The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study

KIM DELBAERE1,2,3, JACQUELINE C. T. CLOSE1,4, A. STEFANIE MIKOLAIZAK1, PERMINDER S. SACHDEV5,6, HENRY BRODATY5,6, STEPHEN R. LORD1

1Falls and Balance Research Group, Prince of Wales Medical Research Institute, University of New South Wales, Randwick, Sydney, Australia
2Department of Experimental-Clinical and Health Psychology, Faculty of Psychology and Educational Sciences, Ghent University, Belgium
3Department of Rehabilitation Sciences and Physiotherapy, Faculty of Medicine and Health Sciences, Ghent University, Belgium
4Prince of Wales Hospital, Clinical School, University of New South Wales, Randwick, Sydney, Australia
5School of Psychiatry, University of New South Wales, Prince of Wales Hospital, Randwick, Sydney, Australia
6Neuropsychiatric Institute, Prince of Wales Hospital, Randwick, Sydney, Australia

Address correspondence to: S. R. Lord. Email: s.lord@powmri.edu.au

Abstract

Objective: this study aimed to perform a comprehensive validation of the 16-item and 7-item Falls Efficacy Scale International (FES-I) by investigating the overall structure and measurement properties, convergent and predictive validity and responsiveness to change.

Method: five hundred community-dwelling older people (70–90 years) were assessed on the FES-I in conjunction with demographic, physiological and neuropsychological measures at baseline and at 12 months. Falls were monitored monthly and fear of falling every 3 months.

Results: the overall structure and measurement properties of both FES-I scales, as evaluated with item response theory, were good. Discriminative ability on physiological and neuropsychological measures indicated excellent validity, both at baseline (n = 500, convergent validity) and at 1-year follow-up (n = 463, predictive validity). The longitudinal follow-up suggested that FES-I scores increased over time regardless of any fall event, with a trend for a stronger increase in FES-I scores when a person suffered multiple falls in a 3-month period. Additionally, using receiver-operating characteristic (ROC) curves, cut-points were defined to differentiate between lower and higher levels of concern.

Conclusions: the current study builds on the previously established psychometric properties of the FES-I. Both scales have acceptable structures, good validity and reliability and can be recommended for research and clinical purposes. Future studies should explore the FES-I’s responsiveness to change during intervention studies and confirm suggested cut-points in other settings, larger samples and across different cultures.

Keywords: fear of falling, Rasch analyses, accidental falls, ageing, sensitivity to change, elderly

Introduction

Fear of falling is an important psychological factor associated with falls in older people and has been reported in both risk factor [1, 2] and intervention studies [3, 4]. It has been inconsistently defined in studies [5], and these inconsistencies have resulted in a large variability in reported prevalence of fear of falling ranging from 12 to 92% [6]. In order to maximise interpretation of findings from epidemiological and intervention studies on fear of falling, a consistent measure is necessary.
This study aimed to further evaluate the psychometric properties of the FES-I. The structure and measurement properties were evaluated with item response theory [10]. This approach is becoming the chosen approach amongst psychometricians and is arguably the best method for developing scales and questionnaires [11]. We explored convergent and predictive validity of the scale by investigating discriminative ability of the FES-I on a range of physiological and neuropsychological measures. A longitudinal follow-up allowed us to explore the stability over time and responsiveness to change of the FES-I.

**Methods**

**Participants**

A total of 500 people aged 70–90 years were recruited from a cohort of 1,037 community-dwelling people living in Sydney and participating in the Sydney Memory and Ageing Study (see Acknowledgements). Exclusion criteria were neurological, cardiovascular or major musculoskeletal impairments that precluded participants from completing assessments. Approval for the study was obtained from the University of New South Wales Human Research Ethics Committee.

**Measures**

*The FES-I*

The FES-I is a self-report questionnaire, providing information on level of concern about falls for a range of activities of daily living (Figure 1). The original questionnaire contains 16 items scored on a four-point scale (1 = not at all concerned to 4 = very concerned) [7]. The shortened questionnaire contains seven items [8]. The FES-I was assessed at baseline and then every 3 months for 12 months.

