

The Maastricht social participation profile: development and clinimetric properties in older adults with a chronic physical illness

Godelief M. J. Mars · Gertrudis I. J. M. Kempen ·
Marcel W. M. Post · Ireen M. Proot · Ilse Mesters ·
Jacques T. M. van Eijk

Accepted: 24 August 2009 / Published online: 21 September 2009
© Springer Science+Business Media B.V. 2009

Abstract

Purpose To develop and test the Maastricht Social Participation Profile (MSPP), an instrument measuring the actual social participation by older adults with a chronic physical illness, in accordance with their own definition of social participation.

Methods The development process consisted of a number of steps, ending with a field test in two waves ($n = 412$ and $n = 125$) among a random sample of people older than 59 years with either COPD or diabetes mellitus. Reproducibility was evaluated with intraclass correlation coefficients (ICCs) and smallest real differences at group level (SRDs_{group}). Convergent and discriminant validity were evaluated with Pearson correlation coefficients between the MSPP and the Frenchay Activities Index (FAI).

Results The MSPP consists of four indices: consumptive participation, formal social participation, informal social

participation-acquaintances and informal social participation-family. Each index measured diversity and frequency of participation. ICCs ranged between 0.63 and 0.83. SRDs_{group} ranged between 0.05 and 0.09. Convergent and discriminant validity were supported by the correlations between the MSPP_{frequency} and the FAI.

Conclusions The MSPP has good validity and acceptable reproducibility. Its distinguishing features are its focus on actual *social* participation and the possibility to calculate both diversity and frequency scores.

Keywords Chronic disease · Social participation · Instrument · Psychometrics

Introduction

Demographic ageing has drawn the attention of policy makers to the negative effects that diseases and disabilities may have on participation by older people [1, 2]. Participation is defined by the International Classification of Functioning, Disability and Health (ICF) as ‘involvement in life situations’ [3]. Policy makers aim to promote participation in this group because of the expected benefits for society and increased quality of life for the individuals concerned [1, 2, 4]. For quality of life, research suggests that social roles may be a more important aspect of participation than daily activities [4, 5]. For society to function, social contact and exchange between people are imperative. Research would therefore benefit from a measurement instrument, which focuses on the social aspects of participation. We developed such an instrument, the Maastricht Social Participation Profile (MSPP).

The MSPP intends to measure *actual social participation* by older adults with a chronic physical illness. It builds

G. M. J. Mars · G. I. J. M. Kempen · I. M. Proot · I. Mesters ·
J. T. M. van Eijk
School for Public Health and Primary Care, Maastricht
University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

G. M. J. Mars (✉)
Statistics Netherlands, PDV/DOP, Room 1D02, PO Box 4481,
6401 CZ Heerlen, The Netherlands
e-mail: g.mars@cbs.nl

M. W. M. Post
Rehabilitation Centre De Hoogstraat, Rembrandtkade 10, 3583
TM Utrecht, The Netherlands

M. W. M. Post
Rudolf Magnus Institute for Neuroscience, University Medical
Centre Utrecht, Utrecht, The Netherlands

on a definition given by older adults with a chronic physical illness themselves. They define social participation as a positive experience having one or more of the following three characteristics: social contact, contributing to society (like paying a visit) or receiving from society (like receiving a visit) [6]. This definition *excludes* behaviours which do *not* involve an exchange between people (like doing own household chores), which distinguishes social participation from the broader concept of participation. Furthermore, the definition *includes* behaviours which involve receiving from society, while other definitions of *social* participation tend to exclude these behaviours [6–12].

The question of whether social participation involves only contributing or both receiving and contributing is particularly relevant in the case of people with a chronic illness, because their opportunities to contribute may diminish. In response, people may explore alternatives [13]. If social participation involves both receiving and contributing, there are more alternatives, making it easier to maintain a given level of social participation (substitution). For example, instead of paying a painful or fatiguing visit to a friend (contribute), the friend may come to visit (receive).

The MSPP measures *actual* social participation, which refers to the frequency and diversity of social participation: how often do people engage in social participation and in how many different types of social participation do they engage? Information about actual social participation of people with a chronic illness is important from a societal perspective, because it tells us to what extent people are integrated in society. From an individual perspective, actual social participation may play a less important role in quality of life than its subjective experience [5]. For people are autonomous and differ in the frequency and types of social participation they prefer. Nonetheless, information about actual social participation may improve our understanding of subjective social participation and quality of life. This requires the use of additional instruments to measure subjective social participation and quality of life. Are people more satisfied about their social participation if they participate more often or if they participate in several different ways—or are both equally important (or unimportant)? Also, if a measure improves people's subjective social participation, it is important to understand why: did the actual social participation change or did people afterwards feel better about the same actual social participation?

This paper addresses the development and clinimetric properties of the MSPP. The MSPP was developed as a self-administered generic measure for actual social participation by older adults with a chronic physical illness with the purpose of discrimination and evaluation.

Methods

The MSPP was developed with a sample of older adults with either chronic obstructive pulmonary disease (COPD) or diabetes mellitus type 2 (diabetes). COPD and diabetes both take a gradually deteriorating course, but COPD has intermittent exacerbations, while diabetes is characterized by a long stabilization phase followed by chronic complications [14].

The development process consisted of a number of steps, after each of which the MSPP was revised. Figure 1 outlines the sequence and purpose of the steps. In this section, we explain the steps. The results section will focus on the final step, the second field test, in which we evaluated the reproducibility and validity of the semi-final version in order to arrive at a final version of the MSPP.

Medical ethics committee approval was granted.

