Using the Rasch Model to Develop a Measure of Participation Capturing the Full Range of Participation Characteristics for the Patients with Hand Injuries

Maryam Farzad[1]   Fereydoun Layeghi[2]   Ali Hosseini[1]   Gale Whiteneck[3]   Ali Asgari[4]

[1] Department of Occupational therapy, The University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[2] Department of Clinical Science, The University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[3] Department of Research, Craig Hospital, Englewood, Colorado, United States

[4] The University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Address for correspondence Ali Asgari, PhD, The University of Social Welfare and Rehabilitation Sciences, Kodakyar Street, Daneshjo Boulevard, Evin, Tehran, Iran (e-mail: asgarya@ut.ac.ir).

Abstract

Objective   The purpose of this paper was to report on the first step in the development of a new instrument to measure participation including the full range of its characteristics.

Methods   The 30-item participation behavior questionnaire (PBQ) was developed from four main sources (a literature review of the theatrical basis of participation, available participation measures, and interviews with patients and experts about participation). Item selection and the reliability and validity of the measure were explored using Rasch measurement modeling for analysis.

Participants   A total of 404 individuals referred to rehabilitation after hand, wrist, or upper extremity surgery to reduce impairment from trauma, at least 2 months post-injury.

Results   An initial pool of 100 items; reflecting 14 characteristics of participation was initially reduced to 91 items after review by 15 participation experts and then further reduced to 30 items by three rounds of Rasch analysis removing misfitting items. The final PBQ has a person reliability of 0.91 with separation of 3.22, indicating it can reliably differentiate four levels of participation. There are no misfitting items and the instrument is unidimensional. All 14 characteristics of participation were retained in the PBQ, and none of the 30 items refer specifically to upper extremity issues.

Conclusion   The 30 participation behavior items of the PBQ show promise of being a psychometrically sound measure of participation. Further research is needed to validate the PBQ in samples of people with a range of other disabilities.

Keywords
► participation
► Rasch analysis
► outcome measure
► hand injuries

Introduction

Participation is a core concept in the International Classification of Function, Disability, and Health (ICF) of the World Health Organization.[1] As participation is an important long-term outcome of medical, rehabilitation, and social service programs supporting individuals who are disabled because of impairments,[2] it must be evaluated to help clinicians and rehabilitation services measure the participation levels of their clients and track their changes. However, there is no consensus on either the construct or the measurement of participation.[2,3]

Participation is defined in the ICF as "involvement in a life situation," whereas its converse, *participation restrictions*, are defined as "problems an individual may experience in involvement in life situations."[4] Though the ICF blurs the distinction between activity and participation by having only one coding system for both,[3] some have suggested that participation is characterized by specific chapters of ICF, including: Assisting others (d660), Interpersonal interaction and relationships (d730-d779), Education (d810-d839), Work and employment (d840-d859), Economic life (d860-d879), Community life (d910), Recreational and leisure (d920), Religion and spirituality (d930), Political life and citizenship (d950).[5] However, in rehabilitation research, other concepts are considered as participation as well. Though some consider participation as a construct that can be determined objectively, others have developed measures that reflect subjective aspects of participation, including autonomy; satisfaction; perceived difficulty of participation; equality with peers; and experiences of inclusion, independence, and nondiscrimination.[6–8] Conceptual and measurement issues requiring clarification include identifying domains of participation, ascertaining metrics appropriate to each domain, describing typical levels of each metric in each domain, and identifying the key characteristics of participation.[9]

The current process of developing a new measure of participation was based on our recent content analysis study to conceptualize participation after hand injuries. In this content analysis study, the models and theories about participation and also the outcome measures that claimed to measure participation were reviewed, outcome measures specific to hand and upper extremity impairment that had items referring to participation were also included. Semi-structured interviews were also conducted with patients who had received hand rehabilitation and also with hand surgeons and hand therapists. Extensive coding and analysis of these literature and instrument reviews, and patient and professional interviews identified 14 main characteristics or domains of participation (manuscript under review): (1) Assisting others, (2) Interpersonal relationship and communication, (3) Education, (4) Work and employment, (5) Economic life, 6) Religion and spirituality, (7) Political life and citizenship, (8) Role, (9) others: participation cannot be done alone, (10) Domestic life, self-care, looking after ones health, (11) Leisure and recreation, (12) Subjective participation, (13) Environment: social, cultural, and temporal, and (14) Community life (see ►**Table 1** for a description of these categories).

