

Assessment of fatigue among working people: a comparison of six questionnaires

J De Vries, H J Michielsen, G L Van Heck

Occup Environ Med 2003;**60**(Suppl I):i10-i15

See end of article for authors' affiliations

Correspondence to:
Dr J De Vries, Tilburg University, Department of Psychology and Health, PO Box 90153, 5000 LE Tilburg, Netherlands; j.devries@uvt.nl

Accepted 1 March 2003

Aims: To compare the psychometric qualities of six fatigue questionnaires in a sample of working persons.

Methods: Internal consistency and test-retest reliability, content validity, convergent validity, and the dimensionality of the fatigue instruments were explored.

Results: All scales had a satisfactory internal consistency. Furthermore, based on factor analyses and Mokken scale analyses, all scales were unidimensional and appeared to measure an identical construct. The Fatigue Assessment Scale (FAS) had the highest factor loading on the one factor solution obtained in a factor analysis of the total scores of all scales.

Conclusions: All the questionnaires were unidimensional and had good reliability and validity. The FAS was the most promising fatigue measure.

Fatigue is one of the major complaints in primary care settings.^{1,2} Also in general population studies, fatigue is commonly reported (14–22%).³ In a recent extensive study, it was found that about 25% of Dutch employees report fatigue at work.⁴ Another Dutch study showed that over one third of the recipients of work disability benefit are occupationally disabled on mental grounds.⁵ The majority of these individuals suffer from chronic job stress and burnout. The most characteristic component of burnout⁶ is emotional exhaustion, a fatigue related concept. Emotional exhaustion refers to feelings of being overextended and depleted of one's emotional and physical resources.⁷

Fatigue is defined as "an experience of tiredness, dislike of present activity, and unwillingness to continue",⁸ or as a "disinclination to continue to performing the task at hand and a progressive withdrawal of attention" from environmental demands.⁹ As a gradual and cumulative process, fatigue reflects vigilance decrement and decreased capacity to perform, along with subjective states that are associated with this decreased performance. It is a general psychophysiological phenomenon that diminishes the ability of the individual to perform a particular task by altering alertness and vigilance, together with the motivational and subjective states that occur during this transition.¹⁰ As a consequence, there is reduced competence and willingness to develop or maintain goal directed behaviour aimed at adequate performance.¹¹ This view of fatigue is used in the present study.

There is no standard way to assess fatigue. Fatigue can be measured objectively as well as subjectively. Objective fatigue measures focus on physiological processes or performance such as reaction time or number of errors.¹² Subjective ways to assess fatigue include diary studies, interviews, and

questionnaires.^{13–15} Often, questionnaires are used in large scale studies because of their shortness and self report format.

Until about 10 years ago, fatigue questionnaires for particular studies were mainly developed on an ad hoc basis. Two recent reviews^{16,17} showed that most fatigue questionnaires are developed for specific patient groups, such as patients with cancer,^{18–20} or ill persons in general.^{21–23} Little is known, however, about the applicability of these questionnaires in healthy populations, although several fatigue measures are claimed to be useful in patient populations as well as in groups of healthy persons.^{13,24} One of the few questionnaires that has been developed for use in hospital populations as well as community populations is the Fatigue Scale (FS).¹³ Furthermore, fatigue is also frequently measured using subscales of broader measures. The Emotional Exhaustion Scale in burnout questionnaires (e.g. MBI)²⁵ and the Energy and Fatigue Scale of the World Health Organisation Quality of Life assessment instrument (WHOQOL-100)²⁶ are good examples of this approach.

Before the start of the 1990s, fatigue was predominantly seen as a unidimensional construct.²⁷ Nowadays, many authors conceive of fatigue as a multidimensional construct.^{24,28} For instance, Smets and coworkers²⁴ discern five components: general fatigue, physical fatigue, reduction in activity, reduction in motivation, and mental (cognitive) fatigue. Others, for instance, Schwartz and coworkers,²³ have developed a three dimensional scale, distinguishing situation specific fatigue, consequences of fatigue, and response to rest/sleep. The two reviews mentioned^{16,17} state that multidimensional fatigue scales are seen as more comprehensive, and

Main messages

- The examined fatigue scales are reliable and valid.
- Fatigue is unidimensional in a working population.
- For a working population the FAS is the most promising fatigue measure.

Policy implications

- A distinction in types of fatigue could not be found. Hence, employees experience overall fatigue.

