ORIGINAL REPORT

COMPARING THE VALIDITY OF FIVE PARTICIPATION INSTRUMENTS IN PERSONS WITH SPINAL CONDITIONS

Vanessa K. Noonan, PhD PT[1,2], Jacek A. Kopec, MD PhD[2,3], Luc Noreau, PhD[4,5], Joel Singer, PhD[2,6], Louise C. Mâsse, PhD[7], Hongbin Zhang, MSc[1] and Marcel F. Dvorak, MD[1]

From the [1]Division of Spine, Department of Orthopaedics, [2]School of Population and Public Health, University of British Columbia, [3]Arthritis Research Centre of Canada, Vancouver, BC, [4]Rehabilitation Department, Laval University, [5]Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Québec City, OC, [6]Canadian HIV Trials Network, and [7]Department of Pediatrics, University of British Columbia, Vancouver, BC, Canada

Objective: To evaluate and compare the construct validity of 5 participation instruments developed using the International Classification of Functioning, Disability and Health (ICF).

Methods: A total of 545 subjects diagnosed and treated for a spinal condition at an acute hospital were followed-up and consented to complete a questionnaire. Subjects completed 5 participation instruments (Impact on Participation and Autonomy (IPA), Keele Assessment of Participation (KAP), Participation Measure-Post Acute Care (PM-PAC), Participation Objective Participation Subjective (POPS), World Health Organization Disability Assessment Schedule II (WHODAS II)). In addition, each subject completed a health status instrument and a quality of life instrument. The dimensionality, convergent/discriminant validity and known-group validity of the participation instruments were assessed.

Results: A confirmatory factor analysis of the facture structure for the IPA and PM-PAC demonstrated adequate model fit. For convergent/discriminant validity, correlations were generally higher among similar domains of the WHODAS II, IPA, KAP and PM-PAC, and as expected the lowest correlations were observed with the objective domains of the POPS. Most instruments demonstrated known-group validity.

Conclusion: Differences in the construct validity evidence of the POPS compared with the other 4 instruments were noted. To date, there is no gold standard for measuring participation, and clinicians and researchers should consider the type of information required prior to selecting an instrument.

Key words: consumer participation; World Health Organization; rehabilitation; questionnaires; disability evaluation.

Correspondence address: Vanessa K. Noonan, Division of Spine, Department of Orthopaedics, University of British Columbia, Vancouver, BC, Canada. E-mail: Vanessa.Noonan@vch.ca

INTRODUCTION

It is increasingly recognized that a person's ability to participate in life situations is an important rehabilitation outcome that needs to be measured (1). The World Health Organization's revised model of disability, the International Classification of Functioning, Disability and Health (ICF) includes participation as 1 of the 3 major components that comprise functioning and health (2). Participation is defined as "the involvement in life situations", and participation restrictions reflect the problems that an individual may experience in those life situations (2).

A recent review of the literature identified 11 instruments, which were developed using the ICF (3). Although there has been tremendous progress in developing new instruments to measure the concept of participation, it is currently not known how the instruments compare due to differences in the types of health conditions included in the studies. It has been recommended that studies directly comparing the measurement properties of existing instruments are conducted (3).

This study focused specifically on the construct validity of the participation instruments. Validity assesses whether the instrument measures what it intends to measure (4). Validity is not a property of an instrument, but rather it is the meaning or interpretation that can be derived from the instrument scores for a particular purpose (5). Construct validity examines the theoretical relationship of the questions to each other and to hypothesized scales (6). Specifically, construct validity assesses whether the domain measures one underlying construct, which is referred to as dimensionality (6) or evidence based on internal structure. Assessing construct validity also includes examining relationships between hypothesized similar or dissimilar domains in other instruments, referred to as convergent or discriminant validity (6). Relationships can also be examined between groups of individuals based on sociodemographic variables, such as age, or clinical variables, such as diagnosis, which is referred to as known-group validity (6). Other types of validity include face and content validity. Face validity examines whether the instrument appears to measure what it intends to measure, and content validity assesses how well the questions cover the health components being measured (7). Although these are important measurement properties, they will not be addressed in this paper.

Comparing the construct validity of instruments, all purporting to measure participation, in a single sample of persons with spinal conditions will help to determine whether differences in how this concept is operationalized (e.g. asking about dif-

ficulties, limitations or frequency) are captured in the domain scores. In addition, comparisons with the scores on instruments measuring concepts, such as health status or quality of life, will further enhance our understanding of how participation compares with other concepts and will assist clinicians and researchers in selecting an instrument for a given purpose. The purpose of this study was to evaluate the construct validity (unidimensionality, convergent/discriminant validity and known-group validity) of 5 participation instruments: Impact on Participation and Autonomy (IPA) (8), Keele Assessment of Participation (KAP) (9), Participation Measure-Post Acute Care (PM-PAC) (10), Participation Objective and Participation Subjective (POPS) (11), and World Health Organization Disability Assessment Schedule II (WHODAS II) (12).

METHODS

Recruitment and study procedures

A retrospective review of the spine database for the Acute Spine Program at Vancouver General Hospital was performed to identify individuals who were admitted between 2000 and 2005 with a diagnosis of: (*i*) traumatic or non-traumatic spinal cord injury (SCI); (*ii*) a spinal column fracture without neurological involvement; or (*iii*) a spinal degenerative disease (e.g. disc degeneration, spondylosis). Individuals were excluded during the initial database review and the recruitment phase if they were deceased; could not be contacted; did not speak English; had a cognitive deficit; were not physically able to complete the instruments (e.g. ventilator-dependent); or were discharged from hospital in the past 3 months and unable to perform regular activities (e.g. bed rest prescribed due to a pressure sore). A sample size of 200 individuals who completed the questionnaire in each diagnostic group was targeted, and eligible individuals were randomly selected until the final sample size was achieved or until all individuals were contacted. The study was approved by the Research Ethics Board at the University of British Columbia and all individuals provided written informed consent.

