circuit is available. French versions of the WhOM, OCAWUP, and WST are available, and a questionnaire version of the WST 2.4 was validated for a telephone or face-to-face interview.⁵³

Many of the capacity measures require substantial space and equipment (ie, WST 1.0, WST 2.4, Wheelchair Circuit, OCAWUP). For example, the OCAWUP requires a 6-m (20-ft) ramp that can be adjusted to 3 different slope angles.

Summary Appraisal

Table 5 summarizes the properties of all of the measures included in the review. Good reliability evidence is reported for most of the measures, but validity information is only available for 6 of the 11 measures and responsiveness for 3. In this regard, the Wheelchair Circuit, AML, and WST 2.4 have the most extensive psychometric information. The respondent burden is fair for most measures, except the 4 functional tasks, which has the least respondent burden and is considered excellent.

The relationship between the measures and the ICF taxonomy is illustrated in figure 1. Position along the x axis indicates the degree to which the instrument measures activity or par-

ticipation. Position on the y axis indicates whether the instrument measures performance or capacity. Position on the z axis indicates how well the tool isolates the contribution of the wheelchair to activity and participation. For example, the WUFA measures capacity (y axis) to perform activities (x axis), but does not indicate that the type of wheelchair and wheelchair seating should be recorded (z axis). Alternatively, with the PCDA, the type of power chair is recorded (z axis) and the activity of power mobility (x axis) is evaluated in the natural environment of the user, so it is neither a capacity test in a standardized environment nor a performance assessment of what the user normally does. For this reason, it is positioned halfway between capacity and performance (y axis). Measures like the FEW-Q and WUFA are indicated using slightly wider boxes, as they incorporate a small number of nonmobility related self-care activities.

DISCUSSION

Most of the items of the measures included in this review, with the exception of the WhOM, FEW-Q, and WUFA, focus exclusively on the construct of mobility. The FEW-Q and WUFA include only a few self-care items, whereas the WhOM

Table 5: Summary of Measures

Instrument	Measurement Properties				
	Reliability	Validity	Responsiveness	WC Contribution	Respondent Burden
WUFA	A	NR	NR	C	B
FEW-Q	A	NR	NR	C	B
FFT	A+	NR	NR	C	A
WST 1.0	B	B	C	B	B
WST 2.4	A+	B	NR	B	B
AML	A	B	B	C	B
OCAWUP	A	A	NR	B	B
WhOM	A+	B	NR	B	B
WC	A+	B	B	C	B
PIDA	B	NR	NR	B	B
PCDA	NR	NR	NR	B	B

Abbreviations: NR, not reported.

Legend: A, very good; A+, excellent (allows for individual comparisons); B, fair; C, poor.

