

## DEVELOPMENT OF A FATIGUE SCALE

TRUDIE CHALDER,\* G. BERELOWITZ,† TERESA PAWLIKOWSKA,‡ LOUISE WATTS,‡  
S. WESSELY,\* D. WRIGHT§ and E. P. WALLACE||

(Received 8 January 1992; accepted in revised form 15 July 1992)

**Abstract**—A self-rating scale was developed to measure the severity of fatigue. Two-hundred and seventy-four new registrations on a general practice list completed a 14-item fatigue scale. In addition, 100 consecutive attenders to a general practice completed the fatigue scale and the fatigue item of the revised Clinical Interview Schedule (CIS-R). These were compared by the application of Relative Operating Characteristic (ROC) analysis. Tests of internal consistency and principal components analyses were performed on both sets of data.

The scale was found to be both reliable and valid. There was a high degree of internal consistency, and the principal components analysis supported the notion of a two-factor solution (physical and mental fatigue). The validation coefficients for the fatigue scale, using an arbitrary cut off score of 3/4 and the item on the CIS-R were: sensitivity 75.5 and specificity 74.5.

### INTRODUCTION

FATIGUE is both a ubiquitous symptom and is difficult to define. Some definitions attempt to identify the source of the fatigue e.g. muscle dysfunction, while others take a behavioural view, treating it in terms of performance decrements. From a physiological perspective 'weakness is a diminished ability of rested muscle to exert maximal force, whilst fatigue is a loss of maximal force-generating capacity that develops during muscular activity' [1].

Perception of fatigue is as relevant clinically as fatiguability. Many synonyms are used to describe fatigue and worded differently it means different things to different people. For example 'exhaustion' is less common than 'general fatigue' [2], while 'feeling tired all the time' is more common than 'feeling weak' [3]. Fatigue can be represented as a single phenomenon or discrete variable, but it is probably more appropriate to view it as a continuous dimension, that is experienced as a subjective internal feeling.

Today, fatigue remains as difficult to measure as it was in 1921 when Muscio stated 'that the term be absolutely banished from scientific discussion, and consequently that attempts to obtain a fatigue test be abandoned' [4]. This sentiment was reiterated by Forbes in 1943 [5]. More recently attempts have been made to measure fatigue objectively and subjectively. Fatiguability as an objective inability to sustain power can be measured electrophysiologically, but it is not necessarily related to the subjective sensation of fatigue [6]. Several attempts have been made to produce scales that measure both perception and severity of fatigue, without general acceptance [7,8].

---

\*Academic Department of Psychological Medicine, King's College Hospital, Denmark Hill, Camberwell, London SE5 9RS, U.K.

†Queen Mary's University Hospital, Roehampton Lane, London SW15 5PN, U.K.

‡General Practice, 2 Scarsdale Place, Kensington, London W8 5SX, U.K.

§Department of Microbiology, Charing Cross and Westminster Medical School, Fulham Palace Road, London W6 8RF, U.K.

||St. Mary's Hospital Medical School, London W2 1PG, U.K.

Address correspondence to Trudie Chalder.

Lee *et al.* [9] developed a reliable 18-item visual analogue scale (VAS) to measure subjective symptoms of fatigue including items relating to energy. The advantages of using VAS scales are considerable [10], but there appears to be sufficient evidence that other methods may yield more precise measurement and increased levels of satisfaction among respondents. Krupp *et al.* [11] developed a nine-item fatigue severity scale for patients with a clearly defined neurological disorder. However, there was a preponderance of items which related to the behavioural consequences of fatigue rather than symptoms. The Piper Fatigue Scale measures multidimensional aspects of fatigue but because of its length is impractical for routine use [12].