*Other measures*

Physical performance was assessed with (i) maximal isometric quadriceps strength (kg) [12], (ii) postural sway by recording displacements of the body at the level of the waist (mm) while standing on a foam mat with eyes open [12] and (iii) gait speed (in seconds) by walking 3 m, turning and returning at normal pace. The Physiological Profile Assessment (PPA) was used to gain an estimate of physiological falls risk [12]. Levels of disability were assessed using the 12-item World Health Organization Disability Assessment Schedule (WHODAS II, total score range 0–36) [13]. Quality of life was assessed using the 20-item AQOL II (total score range 0–100) [14]. Symptoms of depression were assessed using the self-report 15-item Geriatric Depression Scale (GDS, total score range 0–15, with scores ≥5 indicating possible depression) [15]. Symptoms of anxiety in the past month were assessed using the 9-item Goldberg Anxiety Scale (GAS, total score range 0–9, with scores ≥5 indicating possible anxiety) [16]. Cognitive processing performance was tested using the Trail Making Test (Trails B), which requires subjects to draw lines connecting a number of circles alternating between letters and numbers [17]. A fall was defined as ‘an unexpected event in which the participants come to rest on the ground, floor or lower level’ [18]. The number of falls in the previous year was recorded at baseline. Falls frequency during the 1-year follow-up was monitored with monthly falls diaries [19]. All participants were assessed on each measure at baseline. A total of 494 participants completed the 12-month follow-up for falls, and 463 participants were reassessed after 1 year.

**Analyses**

The questionnaire structure was evaluated by using item response theory, i.e. Rasch modelling (Winsteps©, John M. Linacre). Rasch modelling concentrates on the probability that an individual with a certain level of concern will answer each item in a given way to match that level of concern [11]. Fit statistics were used to examine how well the data from people and items met the model assumptions. The internal structure of the questionnaires was examined by factor analysis using an unrotated principal components analysis. Internal consistency (Cronbach’s alpha) was evaluated by calculating the reproducibility of the responses of each person on the questionnaires [10]. The item-respondent map shows how each participant responded on each item. This map was inspected to evaluate content representation of each item to ensure items and respondents were appropriately targeted [10].

Further analyses were performed using SPSS for Windows (Version 17, SPSS, Inc., Chicago, IL, USA). After a logarithmic transformation of the FES-I, assumptions for parametric analyses were met [20]. Pearson’s correlations were calculated to compare the different questionnaires. Sensitivity to change was assessed by investigating main and interaction effects using a mixed 2-factor Repeated Measures (RM) ANOVA with one within factor (occasion at two levels) and one between factor (falls, injurious falls). Validity of the baseline FES-I was assessed by using independent t-tests to examine between-group differences in total scores according to age, gender and a variety of falls risk factors that have previously been associated with falls and fear of falling [7, 21–25]. Convergent validity used criterion measures collected at baseline, and predictive validity used criterion measures collected at the 1-year follow-up reassessment. Cut-off scores were established for each FES-I based on examination of receiver-operating characteristic (ROC) curves.

**Results**

**Participants**

The mean age of participants was 77.4 years (standard deviation [SD] 6.08), and 279 (55.8%) were women. A small majority of participants (n = 269, 53.8%) completed high school, and 208 (41.6%) previously worked in managerial
or professional occupations. On self-rated health status using a five-point scale, 85% (n = 425) of the sample rated their health as good, very good or excellent. Thirty percent (n = 149) of the participants reported one or more falls in the previous year, and 43.6% (n = 214) reported one or more falls during the 1-year follow-up. Means and SDs for the total FES-I scores are presented in Table 1.

Table 1. Total score ranges, means, SDs and Pearson's correlations between all FES-I scores assessed in three-monthly intervals