Though dozens of participation measures have been published, none were found to accurately reflect all of the aforementioned characteristics of participation.[10] Therefore, the objectives of this research were to develop a participation item pool written to reflect all 14 characteristics, then use initial expert review followed by Rasch analysis to identify the best items in that pool forming a unidimensional instrument, and finally to determine how many of the defining 14 characteristics were retained in the unidimensional measure of participation for hand injured patients.

Methods

Participants and Procedures

In the first phase of the study 15 participation experts recruited from hand surgeons, hand therapists, and the therapists that address patient participation reviewed the initial pool of 100 items to determine whether all item content reflected aspects of participation. The convenience sample for the second phase of the study consisted of 404 patients with hand and upper extremity injury referred by hand surgeons to the outpatient hand clinic and rehabilitation clinics during a 12-month period (February 2013–2014). The sample size was adequate to support the planned Rasch analysis.[11,12] Inclusion criteria were age 18 or older and at least 2 months post-injury. Participants were selected from both urban and rural areas. Exclusion criteria were patients who were unable to write and read, had cognitive dysfunction, neurologic diseases, or declined participation. All the individuals gave research authorization to participate in the study. The medical ethic committee of our university approved the study.

The questionnaire was administered in hospital and clinic waiting rooms by postgraduate hand therapy students who were available to answer any questions raised by the patients. The participants completed the self-report individually, and questionnaires took nearly 10 to 20 minutes to complete.

Instrument

The participation behavior questionnaire (PBQ) was developed with 100 items assessing the 14 main characteristics of participation. Of the 100 initial items, 9 items were deleted after a content analysis review by 15 participation experts, which yielded a consensus that the 9 items (all written to reflect emotional and personal reactions) did not represent participation. The remaining 91 items were administered in a written questionnaire format to assess hand and upper extremity injured patient's reports of their participation behavior after traumatic injury and surgery. Response categories for all items included completely disagree, somewhat disagree, somewhat agree, and completely agree. The four-point scale of 1–4 was reversed for negatively worded items.

Data Analysis

The analyses to assess the psychometric properties of the PBQ were investigated using Rasch rating scale model

Table 1 The participation domains

N	Domain	Description
1	Assisting others	Caring for household objects and assisting others
2	Interpersonal relationship and communication	Communicating and having relationships with family and friends
3	Education	A part of human activities in the educational area
4	Work and employment	All activities that refer to workplace and job, Engaging in programs related to preparation for employment
5	Economic life	Managing economic issues of life
6	Religion and spirituality	Engaging in religious or spiritual activities
7	Political life and citizenship	Engaging in the social, political, and governmental life of a citizen
8	Role	Managing life roles like being a mother
9	Other persons	Activities that need other people and cannot be done alone
10	Domestic life	Engaging in acquisition of necessities, household tasks, self-care, looking after ones health
11	Leisure and recreation	Engaging in any form of play, recreational, or leisure activity
12	Subjective participation	Inner perceive of the participant from so many perspective such as: emotional/psychological, engagement, appearance, adaptation, compensatory behavior physical, independence and autonomy, personal reaction, satisfaction
13	Environment	Social, cultural, and temporal environment
14	Community life	Engaging in all aspects of community and social life, such as engaging in charitable organizations, service clubs, or professional social organizations

(RSM).[13,14] The Rasch model, introduced by Georg Rasch,[15,16] provides a contemporary strategy for developing a psychometrically sound instrument. Items that best measure the full range of a construct are selected. Respondents and items are measured on the same metric based on the probability of how a given individual will respond to a given item, defining an order or hierarchy of both respondents and items.[17] Rasch analysis generates true interval measures from the ordinal ratings made by the client.[18] Rasch analysis offers a detailed means to evaluate whether each item on a scale works well to capture the intended construct. Rasch analysis tests whether the items coalesce to form a single construct or underlying dimension; this property is referred to as unidimensionality.[19] It provides information on unidimensionality, reliability, construct validity, and the hierarchies of the items in the scale.[20]