Chronic fatigue syndrome, sometimes known as post viral syndrome or myalgic encephalomyelitis (ME) is a condition which has recently received much attention from the media and professionals alike. This renewed interest is associated with aetiological, linguistic and methodological difficulties in investigation. Pleas have been proffered for more precise, operational case definitions and for better methodological studies of the epidemiology and symptomatology of fatigue [13]. As a consequence the need for a reliable valid instrument to measure the severity of fatigue has been highlighted. Lloyd and colleagues [14] devised a long questionnaire which accommodates the different symptoms associated with CFS and Ray *et al.* [15] describe a multidimensional measure which incorporates the diverse symptoms linked with the illness. We now describe a short scale which can be used in both hospital and community populations and investigate its reliability and validity.

## METHOD

### *Design of the scale*

A list of 14 questions were generated by various experts in the field to reflect physical and mental fatigue (Appendix). Symptoms not related specifically to fatigue but which are associated with CFS were not included as the intention was to produce a scale that measured fatigue specifically.

Rating scale items were used in which responses lay along a continuum, in the same way as the General Health Questionnaire [16]. Four options were used 'better than usual', 'no more than usual', 'worse than usual', 'much worse than usual'.

Scoring of the questionnaire was carried out using a bimodal response system, now referred to as the GHQ method. This is a simple method of scoring and has the advantage of eliminating errors due to 'end users' and 'middle users'. Alternatively a 'Likert score' can be used where weights are assigned to each position [17].

The scale was based on a similar questionnaire which has been applied to a hospital-based case control study [18], and a study designed to measure response to treatment [19]. These studies showed that the scale had good face validity and was sensitive to change. A shorter modified version was used by David and colleagues [13] in a community study looking at the prevalence and associations of symptoms of fatigue. They removed all of the mental fatigue items, except the item related to concentration.

### *Subjects*

Two-hundred and seventy-four new registrations at a general practice in Kensington were asked to complete the 14-item fatigue scale. These subjects were between the ages of 18 and 45 yr.

In addition, 100 consecutive attenders to the same practice for any reason were asked to fill in the 14-item fatigue scale. They were also seen by the practice nurse and asked the relevant question on fatigue as listed in the Revised Clinical Interview Schedule (CIS-R) [20]. This is a recent modification of the CIS [21], a commonly used psychiatric interview for community and primary care. The question required an affirmative answer to one of two searching questions ('Have you noticed that you've been getting tired recently?' and 'Have you felt that you've been lacking in energy?'). If a positive answer was elicited, four supplementary questions were asked:

1. On how many days have you felt tired or lacking in energy during the past week? (Score 1 for greater than or equal to 4 days.)

2. Have you felt tired/lacking in energy for more than 3 hr on any day in the past week? (Score 1 for greater than 3 hr.)

3. Have you felt so tired that you've had to push yourself to get things done during the past week? (Score 1 for yes.)

4. Have you felt tired or lacking in energy when doing things you enjoy during the last week? (Score 1 for yes.)

This gives a total fatigue score from 0–4. Scores of 2 or above are regarded as fatigued.

#### Analyses

Statistical procedures were performed using the Statistical Package for Social Scientists (SPSS/PC). A principal components analysis [22] was carried out on the 274 new registrations and the 100 consecutive attenders. This is a method used to find a combination of variables that adequately explain the overall observed variation and reduces the complexity of the data. The analysis will yield as many components as items, the first component accounting for the largest proportion of the variance amongst the items, the second component accounting for the second largest proportion that is independent of the first component and so on. It was followed by varimax rotation to aid interpretation of the results.

The homogeneity or internal consistency of the scale was checked by two methods. First, Cronbach's alpha [23] was calculated for all items, taking out different items one at a time. Second, split half reliability was calculated by randomly dividing the items of the scale into two sub-scales and correlating them with each other.

Relative Operating Characteristics (ROC) analysis was used to define a cut-off score on the fatigue scale. It assesses the discriminating ability of an instrument across the total spectrum of morbidity. A R.O.C curve is obtained by plotting sensitivity, the proportion of true cases identified by the scale, against the false positive rate, for all cut-off points on the scale. The curve is thus able to illustrate the ability of the scale to discern between 'cases' and 'non-cases', demonstrated by a  $2 \times 2$  table. It is the area under the ROC curve which assesses the discriminating ability of the instrument. If a questionnaire is unable to discriminate between 'cases' and 'non-cases', then the sensitivity will be the same as the specificity for all cut-off points, and the curve will be on the diagonal i.e. a straight line. Therefore the more sensitive the scale the greater the area under the ROC curve where 1 = total agreement.

