Universiteit Leiden
The Netherlands

A questionable factor structure of the multidimensional fatigue inventory in the general Dutch population

Kieffer, J.M.; Starreveld, D.E.; Boekhout, A.; Bleiker, E.M.

Citation

Kieffer, J. M., Starreveld, D. E., Boekhout, A., & Bleiker, E. M. (2021). A questionable factor structure of the multidimensional fatigue inventory in the general Dutch population. *Journal Of Clinical Epidemiology*, *137*, 266-276. doi:10.1016/j.jclinepi.2021.05.005

Version: Publisher's Version

License: Creative Commons CC BY 4.0 license

Downloaded from: https://hdl.handle.net/1887/3271000

Note: To cite this publication please use the final published version (if applicable).

ELSEVIER

Journal of Clinical Epidemiology 137 (2021) 266–276

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A questionable factor structure of the multidimensional fatigue inventory in the general Dutch population

Jacobien M Kieffer[a,1], Daniëlle EJ Starreveld[a,1], Annelies Boekhout[a], Eveline MA Bleiker[a,b,*]

[a] Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
[b] Department of Clinical Genetics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

Accepted 5 May 2021; Available online 14 May 2021

Abstract

Objective: One of the most commonly used tools to measure fatigue is the Multidimensional Fatigue Inventory (MFI). Studies into the scale structure of the MFI show discrepant findings. The objective of this study was to investigate the scale structure of the MFI in the general Dutch population.

Study design and Setting: Using data from a Dutch probability-based internet panel (n = 2512), the original 5-factor model, a 4-factor, and a 5- and 4-bifactor model of the MFI were tested with confirmatory factor analyses. Additional models were investigated using exploratory factor analysis.

Results: Results neither confirmed a 5-factor (RMSEA = 0.120, CFI = 0.933, TLI = 0.920) nor a 4-factor model (RMSEA = 0.122, CFI = 0.928, TLI = 0.917). The two bi-factor models also showed a poor fit (bi-4-factor: RMSEA = 0.151, CFI = 0.895, TLI = 0.873; bi-5-factor: RMSEA = 0.153, CFI = 0.894, TLI = 0.871). Exploratory factor analysis did not support an alternative model, but seemed to show robustness in the loading of the original *general fatigue* items.

Conclusion: Our results did not provide empirical support for a four or five (bi-)factor structure of the MFI, nor for an alternative model. The most reliable scale of the MFI seems to be the *general fatigue* scale that could be used as a general indicator of fatigue.
© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Keywords: Psychometrics; Fatigue; Factor analysis; Measurement model

What is new?

Key findings

- A psychometric evaluation of the Multidimensional Fatigue Inventory in the Dutch general population (n = 2512) did not confirm the original 5-factor structure, nor an alternative 4-factor or 5- and 4-bifactor model.

What this adds to what is known

- The MFI is an internationally widely used outcome measure to assess fatigue. The lack of a clear factor structure makes it questionable whether the MFI measures multiple dimensions of fatigue.
- The conceptual and structural issues related to the MFI raised in this study question whether the conclusions based on results on the five scales of the MFI are reliable.
- The *general fatigue* scale showed robust loadings and showed the highest correlation with a fatigue rating from 0 (no fatigue) to 10 (worst fatigue) suggesting that the *general fatigue* scale could be a good measure of fatigue.

What should change now

- When the MFI is used, the results of the scales should be interpreted with caution. We suggest to draw conclusions about fatigue on the *general fatigue* scale only.

Conflict of interest: None.

[1] both authors contributed equally to this manuscript.

* Corresponding author. E.M.A. Bleiker, Tel.: (+31) 20 512 9111, fax: (+31) 20 512 2322.

E-mail address: e.bleiker@nki.nl (E.M. Bleiker).

https://doi.org/10.1016/j.jclinepi.2021.05.005
0895-4356/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Introduction

Fatigue is a symptom that is familiar to almost all individuals. There is a high prevalence of fatigue in both the normal population [1] and in individuals with (chronic) illnesses, for example, example given cancer [2]. However, there is a lack of consensus on the definition and multidimensionality of fatigue. For example, a general definition of fatigue is: “aoverwhelming sense of tiredness, lack of energy, and a feeling of exhaustion, associated with impaired physical and/or cognitive functioning; which needs to be distinguished from symptoms of depression” [3]. This general definition ignores the current discussion on the dimensionality of fatigue. Some authors propose that fatigue can be distinguished in mental and physical fatigue [4], while others propose more than two dimensions, for example, the EORTC-FA12 [5] measures three dimensions (physical, emotional and cognitive fatigue) [5,6]. Due to the lack of consensus on the multidimensionality of fatigue, a gold standard to measure fatigue is missing.

One of the most commonly used questionnaires for fatigue in Europe is the Multidimensional Fatigue Inventory (MFI) [7]. It was developed by Smets and colleagues [6] to meet the need for a brief questionnaire that excludes somatic items (such as headache) and measures multiple dimensions of fatigue. A priori defined dimensions based on literature and patient interviews (n = 12) included: general fatigue (general remarks that reflects an individual’s functioning), physical fatigue (feeling of tiredness), reduced activity (often co-occurring with fatigue), reduced motivation (to start with new activities), and mental fatigue (cognitive symptoms related to fatigue) [6]. These dimensions were confirmed in samples of radiotherapy patients (n = 111), chronic fatigued patients (n = 357), psychology students (n = 481), medical students (n = 158), and army recruits (n = 316), using confirmatory factor analyses [6,8].

