J Rehabil Med 2007; 39: 156–162

ORIGINAL REPORT

IMPACT ON PARTICIPATION AND AUTONOMY QUESTIONNAIRE: INTERNAL SCALE VALIDITY OF THE SWEDISH VERSION FOR USE IN PEOPLE WITH SPINAL CORD INJURY

Maria Larsson Lund, PhD[1], Anne G. Fisher, ScD[1], Jan Lexell, MD, PhD[2,3] and Birgitta Bernspång, PhD[1]

From the [1]Department of Community Medicine and Rehabilitation, Occupational Therapy, Umeå University, Umeå, [2]Department of Health Sciences, Luleå University of Technology, Luleå and [3]Department of Rehabilitation, Lund University Hospital, Lund, Sweden

Objective: **To evaluate aspects of internal scale validity and reliability of the Swedish version of the Impact on Participation and Autonomy questionnaire (IPA-S) for use in people with spinal cord injury.**
Subjects: **A total of 161 persons with spinal cord injury.**
Method: **IPA-S psychometric properties were evaluated using Rasch rating scale analysis.**
Results: **The results show that the IPA-S has 27 items for perceived participation and 6 items for perceived problems with participation, indicating 2 underlying unidimensional constructs for use in people with spinal cord injury, after removal of misfitting items. The established hierarchical order of the item calibration values in both scales appeared logical and the distances between the items, with a few exceptions in the problem scale, were appropriate. Both scales demonstrated good separation reliability. The range of item calibration values in both scales did not fully cover the range of measures of persons' perceived participation and problems thereof.**
Conclusion: **The psychometric properties of IPA-S are promising and indicate that IPA-S has potential to be developed further.**

Key words: participation, assessment, outcomes, needs, measurements, Rasch, rehabilitation.

J Rehabil Med 2007; 39: 156–162

Correspondence address: Maria Larsson Lund, Department of Community Medicine and Rehabilitation, Occupational Therapy, Umeå University, SE-901 87 Umeå, Sweden. E-mail: maria.larsson.lund@occupther.umu.se

Submitted February 1, 2006; accepted October 11, 2006

"Participation", that is, persons' own lived experiences of their involvement in a life situation (1) is a new concept in the context of the World Health Organization's (WHO) classification of health outcomes (c.f. 1, 2) and has replaced the concept "handicap" (2). A review of current instruments in the area shows that few measure participation in accordance with the definition in the International Classification of Functioning, Disability and Health (ICF) (3). In addition, in the ICF, participation is coded through an individual's performance, i.e. doing in their current environment, and assessed against population standards (1). This indicates a need to develop instruments that more fully take into consideration persons' own perspectives of involvement in life situations. Otherwise, there is a risk that objective perspectives may overshadow the lived experience of persons with disabilities, and thereby influence the prioritization and design of services, both at the individual level and at the level of society.

One instrument measuring participation is the Impact on Participation and Autonomy questionnaire (IPA). The IPA was developed in the Netherlands and has been translated into English, Swedish and German (4). It is a generic questionnaire intended for adults with chronic conditions associated with a wide range of diagnoses. The instrument was developed based on the standpoint that autonomy adds a personal perspective that is important in understanding participation (5). In the IPA, an individual's perceived participation is covered in 31 items, and problems with participation are covered in an additional 8 items (6). The original version of IPA has been shown to have good construct, convergent and divergent validity, although a few items were found to be psychometrically weak. The IPA also showed good test-retest reliability and homogeneity, and the ability to detect change over time (7–9).

Given the potential advantages and psychometric strengths of the IPA, we decided to use the IPA to describe participation in persons with spinal cord injury (SCI) in Sweden (10). A Swedish version of the IPA (IPA-S) was developed (10) through a forward-backward translation procedure (11, 12) of the original Dutch version. Even if the IPA-S translation is equivalent to the original version, one cannot assume that the properties of the previous version pertain to the next version. Thus, while previous studies support the psychometric properties of the original Dutch version of the IPA, research is still needed to establish the validity and reliability of IPA-S.

The overall aim of the present study was to evaluate aspects of internal scale validity and reliability of the IPA-S for use with people with SCI. Specifically, the aims were to evaluate internal scale validity of each of the 2 IPA-S scales in terms of: (*i*) whether the items defined a single unidimensional construct; (*ii*) whether the items were appropriately spread along the linear continuum of increasing difficulty; and (*iii*) whether the items were well targeted to the sample. Finally, reliability was evaluated in terms of whether the items separated persons into distinct levels of participation.