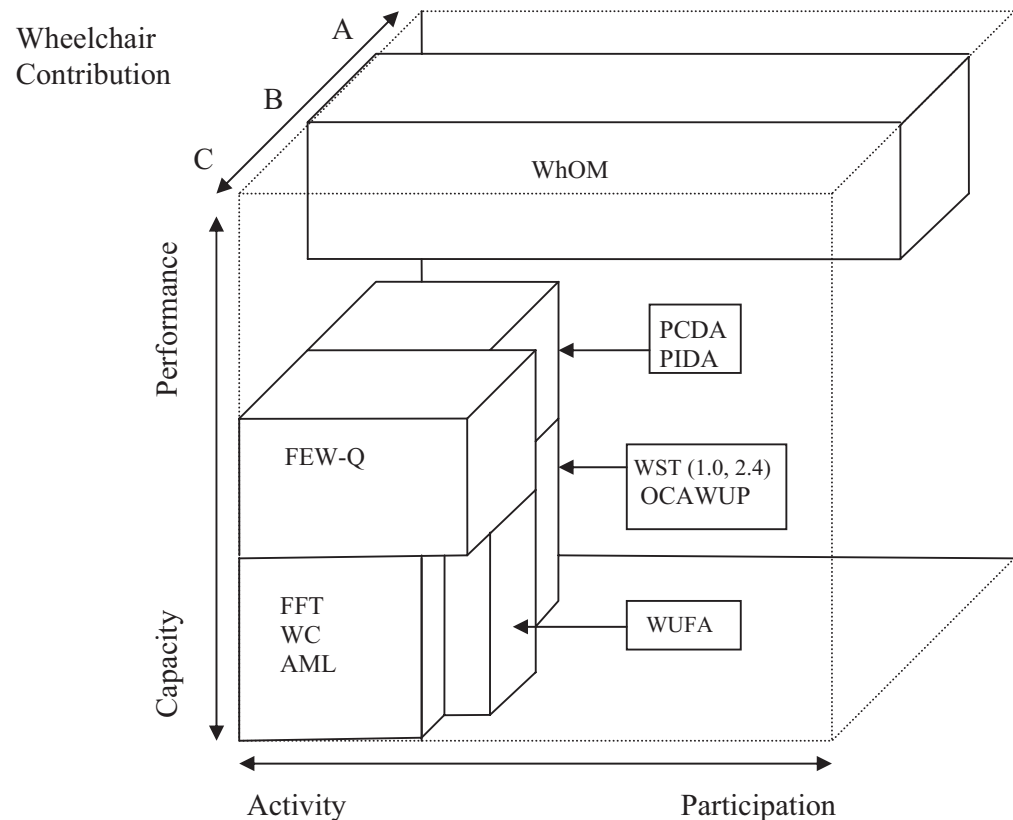


Fig 1. Instruments and the ICF.
Abbreviations: FFT, 4 functional tasks; WC, wheelchair circuit.

is the only measure to potentially include all ICF participation chapters (domains). The preassessment interviews of the OCAWUP and PCDA refer to participation in areas such as work and family roles, but the scoring does not take those aspects into consideration. Moreover, most of the measures tend to measure capacity in standardized environments, rather than performance or capacity in real-life settings. Some measures, such as the PIDA, PCDA, and FEW-Q, assess the capacity of the wheelchair user in his/her natural environment; but this approach does not capture what the person actually does on a daily basis. The WhOM and FEW-Q are the only measures that look at performance subjectively. It has been suggested that both the objective and the subjective measurement of capacity and performance represent complementary rather than competing perspectives.⁵⁴ Therefore, figure 1 can assist researchers and clinicians in the selection of complementary tools.

There are a number of issues with the reliability coefficients of the measures in this review. Two of these problems are the lack of CIs and use of Spearman correlations for inter- and intrarater reliability. These are of particular concern, because they fail to account for within subject and between subject levels of error, which undermines confidence in the accuracy of the reliability coefficients. As well, although most measures have reliability coefficients above .75, only 4 measures (4 functional tasks, WST 2.4, WhOM, Wheelchair Circuit) show reliability coefficients at the level recommended for individual comparisons.²²

Validity is an issue for all of the measures in this review. Many measures only have content validity evidence (WUFA, FEW-Q, PIDA, PCDA). Construct validity testing and responsiveness estimates are needed to give credibility to these mea-

asures. As indicated by generalizability theory,⁵⁵ it cannot be assumed that reliability and validity evidence from 1 population or setting can be applied to another. Many of the measures require validation in various populations of people who use wheelchairs and their applicability to the clinical setting deserves some attention, because the administrative burden is quite demanding. For example, because the Wheelchair Circuit was validated only with people with SCI who used manual wheelchairs, additional research with other populations is warranted. The environment is another aspect that plays an important role in shaping activity and participation outcomes of people who use wheelchairs. Future measures should document more extensively either how those variables interact in real-life situations or how they can be used in conjunction with environmental measures such as the Craig Hospital Inventory of Environmental Factors⁵⁶ or the Measure of the Quality of the Environment,⁵⁷ which could capture the impact of physical and social environmental factors on performance.

By not documenting the type of wheelchair system used, the FEW-Q, 4 functional tasks, Wheelchair Circuit, WUFA, and AML omit a potentially confounding variable. For example, residents in nursing facilities might have limited participation, not because of a lack of desire or capacity but because of their inappropriate wheelchair systems.⁵⁸ As well, FEW-Q, 4 functional tasks, Wheelchair Circuit, WUFA, and AML do not provide guidelines about the subject's use of more than 1 wheelchair during testing. For example, a client may use a power wheelchair to facilitate grocery shopping in the community and a manual chair to maximize participation inside the home; but this variable would potentially be unidentified with these measures. This lack of attention to assistive technology is a problem common to many outcome measures.¹⁹

Measures that ignore the impact of wheeled mobility on activity and participation are limited in their ability to evaluate the effectiveness of wheelchair interventions. To correct this deficiency to some extent, the instructions for administration of these measures could easily be altered by insisting that wheelchair attributes, including model, seating, and configuration and set up, be recorded prior to administering the measure. As well, the use of multiple wheelchairs during testing needs to be documented.⁵⁹ As suggested by our definition of wheelchair contribution, ideally the measure would use a dual rating to differentiate clearly the effects of different models of wheelchairs on activity and participation outcomes. None of the measures satisfied this criterion.