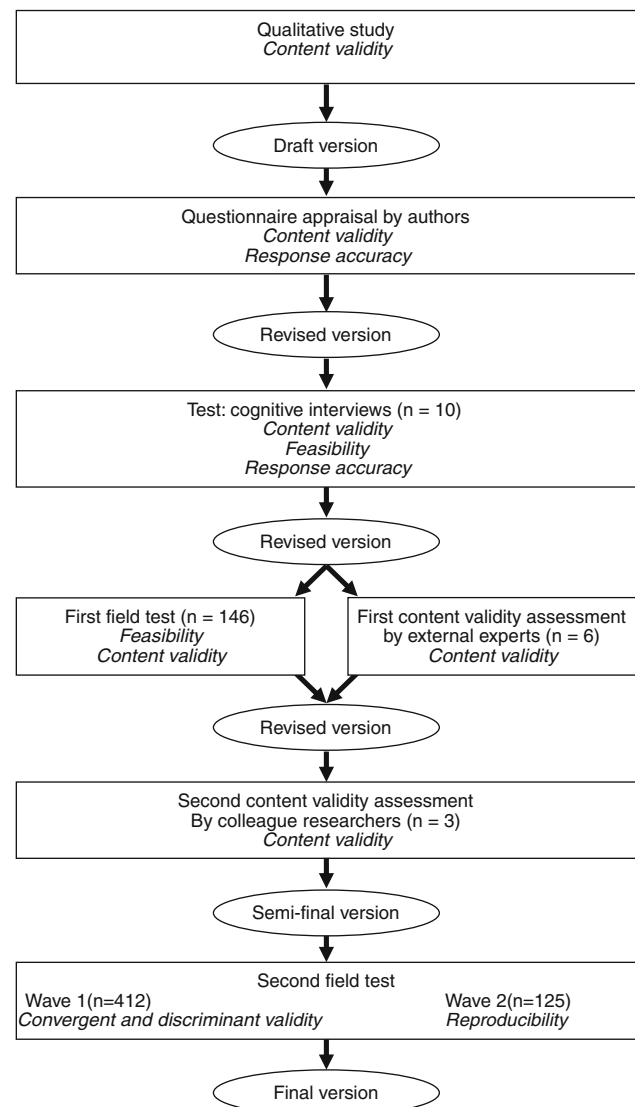


Fig. 1 Development of the Maastricht Social Participation Profile

Development of the semi-final version

Initially, we conducted a qualitative study [6]: a literature search and ten individual interviews resulted in a list of social participation examples, which was used in two focus group sessions. The discussion focused on participants' reasons for classifying certain items as social participation while rejecting others. This resulted in the definition of social participation already presented in the introduction.

Next, following guidelines about questionnaire design [15–19], we constructed three indices, based on the presented definition and using items from the list of social participation examples [6]. We deviated from the definition in two ways. First, we did not operationalize that social participation should be a positive experience, because this involves subjective evaluation, while we wanted to develop a measure for *actual* social participation. Second, we treated 'social contact' as a necessary characteristic to focus the content of the MSPP. Consequently, all three indices included only participation behaviours involving social contact. The first index concerns consumptive participation (CP, nine items), which is characterized as benefiting from society (for example taking a course or visiting a restaurant). The second index concerns formal social participation (FSP, two items), which is characterized as contributing to society (participation in clubs and volunteer work). The third index concerns informal social participation (ISP, nine items), which is characterized as contributing to society, receiving or both (contact with family, friends and acquaintances).

The response format referred to the number of times something was done in the last 4 weeks, but the response key indicated with how often a week this corresponds. The former is easier to answer when counting (rare and salient behaviours), while the latter is easier when estimating (frequent and mundane behaviours) [20]. The MSPP includes both.

The authors GM, GK, IP, IM and JvE systematically evaluated response accuracy and content validity [21]. Content validity refers to 'the extent to which an empirical measurement reflects a specific domain of content' [22: 20]. The ISP was split into separate indices for acquaintances (ISP-A) and family (ISP-F) (identical items).

Next, we conducted ten cognitive interviews to test whether items were interpreted as intended (content validity). Participants were one man and two women with COPD and six men and one woman with diabetes, ranging in age between 65 and 83 years. Participants were asked to formulate retrospectively (or concurrently if they preferred) how they had interpreted items and decided on their answers [23, 24]. Probing techniques were used to check

feasibility and response accuracy. Although revisions were made, the main result of the interviews was that items had been interpreted as intended.

The first field test involved a random sample of adults older than 59 years with either COPD ($n = 71$) or diabetes ($n = 75$) (May 2004). We analysed missing value patterns, frequency distributions, inter-item correlations and comments written on questionnaires. All statistical analyses in this study were done with the SPSS computer program version 12.0.1 [25]. The two highest response categories were combined.

Parallel to the first field test, content validity was assessed by six Dutch experts in the area of participation research, who had not been involved in the project so far. The experts received the MSPP, a schematic representation of its operationalization, and the argumentation behind it. They were asked to comment on both the operationalization and the underlying conceptualization and could freely structure their response or follow a more detailed list of questions which was provided. Generally, the experts were positive about the conceptualization and operationalization, but they also made some critical remarks. In response, items were added, removed and rephrased. Not all issues raised by the experts resulted in revisions, however, either because we could not (item overlap) or would not (for reasons of feasibility and conceptual choices).

In the second assessment of content validity, three colleague researchers not involved in the project sorted, independently of each other, the items of the MSPP to the hypothesized indices. The sortings were compared, and the intended ordering and differences discussed. No revisions were necessary.