The original validation of the MFI provided evidence for the five dimensions of fatigue [6,8]. Several studies investigated the psychometric properties of the MFI. Only two studies [8,9] identified the originally proposed factor structure. Most studies reported different factor structures such as a three [10–12], a four [13–16] or a five-factor structure with different item loadings compared to the original factor structure [17–21]. Multiple studies have presented a combination of the general and physical fatigue scales [6,11,13–17,19,20,22] (see Table 1). Originally, Smets et al. [6] also reported a four factor model in which the general and physical fatigue scales were combined, but chose a 5-factor model because the separate scales of general and physical fatigue might provide additional information for other constructs associated with fatigue.

Considering these discrepant findings, the objective of this study was to further investigate the factor structure of the MFI in the general Dutch population with the aim of generating an optimal scoring algorithm. Therefore, we investigated the original five factor structure, and the alternative four factor structure (general fatigue and physical fatigue combined), and two bi-factor models, in which both the 4- and 5-factor models are modeled as hierarchical structures that include a general factor and specific domain factors.

Methods

Data source

Data collection for this paper was conducted by CentERdata, an institute for online data collection and research located at Tilburg University, the Netherlands (www.centerdata.nl). This institute coordinates the LISS (Longitudinal Internet Studies for the Social Sciences) panel [23,24]. This internet panel is a probability sample of households drawn from the population register by Statistics Netherlands. Approximately 5000 households, representative of the Dutch-speaking population living in the Netherlands, are included in this panel. Households without internet-access are loaned equipment to provide internet-access. Panel members receive a monthly invitation to complete an online questionnaire, which will take 15 to 30 minutes in total. This questionnaire is completed by one member of the household. Panel members are paid for each completed questionnaire. A full description of the recruitment of (new) panel members is described in further detail elsewhere [24].

In December 2017, CentERdata invited 3.590 randomly selected panel members to complete an online questionnaire that included questions on lifestyle (smoking, drinking), chronic disorders, cancer specific health-related quality of life (EORTC-QLQ-C30), and the MFI. These panel members were aged 16 years or older with an oversampling of 18 to 34 years and 75 years and older. After invitation, 2.544 (70.9%) individuals started with the questionnaire battery and 2512 individuals completed the battery including the MFI (70.0%). Our analyses are based on the sample that completed the total battery. Compared to nonresponders, responders were older, more often married, and more often retired (Table 2).

Ethics statement

In the Netherlands, ethical approval for questionnaire research in the general population is not required. Data collection abides the European “General Data Protection Regulation (GDPR)”. All participants gave double consent: first to participate in the LISS panel and second to receive monthly questionnaires.

Measurements

The original Dutch version of the MFI [6] was used to measure fatigue. It contains five scales; general fatigue

Table 1. Overview of validation studies of the MFI

Language	Ref	Population	Factor structure	Factor analysis	Remarks
Dutch	[5]	Patients with: cancer treated with RT (n = 111), chronic fatigue syndrome (n = 395 psychology students (n = 481), medical students (n = 158), junior physicians (before and after first practical training; n = 46), and army recruits (n = 160 and n = 156 after military training)	5 (GF, PF, MF, RA, RM)	CFA	Original validation study, participants completed 24 items.
	[11]	Patients with cancer receiving RT (n = 141)	5 (Original GF, PF, MF, RA, RM)	CFA	
	[16]	Patients with Parkinson's disease (n = 153)	4 (GF and PF combined, MF, RA, RM)	PCA	Correlations between scales, total score might be more valid as a general fatigue score.
German	[13]	Chronically critically ill patients (post-acute ICU;n = 113)	3 (GF, PF, RM)	CFA	MFI is not reliable in this sample, too many irrelevant items for individuals on the post-acute ICU
Polish	[14]	Patients with cancer (n = 340)	3 (PF, MF, RM)	PCA	No good fit to model A: fatigue as a unidimensional factor or model B: original 5 factor structure. Model C is result of post-hoc modifications
French	[18]	Patients with thyroid disease (n = 225)	4 (GF and PF combined, MF, RA, RM)	PCA, varimax	
Korean	[19]	Outpatients visiting university hospital (n = 595)	4 (GF and PF combined, MF, RA [negative phrased], RM [positively phrased])	PCA, varimax	
Brazilian-Portuguese	[20]	Survivors of Hodgkin lymphoma (n = 200)	5 (GF and PF combined, MF, RM (separated over two factors), RA)	Principal axis factoring, Varimax	
Persian	[17]	Patients with chronic hepatitis B (n = 297)	4 (PF, RA, MF, RM)	PCA	
Hindi	[12]	Patients with cancer (n = 200)	5 (Original GF, PF, MF, RA, RM)	CFA	Insignificant correlations between scales
Chinese	[15]	Patients with cancer prior to CT and last week CT (n = 385)	3 (spiritual fatigue, PF, MF)	Exploratory, Varimax	
	[21]	Patients with major depression (n = 137)	5 (physical and mental energy, lack of physical and mental energy, MF, RA, activity planning)	PCA, Varimax	Lower internal consistency compared to patients with cancer, fatigue symptoms and Parkinson's disease.
English	[22]	US adult population (CFS-like n = 292; chronically unwell n = 269; well n = 222)	5 (PF, MF, RA, RM, general/reduced motivation)	PCA, Varimax	All scales discriminated between groups
	[42]	Patients treated with dialysis (n = 470)	No reliable factor model was confirmed	CFA	Poor model fit to 5-factor, 1-factor, and bi-factor model