## RESULTS

Results are first given for the 274 new registrations and second for the 100 consecutive attenders. Principal components analysis suggested a two-dimensional solution, with the first principal component accounting for 44.2% of the variance and the other principal components descending in order from 12.1 to 1.8%. The first two principal components therefore accounted for 56.3% of the variance. The pattern of coefficients in the first two eigen vectors supports the idea that the scale consists of two constructs: physical and mental fatigue (Table I). After rotation, the physical fatigue items had higher factor loadings than the mental fatigue items, confirming the notion of a two-factor solution (Table II).

TABLE I.—RESULTS OF THE PRINCIPAL COMPONENTS ANALYSIS CARRIED OUT ON 274 SUBJECTS USING THE FATIGUE SCALE

Variable	Factor	Eigenvalue	Pct of Var	Cum Pct
B	1	6.19459	44.2	44.2
C	2	1.69318	12.1	56.3
D	3	0.99874	7.1	63.5
E	4	0.86664	6.2	69.7
F	5	0.70155	5.0	74.7
G	6	0.61327	4.4	79.1
H	7	0.54154	3.9	82.9
I	8	0.47656	3.4	86.3
J	9	0.41223	2.9	89.3
K	10	0.39369	2.8	92.1
L	11	0.30465	2.2	94.3
M	12	0.29462	2.1	96.4
N	13	0.26059	1.9	98.2
O	14	0.24814	1.8	100.0

TABLE II.—RESULTS OF THE  
PRINCIPAL COMPONENTS ANALYSIS  
CARRIED OUT ON 274 SUBJECTS USING THE  
FATIGUE SCALE (AFTER ROTATION)

Factor 1	Factor 2
0.77072	0.23677
0.82670	0.02653
0.68675	0.19619
0.52066	0.34611
0.66156	0.19775
0.75452	0.28010
0.64494	0.21431
0.72279	0.29594
0.43201	0.68779
0.29486	0.74073
0.08970	0.81218
0.08467	0.75735
0.25431	0.72110
0.31505	0.47803

Cronbach's alpha [10] was calculated for all items, and after taking out different items one at a time. The result was not significantly changed by this (range 0.88–0.90).

Split half reliability was also calculated. The Pearson product–moment correlation coefficient between the two halves of the questionnaire is used in the Spearman–Brown formula to give an estimate of the reliability for the whole questionnaire. With seven items in part 1 and seven items in part 2 the results were 0.8613 and 0.8466 respectively.

A principal components analysis followed by varimax rotation was also carried out on the 100 consecutive attenders. Again, a two-dimensional solution was suggested, with the first principal component accounting for 38.1% of the variance, and the other principal components descending in order from 16.7 to 1.1%. The first two components accounted for 54.8% of the variance and as with the 274 consecutive attenders the pattern of coefficients in the first two eigen vectors supports the notion of a two factor solution (Table III).

TABLE III.—RESULTS OF THE PRINCIPAL COMPONENTS ANALYSIS CARRIED OUT  
ON THE 100 CONSECUTIVE ATTENDERS

Variable	Factor	Eigenvalue	Pct of Var	Cum Pct
B	1	5.33203	38.1	38.1
C	2	2.33241	16.7	54.7
C	3	1.02675	7.3	62.1
D	4	0.94070	6.7	68.8
E	5	0.82259	5.9	74.7
F	6	0.75777	5.4	80.1
G	7	0.64317	4.6	84.7
H	8	0.53492	3.8	88.5
I	9	0.41358	3.0	91.5
J	10	0.33313	2.4	93.8
K	11	0.30278	2.2	96.0
L	12	0.22694	1.6	97.6
M	13	0.17385	1.2	98.9
N	14	0.15938	1.1	100.0

Comparing the Fatigue Scale with the fatigue question in the CIS-R, ROC analysis was used to define the best Fatigue Scale cut-off score on the 100 consecutive attenders. The results (Table IV) suggest that the optimum cut-off is 3/4 with the area under the ROC curve being 0.85. This is of course a conservative estimate as it assumes that the answer to the CIS question is the 'Gold Standard', and that the CIS is error free. Both of these assumptions are open to question.