Abbreviations: CFS, chronic fatigue syndrome; CIS-20, Checklist Individual Strength; CIS-PA, Checklist Individual Strength-reduced Physical Activity; CIS-SEF, Checklist Individual Strength-Subjective Experience of Fatigue; CON, reduced concentration; EE, emotional exhaustion; EF-WHOQOL-100, Energy and Fatigue scale World Health Organisation Quality of Life assessment instrument; FAS, Fatigue Assessment Scale; FS, Fatigue Scale; H, scalability coefficient; MBI-EE, Maslach Burnout Inventory-Emotional Exhaustion; MBI-NL, Maslach Burnout Inventory-Netherlands; MBI, Maslach Burnout Inventory; MOT, reduced motivation; MS, multiple sclerosis; MSP, Mokken Scale analysis for polytomous items; NRS, Need for Recovery Scale; NWO, Netherlands Organisation for Scientific Work; PA, reduced physical activity; SD, standard deviation; SEF, subjective experience of fatigue; SPSS, Statistical Package for the Social Sciences; VBBA, questionnaire on perception and judgement of work; WHOQOL-100, World Health Organisation Quality of Life assessment instrument; WORC, Work & Organisation Research Centre

hence as more adequate for providing a complete description of an individual's fatigue experience.¹⁶ However, convincing empirical evidence for the multidimension assumption is still lacking.²⁹ Moreover, two studies have recently shown that fatigue is best conceived of as a unidimensional construct.^{29,30}

The aim of the present study was to determine the psychometric qualities of several types of fatigue questionnaires in a sample of working persons. The questionnaires were selected on the basis of their use in working and healthy populations, and most were part of the prescribed set of fatigue questionnaires used in a nationwide project. This was done through a strength-weakness analysis that focused on exploring and testing: (1) internal consistency and test-retest reliability; (2) content validity; (3) convergent validity; and (4) the dimensionality of the fatigue instruments. The selected questionnaires are reliable, valid, and frequently employed.

METHODS

Respondents

Participants ($n = 351$) were recruited through random telephone calls. A random digit dialing method was used to telephone potential participants. If they agreed to participate, the questionnaire was sent by post. All respondents worked at least 20 hours per week (mean 35.3 hours, SD 9.06; 25% worked 30 hours or less and 25% worked 36–40 hours). In total, 183 men (mean age 45 years, SD 8.4) and 166 women (mean age 43 years, SD 9.5) participated (total response 48%). Gender was unknown for two respondents. Twenty seven per cent of the respondents were single, while 638 persons (73%) were married or living with a partner. Forty six per cent ($n = 399$) had a college education. Lower educated people were somewhat underrepresented and highly educated persons slightly overrepresented. However, this is not uncommon for this kind of study.³¹ With respect to gender, marital status, and age, the data are representative for the Dutch working population.⁴ The branch that the participants worked in were: industry/agriculture ($n = 32$), construction ($n = 21$), trade/repairs/hotels ($n = 33$), transport ($n = 8$), financial services ($n = 33$), care sector ($n = 71$), other services ($n = 37$), public sector (government) ($n = 39$), education ($n = 38$), and unknown ($n = 13$). Concerning being ill, the question was asked: were you ill during the past week? People ($n = 25$) who were ill, indicated that they were having a cold ($n = 13$), or had a health problem such as back pain, asthma, or a chronic illness. These individuals were not excluded, because the illness was not severe enough that they had to stop working for more than a week.

Questionnaires

The Checklist Individual Strength-20 (CIS-20)¹⁵ consists of 20 statements and provides a total fatigue score and scores for four components of fatigue: subjective experience of fatigue (SEF; eight items), reduced concentration (CON; five items), reduced motivation (MOT; four items), and reduced physical activity level (PA; three items). Respondents use a seven point rating scale (1, yes, that is true, to 7, no, that is not true). The reliability of the CIS is good.¹⁵ Furthermore, the CIS yielded different scores for CFS patients, multiple sclerosis (MS) patients, and patients with abdominal pain. Moreover, the subscales of the CIS correlated significantly with comparable scales.¹⁵ A total score above 76 is considered high. Although the CIS was developed for CFS patients, the questionnaire is claimed to be also appropriate for healthy populations.³² In a number of recent studies among working persons only the total CIS score has been used, while in other investigations one or more subscales have been employed. In the present study, we evaluated the total score as well as subscale scores in order to provide a complete picture concerning the CIS.

The Emotional Exhaustion subscale (EE scale) of the Dutch version of the Maslach Burnout Inventory (MBI²⁵; MBI-NL⁶),

comprises five items, each with a seven point rating scale ranging from 1, never, to 7, always. The scale has well established validity and high internal consistency.⁶ A score above 2.20 is considered high.