To assess dimensionality, fit of the items to the RSM was explored. In addition, principal component analysis (PCA) of residuals was used to examine whether a substantial factor exists in the residuals after the primary measurement dimension has been estimated.[21,22] Construct validity was examined by estimating Information-weighted fit (Infit) Mean Square (MnSq), Outlier-sensitive fit mean score (Outfit) MnSq, and Z-score standardized (Zstd) fit statistics. Items were considered to misfit the model and not belong to the constructs when MnSq had values >1.4 and were associated Zstd > positive 2.0.[20] Reliability, including the item, and person measure reliability were estimated with values >0.7 suggesting good reliability; item and person separation indices were also estimated with values >2 suggesting good separation. Low

person separation implies that the instrument may not be sensitive enough to distinguish between high and low participation. Low item separation implies that the person sample is not large enough to confirm the item difficulty hierarchy or construct validity of the instrument.[23]

Item maps were used to plot the hierarchy of the items in the scale and compare their distribution to that of the respondents. Linearity of the construct can also be examined by reviewing item hierarchies, or the relative difficulty of each item compared with another on a linear continuum. Item difficulties are given in logits, an interval measure, with the mean of the item difficulties centered at 0. Higher scores represent greater participation.

To obtain an appropriate and valid subset of items, we used systematic steps based on both conceptual and psychometrical inquiry. We excluded items with poor psychometric properties based on the item Infit and Outfit, and item-measure (point-biserial [ptbis]) correlations (corr.). Another Rasch analysis was run after omitting the misfitting items and persons. This process was repeated until the entire remaining respondents and items fit. Winsteps (version 3.68.1)[24] was conducted to evaluate the psychometric properties and hierarchies of item difficulty placed on the item-person map.

Results

Demographics of the sample appear in ►**Table 2**. After the original 100 items draft of the PBQ was reduced to 91 items by the expert panel, further item reduction occurred in three rounds of Rasch analysis to reach an appropriate and valid

Table 2 Demographic characters of participant

Variable		*n*	%
Sex	Male	274	67.8
	Female	130	32.2
Marriage	Single	203	50.2
	Married	201	49.8
Injured part	Shoulder	49	12.1
	Elbow	55	13.6
	Wrist	122	35.1
	Hand	128	31.7
	Total	30	7.4
Injured hand	Right	148	36.6
	Left	256	63.4
Dominant hand	Right	57	14
	Left	347	86

subset of items. Results of the final analysis are shown in ►**Table 3** and ►**Fig. 1**. The 30 remaining items showed good fit to the Rasch model, suggesting that PBQ met the criteria for the unidimensionality assumption of the Rasch model. This is evidenced not only by the fit of items but also by the positive item ptbis corr. coefficients in ►**Table 3**.

Unidimensionality

To examine unidimensionality, the Rasch PCA on the residuals of items was used and resulted in one component with an eigen value >1. This explained 60.3% of the variance in the PBQ and met the criteria for unidimensionality (>40%). That is, participation is the only construct underlying the 30 items of PBQ.

Content Validity

In the first phase to examine the content validity, the experts were asked to rate the items based on relevance to the 14 participation characteristics through use of a rating scale ranged 0 to 10. To estimate the rater reliability, the intraclass correlation (ICC)[25] was calculated at 0.97 (p <.01), indicating there was high agreement that the nine items deleted did not measure participation.

Construct Validity

Misfitting Items

In the second phase, the construct validity of PBQ was examined to reduce the misfitting items based on the Rasch

model. Item reduction happened in three steps on the 91-item draft of the PBQ to reach an appropriate and valid subset of items. The final 30 remaining items in the third step showed a good fit to the Rasch model (►**Table 3**). The Infit and Outfit values for the 30 items in ►**Table 4** suggest that PBQ met the criteria for the construct validity assumption of the Rasch model. Construct validity is also supported by the positive item ptbis corr. coefficients.

Reliability

The reliability of the PBQ was excellent with an item measure reliability of 0.91, suggesting that items can be discriminated from one another based on their difficulty. The person measure reliability was 0.96 showing that the scale can discriminate participants based on their estimated participation level. The item separation index of 4.38 indicates that items in PBQs have the potential to be categorized into four levels. The person separation index of 3.22 shows that the participants can be divided into three groups based on their participation level, and also our sample was big enough to show that PBQ is sensitive to distinguish between patients with different levels of participation.