(*continued on next page*)

Table 1 (*continued*)

	[23]	Patients with Sjogren's syndrome (n = 34) or rheaumatoid arthritis (n = 48)	5 (GF and PF combined, MF, RA, RM separated over two factors)	PCA, Varimax
	[24]	Patients with cancer (n = 210)	5 factor structure was obtained but item loadings were not those proposed and dual loadings were seen.	PCA, Varimax
Swedish	[25]	Cancer patients receiving RT (n = 100); palliative cancer patients (n = 284); outpatients at a medical clinic (n = 145); hospital staff (n = 220)	5 (GF, PF, MF, RA, RM)	Cronbach's alpha

Abbreviations: GF, general fatigue; PF, physical fatigue; MF, mental fatigue; RA, reduced activity; RM, reduced motivation; CFA, confirmatory factor analysis; PCA, principal components analysis; RT, radiotherapy; ICU, intensive care unit; CT, chemotherapy; CFS, Chronic Fatigue Syndrome.

(items 1, 5, 12, 16), mental fatigue (items 7, 11, 13, 19), physical fatigue (items 2, 8, 14, 20), reduced motivation (items 4, 9, 15, 18) and reduced activity (items 3, 6, 10, 17). Items are scored on a 5-point scale on which the participant expressed the degree to which the statement applied to him or her (from agreement "yes, that is true" to disagreement "no, that is not true") in the previous days. Item scores are summed to create a sum score for each scale, ranging between 4 (best condition) and 20 (worst condition). Higher scores indicate more fatigue.

An additional 10-point Visual Analogue Scale (VAS) for fatigue was included. Participants were asked "if you had to mark your fatigue with a score on a scale from 1 (no fatigue at all) to 10 (worst imaginable fatigue), which score would you give your fatigue?"

Statistical analysis

Descriptive statistics were used to report the sociodemographic characteristics of the sample. Pearson correlation analyses were used to calculate the correlation between the scales of the original 5-factor structure and the VAS-fatigue score.

We evaluated the 4- and 5-factor model using confirmatory factor analysis (CFA) using the lavaan package in R [25] and the semTools package [26]. We also modeled the 4- and the 5-factor model as hierarchical structures including a general factor and specific domain factors [27]. This evaluated whether item variation in the MFI reflects variation in a single unidimensional construct or if a questionnaire is multidimensional and scales are needed [28]. This bi-factor model allows items to simultaneously load on a general factor, in our case fatigue, and on a secondary factor of a specific fatigue domain. These specific domain factors account for the residual variance between the items once the contribution to the general factor has been partialed out. All domain factors are uncorrelated and have the same conceptual footing because they all contribute to the general factor [27]. We used the diagonally weighted least squares estimator with the mean- and variance adjustment procedure [29]. A mean- and variance-adjusted scaled chi square was calculated for each model. This is the standard (normal-theory) chi square statistic divided by a scaling correction to better approximate a chi square under non-normality [30]. We also reported the comparative fit index (CFI) and the Tucker-Lewis Index (TLI) (for both, values $\geq$ 0.97 indicate a good fit, and between 0.95 and 0.97 an acceptable fit), and the Root Mean Square Error of Approximation (RMSEA) (values $<$ 0.05 indicating a good fit, and between 0.05 and 0.08 an acceptable fit) [31]. Because these goodness-of-fit statistics are derived from the models using the chi squared test, they too are scaled and become robust to non-normality [32]. All standardized factor loadings were required to be greater than 0.4 and statistically significant [33].

In case of poor model fit, rather than relying on modification indices, we subsequently carried out exploratory factor analyses (EFA). We evaluated models from one- to six factors using EFA with Geomin rotation and diagonally weighted least squares estimator in Mplus [34–36]. We again used the scaled CFI, TLI, and RMSEA as indicators of model fit. All standardized factor loadings were required to be greater than 0.4 and statistically significant. Items were considered unstable if cross-loadings were significant on another factor with a difference between the two highest loadings being smaller than 0.2 [33]. We used the Kaiser criterion and scree plot to determine the number of factors that would yield the best solution [37].

Results

Table 2 summarizes the sociodemographic and clinical characteristics of the respondents. In total, 1165 men (46.4%) and 1347 women (53.6%) with a mean age of 52.1 years (standard deviation = 18.5) completed the questionnaire. Forty percent of the responders reported no comorbidities. The top six of comorbid diseases in the past 12 months were: back pain (28.9%), high blood pressure (20.1%), arthrosis (17.7%), cancer (9.5%), asthma/chronic bronchitis/COPD (8.7%), and heart disease (8.0%). Depression was reported by 5.9% of the participants.