Overall, none of the measures were scored as excellent on all of the ratings, which suggests that all could be improved or evolved in some way. Obviously, most measures would benefit from additional psychometric testing, especially in terms of validity and responsiveness testing. If these tools are to be used clinically, measures like the WUFA, FEW-Q, and PIDA may require modification or additional testing to ensure that their reliability coefficients are sufficient to make individual comparisons.

In terms of using the ICF to classify wheelchair activity and participation measures, 2 primary issues were identified. First, a majority of items were coded as d465 (moving around with equipment), which suggests that this code does not provide adequate descriptive detail. For example, this code, which does not even specify the type of equipment being used, was used to encapsulate a diverse assortment of activities including propelling on different surfaces (eg, carpet, gravel, pavement, grass), negotiating various curbs, inclines, and thresholds and navigating in a variety of settings. Unlike the code for walking (d450), which has additional fourth level codes that allow specification of walking short distances (d4500), walking long distances (d4502), walking on different surfaces (d4502), and walking around obstacles (d4503), the absence of fourth level codes for d465 seriously limits the granularity of this category. Second, elements of items, in the FEW-Q, for example, were coded as "not covered," because the ICF does not include subjective elements of activity and participation, which has previously been noted as a limitation.^{60,61}

There are a number of strengths and limitations that are evident when wheelchair-specific measures are compared with generic measures of activity and participation. A strength of wheelchair-specific measures is that they encourage consideration of factors that are particularly relevant to wheelchair users, such as wheelchair propulsion, terrain, and issues of working at different surface heights. One drawback to the use of wheelchair-specific measures is that they prevent comparisons between wheelchairs users and other populations. It should be noted, however, that such head-to-head comparisons may not be feasible with more generic measures, which may lack either face or content validity for people who use wheelchairs or may be unacceptable to them. This would include measures such as the Barthel Index,⁶² FIM instrument,³⁶ and Assessment of Life Habits⁶³ that reduce scores for people who use assistive technology. As well, the inclusion of ambulation items in instruments such as the Barthel Index, FIM instrument, Late Life Function and Disability Instrument,⁶⁴ and WHO Disability Assessment II⁶⁵ may limit acceptability and thus negatively impact the completion rates of people who use wheelchairs as their primary means of mobility.⁶⁶

Study Limitations and Strengths

This review was limited to wheelchair-specific measures that were described in at least 1 published peer-reviewed article.

For example, the WST versions 3.2 and 4.1 were excluded from the review because they have not been published in a peer-reviewed journal. Measures described only in grey literature, such as conference proceedings or academic theses, were also not included. As the review focused on measures developed exclusively for people who use wheelchairs, it did not include measures developed for other populations that may have applicability for these people, such as (1) the PARTS/M³⁵ that was developed for people with mobility impairments living in the community, (2) diagnosis-specific instruments, like those intended for people with SCI, many of whom use wheelchairs, or (3) more generic measures of participation that do not penalize people who use wheelchairs including the Impact on Participation and Autonomy Questionnaire⁶⁷ or Keele Assessment of Participation.⁶⁸ A number of wheelchair-specific measures were excluded because of their focus on body function and structure, and others were excluded because they were diagnosis rather than wheelchair-specific. The evaluation of the conceptual issues associated with these measures is a particular strength of the study, as many reviews of outcome measures focus exclusively on issues of reliability, validity, and applicability.

CONCLUSIONS

The field of outcome measurement for wheelchair outcomes is still in its infancy. A comprehensive review of activity and participation outcome measures for people who use wheelchairs identified 11 wheelchair-specific measures. Most of the measures were focused on the measurement of wheeled mobility capacity; and only 3 tools, the FEW-Q, WUFA, and WhOM, looked at activity and participation more broadly. The WhOM was the only tool that could potentially measure activity and participation across ICF domains. The review found that the psychometric testing of most of these measures was limited. This study suggests the need for further testing and development of wheelchair-specific outcome measures that will allow clinicians to justify their equipment recommendations and to show the effectiveness of their interventions to their clients, colleagues, and health care administrators.