The Energy and Fatigue subscale from the World Health Organisation Quality of Life assessment instrument (EF-WHOQOL-100²⁶; Dutch version³³) contains four items. Answers are given on a fivepoint Likert scale (1, never, to 5, always): two positively phrased items using the term "energy" and two negatively phrased featuring the word "fatigue". This scale has been found to have a good reliability and to have excellent convergent validity.³⁴

The 11 item Fatigue Scale (FS¹³; Dutch translation³⁵) distinguishes mental fatigue (four items), describing cognitive difficulties, and physical fatigue (seven items). This measure uses a five point rating scale (1, never, to 5, always). It is also possible to calculate a total fatigue score. The scale was found to be both reliable and valid¹³ and has shown sensitivity to treatment changes.³⁶

The Need for Recovery scale (NRS) from the Questionnaire on Perception and Judgement of Work (VBBA³⁷) is designed to measure the short term effects of a day of work. The 11 items are rated on a dichotomous scale, Yes-No. Reliability and validity of the scale are good.³⁷

The 10 item Fatigue Assessment Scale (FAS²⁹) is a new fatigue scale that was developed in large samples of the Dutch working and general population. The items were selected from an initial item pool consisting of 40 items taken out of existing fatigue questionnaires and represent physical (five items) and mental fatigue (five items). Despite this, based on factor analyses and Mokken Scale analyses, the FAS is considered unidimensional and consequently, only a total score is calculated. The instruction of the FAS is directed at how a person usually feels. The five point rating scale varies from 1, never, to 5, always. Reliability and validity appear to be good. The FAS has a reliability of 0.90 and does not measure emotional stability or depression.²⁹

Statistical procedure

Means, standard deviations, and Cronbach's alpha coefficients were calculated for each (sub)scale. Cronbach alpha is a general formula for estimating the reliability of a test by looking at inter-item consistency. It measures how well a set of items measures a single unidimensional latent construct. It is mathematically equivalent to the average of all possible split-half estimates of reliability. Pearson correlations as well as partial correlations, controlling for overlap in items, were used to examine the associations between all questionnaires (convergent validity). In order to scrutinise the dimensionality of fatigue, exploratory principal components analyses were carried out at different levels: (1) the items per questionnaire (this is also used to establish content validity); (2) the pooled items taken from all questionnaires involved; (3) all subscale scores; and (4) all total scores. The scree test³⁸ was used to identify the number of factors. This graphical method involves inspection of the plotted eigenvalues, representing the variances extracted by the factors, against these factors and the detection of discontinuity in the slope of the plotted points. Subsequently, the dimensionality of fatigue was examined with Mokken Scale Analysis,^{39,40} but only for the analyses involving single item scores because of the fact that for Mokken Scale analysis only single item scores can be used, not summated scores. For our analyses we used the computer program Mokken Scale analysis for Polytomous items (MSP⁴¹). This program uses cluster analysis techniques for selecting unidimensional subscales from larger sets of items. Each subscale is selected to optimise the scale H for the subset of items selected (the scale H is a weighted mean of the item pair Hs). For reliably ordering persons on a (sub)scale, the scale H has to be at least 0.3 (default in MSP⁴¹). However, higher values are

Table 1 Mean, standard deviation, and reliability coefficients of the (sub)scales

(Sub)scale	Number of items	Mean	SD	Alpha
CIS-Total	20	53.75	25.58	0.96
CIS-Subjective Experience of Fatigue	8	23.41	12.58	0.96
CIS-Reduction of Concentration	5	12.73	7.24	0.92
CIS-Reduction of Motivation	4	10.42	5.40	0.87
CIS-Reduced Level of Physical Activity	3	7.22	4.42	0.88
MBI-Emotional Exhaustion	5	1.49	1.11	0.88
WHOQOL-Energy and Fatigue	4	10.18	2.91	0.88
FS-Total	11	19.95	5.81	0.87
FS-Mental Fatigue	4	6.89	2.03	0.72
FS-Physical Fatigue	7	13.11	4.39	0.84
Need for Recovery	11	23.59	8.00	0.91
Fatigue Assessment Scale (FAS)	10	19.26	6.52	0.90

desirable because they indicate higher measurement reliability, and a scale $H > 0.5$ is interpreted as indicative of a strong scale. The quality of individual items as contributors to reliable person ordering is guaranteed by only admitting items to a scale if the item scalability coefficient (item H ; a weighted mean of all item pairs in which the studied item figures) is at least 0.3.⁴¹ Based on recommendations,⁴² MSP was used with lower bounds of 0.0, 0.3, 0.4, and 0.5, respectively, for item selection using all 40 items. MSP is one of the few programs for item response theory analysis⁴³ that has an automated item selection procedure. All other analyses were performed using SPSS 9.0.⁴⁴ Missing values were not replaced, because missingness did not exhibit any systematic pattern.