All potentially eligible individuals were mailed a questionnaire. Data was obtained from hospital databases and from a questionnaire completed by the respondents. Clinical data included variables such as diagnoses, neurological impairment (assessed using the International Standards for the Neurological Classification of SCI) (13) and comorbidities at the time of follow-up (one section of the Self-Administered Comorbidity Questionnaire, which asks about the presence or absence of 14 comorbid conditions) (14). Sociodemographic information collected included variables such as age, gender, marital status and living in an urban or rural setting (using methodology from Statistics Canada (15)) at the time of follow-up. Socioeconomic information consisted of variables such as education, employment and compensation status at the time of follow-up. Subjects completed 5 participation instruments, 1 health status instrument specific to their spinal condition and a quality of life instrument. Health status instruments were included to compare the information obtained from instruments designed to measure relevant aspects of the spinal condition (condition specific instrument) with the information obtained from the participation instruments. The health status instruments used were the Neck Disability Index (NDI) (16) for subjects with a cervical spinal column or degenerative condition, the Oswestry Disability Questionnaire (ODQ) (17) for subjects with a thoracic or lumbar spinal column injury or degenerative condition and the Self-Reported Functional Measure (SRFM) (18, 19) for subjects with a SCI. Subjective quality of life was measured using the Life Satisfaction-11 (LiSat-11) (20). Descriptions of all the instruments are included in Table I.

Statistical analysis

The following aspects of construct validity were assessed: dimensionality, convergent and discriminant validity and known-group validity.

Dimensionality was evaluated using item-to-scale correlations and by conducting a confirmatory factor analysis (CFA). The item-to-scale correlations (scale is defined as a domain-level total score), were corrected for overlap by removing the question from the scale when calculating the total score. A correlation of ≥ 0.40 is recommended (6). The question should have a higher correlation with the domain (scale) it belongs to (correlations within a domain or intra-domain) compared with the other domains (correlations among domains or inter-domain). The median and range of the 2 types of correlations were calculated and compared. The CFA tested if the proposed factor structure fit in this study sample (strictly confirmatory approach) and so no modifications were made to the models except for allowing correlated errors within a factor and not across factors (21). Robust maximum likelihood estimation was used to account for the non-normal data distribution (22). All analyses were conducted using Lisrel 8.08 (Scientific Software International, Lincolnwood, IL, USA). Model fit was evaluated using 3 fit indices: the root mean square error of approximation (RMSEA), where a value less than 0.05 is considered to be a close fit and an upper value of 0.080 is considered reasonable fit (23); the comparative fit index (CFI), where a value near 1.0 indicates a close fit of the data to the model and values close to or greater than 0.95 are recommended (24); and the standardized root mean residual (SRMR), where values less than 0.08 are recommended (24). To maximize the sample size for the CFA the domains associated with work and education were excluded. Dimensionality was not assessed in the KAP, POPS or WHODAS II (no CFA). Results for the KAP are presented at a question-level or the number of participation restrictions are reported, and dimensionality has not been previously assessed (9). In the POPS scoring algorithm the questions included in the domains are not necessarily intended to be related, but instead comprise an index, often referred to as a clinimetric approach and assessing dimensionality would not be relevant (25). Finally, for the WHODAS II, very few details are provided on the factor structure and a different version of the WHODAS II was tested, so it would not be difficult to compare the results (12).

Convergent and discriminant validity were assessed by examining: (*i*) the associations among similar participation domains in the instruments (e.g. all domains measuring mobility); (*ii*) the associations between participation domains and scores from health status instruments; and (*iii*) associations between domains in participation instruments and questions in the LiSat-11. Correlations were assessed using Spearman *rho*. Values greater than or equal to ± 0.70 were considered strong, ± 0.50 to ± 0.69 were considered moderate, ± 0.31 to ± 0.49 were considered fair and less than or equal to ± 0.30 were considered weak (26).

A priori hypotheses regarding the expected directions and strengths of the associations were tested. For the participation instrument, since all the instruments used the ICF as a conceptual model, in this study we mapped domains within the instruments to the ICF chapter headings, also referred to as ICF domains (see Appendix I). For convergent validity, it was hypothesized that domains measuring similar constructs in the participation instruments would have a strong or moderate correlation, with the exception of the POPS objective domains scores, where a fair or weak correlation was expected. For the health status instruments, it was hypothesized that there would be a strong or moderate correlation between the participation domains that measure similar constructs as the health status instruments (e.g. SRFM and participation domains that assess mobility), except in the POPS. Similarly, strong or moderate correlations were expected between the participation domains and questions in the LiSat-11 containing similar content.

Relationships between the participation domains and other study variables were hypothesized to assess known-group validity. The study variables assessed were motor score (SCI group), traumatic vs non-traumatic injury (SCI group), level of spinal injury, presence of back pain, age and gender. No differences were expected for type or level of injury, or gender. Subjects in the SCI group with a lower motor score were expected to have worse participation in domains related to mobility (except transportation), self-care, domestic life, and community, social and civic life compared with subjects with a