Table 2. Sociodemographic and clinical characteristics and fatigue scores on the MFI for the total sample (n = 2512)

	Responders (n = 2512)	Non-responders (n = 1078)[a]
Age in years (M, SD)	52.1 (18.5)	39.3 (16.3)
Sex		
Male	1165 (46.4)	469 (43.5)
Female	1347 (53.6)	609 (56.5)
Living situation		
Married (n, %)	1262 (50.2)	436 (40.4)
Not married (n, %)	1250 (49.8)	642 (59.6)
Education		
Primary education (n, %)	186 (7.4)	79 (7.4)
High school and vocational education (n, %)	1407 (56.1)	576 (53.7)
College and university (n, %)	915 (36.5)	417 (38.9)
Missing (n)	4	6
Employment		
Paid job / self-employed (n, %)	1194 (47.5)	661 (61.3)
Unemployed (n, %)	349 (13.9)	115 (10.7)
Student (n, %)	218 (8.7)	195 (18.1)
Retired (n, %)	636 (25.3)	67 (6.2)
Work disabled (n, %)	98 (3.9)	30 (2.8)
Other (n, %)	17 (0.7)	10 (0.9)
Self-reported comorbidities[b] (in past 12 months)		
0 (n, %)	914 (39.7)	
1 (n, %)	659 (28.6)	
$\geq$ 2 (n, %)	732 (31.8)	
Missing[c]	207	
Fatigue		
General fatigue (M, SD)	9.8 (4.4)	
Physical fatigue (M, SD)	8.8 (4.2)	
Reduced activity (M, SD)	9.3 (3.9)	
Reduced motivation (M, SD)	8.7 (3.6)	
Mental fatigue (M, SD)	8.3 (3.7)	
Sum score (M, SD)	44.9 (16.7)	
VAS (M, SD)	4.1 (2.3)	

Note the reported percentages refer to valid cases.
Abbreviations: M, Mean; SD, standard deviation
[a] Responders differed significantly from non-responders on age, living situation, and employment (all $P < 0.001$).
[b] Comorbidities as measured by an adapted version of the Self-Administered Comorbidity Measure [39], including heart disease, stroke, high blood pressure, lung disease, diabetes mellitus, gastric ulcer, kidney disease, liver disease, anemia or other blood disease, thyroid disease, depression, arthrosis, back pain, rheumatoid arthritis, and other medical conditions.
[c] Data of 207 responders was missing because the labels ‘yes’ or ‘no’ were not shown to responders who completed the questionnaire on their smartphone. This data was considered unreliable and not included.

The responses to the individual items of the MFI are depicted in Table 3 and show that the majority of the participants reported none to mild fatigue. Based on the VAS-fatigue, 49% of the participants reported mild fatigue (VAS 3 or lower), 31% reported moderate fatigue (VAS 4 to 6), and 20% reported severe fatigue (VAS 7 or higher) [38].

We found a strong correlation between the VAS score and general fatigue (r= 0.77). Moderate correlations were found between the VAS and the remaining scales (range: r= 0.52 to 0.65).

Confirmatory Factor Analyses (CFA)

Standardized factor loadings for the original 5-factor model and the 4-factor model are presented in Supplemental Table S1 and S2. Although, both the original 5-factor model, and the 4-factor model revealed statistically significant standardized factor loadings greater than 0.4 on all factors, both model showed a poor model fit according the fit indices (Table 4). We also observed high correlations between the factors in the original 5-factor model (rang-

Table 3. Distribution of responses on the single items of the MFI in the total sample (N = 2512)

MFI	N (%)				
	1 Yes, this is true	2	3	4	5 No, this is not true
General Fatigue					
1. I feel fit	1049 (41.8)	637 (25.4)	443 (17.6)	221 (8.8)	162 (6.4)
5. I feel tired	246 (9.8)	478 (19.0)	519 (20.7)	543 (21.6)	726 (28.9)
12. I feel rested	559 (22.3)	708 (28.2)	608 (24.2)	418 (16.6)	219 (8.7)
16. I tire easily	209 (8.3)	374 (14.9)	518 (20.6)	584 (23.2)	827 (32.9)
Physical Fatigue					
2. Physically I feel only able to do a little	139 (5.5)	216 (8.6)	333 (13.3)	457 (18.2)	1367 (54.4)
8. Physically I can take on a lot	870 (34.6)	786 (31.3)	464 (18.5)	238 (9.5)	154 (6.1)
14. Physically I feel I am in a bad condition	160 (6.4)	253 (10.1)	466 (18.6)	587 (23.4)	1046 (41.6)
20. Physically I feel I am in an excellent condition	643 (25.6)	748 (29.8)	519 (20.7)	326 (13.0)	276 (11.0)
Reduced Activity					
3. I feel very active	654 (26.0)	789 (31.4)	597 (23.8)	308 (12.3)	164 (6.5)
6. I think I do a lot in a day	664 (26.4)	665 (26.5)	658 (26.2)	303 (12.1)	222 (8.8)
10. I think I do very little in a day	144 (5.7)	308 (12.3)	481 (19.1)	571 (22.7)	1008 (40.1)
17. I get little done	116 (4.6)	251 (10.0)	491 (19.5)	629 (25.0)	1025 (40.8)
Reduced Motivation					
4. I feel like doing all sorts of nice things	922 (36.7)	781 (31.1)	497 (19.8)	209 (8.3)	103 (4.1)
9. I dread having to do things	133 (5.3)	290 (11.5)	463 (18.4)	655 (26.1)	971 (38.7)
15. I have a lot of plans	681 (27.1)	750 (29.9)	717 (28.5)	250 (10.0)	114 (4.5)
18. I don't feel like doing anything	100 (4.0)	248 (9.9)	488 (19.4)	602 (24.0)	1074 (42.8)
Mental Fatigue					
7. When I am doing something, I can keep my thoughts on it	1243 (49.5)	688 (27.4)	363 (14.5)	150 (6.0)	68 (2.7)
11. I can concentrate well	1084 (43.2)	724 (28.8)	476 (18.9)	173 (6.9)	55 (2.2)
13. My thoughts easily wander	143 (5.7)	255 (10.2)	432 (17.2)	647 (25.8)	1035 (41.2)
19. It takes a lot of effort to concentrate on things	156 (6.2)	338 (13.5)	515 (20.5)	637 (25.4)	866 (34.5)

MFI, Multidimensional Fatigue Inventory physical fatigue forming one factor, high correlations were observed between factors (ranging between 0.69-0.97; Table 5), with the highest correlations again being between the reduced motivation and reduced activity scale.