Factor analysis uses the correlations or covariances between items, and is vulnerable to the influence of differences in the items' frequency distributions, which may produce artificial "difficulty factors".⁴⁵ Mokken Scale analysis is based on the scalability coefficient, H ,⁴⁶ that equals the ratio of the items' covariance and their maximum covariance given the items' univariate frequency distributions. This way, the effect of different frequency distributions is eliminated. Thus, Mokken Scale analysis does not produce artefacts because of differences in frequency distributions.

RESULTS

Table 1 shows mean and standard deviation of the fatigue questionnaires. Inspection of these results reveals that no excessive high or low scores were found in this sample.

Reliability

Table 1 shows that the internal consistency of the fatigue (sub)scales is satisfactory. The alphas ranged from 0.72 (Fatigue Scale-Mental Fatigue) to 0.96 (CIS total).

Content validity

Content validity is concerned with whether the content of a particular fatigue questionnaire supports the assertion that it is a measure of fatigue. Showing that the questionnaire at hand measures one area of content and not a mixture of all sorts of things contributes to its content validity. Six exploratory factor analyses were performed separately for each of the fatigue questionnaires. The scree plot of each factor analysis clearly revealed one factor. The single factor extracted from each questionnaire explained between 43% (Fatigue Scale) and 75% (EF-WHOQOL-100) of the (observed) variance. Because the FS and the CIS are assumed to reflect a multidimensional structure, additional factor analyses were done for each of these scales with a forced number of factors (two for the FS and four for the CIS). The percentage of explained variance increased to 55% and 76%, respectively. However, all CIS factor loadings were higher on the first factor in comparison to the other factors. In general, the other factor loadings

differed by 0.3 from the loading on the first factor. The loadings on the first factor ranged from 0.60 to 0.83, on the second from -0.43 to 0.31, on the third from 0.51 to 0.31, and on the fourth from 0.32 to 0.48. This was also true for the forced factor solution of the FS items. Item 11 ("Do you make slips of the tongue when speaking?") was an exception to this outcome.

Convergent validity

From a theoretical standpoint, the most important kind of validity is construct validity, reflecting the relation of the questionnaire results to the theoretical concept that the questionnaire is trying to measure. For the construct validity of a particular questionnaire convergent validity is important.⁴⁷ Convergent validity refers to the degree to which the operationalisation of a particular construct is similar to other operationalisations that it theoretically should be similar to.

The correlations, corrected for item overlap, between the (sub)scores of the fatigue questionnaires were moderate to strong, ranging from 0.43 (FS-Mental Fatigue with CIS-PA) to -0.81 (CIS-SEF with EF-WHOQOL-100) (all $p < 0.001$; see table 2). When looking at associations between subscale scores and total scores for the CIS and the FS, it appeared that the subscales correlated at least 0.78 ($p < 0.001$) with their total score.

Dimensionality

The factor structure and the scalability, using coefficient H of the six questionnaires, were explored separately. As already mentioned, when examining the content validity, the scree plots of the exploratory factor analyses revealed that all questionnaires consisted of one factor. In order to check these findings, Mokken Scale analyses were conducted.

Based on the MSP analyses, again all scales were found to be reliable and unidimensional. The scale H ranged from 0.48 (Fatigue Scale) to 0.78 (EF-WHOQOL-100), as depicted in table 3. Values between 0.4 and 0.5 usually are interpreted as "medium" results; values above 0.5 as good. Only one item, from the Fatigue Scale (FS), had to be excluded from the set of FS items, identical to the results of the factor analysis.

Exploratory factor analysis at the item level, using the total set of pooled items, yielded one factor, that explained 44% of the total variance. Moreover, the MSP analysis in which all fatigue items were taken together, showed that 60 items appeared to form one reliable scale (see table 3). In other words, not only do the questionnaires separately measure one construct, they also measure a similar construct, fatigue.

In addition, a factor analysis on all subscales of the six fatigue questionnaires also revealed one factor, explaining 66% of the variance. In the latter analysis, the FAS had the highest loading on this factor.