Semi-final version of the MSPP

The semi-final version of the MSPP consisted of 26 social participation items in four indices: CP (seven items), FSP (three items), ISP-A (eight items) and ISP-F (eight items). All items had the same response format: did not do this in last 4 weeks (zero times), did this less than once a week (one to three times), did this once to twice a week (four to eight times), did this more than twice a week (nine times or more). Two types of scores could be calculated for each index: diversity and frequency. Diversity scores refer to the number of items on which a respondent had a score of at least one. Frequency scores reflect the mean score of the items. In addition, the total diversity score refers to the number of indices on which a respondent had a score of at least one. Higher scores indicate more diverse or more frequent social participation. In the present study, scores were only calculated if there were no missing values in a given index. The MSPP is included in [Appendix](#).

Methods of the second field test

Sample and data collection

The semi-final version of the MSPP was sent out in two waves to a new random selection of people older than 59 years with either COPD or diabetes. They had previously been screened for a study about chronic illness and depression carried out by the School for Public Health and Primary Care of Maastricht University, the Netherlands (Delta study) [26], which had also asked consent to participate in the present study. Those included in the Delta study (criterion: minor or mild to moderate depression) were not invited to participate in the present study to prevent high respondent burden. The Delta study recruited participants through 89 family practices in the south of the Netherlands.

In wave one (October 2004), 600 questionnaires were sent out to people with either diabetes ($N = 300$) or COPD ($N = 300$). To increase response, a telephone reminder was issued after 2 weeks. Respondents who returned questionnaires with missing values were also followed up by telephone. One-third of the participants who returned the questionnaire (random selection stratified by disease) received the questionnaire again 4 weeks after their first response to assess reproducibility (wave two). A period of 4 weeks was chosen, because the time frame of the items was ‘last 4 weeks’ and we wanted to avoid partially overlapping time frames. We considered 4 weeks long enough to prevent recall bias.

Instruments

Besides the MSPP and questions about background characteristics (socio-demographics and health), the questionnaire included parts of the RAND-36 [27, 28] to measure general health perception and physical functioning. The Frenchay Activities Index (FAI) [29, 30] was included to assess construct validity. It consists of fifteen activities (scored on a four-point scale) in three subscales: leisure/work, outdoors and domestic domain. Higher scores indicate that people are more active. The FAI has been validated in a Dutch sample of stroke patients and a control group of older adults. Construct validity was acceptable and Cronbach’s alpha > 0.60 in both groups for all three subscales [29].

Analyses

Reproducibility Reproducibility of the MSPP was evaluated with intraclass correlation coefficients (ICCs) [31] and with smallest real differences at group level

($SRDs_{\text{group}}$) [32]. ICCs were computed for each index and each item separately using a two-way random effects model with absolute agreement between the scores of wave one and two [33]. ICCs are relevant if the MSPP is used for discrimination purposes and should be at least 0.70 [34]. $SRDs_{\text{group}}$ were computed for each index according to the following formula [35]:

$$SRD_{\text{group}} = \frac{SD_{\text{wave2-1}}}{\sqrt{n}} \times 1.96$$

SRD_{group} is relevant if the MSPP is used for evaluation purposes, because it indicates the magnitude of difference that may, with 95% confidence, be expected between two measurements on the same, stable group of participants (‘noise’). The SRD is expressed in the same units as the indices and should be smaller than the minimal amount of change that is considered to be important (MIC) [34]. As we do not know yet which amount of change researchers and/or patients may consider important, readers should judge the SRD levels for themselves. To facilitate interpretation, we here define the MIC as the amount of change in the mean scores if half of the sample remains stable and the other half scores one point higher on one item (frequency scores) or scores on one item more (diversity scores).

Convergent and discriminant validity To evaluate convergent and discriminant validity of the MSPP, we used the FAI [29, 30], because it is a measure for *actual* participation, like the MSPP. To our knowledge, the FAI is the only concise instrument for actual participation validated in a Dutch sample. The FAI measures the broad concept of participation, rather than *social* participation and could, therefore, be used for convergent as well as discriminant validation.

The FAI domestic domain (preparing meals, washing up, washing clothes, light housework, heavy housework) does not measure social participation, but only activities which do not involve an exchange between people (not related to MSPP indices). By contrast, the FAI leisure/work domain (social outings, pursuing hobby, outings/car rides, house/car maintenance, gainful work) covers all three characteristics of social participation: social contact, contributing to society and receiving from society (positively related to all four MSPP indices). Finally, the FAI outdoors domain (local shopping, walking outdoors, driving/bus travel, gardening, reading books) includes items which may involve social contact and receiving from society, but not contributing (positively related to MSPP consumptive participation and MSPP informal social participation, not related to MSPP formal social participation).

We hypothesized that MSPP $CP_{\text{frequency}}$ should correlate positively (Pearson correlation coefficient) with FAI leisure/work and outdoors (convergent validity), and those correlations should be higher than the correlation with FAI domestic (discriminant validity) [36], tested with Steiger's [37] (four hypotheses). The correlation with FAI domestic should be lower rather than absent, because the MSPP and FAI might correlate for other reasons, like physical functioning. We hypothesized the same for MSPP $ISP-A_{\text{frequency}}$ and $ISP-F_{\text{frequency}}$ (eight hypotheses). MSPP $FSP_{\text{frequency}}$ should correlate positively with FAI leisure/work, and this correlation should be higher than the correlations with FAI outdoors and domestic (three hypotheses). Twelve of the fifteen hypotheses should find empirical support [34].

Comparison between COPD and diabetes Reproducibility and validity analyses were carried out for COPD and diabetes separately.