 ISSN 1650-1977
DOI: 10.2340/16501977-0031

METHODS

Procedure and participants

Potential participants were identified from a database at a SCI rehabilitation unit in the southern part of Sweden, whereupon the IPA-S was posted to 278 persons. A detailed description of the sampling and data collection procedure is presented elsewhere (10). In total, 161 persons with SCI, age range 17–84 years (mean age 52 years, standard deviation (SD) 18.2), participated in the study. As shown in Table I, most had paraplegia. Each participant's SCI was either complete or incomplete, which was reflected by their preferred form of mobility. Most needed social support in their activities of daily living, but 41% needed no social support.

Instrument

The IPA questionnaire contains 39 items in 8 areas: (*i*) mobility (5 items), (*ii*) self-care (6 items), (*iii*) household tasks and family role (7 items), (*iv*) spending money (2 items), (*v*) leisure (2 items), (*vi*) social relations (8 items), (*vii*) paid work and voluntary work (6 items), and (*viii*) education and learning (2 items) (6). Each of the 8 areas begins with one or several questions, each having the same format, about perceived participation in different life situations (for example: "My chances of getting around in my home are…"). For each question there were 5 response options: very good, good, fair, poor or very poor. At the end of each section of questions, there was a final question addressing the person's perceived extent of problems with participation within that area (for example: "With regard to your mobility, to what extent does your health or disability cause problems?"). The response options were: no problem, minor problems or severe problems. In accordance with the IPA manual (6), the items of perceived participation and perceived problems herein were coded in 2 separate sets of scores, with a range per item of zero to 4 and zero to 2, respectively. A higher score represented greater restriction in participation or greater perceptions of problems in participation. The IPA manual recommends that respondent's scores (self-ratings) on the items of perceived participation are summed for each of the domains: autonomy indoors, autonomy outdoors, family role, social relationships and work and education (6). Since we viewed all 5 domains as representing the same underlying construct, our plan was to analyse them together – one set for participation and one set for problems.

Rasch rating scale analysis

The data were analysed using a Rasch rating scale model and Winsteps software version 3.58 (13). Rasch analysis computer programs convert ordinal raw scores through logistic transformation into equal interval measures expressed in terms of log-odds probability units, i.e. logits. This analysis makes it possible to position persons and items (in equal units) on the same linear scale continuum according to their estimated item calibration values and person-perceived participation measures. Rasch analyses also generate detailed goodness-of-fit statistics and separation statistics for items and persons (14–16). Goodness-of-fit statistics are used to estimate whether the items and persons fit the assertions of the respective Rasch measurement model. The assertions of the simple Rasch model are: (*i*) when a group of persons is scored on a set of items, persons with more ability have greater probability of accomplishing difficult items than persons with less ability; and (*ii*) all persons are more likely to accomplish easier items than more difficult ones (14). Transformed to the present study, the assertions of the simple Rasch model of the IPA were: (*i*) persons perceiving more participation have greater probability of obtaining lower scores on IPA items than persons perceiving less participation; and (*ii*) all persons are more likely to get lower scores on items perceived as easier for participation than items that are perceived as more difficult for participation.

Table I. *Characteristics of the 161 participants with spinal cord injury (SCI)*

Characteristics	%
Gender (male/female)	63/37
Level of injury (paraplegia/tetraplegia)	62/38
Cause of SCI (traumatic/non-traumatic)	52/48
Marital status (single/cohabiting or married)	36/64
Residence (own home/service house)	97/3
Vocational status	
Working full or part-time	36
Student	2
Unemployed	1
Sick leave	12
Disability pension	22
Retired	27
Mobility	
Wheelchair	52
Walking-aid	18
No mobility aid	30

The item calibration values represent the item's position on the linear scale continuum. The item calibration values should progress from easy to difficult, with an appropriate spread of items. Large distances between the items result in a test that may not discriminate well between persons within or near those gaps. If a test is well targeted for a sample, the persons' perceived participation measures should fall between the boundaries of the most difficult and the easiest items. A lack of items at the upper and lower ranges of the persons may also decrease the ability of the test to separate persons into levels of participation. The test information function is the sum of the item information functions (17). Dips in the function indicate less sensitivity. Person separation statistics can be used to evaluate the number of statistically distinct strata of person measures that are distinguishable (14, 16). Person standard errors (SE) can be used to examine sensitivity to changes and high SE indicates lack of sensitivity (14).