Table 4. Scaled fit indices; confirmatory factor analyses and Bi-factor analyses on the MFI.

	Original 5-factor model	5-BI factor model	4-factor model	4-BI factor model
CFI	0.933	0.895	0.928	0.894
TLI	0.920	0.873	0.917	0.871
RMSEA	0.120	0.151	0.122	0.153

Note: 4-Factor Model, model consisting of 20 indicators and four factors: general and physical fatigue combined, reduced activity, reduced motivation, and mental fatigue; 5-Factor Model, model consisting of 20 indicators and five factors: general fatigue, physical fatigue, reduced activity, reduced motivation, and mental fatigue; Bi-factor model, a hierarchical structure that includes a general factor and specific domain factors.

Abbreviations: CFI, comparative fit index; MFI, Multidimensional Fatigue Inventory; RMSEA, root mean square error of approximation; TLI, Tucker-Lewis Index.

Table 5. Between-factor correlations of the Multidimensional Fatigue Inventory.

Original 5-factor model					
	GF	PF	MF	RA	RM
GF	1				
PF	0.920	1			
MF	0.719	0.625	1		
RA	0.824	0.878	0.704	1	
RM	0.818	0.825	0.756	0.966	1
4-Factor Model					
	GPF	MF	RA	RM	
GPF	1				
MF	0.692	1			
RA	0.869	0.704	1		
RM	0.838	0.756	0.966	1	

Note: 4-Factor Model, model consisting of 20 indicators and four factors: general and physical fatigue combined (GPF), reduced activity (RA), reduced motivation(RM), and mental fatigue (MF); 5-Factor Model, model consisting of 20 indicators and five factors: general fatigue (GF), physical fatigue (PF), reduced activity (RA), reduced motivation(RM), and mental fatigue (MF).

ing between 0.63–0.97; Table 5), with the highest correlations being between the general and physical fatigue scale (r = 0.92) and the reduced motivation and reduced activity scale (r= 0.97). Similarly, for the 4-factor model with general and physical fatigue forming one factor, high correlations were observed between factors (ranging between 0.69-0.97; Table 5), with the highest correlations again being between the reduced motivation and reduced activity scale.

When modeling these models as hierarchical structures including a general factor and specific domain factors, we found a poor fit for both bi-factor models (Table 4). Additionally, results showed small non-significant factor loadings of items 6 on RA ($P = 0.458$), and of item 9 on RM ($P= 0.511$), and negative residual variances for items 1, 2, 7, and 19 when modeling the hierarchical 5-Factor Model. Similarly, results showed small non-significant factor loading of items 6 on RA ($P= 0.121$), and of item 9 on RM ($P= 0.938$), and negative residual variances for items 1, 3, 4, and 7 when modeling the hierarchical 4-factor model. This indicates identification problems suggesting the inappropriateness of both models for this data.

Exploratory Factor Analyses (EFA)

Due to the lack of evidence of an adequate model from the CFA, we further investigated the scale structure of the MFI using EFA. EFA identified a 4-factor solution, reflecting one factor combining physical and general fatigue, a mental fatigue factor, and two factors both having a combination of reduced activity and reduced motivation indicators. Table 6 shows the standardized factor loading per indicator, with the largest loading in bold. Model fit was poor to moderate (CFI = 0.965 and TLI = 0.943, RMSEA = 0.101). Factor correlations were low to moderate, ranging from 0.23 to 0.58. Although items loaded significantly on their factors, half of the items of the MFI cross-loaded significantly on other factors (see Table 6), indicating that these items are unstable. The only items appearing to be more robust in their loading are the original general fatigue items. When evaluating the factor loadings in bold, and taking the cross loading into account, we found that eight out of 10 negatively worded items (bold underlined, or bold in Table 6) tended to cluster together on Factor 4. Although less pronounced, a similar trend was found for the positively worded items (bold cursive or bold in Table 6), of which six out of 10 tended to cluster on factor 1 (see Table 6).

Discussion

The MFI has been used in numerous studies to measure multiple dimensions of fatigue, but consensus about its scale structure or scoring procedure is lacking. In this study, we were unable to replicate the original 5-factor model as proposed by Smets et al. [6], nor was there support for a 4-factor model (combining general and physical fatigue). Adding a general factor to the 5-factor and 4-factor model (ie, creating a bi-factor model) also did not yield satisfactory results. With additional explorative analyses, we were unsuccessful in identifying an alternative model.

Most other similarly conducted studies have not demonstrated empirical support for the original 5-factor structure of the MFI. Instead, models with different structures were found [10–21]. Chilcot et al. [40], also evaluated the bi-factor structure of the 5 factor model and like us, were unable to confirm it. Similar to the results of Smets et al. [8], we found correlations between the original five factors to be high. This generally indicates an overlap in variation,

Table 6. Single item (cross-)loadings on the four factor solution of exploratory factor analyses.