Results of the second field test

Response and sample characteristics

Of the 600 questionnaires sent out in wave one, 412 (69%) were returned (206 COPD and 206 diabetes). Four weeks later, in wave two, 125 of 137 questionnaires were returned (91%). The percentage of respondents in wave one without missing values on an index was 93% for CP, 97% for FSP, 91% for ISP-A and 91% for ISP-F (before telephone follow-up). Mean age was 70 (range 60–87). More men than women participated, as a result of a skewed sex distribution in the sampling frame. General health perception and physical functioning [27, 28] were significantly worse in participants with COPD than in participants with diabetes. Co-morbidity was common in both. Table 1 presents various characteristics of participants.

Scores on the MSPP

Table 2 presents the scores on the MSPP for participants with COPD and diabetes separately. Observed scores on the MSPP covered the entire range of theoretically possible scores for all indices except $CP_{\text{frequency}}$, $ISP-A_{\text{frequency}}$ and total diversity. On this last score, the observed score range reveals that all participants engaged in at least one type of social participation as measured by the MSPP. The results further suggest that people with diabetes tended towards a more diverse and more frequent social participation than people with COPD, but differences were small. Only total diversity ($P = 0.02$) and $FSP_{\text{frequency}}$ ($P = 0.02$) were significant at 0.05 level.

Table 1 Characteristics of the study population in the validation study (as measured in wave 1)

	COPD (<i>n</i> = 206)	Diabetes (<i>n</i> = 206)
Mean age (SD)	70.5 (6.6)	70.0 (6.2)
Sex (%)		
Male	67.5	68.4
Female	32.5	31.6
Education level (%)	*	*
Primary	28.4	18.6
Secondary	61.3	63.7
Tertiary	10.3	17.6
Paid employment (%)	3.9	4.4
Living together with partner or other person(s) (%)	75.1	73.2
RAND-36 (scale 0–100, higher is better)		
Mean general health perception (SD)	46.0 (18.7)*	50.1 (18.1)*
Mean physical functioning (SD)	55.7 (27.1)*	62.6 (28.9)*
Mean duration of disease in years (SD)	18.9 (18.1)*	9.6 (8.6)*
Duration of disease in years (min–max)	0–79	0–44
Mean number of chronic conditions (SD)	2.8 (2.1)	3.0 (2.0)

* Significant difference between COPD and diabetes at 0.05 level (Pearson chi-square or one-way analysis of variance)

Reproducibility

Tables 3 and 4 show reproducibility results. Index ICCs ranged from 0.63 for $CP_{\text{frequency}}$ to 0.83 for $FSP_{\text{diversity}}$ (should be 0.70). Item ICCs were partly low, except in FSP. One might expect low ICCs to be found in particular in items referring to irregular types of participation, but this was not evident. $SRDS_{\text{group}}$ were smaller than the MICs (as they should), except for $ISP-A_{\text{frequency}}$ and $ISP-F_{\text{frequency}}$.

Convergent and discriminant validity

Convergent and discriminant validity of the MSPP were supported by the correlations between the $MSPP_{\text{frequency}}$ and the FAI, which Table 5 shows. Convergent correlations were higher than discriminant correlations, but not very high. Differences between correlations were significant except for one, which means that 14 of 15 hypotheses found significant empirical support.

Reproducibility and validity for COPD and diabetes separately

Separate analyses for COPD and diabetes suggested better reproducibility of $CP_{\text{diversity}}$, $CP_{\text{frequency}}$, $ISP-A_{\text{diversity}}$ and $ISP-A_{\text{frequency}}$ in diabetes than in COPD (e.g. $ISP-A_{\text{frequency}}$: ICC diabetes = 0.80, ICC COPD = 0.64), while

Table 2 Scores on the MSPP by disease

	Number of cases	Theoretical score range	Observed score range	COPD mean (SD)	Diabetes mean (SD)
Total diversity ^a	382	0–4	1–4	3.29 (0.80)*	3.47 (0.73)*
Diversity ^b					
CP	410	0–7	0–7	2.40 (1.61)	2.63 (1.57)
FSP	410	0–3	0–3	0.81 (0.96)	1.00 (0.99)
ISP-A	394	0–8	0–8	4.83 (2.28)	4.90 (2.14)
ISP-F	398	0–8	0–8	4.50 (2.18)	4.72 (2.08)
Frequency ^c					
CP	410	0–3	0–1.86	0.45 (0.34)	0.51 (0.37)
FSP	410	0–3	0–3	0.44 (0.61)*	0.60 (0.70)*
ISP-A	394	0–3	0–2.75	0.94 (0.61)	1.00 (0.56)
ISP-F	398	0–3	0–3	0.88 (0.57)	0.95 (0.56)

* Difference between COPD and diabetes significant at 0.05 level (one-way analysis of variance)

^a Number of indices on which respondents had a score of at least one

^b Number of items in the index on which respondents had a score of at least one

^c Mean score on the items in the index

Table 3 Reproducibility of the MSPP indices

	Index ICC (95% confidence interval)	Mean item ICC	Range item ICC	SRD group ^a	Minimal important change ^b
Diversity					
Total	0.74 (0.65–0.82)				
CP	0.68 (0.58–0.77)	–	–	0.22	0.5
FSP	0.83 (0.76–0.87)	–	–	0.11	0.5
ISP-A	0.72 (0.62–0.80)	–	–	0.28	0.5
ISP-F	0.71 (0.61–0.79)	–	–	0.30	0.5
Frequency					
CP	0.63 (0.51–0.73)	0.58	0.38–0.76	0.05	0.07
FSP	0.81 (0.74–0.86)	0.72	0.61–0.79	0.08	0.17
ISP-A	0.72 (0.62–0.79)	0.52	0.37–0.61	0.07	0.06
ISP-F	0.64 (0.52–0.73)	0.46	0.31–0.62	0.09	0.06

^a The SRDs are expressed in observed units and should be interpreted against the theoretical score ranges: 0–3 for the frequency scores, 0–7 for CP diversity, 0–3 for FSP diversity and 0–8 for ISP-A and ISP-F

^b The minimal important change (MIC) is here defined as the amount of change if half of the sample remains stable and the other half scores one point higher on one item (frequency scores) or scores on one item more (diversity scores)

ISP-F_{diversity} and ISP-F_{frequency} yielded worse reproducibility results in diabetes than in COPD (e.g. ISP-F_{diversity}: ICC diabetes = 0.62, ICC COPD = 0.79). For FSP, results were similar in diabetes and COPD.