The goodness-of-fit statistics, the mean square (MnSq) and standardized goodness-of-fit statistics (z), for the items were examined to determine whether the items met the measurement requirements for unidimensionality. The expected MnSq is 1.00 and the expected value of z is zero. For example, if an item that had a low likelihood of being endorsed, and a person with low perceived participation reported unexpectedly high perceived participation on that item, that item would have high decreased goodness-of fit. When an item misfits, this can indicate that the item is deficient related to the rest of the items on the scale and represents another construct (14, 16). A test is considered to have acceptable unidimensionality when at least 95% of the items fit the Rasch measurement model (14).

The psychometric properties of IPA-S were evaluated in 2 phases. One part included the 31 items covering perceived participation and the other part included the 8 items covering perceived problems with participation. The evaluation of the 2 parts followed the same steps of analysis. In the first step, the structure of the rating scale was evaluated. Since no evidence of average measure or step calibration disordering was found (18), we proceeded to evaluate the psychometric properties of the scale: (*i*) evidence of internal scale validity in terms of unidimensionality; (*ii*) evidence of targeting by comparison of person mean to item mean and by visual inspection of the person-item variable map of the item calibration values in relation to the persons' perceived participation measures on the same linear continuum; and (*iii*) evidence of separation reliability.

If we found misfit among the items our plan was to omit one item at a time starting with the item that had the highest MnSq and z until either: (*i*) no more items misfit; or (*ii*) separation values began to decrease. Then, if items were removed, our plan was to compare the measures estimated based on the inclusion of all items to the measures estimated based on the removal of the misfitting items in order to determine the extent to which the misfitting items disrupted the measurement system. We pre-set our criteria for misfit at MnSq $\geq$ 1.4 (19) with an associated $z \geq 2.0$ (20). In addition, a principal component analysis

of the standardized residuals was planned to be undertaken after all the misfitting items had been removed to confirm unidimensionality of the scales.

Finally, it is important to note that a rating scale model is generally preferred when the definition of the rating scale is the same for all items. An Andrich partial credit model should be used when the definition of the rating scale differs from item to item (14). Since there may have been a risk that the rating scale steps of one or both scales differed among items, we also evaluated whether the use of a partial credit rating scale model would result in enhanced goodness-of-fit. The results of our analysis of the data using an Andrich paritial credit model (where each item had its own rating scale structure) were similar in terms of item misfit and need for subsequent item removal. Since the overall results were slightly better with the original rating scale model than with the Andrich model, we retained the original rating scale model, and report only those results below.

RESULTS

The analysis of all of the 31 items of perceived participation scale showed that 4 of the items misfit (Table II). The hierarchy of the items showed that the 3 least challenging items, placed at the upper end of the hierarchy, related to different aspects of participation in interactions with nearest person. Three of the 5 most challenging items at the lower end of the hierarchy were related to different aspects of participation in work.

Fig. 1 shows the distribution of all items and all persons on a common linear continuum of perceived participation, ranging from persons perceiving less participation and items that are easier to report high perceptions of at the top, to persons perceiving more participation and items that are more difficult to report high perceptions of at the bottom. In the left-hand column of Fig. 1 persons are plotted at their perceived participation measure and in the 3 columns on the right, the items are plotted. In the middle column, the items are plotted at the rating scale midpoint. In the first and third columns are the first and last rating scale steps.

The distribution of persons ranged from 2.45 to –6.35 logits. When the full range of the item difficulties is considered in relation to the person distribution (Fig. 1), it reveals that persons' perceived participation is not fully covered by the range of the item calibration values, and that items that might better differentiate persons perceiving most participation are missing. The mean of the persons' perceived participation measures (–0.61 logits) was lower (indicating higher participation) than the mean of the item calibration values (zero

Table II. *Measurement report of items of perceived participation in the Swedish version of the Impact on Participation and Autonomy questionnaire (IPA-S). The items at the top are easier to report high perception of participation in than those at the bottom of the list*