<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>Mean ± SD</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>16-item FES-I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. FES-I at 0 months</td>
<td>16–57</td>
<td>22.6 ± 6.4</td>
<td>0.708</td>
<td>0.711</td>
<td>0.698</td>
<td>0.715</td>
</tr>
<tr>
<td>2. FES-I at 3 months</td>
<td>16–55</td>
<td>21.7 ± 6.6</td>
<td>–</td>
<td>0.780</td>
<td>0.736</td>
<td>0.701</td>
</tr>
<tr>
<td>3. FES-I at 6 months</td>
<td>16–64</td>
<td>22.8 ± 7.8</td>
<td>–</td>
<td>–</td>
<td>0.771</td>
<td>0.763</td>
</tr>
<tr>
<td>4. FES-I at 9 months</td>
<td>16–62</td>
<td>23.6 ± 8.2</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.830</td>
</tr>
<tr>
<td>5. FES-I at 12 months</td>
<td>16–64</td>
<td>23.6 ± 8.1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>7-item FES-I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. FES-I at 0 months</td>
<td>7–25</td>
<td>9.4 ± 3.0</td>
<td>0.694</td>
<td>0.694</td>
<td>0.692</td>
<td>0.683</td>
</tr>
<tr>
<td>2. FES-I at 3 months</td>
<td>7–23</td>
<td>9.2 ± 3.0</td>
<td>–</td>
<td>0.747</td>
<td>0.693</td>
<td>0.658</td>
</tr>
<tr>
<td>3. FES-I at 6 months</td>
<td>7–28</td>
<td>9.7 ± 3.3</td>
<td>–</td>
<td>–</td>
<td>0.750</td>
<td>0.730</td>
</tr>
<tr>
<td>4. FES-I at 9 months</td>
<td>7–27</td>
<td>10.1 ± 3.7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.803</td>
</tr>
<tr>
<td>5. FES-I at 12 months</td>
<td>7–28</td>
<td>10.0 ± 3.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**FES-I items (Shortened FES-I*)**

1. Cleaning the house (e.g. sweep, vacuum or dust)
2. Getting dressed or undressed (*)
3. Preparing simple meals
4. Taking a bath or shower (*)
5. Going to the shop
6. Getting in or out of a chair (*)
7. Going up or down stairs (*)
8. Walking around in the neighbourhood
9. Reaching for something above your head or on the ground (*)
10. Going to answer the telephone before it stops ringing
11. Walking on a slippery surface (e.g. wet or icy)
12. Visiting a friend or relative
13. Walking in a place with crowds
14. Walking on an uneven surface (e.g. rocky ground, poorly maintained pavement)
15. Walking up or down a slope (*)
16. Going out to a social event (e.g. religious service, family gathering or club meeting) (*)
Questionnaire structure of the FES-I

The principal component analysis revealed the greatest eigenvalue was <3 (1.9), supporting the uni-dimensionality of the scale [10]. Internal consistency of the FES-I as a whole was excellent with Cronbach’s alpha of 0.79 [26].

A bubble chart was inspected to ascertain the overall fit of the 16-item FES-I (Figure 1, panel A). Bubble charts display three-dimensional data in two dimensions, with the position of the centre of the bubble indicating two of the values (weighted t statistics [x-axis] and average measure [y-axis]) and the size of the bubble indicating the standard error. The weighted t statistics report the significance of the chi-square statistics occurring by chance when the data fit the Rasch model. The values reported are unit-normal deviates, in which 5% significance corresponds to 1.96. Overfit (i.e. redundancy of individual items) is reported with negative values. The weighted t statistic was not acceptable for four items of which items 5 and 12 showed values >−2.00 and items 11 and 13 values >2.00 [10]. The average measure is the location of an item (expressed in logits) on the latent variable and is the average ‘difficulty’ of that item in the tested population. Items rated as assessing higher levels of concern will have higher average measures. The item-respondent map showed that all items were located between −4.1 and +1.6 logits. Most items (1, 2, 3, 4, 6, 8, 10 and 16) assessed high levels of concern in a small distribution (positive logits, 0 to +1.6), suggesting some redundancy of items. The distribution of items (7, 9, 14 and 15) assessing lower levels of concern was wider (negative logits, 0 to −4.1). In order to reduce the floor effect, these items are crucial for the diversity of the scale.

The analyses of the 16-item FES-I support the item selection for the shortened FES-I [8] (Figure 1, panel B). The weighted t statistic of the 7-item FES-I was not excellent, with item 15 slightly outside the acceptable range (−2.31) [10]. The shorter version was found to be unidimensional, but as expected the Cronbach’s alpha decreased to 0.63, which is still acceptable. The item-respondent map showed a smaller distribution of items on the continuum than the longer version towards the assessment of lower levels of (−1.5 to +1.6 logits). One item assessed very low levels of concern, two moderately low, two moderately high and two very high. This distribution of items indicates that the scale has a good content representation of the construct and will allow scoring of older people with different levels of concern.

Convergent and predictive validity

Significant between-group differences in total scores demonstrated that both questionnaires were sensitive to group differences relating to demographic characteristics and fall risk factors (Table 2). Scores for both versions of the FES-I were significantly higher in women and older participants; participants with a falls history, increased physiological falls risk, poor balance, low muscle strength and slower gait speed; and participants with depressive symptoms, lower quality of life and poor cognitive processes. These group differences remained after the 1-year follow-up (n = 463) for future falls, physiological falls risk, muscle weakness, overall disability and depressive symptoms.