Table I. *Description of the instruments*

Instrument	Questions	Domains	Score generated for study
Participation instruments*			
IPA (8)	39 total; 31 perceived participation, 8 perceived problem	Autonomy Indoors; Autonomy Outdoors; Family Role; Social Life & Relationships; Work/Education	Perceived participation score was calculated for each domain, with a lower score indicating better perceived participation; domains scores range from 0 to 4.
KAP (9)	11 questions plus 4 screening questions	Mobility; Self-Care; Domestic Life; Interpersonal Interactions and Relationships; Major Life Areas; Community Social & Civic life	Mean response for each question was reported, with a lower score indicating better perceived participation; question scores range from 1 to 5.
PM-PAC (10)	51 questions total and 42 are scored	Communication; Mobility; Domestic Life; Interpersonal Relationships; Role Functioning; Work/Employment; Education; Economic Life; Community, Social & Civic life	Participation scores for each domain was calculated, with a higher score indicating better participation; domain scores range from 1 to 5.
POPS (11)	78 questions covering 26 life areas	Domestic Life; Major Life Areas; Transportation; Interpersonal Interactions & Relationships; Community, Recreational & Civic life	Objective scores are based on z scores, which represent the difference between the frequency for each activity compared with reference data and each activity is weighted based on reference data regarding perceived importance, with a higher score indicating greater frequency compared with reference data; objective scores vary for each domain. Subjective scores are obtained by multiplying the individual's importance score by the satisfaction score and range from –4 (important area that a person wants to do more or less of) to +4 (important area that a person is satisfied with the amount of activity).
WHODAS II (12)	36	Understanding & Communicating; Getting Around; Self-Care; Getting Along with People; Life Activities (household/work activities); Participation in Society	Measures both the concepts of activity and participation; a scoring algorithm provided by the World Health Organization produced domain and total scores; a lower score indicates better reported activity or participation; separate scores were calculated for individuals who were working and not working for the Life Activities Domain and the total score; domain and total scores range from 0 to 100.
Health Status Instruments†			
NDI (16)	10	Pain intensity; Personal care; Lifting; Reading; Headaches; Concentration; Work; Driving; Sleeping; Recreation	An overall score ranging from 0 to 50 is produced by summing the questions, with a lower score indicating less pain/disability.
ODQ Version 2.0 (17)	10	Pain intensity; Personal care; Lifting; Walking; Sitting; Standing; Sleeping; Sex life; Social life; Travelling	An overall score ranging from 0 to 50 is produced by summing the questions, with a lower score indicating less pain/disability.
SRFM (18, 19)	13	Moving around inside; Stairs; Transfer bed/chair; Transfer toilet; Eating; Grooming; Bathing; Dressing upper body; Dressing lower body; Toileting; Managing bladder; Managing bowels	An overall score ranging from 0 to 52 is produced by summing the 13 questions, with a higher score indicating the person is more independent.
Quality of Life Instrument*			
LiSat-11(20)	11	Life satisfaction in general; Vocation; Financial situation; Leisure; Social/friends/family; Sexual life; Self care; Family life; Partner relationship; Physical health; Mental health	Mean response for each question was reported, with a higher score indicating greater satisfaction; question scores range from 1 to 6.

*All subjects completed the 5 participation instruments and the quality of life instrument.
†The NDI was completed by subjects with a degenerative or spinal column injury of the cervical spine; the ODQ was completed by subjects with a degenerative or spinal column injury of the thoracic or lumbar spine; the SRFM was completed by subjects with a traumatic or non-traumatic spinal cord injury.
IPA: Impact on Participation and Autonomy; KAP: Keele Assessment of Participation; NDI: Neck Disability Index; ODQ: Oswestry Disability Questionnaire; PM-PAC: Participation Measure-Post Acute Care; POPS: Participation Objective Participation Subjective; SRFM: Self-Reported Functional Measure; WHODAS II: World Health Organization Disability Assessment Schedule II.

higher motor score. Subjects in the spinal column and degenerative group with back pain were expected to have worse participation in domains related to interpersonal interactions and community, social and civic life compared with subjects without back pain. Subjects over 65 years of age were expected to have worse participation in domains related to mobility (except transportation), self-care and domestic life compared with subjects aged 65 years and under. Hypotheses were tested using either linear or ordinal regression with backward stepwise variable selection to adjust for relevant covariates and a *p*-value < 0.05 was considered statistically significant. A hypothesis was considered to be supported if the effect was statistically significant in both the unadjusted and adjusted analysis and in the correct direction (increase or decrease in score as expected).

An index was created for each instrument comparing the number of hypotheses supported out of the total number assessed. It has been recommended that 75% of hypotheses should be supported (27). Calculations for the item and domain correlations were performed using SPSS 16.0 (Chicago, IL, USA) and the known-group hypothesis testing was conducted using SAS 9.1.3 (Cary, NC, USA).

Details of how missing data was managed for this study has been described in another paper (28). The percentages of missing data for the participation instruments (IPA, KAP, PM-PAC, POPS, WHODAS II) were all less than 10%. For the health status instruments (NDI, ODQ, SRFM) and quality of life (LiSat-11) the amount of missing data at the level of the questions was also less than 10%, except for the SRFM, where it was 10–12% because 12 subjects received the wrong version of the questionnaire due to an administrative error.

RESULTS

A total of 545 individuals participated in the study. There were 145 in the SCI group, 187 in the spinal column group and 213 in the spinal degenerative group. The response rates for all eligible individuals ranged from 58% (187/320) in the spinal column group to 62% (213/345) in the spinal degenerative group. Individuals were contacted approximately 4 years after discharge from hospital.

A description of the sample has been described (28). Briefly, 67% of the sample was male (367/545). The mean age and standard deviation (SD) at the time of follow-up was 51.5 (16.6) years. A comparison of individuals who participated in this study, and those who were eligible but did not participate, revealed the sample was older (47.0 vs 40.0 years) on admission to hospital and there were fewer men (67% vs 73%).

Scores for the participation instruments have been described (28). For the health status instruments, the SRFM score (SD) in the SCI group was 1.72 (0.71). Data on the ODQ were available for 272 subjects in the spinal column and spinal degenerative group, and the mean and SD was 1.14 (0.89). The mean NDI score was 1.13 (0.84) (*n* = 128).

Measurement properties

Dimensionality was assessed in 3 of the 5 instruments. It was not assessed in the KAP or the POPS. The item-to-scale (item intra-domain) correlations were all greater than 0.40 in the IPA, PM-PAC and WHODAS II, which suggests the questions were strong indicators of the domains (Table II). Results also indicated there were questions that had stronger correlations with other domains as opposed to their own domain (item inter-domain). In both the IPA and the WHODAS II the questions

asking about sexual or intimate relationships, which are part of domains assessing interpersonal relationships, correlated with domains assessing community, social and civic life as well as work/education.