Perceived participation			Infit*		Outfit*	
Items (domain)†	Measure	SE	MnSq	*z*	Mnsq	*z*
Communication with nearest (SR)	1.65	0.13	1.24	1.6	1.16	0.7
Respect from nearest (SR)	1.61	0.13	1.26	1.7	1.02	0.2
Relationship with nearest (SR)	1.48	0.12	1.28	1.8	1.23	1.1
Eating and drinking when one wants (AI)	1.33	0.12	0.89	–0.8	0.68	–1.7
Respect from acquaintances (SR)	0.64	0.10	0.93	–0.5	1.11	0.7
Getting around indoors when one wants (AI)	0.59	0.10	1.04	0.4	0.90	–0.6
Relationship with acquaintances(SR)	0.57	0.10	0.92	–0.6	0.99	0.0
Getting around indoors where one wants (AI)	0.50	0.10	0.86	–1.2	0.80	–1.4
Going to bed when one wants (AI)	0.44	0.10	1.13	1.1	0.95	–0.3
Minor housework jobs the way one wants (FR)	0.44	0.10	1.01	0.1	0.90	–0.6
Washing, dressing, grooming when one wants (AI)	0.38	0.09	0.74	–2.4	0.64	–2.8
Washing, dressing, grooming the way one wants (AI)	0.34	0.09	0.81	–1.7	0.70	–2.3
Going to the toilet when one needs (AI)	0.30	0.10	1.18	1.5	1.07	0.6
Spending income as wished (FR)	0.21	0.09	1.34	2.7	**1.45**	**2.8**
Contact with colleagues (WE)	0.20	0.13	**1.49**	**2.7**	1.22	1.1
Spending leisure time the way one wants (AU)	–0.02	0.09	0.93	–0.6	0.93	–0.5
Frequency of social contacts (AU)	–0.15	0.09	0.86	–1.3	0.86	–1.1
Visiting friends when one wants (AU)	–0.18	0.09	0.90	–0.9	0.86	–1.1
Contributing to looking after the home (FR)	–0.18	0.09	0.78	–2.1	0.77	–1.9
Getting housework done when one wants (FR)	–0.21	0.09	0.93	–0.6	0.92	–0.6
Fulfilling the one's role at home (FR)	–0.34	0.09	0.69	–3.1	0.65	–3.1
Achieving and maintaining a job one wants (WE)	–0.49	0.13	**1.42**	**2.4**	**1.77**	**3.6**
Living life the way one wants (AU)	–0.83	0.09	0.77	–2.2	0.82	–1.5
Heavy housework jobs the way one wants (FR)	–0.84	0.09	1.26	2.2	1.36	2.7
Getting training or education one wants (WE)	–0.85	0.13	1.07	0.5	1.21	1.2
Intimate relationships (SR)	–0.89	0.09	**1.62**	**4.7**	**1.85**	**5.5**
Doing work the way one wants (WE)	–0.92	0.12	1.09	0.6	1.07	0.5
Going on trips and holiday one wants (AU)	–0.93	0.09	0.90	–0.9	0.90	–0.8
Doing work one wants to (WE)	–0.97	0.13	0.97	–0.2	1.10	0.6
Repairs and upkeep the home (FR)	–1.21	0.10	1.09	0.8	1.20	1.4
Getting a different job (WE)	–1.66	0.14	1.11	0.7	1.06	0.4

*Bold type marks items that misfit.

†IPA domains = AI: autonomy indoors; AU: autonomy outdoors; FR: family role; SR: social relations; WE: work and education.

SE: standard deviation; MnSq: mean square; *z*: standardized goodness-of-fit statistics.

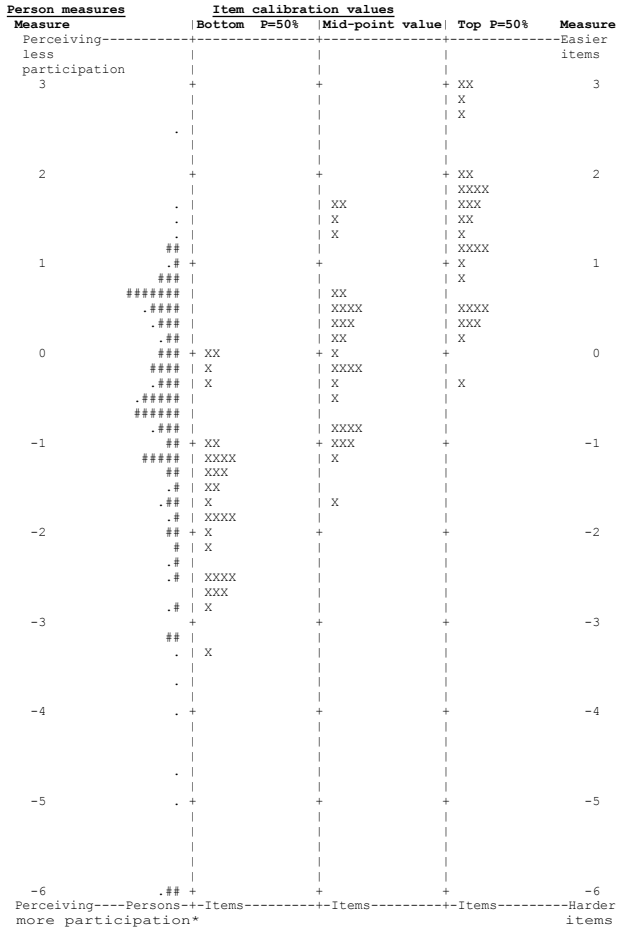

Fig. 1. Map of person-perceived participation measures and item calibration values. In the left-hand column persons are plotted at their perceived participation measure. Each "#" represents two persons and each "." represents one person. The next 3 columns each include all items (X) at different points along the range of the rating scale (see text for more details). *Persons at the bottom perceived no restrictions in their participation (zero scores).