		F1	F2	F3	F4
General fatigue					
1. I feel fit	POS	0,620	**0,878**	0,359	0,583
5. I feel tired	NEG	0,321	**0,801**	0,551	0,448
12. I feel rested	POS	0,506	**0,763**	0,547	0,434
16. I tire easily	NEG	0,418	**0,836**	0,518	0,615
Physical Fatigue					
2. Physically I feel only able to do a little	NEG	0,456	**0,743**	0,331	**0,659**
8. Physically I can take on a lot	POS	***0,654***	**0,747**	0,229	***0,639***
14. Physically I feel I am in a bad condition	NEG	0,444	**0,777**	0,281	**0,649**
20. Physically I feel I am in an excellent condition	POS	0,636	**0,845**	0,236	0,601
Reduced Activity					
3. I feel very active	POS	**0,818**	***0,694***	0,323	0,599
6. I think I do a lot in a day	POS	***0,557***	0,300	0,069	**0,643**
10. I think I do very little in a day	NEG	0,512	0,473	0,246	**0,853**
17. I get little done	NEG	0,534	0,616	0,414	**0,852**
Reduced Motivation					
4. I feel like doing all sorts of nice things	POS	**0,802**	0,510	0,292	0,533
9. I dread having to do things	NEG	0,454	**0,598**	0,469	**0,694**
15. I have a lot of plans	POS	**0,691**	0,375	0,226	0,475
18. I don't feel like doing anything	NEG	**0,609**	0,600	0,457	**0,795**
Mental Fatigue					
7. When I am doing something, I can keep my thoughts on it	POS	***0,526***	0,370	**0,726**	0,432
11. I can concentrate well	POS	0,571	0,456	**0,820**	0,452
13. My thoughts easily wander	NEG	0,360	0,457	**0,736**	**0,546**
19. It takes a lot of effort to concentrate on things	NEG	0,362	0,440	**0,696**	**0,584**

Abbreviation: F, factor.

and brings into question whether these factors are unique and truly represent distinct domains of fatigue. We found one of the largest correlations between general and physical fatigue. Other studies found similar results where the physical and general aspects of fatigue could not be distinguished as separate domains [13–15,17,20].

In our study, we tested various factor structures for the MFI, but to no avail. Although results were highly inconsistent, other studies were able to find evidence for certain factor structures of the MFI. We have conducted our analyses on data from a sample of the general Dutch population. We argue that the factor structures found in other studies might be sample specific (ie, cancer, thyroid disease, Sjogren's syndrome, rheaumatoid arthritis, Parkinson's disease, major depression, post-polio syndrome, chronic hepatitis B, dialysis patients [Table 1]), although no consistent factor structure was proposed. In addition to the use of a heterogeneous sample from the general Dutch population, our study has by far the largest sample size. The sample sizes in most other studies were relatively small for these kinds of factor analytical approaches (Table 1). Rule of thumb dictates a bare minimum of five respondents per parameter estimated to conduct factor analysis [41]. For evaluating the original 5-factor structure of the MFI, this would require a minimum of 350 respondents. If we could assume that the items of the MFI are reliable indicators of the underlying constructs, then a smaller sample size might do. However, in the case of the MFI we would argue that the sparse data might have led to unjust inferences in the past.

The current discussion on the definition and dimensionality of fatigue might also explain the lack of evidence for a robust factor structure and discrepant findings in the literature. Originally, fatigue was originally seen as a unidimensional construct but increased research has suggested a multidimensional construct of fatigue [42]. Michielsen and colleagues [42], showed that four different fatigue assessments claiming to measure one, two or five dimensions of fatigue (excluding the MFI) all measured one unidimensional concept of fatigue. This raises questions about whether the MFI covers the concept it intends to measure. Besides the general fatigue domain, the other domains may reflect constructs that can be, but may not necessarily be influenced by fatigue (ie, the physical fatigue domain rather represents physical functioning and the mental fatigue domain, cognitive functioning). We also found that the general fatigue scale correlates highly with the VAS scale measuring fatigue, supporting the idea that the other scales of the MFI might measure concepts related to or influenced by fatigue instead of fatigue itself. However, it is important to note that the suggested unidimensionality of fatigue might be instrument-specific. Validation studies for other instruments were able to replicate different dimensions of fatigue. For example, the three dimension of fatigue assessed with the EORTC-FA12 have been successfully replicated in the general German population [43] and young adults with cancer [44].

The above pertains to a conceptual approximation of the problem with the MFI. However, (part of) the problem may lie in the semantics of the items. When developing an instrument, the intention is to develop scales that resemble unidimensional constructs. The argument for including both positively and negatively worded items is to prevent response bias, that is, to avoid a respondents' tendency to agree (acquiescence) or disagree (counter-acquiscence) with a question despite its content [45]. Although this response tendency can have an effect on the validity of a questionnaire, reversing items can also lead to mistakes and confusion and may be an even bigger threat to the validity [46]. One study showed that using the original twenty items of the MFI, with 10 positively and 10 negatively worded items, did not prevent response bias. Instead, it facilitated more mistakes than when items were posed in the same direction [47]. Moreover, the reverse wording of items in a questionnaire may inadvertently lead to two distinct factors: one for positive, and one for negative items, purely based on semantics [47]. This was also seen in our exploratory analysis, again with the exception of the general fatigue items. Other studies found similar trends [16–18]. This can be a methodological artefact, or these positively and negatively worded items may simply mirror two separate constructs on different continua. Nevertheless, this is an unintended and unwanted effect of the MFI.