The Person-Item Maps

The Rasch person-item map in ►**Fig. 1** demonstrates a continuum of participation created from measurements of individuals' ratings of each item. This map indicates items and respondents are appropriately targeted.[20] The items are located on the right against the respondents on the left. Items closer to the bottom are easy to participation in and moving up the continuum indicates items that are harder to participate in. Persons with the highest participation on the scale were at the top of the continuum, whereas those with lowest participation were at the bottom of the continuum. ►**Fig. 1** indicates how the participants reported the greatest participation for the items (I feel my family members shun me) and (I feel that others are less interested in socializing with me), whereas the items (my participation in family tasks is reduced) and (I can no longer support my family) were the most difficult activities to participate. The mean of the sample measures was below the mean of the participation items.

Category Function

The participation rating scales function as intended (►**Table 4**). For the rating scales, the observed average measures were ordered for each response category, indicating that as clients' participation increased, the use of a higher response category became more likely. Step calibrations

Table 3 The results of item reduction steps based on Rasch analysis

Steps	Misfitting items	Misfitting persons	Range of ptbis corr.	Variance explained (%)
First	39	15	0.46–0.71	76.4
Second	18	–	0.44–0.71	73.2
Third	5	45	0.50–0.76	60.3

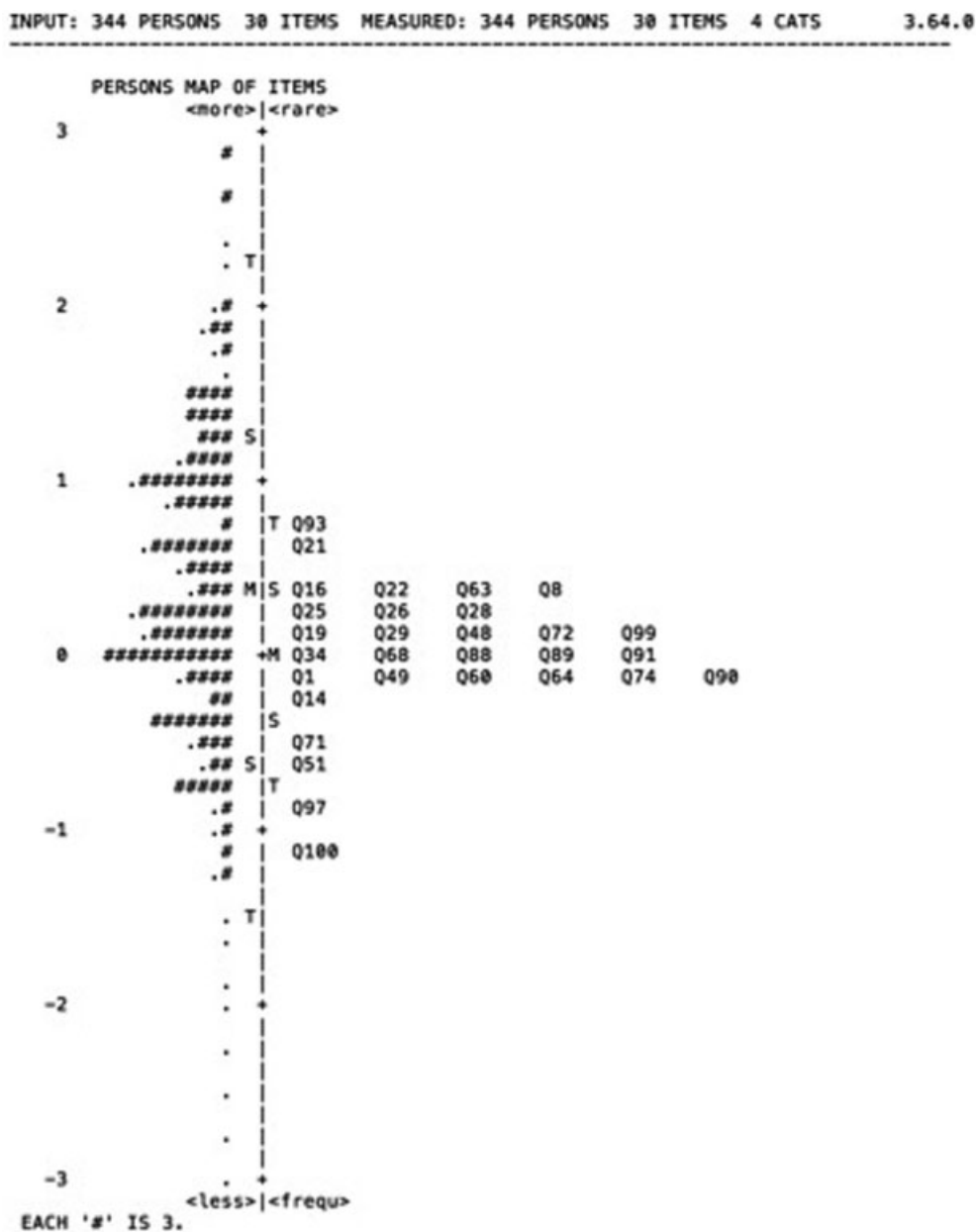Abbreviation: ptbis corr., point-biserial correlation.