logits). Further analysis undertaken to evaluate whether any systematic patterns were found among the 14 persons with the highest perceived participation measures (< –3.0 logits) revealed that 12 of them walked without using any aids and that 2 of them used a wheelchair. Twelve of the 14 persons needed no support in activities of daily living, 12 worked or studied and 2 had retired. The test information function (Fig. 2) indicated good sensitivity, and the person SEs that ranged from 0.19 to 0.33 for persons between 2.45 and –2.06 logits and increased to ≥ 1.00 logits for persons with measures above –4.68 logits. The person separation index was 4.14 (separation reliability = 0.94), indicating that the distribution of persons could be separated into 5 statistically distinct levels (21).

In accordance with our plan, misfitting items were removed one at a time until no more item misfits remained. The analysis resulted in the exclusion of the following 4 items: "Intimate relationship", "Spending income as wished", "Achieving or maintaining a job one wants" and "Contacts with colleagues".

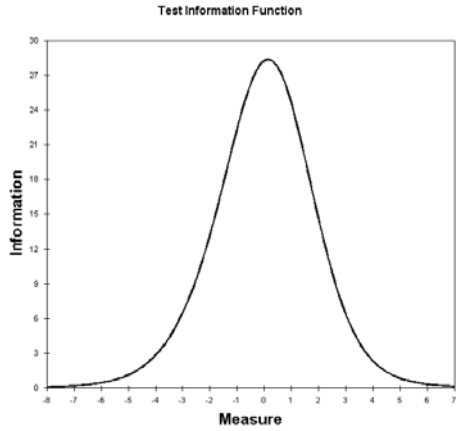

Fig. 2. Test information function of the perceived participation scale.

The separation index for persons remained essentially stable during the removal of the 4 items (varying between 4.14 and 4.06), continuing to indicate that the person distribution could be separated into at least 5 statistically distinct strata. The remaining 27 items demonstrated acceptable goodness-of-fit and thereby constituted a valid unidimensional scale. The distribution of the items and persons showed that the range of both persons and items increased after the exclusion of the 4 items. The items ranged from 1.68 to –1.78 logits and the persons from 2.58 to –6.34 logits. The mean of the persons' perceived participation measures (–0.66 logits) also increased after the item exclusion was completed, and the number of persons who had minimum estimated measure after the exclusion process was completed increased from 5 to 6. A principal component analysis of the standardized residuals revealed that only 2.3% of the unexplained variance was explained by the first factor, further supporting the unidimensionality of the perceived participation scale.

The second phase of the analyses comprised the evaluation of IPA-S perceived problems with participation scale. The item calibration values and the item goodness-of-fit statistics are shown in Table III. The item calibration values ranged from 1.08 to –1.13 logits (mean zero), with the item that was easiest to report a low perception of problems with, financial situation, at the top and the item that was most difficult to report a low perception of problems, work, at the bottom. The person measures ranged from 5.89 to –5.89 logits (mean = 0.26). Thirty-three person measures (20%) were below or above of the range of the items calibration values, indicating that items differentiating between persons perceiving both more and less problems with participation are lacking (Fig. 3). Five persons had maximum estimated measures and 14 persons had minimum estimated measures. The test information function (Fig. 4) supports lack of sensitivity of the scale, likely due to the limited range and number of gaps (especially at mid-range of the scale) between items (Fig. 3). SEs for persons generally ranged from 0.70 to 0.87 for persons between 3.16 and –3.16 logits, but for persons with measures ≥ 3.87 and ≤ –3.49, the person SEs increased to ≥ 1.00 logits, supporting lack of sensitivity. The separation of the persons was 2.13 (separation reliability = 0.82). This

Table III. *Measurement report of items of perceived problems in the Swedish version of The Impact on Participation and Autonomy questionnaire (IPA-S). The items at the top are easier to report low perception of problems with participation in than those at the bottom of the list*