Stability over time and responsiveness to change

A Spearman rho correlation of the serial FES-I measurements suggests that it was only moderately stable over time with correlation coefficients between 0.66 and 0.83 across all versions (Table 1). To assess responsiveness to change, participants were selected who had at least one period of 3 months without suffering a fall (n = 404, Figure 2). The FES-I at the start of the next 3-month period was then compared to the FES-I at the end of that same 3-month period. Participants were classified with respect to two different falls outcome measures: falls (three groups: no falls, n = 294; one fall, n = 93; two or more falls, n = 17) and injurious falls (two groups: no injury, n = 330; injury, n = 74). An RM ANOVA revealed a main effect of time indicating that the level of concern increased over time in all participants (Wilks = 0.96, F1,402 = 17.36, P < 0.001). The group x time interaction was not significant for number of falls (Wilks = 0.99, F2,401 = 1.95, P = 0.143) or injurious falls (Wilks = 0.99, F1,402 = 0.69, P = 0.406). A trend for an interaction effect was found between no falls and multiple falls (Wilks = 0.99, F1,402 = 3.51, P = 0.062).

Identification of cut-points

ROC plots were used to define cut-points for the FES-I scales. Each of the previously described differentiating factors was used separately as a state variable. We defined cut-points as the best trade-off between sensitivity and specificity and aimed to have identical areas under the curve before and after recoding of the FES-I scales. The area under the curve for the 16-item FES-I ranged between 0.58 for balance, 0.67 for previous falls, up to 0.74 for depressive symptoms and was similar for the shortened version of the FES-I. We defined a cut-point to differentiate between low and high concern (16-item FES-I: 16–22 and 23–64; 7-item FES-I: 7–10 and 11–28) and between low, moderate and high concern (16-item FES-I: 16–19, 20–27 and 28–64; 7-item FES-I: 7–8, 9–13 and 14–28).

Discussion

The aim of this study was to address limitations of the initial validation study [7] and to provide evidence for the predictive validity of the FES-I. Overall, the current study was able to confirm the good validity and reliability of the FES-I found in the previous study.

The overall structure of both FES-I versions is acceptable but skewed towards assessing people with higher levels of concern about falling, resulting in a floor effect [27]. To increase the ability of the FES-I to track the full range of concern about falling in healthy older people, add-
itional more demanding balance-related activities should be explored. In the 16-item version, two out of the six more demanding items had an unacceptable fit. The poor fit of the item ‘walking on a slippery surface’ is possibly a result of the current study being undertaken in a temperate climate in which it was difficult for participants to relate to negotiating icy surfaces. Another possibility is that these two contexts are not related to fear of falling but to other psychological constructs such as quality of life (item 13) or risk-taking behaviour (item 11). Our analyses of the 16-item FES-I support the item choice for the shortened version [8].

Convergent and predictive validity of the FES-I were explored extensively by including both physical and psychological measures in a longitudinal design. Convergent

![Figure 2](image.png)

**Figure 2.** Responsiveness to change to falls of the 16-item FES-I (mean ± SEM) over a period of 3 months (in participants who had no falls ($n = 294$), one fall ($n = 93$) or multiple falls ($n = 17$) in the same 3-month period).
The Falls Efficacy Scale International (FES-I)

**Key points**

- The measurement properties of the 16- and 7-item FES-I scales were good when assessed in 500 older community living people.
- Both FES-I scales had excellent validity with regard to external physiological and neuropsychological measures.
- Scores $>23$ for the 16-item scale and scores $>10$ for the 7-item scale indicated high concern about falling.
- Both 16-item and 7-item FES-I scales can be recommended for research and clinical purposes.

**Acknowledgements**

This research was conducted as part of a study on understanding fear of falling and risk-taking in older people, which has been funded by an Australian NHMRC grant (no. 400941). Professor S.R.L. is currently a NHMRC Senior Principal Research Fellow. The participants in this study were drawn from the Memory and Ageing Study of the Brain and Ageing Program, School of Psychiatry, UNSW, funded by a NHMRC Program Grant (no. 350833) to Professors P.S.S., H.B. and G.A. We are grateful to the many researchers involved in this study.

**Conflicts of interest**

None declared.

**References**