A first-order CFA model was assessed in the IPA and PM-PAC to replicate the factor structure previously reported (10, 29). Overall, the models demonstrated adequate fit. All of the models had a RMSEA value less than 0.08 including the upper limit of the 90% confidence interval (CI); however, only the PM-PAC had a value less 0.05 including the lower limit of the 90% CI. The CFI were 0.99 for both the IPA and PM-PAC, suggesting good fit. The SRMR values were 0.060 in the IPA and 0.064 in the PM-PAC. Three correlated error terms were added within a factor for the IPA and only one was added in the PM-PAC. The standardized factor loadings were all greater than 0.40, which is recommended (Table III) (30).

The correlations among similar participation domains are summarized in Table IV. Overall, correlations were higher among the WHODAS II, IPA, KAP and PM-PAC. As expected, the lowest correlations were observed between the objective domains of the POPS and the other instruments. Correlations were lower than expected between the subjective POPS domains and domains in the IPA, KAP, PM-PAC and WHODAS II. Correlations between the participation domains and the health status instruments generally supported our hypotheses. Higher correlations were observed between the health status instruments and the domains related to mobility, self-care, domestic life, work or education and community, social and civic life (Table V). Correlations were highest among the ODQ and the participation domains. The association between the participation domains and questions in the LiSat-11 measuring similar content (interpersonal relationships) was as expected (Table VI), except a higher correlation was observed among the PM-PAC domain economic life and the LiSat-11 question asking about satisfaction with finances (*rho* = –0.51).

The known-group validity indices (number of hypotheses supported/number hypotheses tested) for each the participation instruments were: IPA = 95% (20/21); KAP = 77% (24/31); PM-PAC = 94% (17/18); POPS Objective = 67% (10/15); POPS Subjective = 73% (11/15); and WHODAS II = 84% (21/25). A summary of the results is included in Table VII.

DISCUSSION

The purpose of this study was to compare the construct validity of 5 participation instruments. Results from this study indicate that, given the challenges in measuring a broad concept such as participation, these instruments demonstrate good construct validity in individuals with spinal conditions. The measurement of participation is in the developmental stages, and results from this study may help to explain whether these instruments are measuring similar or different things.

Unidimensionality was assessed by examining the item-to-scale correlations and conducting a CFA on the IPA and PM-PAC. All the instruments demonstrated good item-to-scale correlations. For the IPA, results in this study were generally better than those reported by Sibely et al. (29). The question

Table II. *Results of the item-to-scale tests for the entire sample (*n = 545)

Instruments (range)	No. of questions	Median item intra-domain correlation	Item intra-domain correlation range*	Item inter-domain correlation range†	# Item intra-domain correlations > item inter-domain correlations (%)‡
IPA (0–4)					
Autonomy Indoors	7	0.82	0.73–0.88	0.52–0.71	28/28 (100)
Family Role	7	0.85	0.66–0.87	0.55–0.80	27/28 (96)
Autonomy Outdoors	5	0.88	0.84–0.89	0.65–0.80	20/20 (100)
Social Life & Relationships	6	0.79	0.60–0.83	0.45–0.70	22/24 (92)
Work & Education (*n*=356)	6	0.87	0.81–0.92	0.61–0.80	24/24 (100)
PM-PAC (1–5)					
Communication	6	0.76	0.65–0.85	0.20–0.66	48/48 (100)
Mobility	5	0.80	0.72–0.89	0.37–0.72	39/40 (98)
Domestic Life	3	0.72	0.71–0.74	0.41–0.66	24/24 (100)
Interpersonal Relationships	3	0.74	0.64–0.80	0.22–0.57	24/24 (100)
Role Functioning	4	0.83	0.75–0.88	0.30–0.74	32/32 (100)
Work & Employment (*n*=299)	5	0.77	0.67–0.81	0.26–0.73	40/40 (100)
Education (*n*=63)	4	0.70	0.67–0.78	0.23–0.78	30/32 (94)
Economic Life	3	0.75	0.59–0.77	0.23–0.58	32/32 (100)
Community, Social & Civic Life	9	0.69	0.43–0.80	0.16–0.75	69/72 (96)
WHODAS II (0–100)					
Understanding & Communicating	6	0.74	0.69–0.82	0.26–0.59	30/30 (100)
Getting Around	5	0.73	0.62–0.81	0.30–0.62	29/30 (97)
Self-Care	4	0.77	0.60–0.83	0.37–0.65	24/24 (100)
Life Activities (Non-working; *n*=162)	4	0.85	0.83–0.90	0.36–0.57	20/20 (100)
Life Activities (Working; *n*=383)	8	0.80	0.78–0.85	0.45–0.68	40/40 (100)
Getting Along with People	5	0.70	0.45–0.76	0.24–0.57	27/30 (90)
Participation in Society	8	0.71	0.64–0.78	0.38–0.71	47/48 (98)

*Corrected item-total correlation; for example, in the IPA Autonomy Indoors domain it is the 7 questions with the total score for that domain (7 item intra-domain correlations).
†This includes the item with all the other total domain scores; for example in the IPA Autonomy Indoors domain it includes the 7 questions within the domain with the other 4 IPA domains (28 item inter-domain correlations).
‡This includes the number of corrected item-total correlations that are greater than the question correlations with other domains; for example, in the IPA Autonomy Indoors domain it includes the 7 questions with the other 4 IPA domains (a total of 28 correlations).
See Table I for participation instrument abbreviations.