Acknowledgment: We thank Jacek Kopec, MD, PhD, for his suggestions on an early draft of the review.

References

1. Sawatzky B. Wheeling in the new millennium: the history of the wheelchair and the driving force in wheelchair design today. Available at: http://www.wheelchairnet.org/WCN_WCU/SlideLectures/Sawatzky/WC_history.html. Accessed March 21, 2006.
2. Edlich RF, Nelson KP, Foley ML, Buschbacher RM, Long WB, Ma EK. Technological advances in powered wheelchairs. *J Long Term Eff Med Implants* 2004;14:107-30.
3. DiGiovine CP. Advances in manual wheelchair technology. *Top Spinal Cord Inj Rehabil* 2006;11:1-14.
4. Kaye HS, Kang T, LaPlante MP. Mobility device use in the United States. Disability statistics report. Washington (DC): U.S. Department of Education, National Institute on Disability and Rehabilitation Research; 2000. p 14.
5. McClain L, Medrano D, Marcum M, Schukar J. A qualitative assessment of wheelchair users' experience with ADA compliance, physical barriers, and secondary health conditions. *Top Spinal Cord Inj Rehabil* 2000;6:99-118.
6. Hoenig H, Landerman LR, Shipp KM, George L. Activity restriction among wheelchair users. *J Am Geriatr Soc* 2003;51:1244-51.
7. Meyers AR, Anderson JJ, Miller DR, Shipp K, Hoenig H. Barriers, facilitators, and access for wheelchair users: substantive and methodologic lessons from a pilot study of environmental effects. *Soc Sci Med* 2002;55:1435-46.