Regarding convergent and discriminant validity, analyses for COPD and diabetes separately yielded similar results, except that fewer differences between correlations were significant due to a lower power (data not shown).

Final version of the MSPP

The results of the second field test did not cause us to change the MSPP. The final version of the MSPP is, therefore, identical to the semi-final version (see [Appendix](#)).

Conclusion and discussion

Existing instruments for participation (in the broad sense) in the field of health and disability measure its performance [38–41], frequency [30, 42] or subjective experience [42, 43]. The MSPP also measures frequency of participation, but distinguishes itself, because it can yield both frequency and diversity scores and focuses on the social aspects of participation. Furthermore, it builds on a definition of social participation of older adults with a chronic illness themselves. We first discuss the development, validity and reproducibility of the MSPP and then compare with the development and validation of two other instruments measuring frequency of participation (in the broad sense),

Table 4 Reproducibility of the MSPP items

	ICC
CP	
Organised sport or physical activity	0.69
Cultural or educational event	0.53
Eaten out	0.58
Pub, café or tearoom	0.63
Public event	0.38
Organised games afternoon or evening	0.76
Organised day trip	0.52
FSP	
Club or similar activity	0.61
Committee work	0.79
Organised voluntary work	0.77
ISP-A	
Phoned, written, e-mailed or chatted	0.60
They called in to see you	0.50
You called in to see them	0.57
Away from home, considerable physical effort	0.51
Away from home, little physical effort	0.49
Offered practical help	0.37
Given advice or tips	0.61
Someone to talk to	0.53
ISP-F	
Phoned, written, e-mailed or chatted	0.62
They called in to see you	0.59
You called in to see them	0.52
Away from home, considerable physical effort	0.50
Away from home, little physical effort	0.40
Offered practical help	0.37
Given advice or tips	0.31
Someone to talk to	0.39

Table 5 Convergent and discriminant validity of the MSPP (Pearson correlations with FAI)

MSPP	FAI Leisure/work	FAI Domestic	FAI Outdoors	
CP _{frequency}	0.41*	>*	0.15* <*	0.31*
FSP _{frequency}	0.40*	>*	0.01 >*	- 0.22*
ISP- A _{frequency}	0.48*	>*	0.14* <*	0.30*
ISP- F _{frequency}	0.39*	>*	0.22* <	0.29*

* Correlation or difference between correlations significant at 0.01 level

namely the FAI [30] and (the objective part of) the Participation Objective Participation Subjective (POPS) [42].

The development process of the MSPP did not include the use of standard techniques based on associations between items, like internal consistency and factor analysis.

We decided against these techniques because the items of the MSPP are causal variables rather than indicator variables [44]. Indicator variables reflect an underlying concept, which completely explains the correlations between the indicator variables. In this case, techniques based on associations between items are appropriate. In contrast, causal variables 'are part of the definition of what the concept being measured means. (...) if they are present (...) then the concept in question is present.' [44: 237] There is, for instance, no underlying degree of consumptive participation, which instigates people to go to the cinema. Rather, people engage in consumptive participation *because* they go to the cinema. Causal variables may be associated irrespective of the relationship with the concept they are measuring (e.g. social participation items that are impeded by fatigue). This makes techniques based on associations inappropriate, because these techniques may suggest removing items at the cost of content validity [45], or may suggest grouping items together based on other factors (e.g. fatigue) than the concept in question (social participation). We, therefore, decided not to use these techniques and instead paid close attention to content validity. The results from the first field test show that inter-item Pearson correlations were partly low or even negative, which supports our decision not to use techniques based on associations.

Content validity of the MSPP was scrutinized by experts in the area of participation research and, after amendments, tested again by other researchers. Convergent and discriminant validity were supported by correlations between the indices of the MSPP and the Frenchay Activities Index, but differences between convergent and discriminant correlations were small. One reason might be that the FAI is not an optimal match for convergent validation. The social activities in the FAI are spread in the subscales that also include daily activities. Another reason might be that circumstances like physical functioning produced correlations between the MSPP and FAI indices (convergent and discriminant).

Reproducibility of the MSPP is moderate rather than good for the purpose of discrimination, because two of nine ICCs were lower than the threshold of 0.70 (seven if using lower limit of the 95% confidence intervals). Reproducibility is good for the purpose of evaluation (SRD_{group}), but it is a limitation of the present study that the MIC was rather arbitrarily defined.

Furthermore, reproducibility of the MSPP differs for COPD and diabetes. It is unclear whether this is a limitation of the MSPP or rather a limitation of the present study. As the MSPP measures actual social participation in the last 4 weeks and the interval between waves one and two in our study was likewise 4 weeks, social participation may really have been different between waves one and two. This is not unlikely, considering that reproducibility results

of $CP_{\text{diversity}}$, $CP_{\text{frequency}}$, $ISP-A_{\text{diversity}}$ and $ISP-A_{\text{frequency}}$ were worse in COPD, which is characterized by intermittent exacerbations. In times of exacerbations, people may be forced to, or choose to, restrict social participation, causing social participation to fluctuate more in COPD than in diabetes. To test whether reproducibility really differs for COPD and diabetes, waves one and two would have to take place within the closest possible time, for example on the same day or on two consecutive days.