Table III. *Results of the confirmatory factor analysis for the entire sample*

Instruments	Standardized loadings on first-order factor	RMSEA (90% CI)*	CFI*	SRMR*
IPA (*n*=545)				
Autonomy Indoors	0.73–0.91			
Family Role	0.70–0.91			
Autonomy Outdoors	0.88–0.91			
Social Life & Relationships	0.63–0.89			
Work & Education	NA			
Model Fit		0.071 (0.066, 0.075)	0.99	0.060
PM-PAC (*n*=512)				
Communication	0.56–0.87			
Mobility	0.76–0.94			
Domestic Life	0.78–0.82			
Interpersonal Relationships	0.71–0.92			
Role Functioning	0.76–0.92			
Work & Employment	NA			
Education	NA			
Economic Life	0.67–0.88			
Community, Social & Civic Life	0.74–0.80			
Model Fit		0.054 (0.049, 0.059)	0.99	0.064

*The RMSEA, CFI and SRMR are estimates of the overall model fit.
CFI: comparative fit index; NA: not applicable; RMSEA: root mean square error of approximation; CI: confidence interval; SRMR: standardized root mean square residual; see Table I for participation instrument abbreviations.

Table IV. *Proportion of strong or moderate correlations among similar participation instrument domain scores using the entire sample* (n = 545)

ICF domains	IPA	KAP	PM-PAC	POPS-OBJ	POPS-SUBJ	WHODAS II
	#Strong or moderate correlations/#correlations assessed*					
Communication	NA	NA	0/1	NA	NA	0/1
Mobility	4/6‡	6/10	4/6	0/5†	0/5†	4/6‡
Self-Care	2/2	2/2	NA	NA	NA	2/2
Domestic Life	6/8	10/18	6/8	0/7†	1/7	9/14
Interpersonal Interactions & Relationships	3/5	3/5	4/5	0/4†	1/4	3/5
Major Life Areas – Work/Education	6/8	8/14	11/18	0/7†	0/7	5/8
Major Life Areas – Economic Life	NA	0/1	0/1	NA	NA	NA
Community, Social & Civic Life	3/5	3/5	3/5	0/4†	0/4	3/5

*Strong correlation ≥ ± 0.70; Moderate correlation = ± 0.50 to ± 0.69; Fair correlation = ± 0.31 to ± 0.49; Weak correlation ≤ ± 0.30 and Spearman's *rho* correlation was used. The numbers of correlations vary among the instruments depending on the domains or questions (subdomains) relevant to the ICF chapters in the activities and participation component. See the appendix for a listing of the participation domains mapped to the ICF chapters. Not all instruments cover each content area in the ICF (e.g. self-care, economic life) and are therefore not applicable. Correlations among domains within instruments (e.g. PM-PAC's education and work/employment domains) were not counted.

†Strong or moderate correlations were not expected.

‡Example: the WHODAS II Getting Around domain was compared with a total of 6 domains (IPA Autonomy Indoors, KAP Mobility #1, KAP Mobility #2, POPS Objective Transportation, POPS Subjective Transportation, PM-PAC Mobility) and the correlation was strong or moderate for 4 of the 6 domains.

ICF: International Classification of Functioning, Disability and Health; NA: not applicable; see Table I for participation instrument abbreviations.

regarding "spending my own money" had an item-to-scale correlation of only 0.34 in their study, whereas in our study a value of 0.65 was obtained, which may be due to differences in the distribution of the data. In the IPA and WHODAS II the question asking about sexual/intimate relationships had cross-correlations with domains related to community, social and civic life as well as work. Based on other studies (9, 29) it is not surprising that areas of participation overlap. A recent study by Anderson et al. (31) reported that sexual function is a priority for individuals living with SCI, which further supports the need to include these types of questions. Since there is only 1 question included in each instrument it is not possible to develop a separate domain. The measurement properties of questions asking about sexual relationships should be assessed in individuals with different types of health conditions before suggesting any changes.

Results from the CFA provided additional information pertaining to the factor structure. Confirmatory factor analysis is recommended over exploratory factor analysis when the factor structure has been established since the hypothesized factor structure can be tested empirically (30). In this study the standardized factor loadings for the PM-PAC were similar to the results reported by Gandek et al. (10). The lowest factor loading (0.53) in their study was for a question in the community, social and civic life domain, whereas in our study it was for the question "watching or listening to television and radio" (0.56). Sibley et al. (29) conducted a CFA on the IPA and reported 7 factor loadings less than 0.60, whereas in our study all the factor loadings were greater than 0.63. The differences observed in the factor loadings observed in this study and previous results for the PM-PAC (10) and IPA (29) are likely due to variations in the samples (e.g. age, diagnosis).

This study highlights some of the broader issues surrounding dimensionality for the concept of participation. The 5 participation instruments evaluated in this study were developed using different approaches. The POPS was developed using a clinimetric approach, which does not require the questions to be highly correlated and multiple aspects of participation can be combined to form an overall score or index. Other instruments, such as the IPA and PM-PAC, were developed using psychometric methods such as factor analyses, which rely on associations among the items to create factors that form the domains. In the IPA and PM-PAC the relationships among the factors (domains) are not specified and dimensionality is assessed only within the domains, suggesting that the domains do not necessarily form one single dimension. Until issues related to dimensionality in the concept of participation have been addressed, it will be difficult to determine the best approach for developing new measures and the role of modern measurement methods such as item response theory (32). Furthermore, until there is clarity regarding the conceptualization of participation it will be difficult to resolve these dimensionality issues (32).

Convergent validity was assessed by examining the relationship between similar domains. Overall the correlations were strong (*rho* > ± 0.70) to moderate (± 0.50 to ± 0.69) between similar domains within the IPA, KAP, PM-PAC and WHODAS II. Since the IPA and KAP are both designed to assess autonomy in participation it was surprising that the correlations between these two instruments were not higher in comparison with the others, such as the WHODAS II, which asks about difficulty. The KAP only has one question on self-care compared with 7 in the IPA and 4 in the WHODAS II, and so the use of broad or general questions may explain the lower correlation.