8. Cahill SE, Eggleston R. Reconsidering the stigma of physical disability: wheelchair use and public kindness. *Sociol Q* 1995;36: 681-98.
9. Cahill SE, Eggleston R. Managing emotions in public: the case of wheelchair users. *Soc Psychol Q* 1994;57:300-12.
10. LaPlante MP, Hendershot GE, Moss AJ. Assistive technology devices and home accessibility features: prevalence, payment, need, and trends. *Adv Data* 1992;217:1-11.
11. Harris F. Conceptual issues in the measurement of participation among wheeled mobility device users. *Disabil Rehabil Assist Technol* 2007;2:137-48.
12. Andresen EM. Criteria for assessing the tools of disability outcome research. *Arch Phys Med Rehabil* 2000;81(12 Suppl 2): S15-20.
13. Bergner M, Rothman ML. Health status measures: an overview and guide for selection. *Annu Rev Public Health* 1987;8:191-210.
14. Finch E, Brooks D, Stratford PW, Mayo N. Physical rehabilitation outcome measures: a guide to enhanced clinical decision making. Hamilton: BC Decker; 2002.
15. Law M, Baum CM, Dunn W. Measuring occupational performance: supporting best practice in occupational therapy. Thorofare: Slack; 2001.
16. World Health Organization. Towards a common language of function, disability and health: ICF. Geneva: WHO; 2002. p 1-23.
17. World Health Organization. ICF: international classification of functioning, disability and health. Geneva: WHO; 2002.
18. Perenboom RJ, Chorus AM. Measuring participation according to the international classification of functioning, disability and health (ICF). *Disabil Rehabil* 2003;25:577-87.
19. Rust KL, Smith RO. Assistive technology in the measurement of rehabilitation and health outcomes: a review and analysis of instruments. *Am J Phys Med Rehabil* 2005;84:780-93.
20. Cieza A, Brockow T, Ewert T, et al. Linking health-status measurements to the international classification of functioning, disability and health. *J Rehabil Med* 2002;34:205-10.
21. Jette AM, Haley SM, Kooyoomjian JT. Are the ICF activity and participation dimensions distinct? *J Rehabil Med* 2003;35:145-9.
22. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development. 3rd ed. New York: Oxford Univ Pr; 2003.
23. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955;52:281-302.
24. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;53:459-68.
25. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:1-74.
26. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF activity. *Disabil Rehabil* 2005;27:315-40.
27. Wade DT. Assessment, measurement and data collection tools. *Clin Rehabil* 2004;18:233-7.
28. Auger C, Demers L, Swaine B. Making sense of pragmatic criteria for the selection of geriatric rehabilitation measurement tools. *Arch Gerontol Geriatr* 2006;43:65-83.
29. Andresen EM, Rothenberg BM, Kaplan RM. Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults. *Med Care* 1998;36: 1349-60.
30. Bureau-Chalot F, Novella J, Jolly D, Ankri J, Guillemin F, Blanchard F. Feasibility, acceptability and internal consistency reliability of the Nottingham health profile in dementia patients. *Gerontology* 2002;48:220-5.
31. Nakayama T, Toyoda H, Ohno K, Yoshiike N, Futagami T. Validity, reliability and acceptability of the Japanese version of the General Well-Being Schedule (GWBS). *Qual Life Res* 2000;9:529-39.
32. Goodgold S, Kiami S, Ule D, Schoenberg A, Forman G. Applicability of the functional reach and timed up and go tests for elderly individuals with Alzheimer's disease: pilot investigation. *Phys Occup Ther Geriatr* 2001;19:21-36.
33. Washburn RA, Zhu W, McAuley E, Frogley M, Figoni SF. The physical activity scale for individuals with physical disabilities: development and evaluation. *Arch Phys Med Rehabil* 2002;83: 193-200.
34. Cress ME, Kinne S, Patrick DL, Maher E. Physical functional performance in persons using a manual wheelchair. *J Orthop Sports Phys Ther* 2002;32:104-13.
35. Gray DB, Hollingsworth HH, Stark SL, Morgan KA. Participation survey/mobility: psychometric properties of a measure of participation for people with mobility impairments and limitations. *Arch Phys Med Rehabil* 2006;87:189-97.
36. Marino RJ, Huang M, Knight P, Herbison GJ, Ditunno JF, Segal M. Assessing selfcare status in quadriplegia: comparison of the quadriplegia index of function (QIF) and the functional independence measure (FIM). *Paraplegia* 1993;31:225-33.
37. Stanley RK, Stafford DJ, Rasch E, Rodgers MM. Development of a functional assessment measure for manual wheelchair users. *J Rehabil Res Dev* 2003;40:301-7.
38. Mortenson WB, Miller WC, Miller-Polgar J. Measuring wheelchair intervention outcomes: development of the wheelchair outcome measure. *Disabil Rehabil Assist Technol* 2007;2:275-85.
39. Mills T, Holm MB, Treffer E, Schmeler M, Fitzgerald S, Boninger M. Development and consumer validation of the Functional Evaluation in a Wheelchair (FEW) instrument. *Disabil Rehabil* 2002; 24:38-46.
40. Mills TL. Functioning Everyday with a Wheelchair (FEW): development and validation of self-report and performance-based observation instruments to measure functional outcomes of seating-mobility interventions [PhD dissertation]. Pittsburgh: Univ Pittsburgh; 2003.
41. Dawson D, Chan R, Kaiserman E. Development of the power-mobility indoor driving assessment for residents of long-term care facilities: a preliminary report. *Can J Occup Ther* 1994;61:269-76.
42. Letts L, Dawson D, Kaiserman-Goldenstein E. Development of the power-mobility community driving assessment. *Can J Rehabil* 1998;11:123-9.
43. May LA, Butt C, Minor L, Kolbinson K, Tulloch K. Measurement reliability of functional tasks for persons who self-propel a manual wheelchair. *Arch Phys Med Rehabil* 2003;84:578-83.
44. Kirby RL, Dupuis DJ, Macphee AH, et al. The Wheelchair Skills Test (version 2.4): measurement properties. *Arch Phys Med Rehabil* 2004;85:794-804.
45. Garden J, Miller W, Mortenson W. Reliability and validity of the wheelchair outcome measure (WhOM). Presented to: The 22nd International Seating Symposium; 2006 Mar 2-4; Vancouver (BC).
46. Kilkens OJ, Dallmeijer AJ, de Witte LP, van der Woude LH, Post MW. The Wheelchair Circuit: construct validity and responsiveness of a test to assess manual wheelchair mobility in persons with spinal cord injury. *Arch Phys Med Rehabil* 2004;85:424-31.
47. Kilkens OJ, Post MW, van der Woude LH, Dallmeijer AJ, van den Heuvel WJ. The Wheelchair Circuit: reliability of a test to assess mobility in persons with spinal cord injury. *Arch Phys Med Rehabil* 2002;83:1783-8.
48. Kirby RL, Swuste J, Dupuis DJ, MacLeod DA, Monroe R. The Wheelchair Skills Test: a pilot study of a new outcome measure. *Arch Phys Med Rehabil* 2002;83:10-8.
49. Routhier F, Vincent C, Desrosiers J, Nadeau S, Guerette C. Development of an obstacle course assessment of wheelchair user