Perceived problems			Infit*		Outfit*	
Items	Measure	SE	MnSq	*z*	Mnsq	*z*
Finances	1.08	0.18	**1.77**	**5.4**	**1.78**	**4.6**
Social relations	0.96	0.18	0.85	–1.3	0.88	–0.9
Family role	0.12	0.18	0.75	–2.3	0.72	–2.4
Self-care	–0.01	0.18	0.87	–1.1	0.90	0.8
Education	–0.01	0.22	1.27	1.8	1.24	1.5
Leisure	–0.17	0.17	0.80	–1.8	0.76	–2.1
Mobility	–0.85	0.18	0.77	–2.2	0.87	–0.9
Work	–1.13	0.23	0.93	–0.5	0.97	–0.1

*Bold type marks items that misfit. SE: standard error; MnSq: mean square; *z*: standardized goodness-of-fit statistics.

indicated that the person distribution could be separated into at least 3 statistically distinct strata (21).

The initial analysis of the perceived problems with participation scale revealed that one of the problem items misfit:

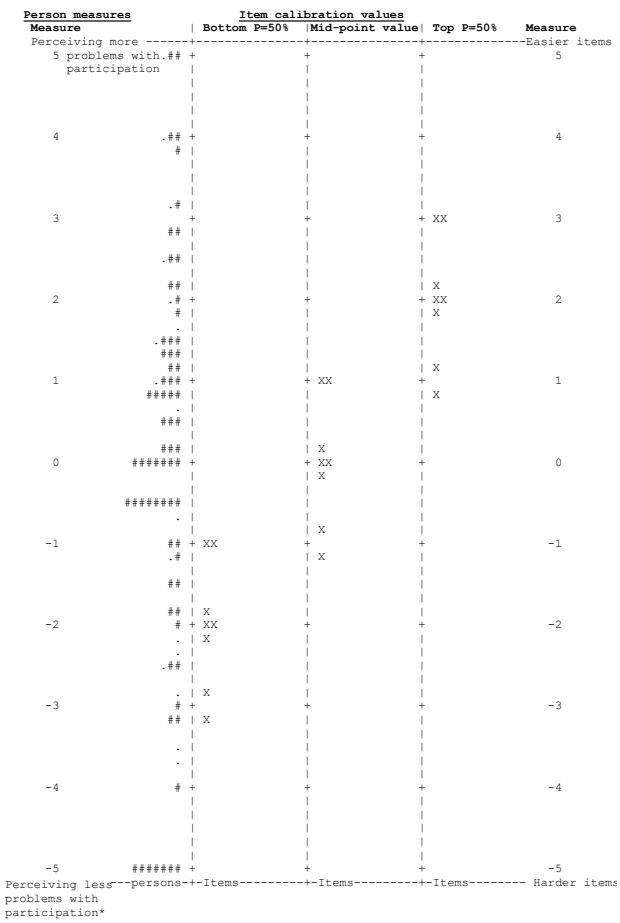

Fig. 3. Map of person-perceived problems with participation measures and item calibration values. In the left-hand column persons are plotted at their perceived problems measure. Each "#" represents two persons and each "." represents one person. The next 3 columns each include all items (X) at different points along the range of the rating scale (see text for more details). *Persons at the bottom perceived no problems with their participation (zero scores).

financial situation (Table III). When this item was removed, an additional item, education, misfitted. The final analysis resulted in 2 excluded items with the remaining 6 items showing acceptable goodness-of-fit. The person separation index remained stable during the analysis (varying between 2.13 and 2.16). Thus, the distribution of persons continued to be separated into 3 statistically distinct levels even after the exclusion of the 2 items. The distribution of the items calibration values increased after the exclusion process was completed and ranged from 1.54 to –1.29 logits (mean zero). After the exclusion, the item "social relationships" was the item that it was easiest to report low perceptions of problem with. The 2 gaps between items on the linear continuum increased to be close to or more than 1.00 logits after the exclusion process was completed. The spread of person measures remained essentially unchanged after the exclusion of the 2 items, ranging from 5.89 to –5.86 logits (mean = 0.36). The number of persons having maximum estimated measures increased to 12 persons and those having minimum estimated measures increased with one person to 15. A principal component analysis of the standardized residuals revealed that 10.4% of the unexplained variance was explained by the first factor, tentatively supporting unidimensionality.