There might be reservations about the sample used, because the MSPP is intended to be a generic instrument for older adults with a chronic physical illness, but the development process involved only two types of chronic disease. By using only two types, we could compare measurement properties. A generic instrument should be robust across different types of disease, meaning that scores may differ, but measurement properties should be the same.

Comparison with the development of the FAI and the objective part of the POPS shows that the former was developed using factor analysis [30], while the latter, like the MSPP, was developed using methods for causal variables. These seem more appropriate for measures of observable activity, like frequency of participation [42]. Particularly, validity of the POPS was explored by comparing results with expectations about differences between groups and correlations between subscale scores. Results and expectations did not match well [42].

Regarding reproducibility, ICCs of the FAI, POPS and MSPP are similar [42, 46]. The ICCs of the POPS subscales vary considerably. The POPS authors suggest as an explanation that participation behaviours that are ‘not scheduled into an invariant behaviour’ may vary between measurements [42]. Likewise, we suggested true variability as an explanation for the differences between COPD and diabetes in reproducibility of the MSPP.

Future research should try to establish the minimal change in MSPP scores deemed important by people with a

chronic illness to facilitate the evaluation of reproducibility. This would also allow the assessment of responsiveness, which is important for evaluation purposes. Furthermore, the MSPP still needs to be tested in other patient groups. Since the items are not specific to people with COPD or diabetes, it might also be worthwhile to test the MSPP in a general population of older adults and to use it to compare healthy older adults with those with a chronic illness. Given the social participation behaviours it covers, the MSPP does not appear to be valid for use in younger age groups.

Although there are some unresolved issues, we conclude that the Maastricht Social Participation Profile is a measure for actual social participation by older adults with a chronic physical illness, which appears to have good validity and acceptable reproducibility for discrimination purposes.

Acknowledgments We would like to thank the participants in the study and those who invited them: Karianne Jonkers, Femke Lamers, Jan van Rooij, the team at the Diabetes Care Centre Maastricht and the Dutch Diabetes Organization. Furthermore, we thank Trudi van den Bos, Mieke Cardol, Ingrid Friesema, Mieke le Granse, Marike Hendriks, Femke Lamers, Tineke Schoot and Leontine van de Ven for assessing the content validity of the MSPP.

Appendix

Below is the English version of the MSPP. It was translated in a translation - back translation procedure involving two professional translators (one native English and one native Dutch) and authors GM and JvE. The Dutch version is available from the authors.

Consumptive Participation: items A1, A3-8

Formal Social Participation: items A2, A9, A10

Informal Social Participation—Acquaintances: items B1-8

Informal Social Participation—Family: items C1-8

A. UNDERTAKING ACTIVITIES

The following questions are about things you have undertaken in the past four weeks.

If a question refers to several activities, you can say how often you have done these things in total.

How to answer
 Not done in the past four weeks = **0**
 Done 1, 2 or 3 times (less than once a week) = **1-3**
 Done 4 to 8 times (once or twice a week) = **4-8**
 Done 9 times or more (more than twice a week) = **9+**

Please circle one answer for each question

HOW OFTEN IN THE PAST FOUR WEEKS...	Number of times done in the past four weeks			
	Not at all	Less than once a week	Once or twice a week	More than twice a week
1. have you taken part in organised sport or physical activity such as an exercise class or swimming session at local pool?	0	1-3	4-8	9+
2. have you taken part in a club, interest group or activity group, church or other similar activity?	0	1-3	4-8	9+
3. have you been to a cultural or educational event such as the cinema, theatre, museum, talk or course.	0	1-3	4-8	9+
4. have you eaten out?	0	1-3	4-8	9+
5. have you been out to a pub, café or tearoom?	0	1-3	4-8	9+
6. have you been to a public event?	0	1-3	4-8	9+
7. have you taken part in an organised games afternoon or evening? For instance, bingo, quiz or card games.	0	1-3	4-8	9+
8. have you been on a day trip organised by a club or society?	0	1-3	4-8	9+
9. have you carried out committee work for a club, society or other group?	0	1-3	4-8	9+
10. have you done any organised voluntary work?	0	1-3	4-8	9+

B. CONTACT WITH FRIENDS AND ACQUAINTANCES

These questions are about how often you do things with people you know. By people you know we mean friends, acquaintances, (ex) colleagues and neighbours, people who do not live with you.

If a question refers to different activities, you can say how often you have done these things in total.

How to answer
 Not done in the past four weeks = **0**
 Done 1, 2 or 3 times (less than once a week) = **1-3**
 Done 4 to 8 times (once or twice a week) = **4-8**
 Done 9 times or more often (more than twice a week) = **9+**

Please circle one answer for each question

HOW OFTEN IN THE PAST FOUR WEEKS...	Number of times done in the past four weeks			
	Not at all	Less than once a week	Once or twice a week	More than twice a week
1. have you phoned, written, e-mailed or chatted to friends or acquaintances?	0	1-3	4-8	9+

2. have any friends or acquaintances called in to see you?	0	1-3	4-8	9+
3. have you called in to see any friends or acquaintances?	0	1-3	4-8	9+
4. have you done something away from home with friends or acquaintances that required considerable physical effort, such as walking or cycling?	0	1-3	4-8	9+
5. have you done something away from home with friends or acquaintances that required little physical effort, such as taking a car trip or going out for a cup of tea or drink?	0	1-3	4-8	9+
6. have you offered any of your friends or acquaintances practical help, such as doing the shopping, giving them a lift, doing odd jobs or filling in forms?	0	1-3	4-8	9+
7. have you given any of your friends or acquaintances advice or tips?	0	1-3	4-8	9+
8. have you supported a friend or acquaintance when they needed someone to talk to?	0	1-3	4-8	9+

C. CONTACT WITH FAMILY

Perhaps you do things with the family as well as friends and acquaintances. In these questions, when we refer to family we mean all members of your family who do not live with you.