In conclusion, our results did not provide empirical support for the two hypothesised measurement models for the MFI, nor for an alternative model in a large sample of the general Dutch population. Results did indicate that the general fatigue scale could be a good measure of fatigue. Nevertheless, the conceptual and structural issues surrounding the MFI which have been raised in this paper warrant considerable cognisance and caution when choosing a (multidimensional) questionnaire to measure fatigue.

Author statement

Jacobien Kieffer: conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review & editing; Danielle Starreveld: conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review & editing; Annelies Boekhout: resources, writing – review & editing; Eveline Bleiker: writing – review & editing, supervision

Acknowledgments

The LISS panel data were collected by CentERdata (Tilburg University, The Netherlands) through its MESS project funded by the Netherlands Organization for Scientific Research.

Author contributions

All authors contributed to the study conception and design. Material preparation and analysis were performed by Jacobien Kieffer and Daniëlle Starreveld. The first draft of the manuscript was written by Jacobien Kieffer and Daniëlle Starreveld and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This study was funded by the Dutch Cancer Society (Grant No. NKI 2015-7909).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jclinepi.2021.05.005.

References

[1] Nielen MMJ B-vDM, Schermer TRJ. Incidentie en prevalentie van gezondheidsproblemen in de Nederlandse huisartsenpraktijk in 2017. Nivel Zorgregistraties eerste lijn [internet] 2019.
[2] Daniëls LA, Oerlemans S, Krol AD, van de Poll-Franse LV, Creutzberg CL. Persisting fatigue in Hodgkin lymphoma survivors: a systematic review. Ann Hematol 2013;92:1023–32.
[3] Shen J, Barbera J, Shapiro CM. Distinguishing sleepiness and fatigue: focus on definition and measurement. Sleep Med Rev 2006;10:63–76.
[4] Hann D, Jacobsen P, Azzarello L, Martin S, Curran S, Fields K, et al. Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. Qual Life Res 1998;7:301–10.
[5] Weis J, Tomaszewski KA, Hammerlid E, Ignacio Arraras J, Conroy T, Lanceley A, et al. International psychometric validation of an EORTC quality of life module measuring cancer related fatigue (EORTC QLQ-FA12). J Natl Cancer Inst 2017;109.
[6] Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995;39:315–25.