Separate analyses revealed that the same strong one factor solution was found, when the sample was split according to

Table 2 Correlations between the (sub)scales

(Sub)scale	1	2	3	4	5	6	7	8	9	10	11	12
1 CIS-Total	–	0.92	0.83	0.79	0.77	0.67	0.71	0.78	0.58	0.73	0.66	0.83
2 CIS-Subjective Experience of Fatigue		–	0.65	0.65	0.58	0.65	0.74	0.75	0.49	0.75	0.61	0.76
3 CIS-Reduction of Concentration			–	0.54	0.61	0.56	0.51	0.66	0.64	0.52	0.57	0.71
4 CIS-Reduction of Motivation				–	0.61	0.50	0.51	0.56	0.41	0.52	0.50	0.63
5 CIS-Reduced Level of Physical Activity					–	0.44	0.45	0.54	0.39	0.47	0.42	0.60
6 MBI-Emotional Exhaustion						–	0.61	0.71	0.51	0.66	0.80	0.78
7 WHOQOL-Energy and Fatigue							–	0.69	0.45	0.70	0.57	0.71
8 Fatigue Scale-Total								–	0.78	0.94	0.68	0.82
9 FS-Mental Fatigue									–	0.55	0.48	0.62
10 FS-Physical Fatigue										–	0.62	0.75
11 Need for Recovery											–	0.79
12 Fatigue Assessment Scale (FAS)												–

All $p < 0.001$.**Table 3** Results of Mokken Scale analyses per scale (lowerbound = 0.3)

Scale	No. items	No. subjects	H	Min (item H) to max (item H)
Checklist Individual Strength	20	338	0.56	0.45 to 0.63
MBI-Emotional Exhaustion	5	346	0.63	0.54 to 0.68
WHOQOL-Energy and Fatigue	4	340	0.78	0.75 to 0.80
Fatigue Scale (FS)	10 (item 11 removed)	342	0.48	0.38 to 0.56
Fatigue Assessment Scale (FAS)	10	338	0.53	0.39 to 0.64
Need for Recovery (VBBA)	11	343	0.54	0.45 to 0.62
All items	60 (item 11 FS removed)	303	0.50	0.32 to 0.59

H, scalability coefficient; item H, item scalability coefficient.

gender and age. The same results were also obtained when only the total scores of the four scales, ignoring subscales, were used (78% of the variance explained).

When looking more closely at the relation between the demographic variables gender, age, and education level and the scores on the fatigue questionnaires, only a very few associations emerged. The EF-WHOQOL-100 scale revealed a significant difference between men and women ($t = -2.267$, $p = 0.024$), with women scoring lower (that is, have more energy). In addition, a positive relation was found between age and the EF-WHOQOL-100 scale ($r = 0.12$, $p = 0.037$). Older participants were more fatigued. Finally, educational level was related to the FS-mental subscale ($F(7, 305) = 2.79$, $p = 0.008$). No other relations were found for the other fatigue (sub)scales.

DISCUSSION

All six fatigue questionnaires had a good reliability. In addition, the content validity of the unidimensional measures was good, while the multifaceted structure of the other instruments (CIS and FS) could not be replicated. The convergent validity of all measures was good. The analyses further consistently showed that the six fatigue questionnaires were unidimensional. An exploratory factor analysis with all fatigue (sub)scales showed that the FAS had the highest factor loading.

The finding that the fatigue measures used in the present study have a good internal consistency is in accordance with previous studies employing one or more of these measures. Furthermore, in line with a previous study employing these questionnaires,²⁹ the content validity of the unidimensional questionnaires—that is, the MBI-EE, EF-WHOQOL, and the FAS, was good. Previous studies had found that the CIS and FS were multidimensional measures assessing four and two aspects, respectively.^{13–15} When forcing the number of factors that these questionnaires are presumed to contain, the original factor structure could not be replicated. This might be caused by the fact that neither measure was especially developed for use in the working population. In relation to this, our

finding underlines the importance of the choice for a single cut off point for the multidimensional CIS-20, indicating a fatigue level that shows that someone is at risk for sick leave or work disability, based on the total score.⁴⁸ Our analyses reveal that such a single cut off score is indicated more than a combination of cut off points for the four dimensions.