Fig. 1 Item maps of the PBQ items. PBQ, participation behavior questionnaire.

were also ordered, inferring that higher rating responses represented more of the construct being measured. All rating response categories had outfit MnSq statistics <1.4. Finally, patients used the lower rating scale category less frequently than the other response categories.

Discussion

The PBQ developed in this research includes 30 items assessing participation behaviors that have good psychometric properties. The PBQ has a person reliability of 0.91 with separation of

Table 4 The participation rating scale functioning

Response category	Observed count	Observed average measure	Outfit MnSq	Step measure
Completely disagree	1,840	−0.57	1.31	None
Relatively disagree	2,398	−0.13	0.78	−0.68
Relatively agree	2,592	0.47	0.87	0.29
Completely agree	3,489	1.10	1.08	0.39

Abbreviation: MnSq, Mean Square.

Table 5 Item characteristics of the final participation item subset

Original item number	Main characteristics	Item content	Measure	Mode	Infit		Outfit		Item
									measure
				SE	MnSq	Zstd	MnSq	Zstd	Correlation
93	2	My participation in family tasks is reduced.	0.75	0.60	0.92	−1.10	0.92	−1.00	0.54
21	2, 10, 1	I can no longer support my family.	0.60	0.60	0.97	−0.40	0.98	−0.20	0.59
63	11	I have some time for fun, but less than before.	0.43	0.60	0.87	−2.00	0.95	−0.70	0.65
1	11	My spare time activities are limited.	0.41	0.60	1.23	1.80	1.32	2.00	0.56
22	12	I feel I have lost independence.	0.35	0.60	1.02	0.30	1.21	1.50	0.60
16	12	I feel I am dependent on my family for most of my tasks.	0.33	0.60	0.91	−1.30	0.97	−0.40	0.61
25	12, 2, 14	I cannot tolerate participating in social activities as before.	0.27	0.60	1.03	0.40	1.03	0.40	0.60
28	7	Using public transport is hard for me.	0.20	0.60	0.94	−0.90	0.98	−0.20	0.66
26	8, 2	I rarely tolerate strolling with friends.	0.19	0.60	0.69	−5.20	0.67	−4.70	0.76
99	12	I feel bad about not being able to behave normally.	0.16	0.60	1.18	1.60	1.19	1.30	0.57
72	12, 1, 10, 8	I can play my role as a member of the family.	0.16	0.60	0.98	−0.30	0.96	−0.50	0.60
29	12, 9	It is difficult to tolerate others' reactions	0.13	0.60	0.89	−1.60	1.05	0.60	0.63
19	2	I have restricted my social participation.	0.13	0.60	0.86	−2.20	0.86	−1.80	0.66
48	2, 9, 12	I barely tolerate being in public places.	0.07	0.60	0.95	−0.80	1.27	1.40	0.59
34	11	I try to attend recreational activities.	0.03	0.60	1.06	0.90	1.08	1.00	0.57
91	8, 9, 10	I am able to host family and friends.	0.01	0.60	1.19	1.70	1.32	1.90	0.48
88	13, 10, 7	I rarely accompany my family to a park or public places.	0.00	0.60	1.03	0.50	1.00	0.00	0.59
68	10, 1, 13, 9	I take care of my family the same as before.	−0.02	0.60	1.12	1.70	1.22	1.60	0.53
89	6	I feel my religious activity has decreased.	−0.04	0.60	1.27	1.70	1.19	1.20	0.56
90	2, 7, 8, 14	I try to communicate with others indirectly (by mail, phone, etc.).	−0.08	0.60	0.93	−1.10	1.00	0.10	0.62
74	4, 2, 13	Collaboration with my colleagues has decreased.	−0.09	0.60	1.10	1.50	1.18	1.00	0.55
60	2, 13, 7	I cannot have the same relationship with my family as before.	−0.12	0.60	0.94	−0.90	1.03	0.40	0.64
64	12	I feel I cannot behave like the past.	−0.12	0.60	0.93	−1.00	0.97	−0.30	0.65
1	12, 2, 8	I rarely meet my friends these days.	−0.16	0.60	1.15	1.60	1.31	1.90	0.54
49	12, 5	I cannot deal with my problem.	−0.18	0.60	1.03	0.40	1.11	1.30	0.59
14	12, 8, 13	I think I cannot play my role in my family.	−0.23	0.60	1.10	1.50	1.28	1.70	0.53
71	2	I try to keep my relationships with my friends.	−0.53	0.70	0.99	−0.10	1.02	0.20	0.62
51	12	I feel my family abandons me.	−0.68	0.70	0.99	−0.10	1.00	0.00	0.55
97	12	I feel that others are less interested in socializing with me.	−0.87	0.70	0.82	−2.30	0.76	−2.20	0.65
100	12	I feel my family members avoid me.	−1.09	0.80	0.95	−0.60	0.88	−0.90	0.54
		Mean	0.00	0.60	1.00	1.10	0.70		
		SD	0.39	0.00	0.13	0.20	1.90		