- performance (OCAWUP): a content validity study. *Technol Disabil* 2004;16:19-31.
50. Routhier F, Desrosiers J, Vincent C, Nadeau S. Reliability and construct validity studies of an obstacle course assessment of wheelchair user performance. *Int J Rehabil Res* 2005;28:49-56.
 51. Harvey LA, Batty J, Fahey A. Reliability of a tool for assessing mobility in wheelchair-dependent paraplegics. *Spinal Cord* 1998;36:427-31.
 52. Middleton JW, Harvey LA, Batty J, Cameron I, Quirk R, Winstanley J. Five additional mobility and locomotor items to improve responsiveness of the FIM in wheelchair-dependent individuals with spinal cord injury. *Spinal Cord* 2006;44:495-504.
 53. Mountain AD, Kirby RL, Smith C. The Wheelchair Skills Test, version 2.4: validity of an algorithm-based questionnaire version. *Arch Phys Med Rehabil* 2004;85:416-23.
 54. Noreau L, Fougereyrollas P, Post M, Asano M. Participation after spinal cord injury: the evolution of conceptualization and measurement. *J Neurol Phys Ther* 2005;29:147-56.
 55. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 3rd ed. Upper Saddle River: Prentice Hall; 2008.
 56. Whiteneck GG, Harrison-Felix CL, Mellick DC, Brooks CA, Charlifue SB, Gerhart KA. Quantifying environmental factors: a measure of physical, attitudinal, service, productivity, and policy barriers. *Arch Phys Med Rehabil* 2004;85:1324-35.
 57. Boschen K, Noreau L, Fougereyrollas P. A new instrument to measure the quality of environment for persons with physical disabilities [abstract]. *Arch Phys Med Rehabil* 1998;79:1331.
 58. Fuchs RH, Gromak PA. Wheelchair use by residents of nursing homes: effectiveness in meeting positioning and mobility needs. *Assist Technol* 2003;15:151-63.
 59. Hoenig H, Giacobbi P, Levy CE. Methodological challenges confronting researchers of wheeled mobility aids and other assistive technologies. *Disabil Rehabil Assist Technol* 2007;2:159-68.
 60. Schuntermann MF. The implementation of the international classification of functioning, disability and health in Germany: experiences and problems. *Int J Rehabil Res* 2005;28:93-102.
 61. Hemmingsson H, Jonsson H. An occupational perspective on the concept of participation in the international classification of functioning, disability and health—some critical remarks. *Am J Occup Ther* 2005;59:569-76.
 62. Mahoney FI, Barthel D. Functional evaluation: the Barthel Index. *Md State Med J* 1965;14:56-61.
 63. Noreau L, Desrosiers J, Robichaud L, Fougereyrollas P, Rochette A, Viscogliosi C. Measuring social participation: reliability of the LIFE-H in older adults with disabilities. *Disabil Rehabil* 2004;26:346-52.
 64. Sayers SP, Jette AM, Haley SM, Heeren TC, Guralnik JM, Fielding RA. Validation of the Late-Life Function and Disability Instrument. *J Am Geriatr Soc* 2004;52:1554-9.
 65. Epping-Jordan J, Ustun T. The WHODAS-II: leveling the playing field for all disorders. *WHO Mental Health Bull* 2000;6:5-6.
 66. Oliver M. *The politics of disablement: a sociological approach*. New York: St. Martin's Pr; 1990.
 67. Cardol M, de Haan RJ, de Jong BA, van den Bos GA, de Groot IJ. Psychometric properties of the impact on participation and autonomy questionnaire. *Arch Phys Med Rehabil* 2001;82:210-6.
 68. Wilkie R, Peat G, Thomas E, Hooper H, Croft PR. The Keele Assessment of Participation: a new instrument to measure participation restriction in population studies. Combined qualitative and quantitative examination of its psychometric properties. *Qual Life Res* 2005;14:1889-99.