All six fatigue measures were moderately to strongly related to each other, indicating good convergent validity. This finding coincides with a previous study that investigated the psychometric properties of these questionnaires.²⁹

Lewis and Wessely⁴⁹ assume that, when fatigue is measured with emotional, behavioural, and cognitive components, it is likely that the concept is multidimensional. This view also reflects the ideas of Smets and colleagues²⁴ and Gawron and coworkers,²⁸ who have stated that nowadays there is general agreement to measure fatigue as a multidimensional concept. The present findings, however, suggest that fatigue should be assessed as a unidimensional phenomenon, at least in a working population. Thus far, statements regarding the multidimensionality of fatigue were based predominantly on the outcomes of factor analyses with the employment of the criterion of eigenvalues greater than 1.0 as indicator for the choice of the number of factors.^{13–50–51} However, this particular criterion greatly overestimates the number of factors and often causes factors to split into bloated specifics.^{32–33} In the analysis of all fatigue items the percentage of the variance explained was 44%. Additional factors do not increase the percentage substantially. Based on the scree plot this is the best exchange between explained variance and adding more factors. Other studies have used confirmatory factor analyses to examine the dimensionality of fatigue,^{24–34} and claim a good fit for a multidimensional model. Smets and coworkers,²⁴ however, did not examine whether a one factor solution would have fit their data equally well. In the present study, where scrutinising the dimensionality was of prime interest, it was clearly shown that all six questionnaires were essentially unidimensional. Studts and colleagues³⁰ also found a one factor solution in data obtained with several other multidimensional fatigue questionnaires.

A possible reason why the results do not support multidimensionality could be that, compared with groups of predominantly healthy persons, patients focus more on symptoms and, therefore, distinguish more aspects of fatigue. Fatigue may be unidimensional for non-patient groups and multidimensional for patients. When looking at scores of chronic fatigue syndrome patients and multiple sclerosis patients on the CIS subscales, these patients did score higher on all subscales. However, the differences in mean scores between those patient populations and the population in the present study were systematically the same for each subscale—that is, the patients did not show a different pattern of scores on the various subscales. Moreover, Studts and colleagues³⁰ found no difference in dimensionality between chronic pain patients and healthy controls. Hopefully, the outcomes of this study will reopen the discussion about the dimensionality of fatigue.

For a number of questionnaires used in the present study norm scores are available to indicate problematic fatigue. With regard to the CIS, the cut off score for the total score is 76. In the present study 17% had high fatigue scores according to this cut off point. On the MBI-EE scale 19.4% scored high or very high, which indicates that fatigue is a severe health problem. In contrast, 8.9% scored very low and 21.5% low on the MBI-EE scale. The percentages for the other questionnaires are quite similar. Moreover, the response categories were all used. No answer was never mentioned. This indicates that the group of participants was not biased in one direction—that is, the non-responders are mainly non-fatigued individuals or mainly fatigued individuals.

With regard to the demographic variables age, gender, and educational level, hardly any associations were found with fatigue scores. Although this is in contradiction to the observation by Lewis and Wessely,⁴⁹ who claimed that women reported two or three times more fatigue than men, it was in line with the results of De Rijk and colleagues.⁵⁵ A possible explanation for this phenomenon might be that the studies cited by Lewis and Wessely often measured fatigue with only a single item or scale and/or a dichotomous response format. The finding that different age groups reported similar fatigue experiences might be explained by the healthy worker effect: the phenomenon that people who stay healthy are able to work until their retirement.^{56, 57}

The present study has some limitations. For example, it was impossible to include all relevant fatigue questionnaires. Therefore, a selection of questionnaires had to be made. The six instruments that were chosen are reliable, valid, and frequently used in Western countries. To our knowledge, this selection of measures forms a good representation of the available fatigue instruments. The use of other measures might have led to different results. Furthermore, the number of questionnaires could influence the scores on the later completed items. However, the comparable percentages of severe fatigue for all measures do not support this idea.

In conclusion, all fatigue questionnaires used in the Fatigue at Work programme measure fatigue unidimensional in a reliable and valid way. The FAS is the most promising questionnaire to measure fatigue in a working population.

ACKNOWLEDGEMENTS

The present study was supported by a grant from the Netherlands Organisation for Scientific Research (NWO), Research Programme "Fatigue at Work" grant no: 580-02-204, and by WORC, research institute of Tilburg University.

Authors' affiliations

J De Vries, H J Michielsen, G L Van Heck, Department of Psychology and Health, Tilburg University, and Research Institute for Psychology and Health, Netherlands