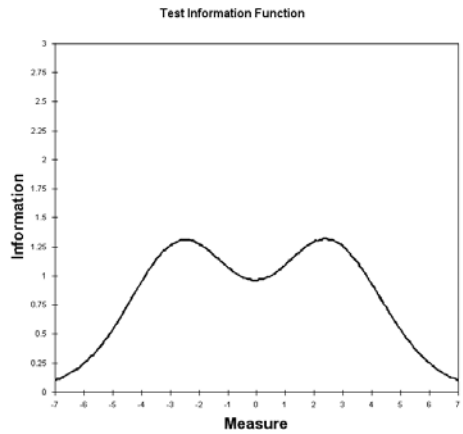

Fig. 4. Test information function of the perceived problem with participation scale.

DISCUSSION

The combined results of the goodness-of-fit evaluation and the principal component analysis revealed that the IPA-S, when used to evaluate persons with SCI, is comprised of 2 unidimensional scales. The final perceived participation scale (after removal of the misfitting items) had 27 items and the final perceived problems with participation scale had 6 items. The hierarchy of the calibration ordering of the items in the scales appears logical, as for example items related to interactions with nearest person would be expected to be easier to report high perceptions of than items related to interactions with acquaintances. In the same way, items related to work would be expected to be the most challenging. Thereby, the hierarchy of items can also be said to support the construct validity of the scale.

Four of the items for perceived participation did not demonstrate acceptable goodness-of-fit. Likewise, the goodness-of-fit statistics for 2 of the items for perceived problems with participation were not acceptable. Thus, the number of items that did not demonstrate acceptable goodness-of-fit exceeded the pre-set criteria of no more than 5% misfit. Possible reasons why the item "intimate relationship" misfit can be related to the population studied, as a common consequence of SCI is sexual dysfunction (22). The misfit of the items for perceived participation: "spending income", "achieving or maintaining a job one wants" and "contact with colleagues" and the misfitting items in the scale for perceived problems with participation: "financial situation" and "education" may be due to the influence of other external factors, such as local job markets and resultant job availability. For example, it is reasonable to assume that people with SCI have even greater difficulties obtaining or maintaining jobs now because of the frequent notices of dismissals and layoffs in the current labour market. Likewise, it is reasonable to assume that their resultant financial positions limit their possibilities of spending money and perhaps becoming more educated.

The fact that the revision of both IPA-S scales showed that the range of the item calibration values, as well as person measures, increased or remained essentially stable indicated that the misfitting items may have disrupted the measurement system. This confirms the idea that a higher number of misfits than expected can be a threat to the validity of the scale (14). Therefore, it is important that the fit of these items are monitored in future studies with new groups of persons with different type of disabilities. If the same items continue to misfit when used with other groups of persons with disabilities, the items should be omitted from the scale (14). However, if the same items fit other groups of persons with disabilities and continue to misfit for persons with SCI, the items can be divided into 2 items or omitted and an alternative version of IPA-S for persons with SCI can be developed.

The item calibration values showed that there were large distances (gaps) between some of the items in the perceived problems with participation scale of IPA-S and the effective logit range of the items was limited. There were notable floor (12 persons) and ceiling (15 persons) effects. In accordance with this, the test information function and SEs for persons indicated insufficient sensitivity. While low sensitivity is a disadvantage when measuring outcomes of rehabilitation interventions (23), low sensitivity was not apparent with the perceived participation scale. Slight ceiling effects were noted (6 persons). In contrast, the perceived problems scale may only be sensitive enough to identify those with and without perceived problems (or alternatively, those with mild vs severe perceived problems).

A plausible explanation of the ceiling effects noted in relation to our sample is that the postal questionnaire was sent to persons who did not fulfil the prerequisite for answering the IPA. That is, the instrument is aimed at persons with various chronic conditions and the analysis clearly showed that many of the respondents receiving perceived participation measures below the items calibration values had incomplete SCI with minimal disability. More specifically, of the 14 persons with the highest levels of perceived participation, 12 walked and were not in need of support from others in their daily activities. In addition, all of these 14 persons, except those who had retired, had returned to work or education. This may mean that new, more difficult items are not needed in the perceived participation scale. However, there were also persons at the lower end of the problem scale, indicating that their degrees of perceived problems were not captured by the range of the item difficulties. In accordance with previous research (7), it became evident here that the scale of perceived problems with participation is less sensitive to change than the scale for perceived participation. This can be seen as a disadvantage in relation to the fact that an absence of a persons' own perceived problems with participation are described as the optimum outcome measure following rehabilitation (7).