TABLE IV.—COMPARISON OF THE FATIGUE SCALE WITH THE CIS-R (FATIGUE ITEM)

Cut off	Sensitivity	Specificity
0/1	97.7	41.8
1/2	88.8	56.4
2/3	81.4	64.9
3/4	75.5	74.5
4/5	66.6	83.6

Based on the above mentioned analysis, three items (Do you start things without difficulty but get weak as you go on?; Do you find it more difficult to find the correct word?; Have you lost interest in the things you used to do?;) were eliminated from the scale, leaving 11 items. Cronbach's alpha for this revised version was 0.8903. As the principal components analysis revealed two possible solutions Cronbach's alpha was also calculated for the two sets of items, (physical fatigue items 1–8; mental fatigue items 9–14). Results were 0.845 and 0.821 respectively.

#### DISCUSSION

Our aim was to produce a short, easy to administer scale which was both reliable and valid. Many synonyms are used to describe fatigue and this was reflected in the design of the questionnaire. It has been suggested that the shorter the scale the less reliable and valid it becomes, however, the revised 11-item scale was found to be both reliable and valid, despite its brevity. Items were chosen for their simplicity and unambiguousness. Like most symptoms, fatigue is better viewed as a dimension [24] as opposed to a category, and response options were chosen accordingly, to accommodate two different scoring methods.

The principle components analyses provided good evidence for the distinction made between the two constructs, physical and mental fatigue. This has implications for the scoring of the questionnaire. To date, a total fatigue score has been obtained by adding up all the items. However, the analyses demonstrate that it would probably be more useful to have two scores, one for physical fatigue and one for mental fatigue.

It is generally accepted that items on a scale are chosen for their homogeneity (how well items correlate with one another and the overall score). Tests of internal consistency were carried out and results showed a high level of internal reliability. As a result of the principal components analyses, measures of homogeneity were applied to both sub-scales. The results also confirmed good internal reliability.

Validity is concerned with whether the test measures what it purports to measure. Usually results of a validation study are presented by a single value for the sensitivity and specificity. ROC analysis, however, allows us to observe the discriminating

ability of the instrument across the total spectrum of morbidity. Several measures of discriminating ability can be obtained, the most useful being the area under the curve. The value was acceptable (0.845) for the fatigue scale. The specificity/sensitivity values were also fairly high for the scale, yielding few false positives, ensuring its usefulness as a case-finding mechanism.

In conclusion a brief, easy to administer self-rating fatigue scale was developed. The intended purpose of the scale is the assessment of symptom severity, the detection of fatigue cases in epidemiological studies and as a valid estimator of change. It is recommended, however, that the scale is not used alone to detect cases, but should be used as an adjunct to a thorough clinical assessment. The scale has good face validity, and reasonable discriminant validity. Although evidence of validity as an estimator of change has been established in an open-treatment trial further evidence could be obtained by using the scale before and after treatment in a controlled trial.

*Acknowledgements*—We would like to thank Graham Dunn for assisting with the interpretation of the results and Professor Hirsch for helpful comments on an earlier draft of the paper. Teresa Pawlikowska was supported by the North West Thames Regional Health Authority and Trudie Chalder was funded by The Linbury Trust.