If a question refers to different activities, you can say how often you have done these things in total.

How to answer	
Not done in the past four weeks =	0
Done 1, 2 or 3 times (less than once a week) =	1-3
Done 4 to 8 times (once or twice a week) =	4-8
Done 9 times or more often (more than twice a week) =	9+

Please circle one answer for every question

HOW OFTEN IN THE PAST FOUR WEEKS ...	Number of times done in the past four weeks			
	Not at all	Less than once a week	Once or twice a week	More than twice a week
1. have you phoned, written, e-mailed or chatted to family?	0	1-3	4-8	9+
2. have any of your family called in to see you?	0	1-3	4-8	9+
3. have you called in to see anyone from your family?	0	1-3	4-8	9+
4. have you done something away from home with a family member that required considerable physical effort, such as walking or cycling?	0	1-3	4-8	9+
5. have you done something away from home with a family member that required little physical effort, such as taking a car trip or going out for a cup of tea or drink?	0	1-3	4-8	9+
6. have you offered any of your family practical help, such as doing the shopping, giving them a lift, doing odd jobs, or filling in forms.	0	1-3	4-8	9+
7. have you given a family member advice or tips?	0	1-3	4-8	9+
8. have you supported a family member when they needed someone to talk to?	0	1-3	4-8	9+

References

1. Hoeymans, N., Timmermans, J. M., De Klerk, M. M. Y., De Boer, A. H., Deeg, D. J. H., Poppelaars, J. L., et al. (2005). *Gezond actief: de relatie tussen ziekten, beperkingen en maatschappelijke participatie onder Nederlandse ouderen [Healthy and active: the relationship between disease, disability and social participation among elderly in the Netherlands]*. Bilthoven: RIVM.
2. De Boer, A. H. (Ed.). (2006). *Rapportage ouderen 2006 [Report on the Elderly 2006]*. Den Haag: SCP.
3. World Health Organization. (2001). *International classification of functioning, disability and health (ICF)*. Geneva: World Health Organization.
4. van Campen, C. (Ed.). (2006). *Gezond en wel met een beperking: Ervaren kwaliteit van leven en functioneren van mensen met langdurige lichamelijke beperkingen [Safe and sound with a disability: Experienced quality of life and functioning of people with chronic physical impairments]*. Den Haag: SCP.
5. Levasseur, M., Desrosiers, J., & Noreau, L. (2004). Is social participation associated with quality of life of older adults with physical disabilities? *Disability and Rehabilitation*, 26(20), 1206–1213.
6. Mars, G. M. J., Kempen, G. I. J. M., Mesters, I., Proot, I. M., & Van Eijk, J. T. M. (2008). Characteristics of social participation as defined by older adults with a chronic physical illness. *Disability and Rehabilitation*, 30(17), 1298–1308.
7. Avlund, K., Holstein, B. E., Mortensen, E. L., & Schroll, M. (1999). Active life in old age. Combining measures of functional ability and social participation. *Danish Medical Bulletin*, 46, 345–349.
8. Mookherjee, H. N. (1998). Perceptions of happiness among elderly persons in metropolitan USA. Perceptual and motor skill. doi: 10.2466/PMS.87.7.787-793.
9. De Klerk, M. M. Y., & Eijssink, M. (1999). Maatschappelijke participatie [Social participation]. In M. M. Y. de Klerk & J. M. Timmermans (Eds.), *Rapportage ouderen 1998* (pp. 147–184). Den Haag: SCP.
10. Dalgard, O. S., & Håheim, L. L. (1998). Psychosocial risk factors and mortality: A prospective study with special focus on social support, social participation, and locus of control in Norway. *Journal of Epidemiology and Community Health*, 52, 476–481.
11. Lackner, J. B., Joseph, J. G., Ostrow, D. G., & Eshleman, S. (1993). The effects of social support on Hopkins symptom checklist-assessed depression and distress in a cohort of human immunodeficiency virus-positive and -negative gay men. *The Journal of Nervous and Mental Disease*, 181(10), 632–638.
12. Bukov, A., Maas, I., & Lampert, T. (2002). Social participation in very old age: Cross-sectional and longitudinal findings from BASE. *Journal of Gerontology: Psychological Sciences*, 57B, P510–P517.
13. Mars, G. M. J., Proot, I. M., Janssen, P. P. M., Van Eijk, J. T. M., & Kempen, G. I. J. M. (2007). How do people with COPD or diabetes type 2 experience autonomy? An exploratory study. *Disability and Rehabilitation*, 29(6), 485–493.
14. Van Eijk, J. T. M., & De Haan, M. (1998). Care for the chronically ill: the future role of health care professionals and their patients. *Patient Education and Counseling*, 35(3), 233–240.
15. Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, 29, 65–88.
16. Belli, R. F., Schwarz, N., Singer, E., & Talarico, J. (2000). Decomposition can harm the accuracy of behavioural frequency reports. *Applied Cognitive Psychology*, 14(4), 295–308.
17. Schwarz, N. (1999). Self-reports—how the questions shape the answers. *American Psychologist*, 54(2), 93–105.
18. Porst, R. (2000). *Question wording—Zur Formulierung von Fragebogen-Fragen. ZUMA How-to-Reihe, 2*. Mannheim: ZUMA.
19. Fowler, F. J. J. (2001). Why it is easy to write bad questions. *ZUMA-Nachrichten*, 48, 49–66.
20. Rockwood, T. H., Sangster, R. L., & Dillman, D. A. (1997). The effect of response categories on questionnaire answers; context and mode effects. *Sociological Methods & Research*, 26(1), 118–140.
21. Lessler, J. T., & Forsyth, B. H. (1996). A coding system for appraising questionnaires. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 259–291). San Francisco: Jossey-Bass Publishers.
22. Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Quantitative applications in the social sciences. In M. S. Lewis-Beck (Ed) (Vols. 07–017) Newbury Park: Sage Publications.
23. Esposito, J. L., & Rothgeb, J. M. (1997). Evaluating survey data: Making the transition from pretesting to quality assessment. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 541–571). New York: Wiley.
24. Kurz, K., Prüfer, P., & Rexroth, M. (1999). Zur Validität von Fragen in standardisierten Erhebungen. Ergebnisse des Einsatzes eines kognitiven Pretestinterviews (On the validity of questions in standardised surveys. Results of a cognitive pretest interview). *ZUMA-Nachrichten*, 44(23), 83–107.
25. SPSS for Windows, Rel. 12.0.1. (2003). SPSS Inc: Chicago.
26. Lamers, F., Jonkers, C. C., Bosma, H., Diederiks, J. P., & Van Eijk, J. T. M. (2006). Effectiveness and cost-effectiveness of a minimal psychological intervention to reduce non-severe depression in chronically ill elderly patients: the design of a randomised controlled trial [ISRCTN92331982]. *BMC Public Health*, 6, 161.
27. Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 30(6), 473–483.
28. Van der Zee, K. I., & Sanderman, R. Het meten van de algemene gezondheidstoestand met de RAND-36: Een handleiding [Measuring general health with the RAND-36: a manual], Noordelijk Centrum voor Gezondheidsvraagstukken: Groningen.
29. Schuling, J., De Haan, R., Limburg, M., & Groenier, K. H. (1993). The Frenchay activities index; assessment of functional status in stroke patients. *Stroke*, 24(8), 1173–1177.
30. Holbrook, M., & Skilbeck, C. E. (1983). An activities index for use with stroke patients. *Age and Ageing*, 12(2), 166–170.
31. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
32. De Vet, H. C. W., Bouter, L. M., Bezemer, P. D., & Beurskens, A. J. H. M. (2001). Reproducibility and responsiveness of evaluative outcome measures: Theoretical considerations illustrated by an empirical example. *International Journal of Technology Assessment in Health Care*, 17(4), 479–487.
33. Nichols, D. P. (1998). Choosing an intraclass correlation coefficient. SPSS Keywords, 67.
34. Terwee, C. B., Bot, S. D. M., De Boer, M. R., Van der Windt, D. A. W. M., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42.
35. Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310.
36. Trochim, W. M. (2005). The research methods knowledge base. Online document. <http://www.socialresearchmethods.net/kb/convdisc.htm>. Accessed 4 May 2006.