REFERENCES

- Bates DW, Schmitt W, Buchwald D, *et al.* Prevalence of fatigue and chronic fatigue syndrome in a primary care practice. *Arch Intern Med* 1993;153:193-7.
- Bensing J, Hulsman R, Schreurs K. Vermoeidheid: Een chronisch probleem [Fatigue: A chronic problem]. *Medisch Contact* 1996;51:123-4.
- Loge JH, Ekeberg Ø, Kaasa S. Fatigue in the general Norwegian population: normative data and associations. *J Psychosom Res* 1998;45:53-65.
- CBS. *De levensituatie van de Nederlandse bevolking 1997. Deel 1: Gezondheid en arbeid [Life circumstances of the Dutch population 1997. Volume 1: Health and work]*. Netherlands: Voorburg/Heerlen, CBS, 1999.
- Houtman ILD, ed. *Trends in arbeid en gezondheid 1996 [Trends in work and health 1996]*. Amsterdam: NIA-TNO, 1997.
- Schaufeli W, Van Dierendonck D. Burnout, een begrip gemeten. De Nederlandse versie van de Maslach Burnout Inventory (MBI-NL) [Burnout, the measurement of a concept. The Dutch version of the Maslach Burnout Inventory (MBI-NL)]. *Gedrag Gezondheid* 1994;22:153-72.
- Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol* 2001;52:397-422.
- Bartley SH. The homeostatic and comfort perceptual system. *J Psychol* 1970;75:157-62.
- Brown ID. Driver fatigue. *Human Factors* 1994;36:298-314.
- Thiffault P, Bergeron J. Fatigue and individual differences in monotonous simulated driving. *Pers Individ Diff* 2002;33:1-18.
- Meijman TF, Schaufeli W. Psychische vermoeidheid en arbeid. Ontwikkelingen in de A&O-psychologie [Mental fatigue and work. Developments in work and organizational psychology]. *De Psycholoog* 1996;31:236-41.
- Åkerstedt T. Psychological and psychophysiological effects of shift work. *Scand J Work Environ Health* 1990;16:67-73.
- Chalder T, Berelowitz G, Pawlikowska T, *et al.* Development of a fatigue scale. *J Psychosom Res* 1993;37:147-53.
- Meesters C, Appels A. An interview to measure vital exhaustion. I. Development and comparison with the Maastricht Questionnaire. *Psychol Health* 1996;11:557-71.
- Vercoulen JHMM, Alberts M, Blijenberg G. De Checklist Individual Strength. *Gedragstherapie* 1999;32:131-6.
- Alberts M, Vercoulen JHMM, Blijenberg G. Assessment of fatigue. The application of the subjective feeling of fatigue in different research studies. In: Vingerhoets AJJM, ed. *Assessment in behavioral medicine*. Reading, UK: Harvard, 2001.
- Friedberg F, Jason LA. *Understanding chronic fatigue syndrome. An empirical guide to assessment and treatment*. Washington, DC: American Psychological Association, 1998.
- Hann DM, Jacobsen RB, Azzarello LM, *et al.* Measurement of fatigue in cancer patients: Development and validation of the Fatigue Symptom Inventory. *Qual Life Res* 1998;7:301-10.
- Okuyama T, Akechi T, Kugaya A, *et al.* Development and validation of the Cancer Fatigue Scale: a brief, three-dimensional, self-rating scale for assessment of fatigue in cancer patients. *J Pain Symptom Manage* 2000;19:5-14.
- Piper BF, Dibble SL, Dodd MJ, *et al.* The revised Piper Fatigue Scale: psychometric evaluation in women with breast cancer. *Oncol Nurs Forum* 1998;25:677-84.
- Alberts M, Smets EMA, Vercoulen JHMM, *et al.* "Verkorte vermoeidheidsvragenlijst": een praktisch hulpmiddel bij het scoren van vermoeidheid ["Shortened Fatigue Questionnaire": a practical aid in scoring fatigue]. *Nederlands Tijdschrift Geneeskunde* 1997;141:1526-30.
- Krupp LB, LaRocca NG, Muir-Nash J, *et al.* The Fatigue Severity Scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989;46:1121-3.
- Schwartz JE, Jandorf L, Krupp LB. The measurement of fatigue: a new instrument. *J Psychosom Res* 1993;37:753-62.
- Smets EMA, Garssen B, Bonke B, *et al.* The Multidimensional Fatigue Inventory (MFI). Psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39:315-25.
- Maslach C, Jackson SE. *The Maslach Burnout Inventory. Manual*, 2nd edn. Palo Alto, CA: Sage, 1986.
- WHOQOL group. Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment. *Psychol Med* 1998;28:551-8.
- Lee KE, Hicks G, Nino Murcia G. Validity and reliability of a scale to assess fatigue. *Psychiatr Res* 1991;36:291-8.
- Gawron VJ, French J, Funke D. An overview of fatigue. In: Hancock PA, Desmond PA, eds. *Stress, workload, and fatigue. Human factors in transportation*. Mahwah, NJ: Erlbaum, 2001:581-95.
- Michielsen HJ, De Vries J, Van Heck GL. Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale (FAS). *J Psychosom Res* 2002;54:345-52.
- Studts JL, De Leeuw R, Carlson CR. Symptom structure of fatigue: a multidimensional or unidimensional construct for behavioral medicine? *Psychosom Med* 2001;63:130.
- Saris WE, ed. *Variation in response behaviour. A source of measurement error in survey research*. Amsterdam: The Netherlands: Sociometric Research Foundation, 1988.
- Beurskens AJHM, Bültmann U, Kant I, *et al.* Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000;57:353-7.