Abbreviations: MnSq, Mean Square; SE, standard error; Zstd, Z-score standardized fit statistics.
(1) assisting others, (2) interpersonal relationship and communications, (3) education, (4) work and employment, (5) economic life, (6) religion and spirituality, (7) political life and citizenship, (8) role, (9) others, (10) domestic life, self-care, looking after ones health, (11) leisure and recreation, (12) subjective participation, (13) environment: social, cultural, and temporal, and (14) community life.

3.22, indicating that it can reliably differentiate four levels of participation. There are no misfitting items and the instrument is unidimensional. The PBQ has good construct validity and all items measure participation (►Table 5).

A unique aspect of this participation instrument is that the content of the items was derived from a series of literature and instrument reviews and interviews of people with disability and their rehabilitation providers, where 14 characteristics were identified that defined participation from different perspectives. All 14 characteristics were represented in the original pool of 100 items, and all of them remain in the final 30 selected items. This indicates that the PBQ has a broad base of content that function effectively as a unidimensional instrument.

It should be noted, however, that none of the 30 final items selected have content that specifically relates to hand or upper extremity disability. Therefore, the PBQ has the potential of being applicable to a wider range of disability. The next logical research step is the validation of the PBQ in a second sample including a range of disabilities. Additional areas of further research might include examining the distribution of PBQ scores, along with group differences, and differential item response using Rasch analysis. Better conceptualization of participation and its characteristics (including the 14 characteristics identified in ►Table 1), as well as empirical examination of how they relate, are also needed aspects of future participation measurement and research.

Independence without participation restrictions is the final goal of therapists and surgeons, so measuring participation broadly and distinguishing it from limitations in activities of daily living can help clinicians evaluate their interventions for patients. Until now, the emphasis of therapists and surgeons was more on activity limitation and impairment (opposition limitation after median nerve injury) and less on participation of patients (returning to work and the same role in their family as before injury). Measuring participation can help a surgeon consider the patient's larger priorities for treatment outcome.

Limitations of this research include the sample only including people with hand/upper extremity disability from one outpatient rehabilitation facility, with no second validation sample. The study included participants from one country and culture, and the domains of participation may differ among cultures. Also, use of a convenience sample that may not be representative, and the potential of an item ordering effect because items were not randomized are limitations of this study. Furthermore, the way participants responded to the 91 items in the questionnaire may not be the same as responses to the final 30-item instrument. Finally, it should be acknowledged that no other measures of impairment, disability, or participation were included in the study, and therefore relationships between the PBQ and other measures were not examined.