#### REFERENCES

1. LEWIS SF, HALLER RG. Physiologic measurement of exercise and fatigue with special reference to chronic fatigue syndrome. *Rev Infectious Dis* 1991; **13**: S98–108.
2. TIBBLIN G, BENGTSSON C, FURUNER B, LAPIDUS L. Symptoms by age and sex. *Scand J Primary Hlth Care* 1990; **8**: 9–17.
3. HANNAY D. Symptom prevalence in the community. *J R Coll Gen Pract* 1978; **28**: 492–499.
4. MUSCIO B. Is a fatigue test possible? *Br J Psychol* 1921; **12**: 31–46.
5. FORBES W. Problems arising in the study of fatigue. *Psychosom Med* 1943; **5**: 155–157.
6. MAY J, KLINE P. Problems in using an adjective checklist to measure fatigue. *J Person Individ Diff* 1988; **9**: 831–832.
7. HUETING JE, SARPHATI HR. Measuring fatigue. *J Appl Psychol* 1966; **50**: 535–538.
8. MONK T. A visual analogue scale to measure global vigour and affect. *Psychiat Res* 1989; **27**: 89–99.
9. LEE KE, HICKS G, NINO MERCIA G. Validity and reliability of a scale to assess fatigue. *Psychiat Res Rep* 1991; **36**: 291–298.
10. STREINER DL, NORMAN GR. *Health Measurement Scales. A Practical Guide to Their Development and Use*. Oxford: Oxford Medical Publications, 1989.
11. KRUPP LB, LARocca NG, MUIR NASH J, STEINBERG AD. The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989; **46**: 1121–1123.
12. PIPER BF, LINDSEY AM, DODD AJ. Fatigue: The measurement of a multidimensional concept. Unpublished raw data. 1984.
13. DAVID A, PELOSI A, McDONALD E, STEPHENS D, LEDGER D, RATHBONE R, MANN A. Tired, weak, or in need of rest: fatigue among general attenders. *Br Med J* 1990; **301**: 1199–1202.
14. LLOYD AR, HICKIE I, BOUGHTON RC, SPENCER O, WAKEFIELD D. Prevalence of chronic fatigue syndrome in an Australian Population. *Med J Aust* 1990; **153**: 522–528.
15. RAY C, WEIR WRC, PHILLIPS S, CULLEN S. Development of a measure of symptoms in chronic fatigue syndrome: The profile of fatigue related symptoms (in press).
16. GOLDBERG DP, WILLIAMS P. *A Users Guide to the General Health Questionnaire*. Windsor: NFER-Nelson. 1988.
17. LIKERT RA. A Technique for the development of attitude scales. *Educat Psychol Measur* 1952; **12**: 313–315.
18. WESSELY S, POWELL R. Fatigue syndromes: a comparison of chronic 'postviral' fatigue with neuromuscular and affective disorders. *J Neurol, Neurosurg Psychiat* 1989; **42**: 940–948.
19. BUTLER S, CHALDER T, RON M, WESSELY S. Cognitive behaviour therapy in chronic fatigue syndrome. *J Neurol, Neurosurg Psychiat* 1991; **54**: 153–158.

20. LEWIS G, PELOSI AJ, ARAYA R, DUNN G. Measuring psychiatric disorder in the community: a standardised assessment for lay interviewers. *Psychol Med* 1992; **22**: 465–486.
21. GOLDBERG DP, COOPER B, EASTWOOD MR, KENWARD HB, SHEPHERD M. A standardised psychiatric interview for use in community surveys. *Br J Prevent Soc Med* 1970; **24**: 18–23.
22. CURETON EE, D'AGOSTINO RB. *Factor Analysis: An Applied Approach*. Hillsdale: Lawrence Erlbaum, 1983.
23. CRONBACH LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; **16**: 297–334.
24. GOLDBERG D, HUXLEY P. *Common Mental Disorders. A BioSocial Model*. London: Routledge, 1992.

## APPENDIX

### 14-Item fatigue scale

#### *Physical symptoms*

1. Do you have problems with tiredness?
2. Do you need to rest more?
3. Do you feel sleepy or drowsy?
4. Do you have problems starting things?
5. Do you start things without difficulty but get weak as you go on?
6. Are you lacking in energy?
7. Do you have less strength in your muscles?
8. Do you feel weak?

#### *Mental symptoms*

9. Do you have difficulty concentrating?
10. Do you have problems thinking clearly?
11. Do you make slips of the tongue when speaking?
12. Do you find it more difficult to find the correct word?
13. How is your memory?
14. Have you lost interest in the things you used to do?