Table V. *Correlations* among participation domains and health status instruments for the entire sample (*n *= 545)†*

		SRFM (*n*=145)	ODQ (*n*=272)	NDI (*n*=128)
IPA	Autonomy Indoors	0.59	0.64	0.52
	Family Role	0.50	0.73	0.71
	Autonomy Outdoors	0.49	0.75	0.67
	Social Life & Relationships	0.38	0.66	0.41
	Work & Education	0.57 (*n*=78)	0.69 (*n*=194)	0.66 (*n*=84)
KAP	Mobility #1	0.47	0.60	0.41
	Mobility #2	0.37	0.63	0.46
	Self-Care	0.47	0.50	0.32
	Domestic Life #4	0.29	0.60	0.49
	Domestic Life #5	0.31	0.57	0.43
	Domestic Life #6	0.23‡ (*n*=59)	0.62 (*n*=162)	0.43 (*n*=65)
	Interpersonal Interactions & Relationships	0.28	0.53	0.50
	Economic Life	0.28	0.46	0.31
	Work	0.45 (*n*=76)	0.51 (*n*=175)	0.46 (*n*=76)
	Education	0.38 (*n*=45)	0.45 (*n*=102)	0.47 (*n*=46)
	Community, Social & Civic Life	0.33	0.54	0.51
POPS Objective	Domestic Life	−0.36	−0.21	−0.01‡
	Major Life Areas	−0.29	−0.33	−0.33
	Transportation	−0.19	−0.05‡	−0.17
	Interpersonal, Interactions & Relationships	−0.16‡	−0.27	−0.16‡
	Community, Recreational & Civic Life	−0.17	−0.18	−0.08‡
POPS Subjective	Domestic Life	−0.30	−0.42	−0.31
	Major Life Areas	−0.20	−0.26	−0.23
	Transportation	−0.14‡	−0.19	0.01‡
	Interpersonal Interactions & Relationships	−0.19	−0.31	−0.14‡
	Community, Recreational & Civic Life	−0.05‡	−0.27	−0.12‡
PM-PAC	Communication	−0.22	−0.49	−0.55
	Mobility	−0.51	−0.68	−0.52
	Domestic Life	−0.38	−0.67	−0.61
	Interpersonal Relationships	−0.30	−0.54	−0.44
	Role Functioning	−0.21	−0.73	−0.63
	Work & Employment	−0.37 (*n*=53)	−0.60 (*n*=168)	−0.65 (*n*=78)
	Education	−0.39 (*n*=24)	−0.51 (*n*=26)	−0.52 (*n*=13)
	Economic Life	−0.22	−0.48	−0.31
	Community, Social & Civic Life	−0.44	−0.78	−0.68
WHODAS II	Understanding & Communicating	0.10‡	0.46	0.45
	Getting Around	0.56	0.79	0.57
	Self-Care	0.67	0.51	0.40
	Life Activities (Non-working)	0.21‡ (*n*=58)	0.62 (*n*=69)	0.57 (*n*=35)
	Life Activities (Working)	0.45 (*n*=87)	0.66 (*n*=203)	0.58 (*n*=93)
	Getting Along with People	0.22	0.54	0.43
	Participation in Society	0.47	0.73	0.58

*Strong correlation ≥±0.70; Moderate correlation =±0.50 to ±0.69; Fair correlation =±0.31 to ±0.49; Weak correlation ≤±0.30 and Spearman's *rho* correlation was used.

†Convergent validity: high correlations (strong/moderate) were expected between the participation domain scores related to (1) mobility, (2) self-care, (3) domestic life, (4) major life areas (work/education) and (6) community, social and civic life and the health status instrument overall scores. Discriminant validity: low correlations (fair/weak) were expected between the participation domain scores related to: (1) communication, (2) interpersonal interactions and relationships, and (3) major life areas (economic life) and the health status instrument overall scores. Lower correlations were also expected between the POPS objective and subjective domain scores and the health status instrument overall scores.

‡Non-significant correlation.

See Table I for participation instrument abbreviations.

Results from this study also highlight the importance of considering the content of the questions contained within domains. For example, the PM-PAC and the WHODAS II both have domains assessing aspects of communication, and the correlation between these two domains was lower than expected, *rho* = (−0.46). However, in the WHODAS II, the questions are related to comprehension and having conversations, whereas the PM-PAC includes questions asking about keeping in touch with others as well as reading books. Given

the different examples provided in these two instruments, it is not unexpected that the correlation was only moderate.

Similarly, the way in which participation was operationalized greatly impacted the relationships between similar domains. In the POPS it was expected that objective assessment of participation would not correlate highly with subjective estimates based on previous studies (11). Higher correlations were expected among the subjective domain of the POPS and the other instruments than were observed, which may be due

Table VI. *Correlations* among the participation instrument domains and the quality of life instrument using the entire sample (*n = 545*)*

	IPA: Social Life & Relationships	KAP: Interpersonal Interactions & Relationships	PM-PAC: Interpersonal Relationships	POPS-OBJ: Interpersonal Interactions & Relationships	POPS-SUBJ: Interpersonal Interactions & Relationships	WHODAS II: Getting Along with People
LiSat-11						
Contact with Friends†	0.60	0.54	−0.68	−0.28	−0.54	0.51
Financial Situation‡	0.45	0.39	−0.51	−0.11	−0.35	0.37

*Strong correlation (≥±0.70); Moderate correlation (±0.50 to ±0.69); Fair correlation = ±0.31 to ±0.49; Weak correlation ≤±0.30 and Spearman's *rho* correlation was used.

†Convergent validity: a strong/moderate correlation was expected between the participation domains related to interpersonal interactions and relationships and the LiSat-11 question asking about satisfaction with the amount of contact with friends and acquaintances (except for the POPS objective Interpersonal Interactions and Relationships domain).