The developers of IPA have revised some of the items included in the original version and have included 2 new items (24). This means that the psychometric properties of the scale need to be examined again to determine whether the revised items and the new ones fit the underlying constructs. In addition, there is also a need to examine if these changes of the items have a positive impact on item targeting, the length of the scale, sensitivity and floor and ceiling effects. This examination will also reveal whether additional items need to be developed and added to the scales. Future research also needs to address the external construct validity to determine if the self reports on the scales really capture and measure participation.

In conclusion, our results show that the IPA-s comprises 2 unidimensional scales that measure perceived participation in 27 items and perceived problems with participation in 6 items for use in people with SCI. The psychometric properties of IPA-S are promising and indicate that IPA-S has potential to be further developed. Future studies with various groups of disabilities are needed to provide further evidence of the psychometric properties of the instrument.

ACKNOWLEDGEMENT

This study was supported by grants from the Swedish Council for Working Life and Social Research.

REFERENCES

1. World Health Organization (WHO). The international classification of functioning, disability and health – ICF. Geneva: WHO; 2001.
2. World Health Organization (WHO). International classification of impairments, disabilities and handicaps: a manual of classification relating to the consequences of disease (ICIDH). Geneva: WHO; 1980.
3. Cardol M, Brandsma JW, de Groot IJ, van den Bos GA, de Haan RJ, de Jong BA. Handicap questionnaires: what do they assess? Disabil Rehabil 1999; 21: 97–105.
4. NIVEL International [homepage on the Internet]. Utrecht: Netherlands Institute for Health Service Research, The questionnaire Impact on Participation and Autonomy (IPA) [cited 2005 Dec 15]. Available from: http://www.nivel.nl
5. Cardol M, de Jong BA, Ward CD. On autonomy and participation in rehabilitation. Disabil Rehabil 2002; 24: 970–974.
6. Cardol M, de Jong BA. Impact op Participatie en Autonomie (IPA). Handleiding en vragenlijst. Amsterdam: Afdeling Revalidatie, Academisch Medisch Centrum; 2001.
7. Cardol M, Beelen A, van den Bos GA, de Jong BA, de Groot I, de Haan RJ. Responsiveness of the impact on participation and autonomy questionnaire. Arch Phys Med Rehabil 2002; 83: 1524–1529.
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. Cardol M, de Haan RJ, van den Bos GA, de Jong BA, de Groot IJ. The development of a handicap assessment questionnaire: the Impact on Participation and Autonomy (IPA). Clin Rehabil 1999; 13: 411–419.
10. Larsson Lund M, Nordlund A, Nygård L, Lexell J, Bernspång B. Perceptions of participation and predictors of perceived problems with participation in persons with spinal cord injury. J Rehabil Med 2005; 37: 3–8.
11. Hilton A, Skrutkowski M. Translating instruments into other languages: development and testing processes. Cancer Nurs 2002; 25: 1–7.
12. Su CT, Parham LD. Generating a valid questionnaire translation for cross-cultural use. Am J Occup Ther 2002; 56: 581–585.
13. Winsteps. Winsteps Rasch Measurement Software version 3.58, 2005. Available from: http://www. winsteps.com
14. Bond TG, Fox CM, editors. Applying the Rasch Model: fundamental measurement in the human sciences. New Jersey: Lawrence Erlbaum; 2001.
15. Wright BD, Masters GN, editors. Rating scale analysis. Chicago: MESA Press; 1982.
16. Wright DW, Stone MH, editors. Best test design. Rasch measurement. Chicago: MESA Press; 1979.
17. Linacre JM. A user's guide to Winsteps. Rasch-models computer programs. Chicago: Winsteps, 2005 [cited 2006 Jun 5]. Available from: http://www. winsteps.com
18. Linacre JM. Optimizing rating scale category effectiveness. J Appl Meas 2002; 3: 85–106.
19. Wright B, Linacre M. Reasonable mean-square fit values. Rasch Meas Trans 1994; 8: 370.
20. Wilson M, editors. Constructing measures: an item response modelling approach. New Jersey: Lawrence Erlbaum; 2005.
21. Fisher W. Reliability statistics. Rasch Meas Trans 1992; 6: 238.
22. Deforge D, Blackmer J, Moher D, Garritty C, Cronin V, Yazdi F, et al. Sexuality and reproductive health following spinal cord injury. Evid Rep Technol Assess 2004; 109: 1–8.
23. Fisher AG. Functional measures, part 2: Selecting the right test, minimizing the limitations. Am J Occup Ther 1992; 46: 278–281.
24. Cardol M. Impact on Participatie en Autonomie (IPA) Handleiding en vrangenlijst. Amsterdam: Afdeling Revalidatie, Academisch Medisch Centrum; 2004.