Conclusion

The 30 participation behavior items of the PBQ show promise of being a psychometrically sound measure of participation.

None of items refer specifically to hand or upper extremity issues, making the PBQ potentially appropriate for people with a broader range of disability. Further research to validate PBQ in a broader disability sample is needed.

Note
Ethic committee of University of Social Welfare and Rehabilitation Sciences, Tehran, Iran, approved this work, using the Rasch model to develop a measure of participation for patients with hand injuries.

Conflict of Interest
None.

References
1 Organization WH. International Classification of Functioning, Disability and Health: ICF: World Health Organization; 2001
2 Dijkers MP. Issues in the conceptualization and measurement of participation: an overview. Arch Phys Med Rehabil 2010;91 (9, Suppl):S5–S16
3 Whiteneck G, Dijkers MP. Difficult to measure constructs: conceptual and methodological issues concerning participation and environmental factors. Arch Phys Med Rehabil 2009;90(11, Suppl): S22–S35
4 Nordenfelt L. Action theory, disability and ICF. Disabil Rehabil 2003;25(18):1075–1079
5 Post MW, de Witte LP, Reichrath E, Verdonschot MM, Wijlhuizen GJ, Perenboom RJ. Development and validation of IMPACT-S, an ICF-based questionnaire to measure activities and participation. J Rehabil Med 2008;40(08):620–627
6 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(06):459–481
7 Brown M. Participation Objective, Participation Subjective. The Center for Outcome Measurement in Brain Injury; 2006
8 Brown M. Participation Objective–Participation Subjective. Encyclopedia of Clinical Neuropsychology: Springer; 2011: 1875–1877
9 Walker N, Mellick D, Brooks CA, Whiteneck GG. Measuring participation across impairment groups using the Craig Handicap Assessment Reporting Technique. Am J Phys Med Rehabil 2003; 82(12):936–941
10 Farzad M, Layeghi F, Asgari A, Hosseini SA, Rassafiani M. Evaluation of non diseased specified outcome measures in hand injuries to assess activity and participation based on ICF content. J Hand Microsurg 2014;6(01):27–34
11 Chen W-H, Lenderking W, Jin Y, Wyrwich KW, Gelhorn H, Revicki DA. Is Rasch model analysis applicable in small sample size pilot studies for assessing item characteristics? An example using PROMIS pain behavior item bank data. Qual Life Res 2014; 23(02):485–493
12 Wright BD. Sample-Free Test Calibration and Person Measurement. Paper Presented at the National Seminar on Adult Education Research (Chicago, February 11–13, 1968)
13 Andrich D. Application of a psychometric rating model to ordered categories which are scored with successive integers. Appl Psychol Meas 1978;2(04):581–594
14 Wright BD, Masters GN. Rating Scale Analysis. Rasch Measurement. ERIC; 1982
15 Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. ERIC; 1993

16 Rasch G. Probabilistic Models for Some Intelligence and Achievement Tests. Copenhagen, Denmark: Danish Institute for Educational Research; 1960

17 Smith EV Jr, Conrad KM, Chang K, Piazza J. An introduction to Rasch measurement for scale development and person assessment. J Nurs Meas 2002;10(03):189–206

18 Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70(12):857–860

19 Wright BD, Stone MH. Best Test Design. Rasch Measurement; 1979

20 Bond TG, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 3rd ed. AE Enschede, The Netherlands: University of Twente; 2001

21 Linacre JM. Detecting multidimensionality: which residual data-type works best? J Outcome Meas 1998;2(03):266–283

22 Smith EV Jr. Understanding Rasch measurement: detecting and evaluating the impact of multidimenstionality using item fit statistics and principal component analysis of residuals. J Appl Meas 2002;3(02):205–231

23 Linacre JM, Wright B. Winsteps. Available at http://www.winsteps.com/index.htm. Accessed June 27, 2013

24 Linacre JM, Wright BD. A User's Guide to WINSTEP. Chicago, IL: MESA Press; 2009

25 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(02):420–428