‡Discriminant validity: a fair/weak correlation was expected between the participation domains related to interpersonal interactions and relationships and the LiSat-11 question asking about satisfaction with the person's financial situation.

LiSat-11: Life Satisfaction-11; see Table I for participation instrument abbreviations.

to the weighting of satisfaction with importance. Overall, the correlations between the subjective domains in the POPS had a fair correlation (±0.31 to ±0.49) with similar domains in the other instruments. In measuring participation, it is important to consider not only if the person is able to do it, but also his or her interests and values (33). As a result it has been suggested that optimal participation may vary for different individuals (33). The low correlations observed between the subjective domains of the POPS with similar domains in the other instruments support this idea.

There were a few associations that were above *rho* = 0.50; for example, the domain assessing interpersonal, interactions and relationship in the POPS and PM-PAC had a correlation of *rho* = (−0.52). These results reinforce that in evaluating the construct validity of an instrument it is important to consider both the content and how the questions are asked, since these can affect the observed relationship. Results from this study also suggest that it may be important to distinguish between difficulty/limitations, autonomy and satisfaction when measuring participation.

The relationships between the participation domains and instruments measuring health status were also examined. As expected, higher correlations were observed between domains assessing mobility, self-care, domestic life, and major life areas (work and/or school), and lower correlations were observed with

domains assessing communication, interpersonal interactions and relationships, and economic life. The POPS had the lowest correlations with the health status instruments, as expected. Overall, the correlations were higher for the ODQ and the participation instruments compared with the SRFM and NDI.

To our knowledge, the health status instruments used in this study have not previously been compared with participation instruments. The ODQ and NDI measure pain and assess the effect of pain (a body function in the ICF) on aspects of participation. The SRFM assesses the need for assistance, which is considered an environmental factor in the ICF, for aspects of mobility and self-care. So it is possible that the health status instruments assess more the influence of other ICF components (e.g. ICF component body functions) on participation. The participation instruments seem to be more "pure" measures of participation and have a broader coverage of domains. More work is needed to further clarify the concepts of health status and participation and inform users which instrument(s) is best for which purpose.

In terms of the correlations between the 5 instruments measuring participation and the LiSat-11, which measures quality of life, as expected, higher correlations were observed between similar content areas (interpersonal interactions and relationships) and lower correlations between different content areas. None of the correlations were strong (≥ 0.70), even with the POPS subjective domains, which combines questions on

Table VII. *Summary* of the study results for validity*

Criteria	IPA	KAP	PM-PAC	POPS-OBJ	POPS-SUBJ	WHODAS II
Dimensionality						
1) Item	+++	NA	+++	NA	NA	+++
2) CFA	++	NA	+++	NA	NA	NA
Convergent/Discriminant						
1) Participation Instruments	++	++	++	+++	+	++
2) Health Status Instruments	+++	++	+++	+++	+++	+++
3) Quality of Life	+++	+++	++	+++	+++	+++
Known-Group	+++	++	+++	+	++	++

*Ratings: +++ met criteria/results as expected; ++ partially met criteria/results partially as expected; + results primarily did not meet criteria/results primarily not as expected.

CFA: confirmatory factor analysis; NA: not applicable; see Table I for participation instrument abbreviations.

importance and satisfaction. In the POPS, since the rating of importance (range 0–4, with a higher number indicating an important area to a person's satisfaction with life) is weighted by how satisfied a person is with the amount of activity (multiplied by –1 if dissatisfied and +1 if satisfied), the importance factor weights the response more than satisfaction, which may explain why higher correlations were not observed.

The assessment of known-group validity was the final aspect of construct validity assessed. The IPA had the greatest number of hypotheses supported (95%) and the POPS objective domains had the lowest (67%), which is below the expected minimum value of 75%. Other studies have also reported fewer hypotheses supported than expected using the POPS (11, 34). As mentioned previously, the POPS operationalizes participation differently compared with the other instruments, and this must be considered when interpreting these results.

When reviewing these results it is also important to acknowledge the limitations of this study. Only the construct validity was assessed, and future research should assess the ability of these instruments to assess clinically important changes following an intervention. In addition, more analyses should be conducted within each of the 3 types of spinal conditions (e.g. testing whether there is factorial invariance for each of the 3 spinal conditions).

In conclusion, this study examined the construct validity of 5 participation instruments. Based on the criteria used to evaluate construct validity in this study, differences were observed between the PM-PAC, IPA, WHODAS II and the KAP, POPS. The KAP was developed to assess participation at a population-level and, consequently, the level of detail was sacrificed for brevity. For the POPS, results from this study suggest it assesses different aspects of participation compared with the other 4 instruments. However, since quality of life instruments also assess satisfaction (e.g. LiSat-11, Quality of Life Index (35)) and importance (e.g. Quality of Life Index (35)) in various life domains, future research should determine the relationship between participation and quality of life, as well as how these concepts differ. Clinicians and researchers should consider the type of information required about the concept of participation before selecting an instrument. Results for the construct validity of the 5 participation instruments are promising, but more evidence is required in studies testing other health conditions and assessing measurement properties such as minimal important change.

REFERENCES

1. Cicerone KD. Participation as an outcome of traumatic brain injury rehabilitation. J Head Trauma Rehabil 2004; 19: 494–501.

2. World Health Organization. International Classification of Functioning, Disability and Health. Geneva: World Health Organization; 2001.

3. Noonan VK, Kopec JA, Noreau L, Singer J, Dvorak MF. A review of participation instruments based on the International Classification of Functioning, Disability and Health. Disabil Rehabil 2009; 31: 1883–1901.

4. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002; 11: 193–205.

5. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performance as scientific inquiry into score meaning. Am Psychol 1995; 50: 741–749.

6. Fayers PM, Machin D. Quality of life: assessment, analysis and interpretation. Chichester: John Wiley and Sons; 2000.

7. 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.

8. 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: 210–216.

9. 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: 1889–1899.