37. Meng, X.-L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*(1), 172–175.
38. Post, M. W. M., de Witte, L. P., Reichrath, E., Verdonshot, M. M., Wijnhuizen, G. J., & Perenboom, R. J. M. (2008). Development and validation of IMPACT-S, an ICF-based questionnaire to measure activities and participation. *Journal of Rehabilitation Medicine*, *40*(8), 620–627.
39. Van Brakel, W. H., Anderson, A. M., Mutatkar, R. K., Bakirtzief, Z., Nicholls, P. G., Raju, M. S., et al. (2006). The participation scale: Measuring a key concept in public health. *Disability and Rehabilitation*, *28*(4), 193–203.
40. Wilkie, R., Peat, G., Thomas, E., Hooper, H., & Croft, P. R. (2005). The Keele assessment of participation: A new instrument to measure participation restriction in population studies. Combined qualitative and quantitative examination of its psychometric properties. *Quality of Life Research*, *14*(8), 1889–1899.
41. Noreau, L., Desrosiers, J., Robicaud, L., Fougereyrollas, P., Rochette, A., & Viscogliosi, C. (2004). Measuring social participation: Reliability of the LIFE-H in older adults with disabilities. *Disability and Rehabilitation*, *26*(6), 346–352.
42. Brown, M., Dijkers, M. P., Gordon, W. A., Ashman, T., Charatz, H., & Cheng, Z. (2004). Participation objective, participation subjective: A measure of participation combining outsider and insider perspectives. *The Journal of Head Trauma Rehabilitation*, *19*(6), 459–481.
43. Cardol, M., De Haan, R. J., de Jong, B. A., Van den Bos, G. A. M., & De Groot, I. J. M. (2001). Psychometric properties of the 'Impact on Participation and Autonomy' (IPA) questionnaire. *Archives of Physical medicine and Rehabilitation*, *82*(2), 210–216.
44. Fayers, P. M., & Hand, D. J. (2002). Causal variables, indicator variables and measurement scales: An example from quality of life. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *165*(2), 233–261.
45. Streiner, D. L., & Norman, G. R. (1995). *Health measurement scales: A practical guide to their development and use*. New York: Oxford University Press.
46. Turnbull, J. C., Kersten, P., Habib, M., McLellan, L., Mullee, M. A., & George, S. (2000). Validation of the Frenchay activities index in a general population aged 16 years and older. *Archives of Physical medicine and Rehabilitation*, *81*(8), 1034–1038.