- 33 **De Vries J**, Van Heck GL. *Nederlandse WHOQOL-100 [Dutch WHOQOL-100]*. Tilburg: Tilburg University, 1995.
- 34 **De Vries J**, Van Heck GL. The World Health Organization Quality of Life Assessment Instrument (WHOQOL-1000): validation study with the Dutch version. *Eur J Psychol Assess* 1997;**13**:164–78.
- 35 **De Vries J**. *Vermoeidheidschaal [Fatigue Scale]*. Tilburg, Netherlands: Tilburg University, 1998.
- 36 **Deale A**, Chalder T, Marks I, et al. Cognitive behavior therapy for chronic fatigue syndrome: a randomized controlled trial. *Am J Psychiatry* 1997;**154**:408–14.
- 37 **Van Veldhoven M**, Meijman T. *Het meten van psychosociale arbeidsbelasting met een vragenlijst. De Vragenlijst Beleving en Beoordeling van de Arbeid (VBBA)*. Amsterdam: NIA, 1994.
- 38 **Cattell RB**. The scree test for the number of factors. *Multivariate Behav Res* 1966;**1**:140–61.
- 39 **Mokken RJ**, Lewis C. A nonparametric approach to the analysis of dichotomous item scores. *Appl Psychol Meas* 1982;**6**:417–30.
- 40 **Sijtsma K**. Methodology review: Nonparametric IRT approaches to the analysis of dichotomous item scores. *Appl Psychol Meas* 1998;**22**:3–31.
- 41 **Molenaar IW**, Sijsma K. *User's manual. MSP5 for Windows. A program for Mokken Scale Analysis for polytomous items*. Groningen, Netherlands: ProGAMMA, 2000.
- 42 **Hemker BT**, Sijsma K, Molenaar IW. Selection of unidimensional scales from a multidimensional item bank in the polytomous Mokken IRT model. *Appl Psychol Meas* 1995;**19**:337–352.
- 43 **Van der Linden WJ**, Hambleton RK. *Handbook of modern item response theory*. New York: Springer, 1997.
- 44 **SPSS C**. *SPSS base 9.0: user's guide*. Chicago: SPSS, 1999.
- 45 **Nunnally JC**. *Psychometric theory*. New York: McGraw-Hill, 1987.
- 46 **Molenaar IW**. Nonparametric models for polytomous responses. In: Linden WJ Van der, Hambleton RK, eds. *Handbook of modern item response theory*. New York: Springer, 1997:369–80.
- 47 **Campbell DT**, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 1959;**56**:81–105.
- 48 **Bültmann U**, De Vries M, Beurskens AJHM, et al. Measurement of prolonged fatigue in the working population: determination of a cutoff point for the Checklist Individual Strength. *J Occup Health Psychol* 2000;**5**:411–16.
- 49 **Lewis G**, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992;**46**:92–7.
- 50 **Vercoulen JHMM**, Swanink CMA, Fennis JFM, et al. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994;**38**:383–92.
- 51 **Vertommen H**, Leysen J. Vermoeidheid: van onhanteerbaar symptoom tot diagnostisch waardevolle gemoedstoestand [Fatigue: from an unmanageable symptom to a diagnostic valuable state of mind]. *Tijdschrift voor de Klinische Psychologie* 1988;**18**:35–59.
- 52 **Rummel RJ**. *Applied factor analysis*. Evanston, IL: Northwestern University Press, 1970.
- 53 **Kline P**. Factor analysis and personality theory. *Eur J Pers* 1987;**1**:21–36.
- 54 **Åhsberg E**. Dimensions of fatigue in different working populations. *Scand J Psychol* 2000;**41**:231–41.
- 55 **De Rijk AE**, Schreurs KMG, Bensing JM. What is behind "I'm so tired"? Fatigue experiences and their relations to the quality and quantity of external stimulation. *J Psychosom Res* 1999;**47**:509–23.
- 56 **Fletcher AC**, Ades A. Lung cancer mortality in a cohort of English foundry workers. *Scand J Work Environ Health* 1984;**10**:7–16.
- 57 **Fox AJ**, Collier PF. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br J Prev Soc Med* 1976;**30**:225–30.