10. Gandek B, Sinclair SJ, Jette AM, Ware JE Jr. Development and initial psychometric evaluation of the participation measure for Post-Acute Care (PM-PAC). Am J Phys Med Rehabil 2006; 86: 57–71.

11. Brown M, Dijkers MP, Gordon WA, Ashman T, Charatz H, Cheng Z. Participation objective, Participation subjective: a measure of participation combining outsider and insider perspectives. J Head Trauma Rehabil 2004; 19: 459–481.

12. World Health Organization. Geneva: WHODAS II; [updated 2001, November 27; cited 2006 March 10]. Available from: http://www.who.int/icidh/whodas.

13. Marino RJ, Barros T, Biering-Sorensen F, Burns SP, Donovan WH, Graves DE, et al. International standards for neurological classification of spinal cord injury. J Spinal Cord Med 2003; 26 Suppl 1: S50–S56.

14. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003; 49: 156–163.

15. Statistics Canada. 2006 Census Dictionary. Statistics Canada Catalogue no 92-566-XWE 2007 [cited 2008 Feb 16]. Available from: http://www12.statcan.ca/english/census06/reference/dictionary/index.cfm.

16. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther 1991; 14: 409–415.

17. Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine 2000; 25: 2940–2952.

18. Hoenig H, McIntyre L, Sloane R, Branch LG, Truncali A, Horner RD. The reliability of a self-reported measure of disease, impairment, and function in persons with spinal cord dysfunction. Arch Phys Med Rehabil 1998; 79: 378–387.

19. Hoenig H, Branch LG, McIntyre L, Hoff J, Horner RD. The validity in persons with spinal cord injury of a self-reported functional measure derived from the functional independence measure. Spine 1999; 24: 539–543.

20. Fugl-Meyer AR, Melin R, Fugl-Meyer KS. Life satisfaction in 18- to 64-year-old Swedes: in relation to gender, age, partner and immigrant status. J Rehabil Med 2002; 34: 239–246.

21. Joreskog KG. Testing structural equation models. In: Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park: Sage Publications Inc.; 1993, p. 294–316.

22. Muthen B, Kaplan D. A comparison of some methodologies for the factor analysis of non-normal Likert variables. Br J Math Stat Psychol 1984; 38: 171–189.

23. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park: Sage Publications Inc.; 1993, p. 136–162.

24. Hu L, Bentler PM. Cuttoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999; 6: 1–55.

25. Feinstein AR. Clinimetric perspectives. J Chronic Dis 1987; 40: 635–640.

26. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31: 247–263.

27. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60: 34–42.

28. Noonan VK, Kopec JA, Noreau L, Singer J, Mâsse LC, Dvorak MF. Comparing the reliability of five participation instruments in persons with spinal conditions: J Rehabil Med 2010; 42: 730–738.

29. Sibley A, Kersten P, Ward CD, White B, Mehta R, George S. Measuring autonomy in disabled people: validation of a new scale in a UK population. Clin Rehabil 2006; 20: 793–803.

30. Brown TA. Confirmatory factor analysis for applied research. New York: Guildford Press; 2006.

31. Anderson KD, Borisoff JF, Johnson RD, Stiens SA, Elliott SL. The impact of spinal cord injury on sexual function: concerns of the general population. Spinal Cord 2007; 45: 328–337.

32. Whiteneck G, Dijkers MP. Difficult to measure constructs: conceptual and methodological issues concerning participation and environmental factors. Arch Phys Med Rehabil 2009; 90: S22–S35.

33. Rochette A, Korner-Bitensky N, Levasseur M. 'Optimal' participation: a reflective look. Disabil Rehabil 2006; 28: 1231–1235.

34. Cantor JB, Ashman T, Gordon W, Ginsberg A, Engmann C, Egan M, et al. Fatigue after traumatic brain injury and its impact on participation and quality of life. J Head Trauma Rehabil 2008; 23: 41–51.

35. Ferrans CE, Powers MJ. Quality of Life Index: development and psychometric properties. ANS Adv Nurs Sci 1985; 8: 15–24.

APPENDIX I. Participation instrument domains mapped to ICF chapter headings

Communication	Mobility	Self-Care	Domestic Life	Interpersonal Interactions & Relationships	Major Life Areas	Community, Social & Civic Life
PM-PAC Communication	*IPA* Autonomy Indoors	*IPA* Autonomy Indoors	*IPA* Family Role	*IPA* Social Life & Relationships	*IPA* Work & Education	*IPA* Autonomy Outdoors
WHODAS II Understanding & Communicating	*KAP* Mobility #1	*KAP* Self-Care	*KAP* Domestic Life #4	*KAP* Interpersonal Interactions & Relationships	*KAP* Work	*KAP* Community, Social & Civic Life
	KAP Mobility #2	*WHODAS II* Self-Care	*KAP* Domestic Life #5	*POPS* Objective Interpersonal Interactions & Relationships	*KAP* Education	*POPS* Objective Community, Recreational & Civic Life
	POPS Objective Transportation		*KAP* Domestic Life #6	*POPS* Subjective Interpersonal Interactions & Relationships	*POPS* Objective Major Life Areas	*POPS* Subjective Community, Recreational & Civic Life
	POPS Subjective Transportation		*POPS* Objective Domestic Life	*PM-PAC* Interpersonal Relationships	*POPS* Subjective Major Life Areas	*WHODAS II* Participation in Society
	PM-PAC Mobility		*POPS* Subjective Domestic Life	*WHODAS II* Getting Along with People	*PM-PAC* Work & Employment	
	WHODAS II Getting Around		*PM-PAC* Domestic Life		*PM-PAC* Education	
			WHODAS II Life Activities (Non-work)		*PM-PAC* Role Functioning	
			WHODAS II Life Activities (Working)		*WHODAS II* Life Activities (Working)	

See Table I for participation instrument abbreviations.