[7] Schwarz R, Krauss O, Hinz A. Fatigue in the general population. Onkologie 2003;26:140–4.
[8] Smets EM, Garssen B, Cull A, de Haes JC. Application of the multidimensional fatigue inventory (MFI-20) in cancer patients receiving radiotherapy. Br J Cancer 1996;73:241–5.
[9] Chandel P, Sultan A, Khan KA, Choudhary V, Parganiha A. Validation of the Hindi version of the Multidimensional Fatigue Inventory-20 (MFI-20) in Indian cancer patients. Support Care Cancer 2015;23:2957–64.
[10] Wintermann G-B, Rosendahl J, Weidner K, Strauß B, Hinz A, Petrowski KJH, et al. Fatigue in chronically critically ill patients following intensive care - reliability and validity of the multidimensional fatigue inventory (MFI-20). Health Qual. Life Outcomes 2018;16:37.
[11] Buss T, Kruk A, Wisniewski P, Modlinska A, Janiszewska J, Lichodziejewska-Niemierko M. Psychometric properties of the Polish version of the Multidimensional Fatigue Inventory-20 in cancer patients. J Pain Symptom Manage 2014;48:730–7.
[12] Tian J, Hong JS. Validation of the Chinese version of Multidimensional Fatigue Inventory-20 in Chinese patients with cancer. Support Care Cancer 2012;20:2379–83.
[13] Elbers RG, van Wegen EE, Verhoef J, Kwakkel G. Reliability and structural validity of the Multidimensional Fatigue Inventory (MFI) in patients with idiopathic Parkinson's disease. Parkinsonism Relat Disord 2012;18:532–6.
[14] Saffari M, Naderi MK, Piper CN, Koenig HG. Multidimensional Fatigue Inventory in people with hepatitis B infection: cross-cultural adaptation and psychometric evaluation of the Persian version. Gastroenterol Nurs 2017;40:380–92.
[15] Gentile S, Delaroziere JC, Favre F, Sambuc R, San Marco JL. Validation of the French 'multidimensional fatigue inventory' (MFI 20). Eur J Cancer Care (Engl) 2003;12:58–64.
[16] Song SW, Kang SG, Kim KS, Kim MJ, Kim KM, Cho DY, et al. Reliability and validity of the Korean Version of the Multidimensional Fatigue Inventory (MFI-20): A multicenter, cross-sectional study. Pain Res Manag 2018;2018:3152142.
[17] Baptista RL, Biasoli I, Scheliga A, Soares A, Brabo E, Morais JC, et al. Psychometric properties of the multidimensional fatigue inventory in Brazilian Hodgkin's lymphoma survivors. J Pain Symptom Manage 2012;44:908–15.
[18] Chung KF, Yu BY, Yung KP, Yeung WF, Ng TH, Ho FY. Assessment of fatigue using the Multidimensional Fatigue Inventory in patients with major depressive disorder. Compr Psychiatry 2014;55:1671–8.
[19] Lin JM, Brimmer DJ, Maloney EM, Nyarko E, Belue R, Reeves WC. Further validation of the Multidimensional Fatigue Inventory in a US adult population sample. Population health metrics 2009;7:18.
[20] Goodchild CE, Treharne GJ, Booth DA, Kitas GD, Bowman SJ. Measuring fatigue among women with Sjogren's syndrome or rheumatoid arthritis: a comparison of the Profile of Fatigue (ProF) and the Multidimensional Fatigue Inventory (MFI). Musculoskeletal Care 2008;6:31–48.
[21] Meek PM, Nail LM, Barsevick A, Schwartz AL, Stephen S, Whitmer K, et al. Psychometric testing of fatigue instruments for use with cancer patients. Nurs Res 2000;49:181–90.
[22] Hagelin CL, Wengstrom Y, Runesdotter S, Furst CJ. The psychometric properties of the Swedish Multidimensional Fatigue Inventory MFI-20 in four different populations. Acta Oncol 2007;46:97–104.
[23] van de Poll-Franse LV, Horevoorts N, van Eenbergen M, Denollet J, Roukema JA, Aaronson NK, et al. The patient reported outcomes following initial treatment and long term evaluation of survivorship registry: scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. Eur J Cancer 2011;47:2188–94.
[24] Scherpenzeel A, Das M. True" longitudinal and probability-based internet panels: evidence from the Netherlands. In: Das M, Ester P, Kasmirek L, editors. Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies. Boca Raton: Taylor & Francis; 2010. p. 77–104.
[25] Rosseel Y. Lavaan: an R package for structural equation modeling and more. J Stat Softw 2012;48:1–36.
[26] Jorgensen TD, Pornprasertmanit S, Schoemann A, Rosseel Y. semTools: Useful tools for structural equation modeling. R package version 05- 32020.
[27] Reise SP, Moore TM, Haviland MG. Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores. J Pers Assess 2010;92:544–59.
[28] Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Qual Life Res 2007;16(Suppl 1):19–31.
[29] Muthén B, Asparouhov T. Latent variable analysis with categorical outcomes: multiple-group and growth modeling in Mplus. Mplus web notes 2002;4:1–22.
[30] Satorra A. Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In: Heijmans DDH, Pollock DS, Satorra A, editors. Innovations in multivariate statistical analysis: A Festschrift for Heinz Neudecker Dordrecht. Kluwer Academic Publishers; 2000. p. 233–47.
[31] Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods of psychological research online 2003;8:23–74.
[32] Satorra A, Bentler PM. Corrections to test statistics and standard errors in covariance structure analysis. In: Von Eye A, Clogg CC, editors. Latent variable analysis: Applications to developmental research. Thousand Oaks, Ca: Sage; 1994. p. 399–419.
[33] Ferguson E, Cox T. Exploratory factor analysis: a users'guide. Int J Sel Assess 1993;1:84–94.
[34] Muthén L, Muthén BO. Mplus User's Guide. 6 ed. Los Angeles, CA: Muthén & Muthén; 2010.
[35] Muthén B, du Toit S, Spisic D. Robust interference using weighted least squares and quadratic estimating equations in the latent variable modeling with categorical and continuous outcomes. Unpublished manuscript, University of California, Los Angeles, USA. 1997.
[36] Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods 2004;9:466.
[37] Cattell RB, Vogelmann S. A comprehensive trial of the scree and KG criteria for determining the number of factors. Multivariate Behav Res 1977;12:289–325.
[38] Oldenmenger WH, Pleun J, de Klerk C, van der Rijt CC. Cut points on 0–10 numeric rating scales for symptoms included in the Edmonton symptom assessment scale in cancer patients: a systematic review. J Pain Symptom Manage 2013;45:1083–93.
[39] Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The self--administered comorbidity questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Care Res (Hoboken) 2003;49:156–63.
[40] Chilcot J, Guirguis A, Friedli K, Almond M, Davenport A, Day C, et al. Measuring Fatigue using the Multidimensional Fatigue Inventory-20: a questionable factor structure in haemodialysis patients. Nephron 2017;136:121–6.
[41] Bentler PM, Chou C-P. Practical issues in structural modeling. Sociol Methods Res 1987;16:78–117.
[42] Michielsen HJ, De Vries J, Van Heck GL, Van de Vijver FJ, Sijtsma KJEJoPA. Examination of the dimensionality of fatigue. 2004;20:39-48.
[43] Hinz A, Weis J, Brähler E, Mehnert A. Fatigue in the general population: German normative values of the EORTC QLQ-FA12. Qual Life Res 2018;27:2681–9.

[44] Friedrich M, Nowe E, Hofmeister D, Kuhnt S, Leuteritz K, Sender A, et al. Psychometric properties of the fatigue questionnaire EORTC QLQ-FA12 and proposal of a cut-off value for young adults with cancer. Health Qual Life Outcomes 2018;16:1–11.
[45] Rammstedt B, Danner D, Bosnjak M. Acquiescence response styles: a multilevel model explaining individual-level and country-level differences. Pers Individ Differ 2017;107:190–4.
[46] Schriesheim CA, Hill KD. Controlling acquiescence response bias by item reversals: the effect on questionnaire validity. Educ Psychol Meas 1981;41:1101–14.
[47] van Sonderen E, Sanderman R, Coyne JC. Ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain. PLoS One 2013;8:e68967.