ORIGINAL ARTICLE

Development of a Multidimensional Balance Scale for Use With Functionally Independent Older Adults

Debra J. Rose, PhD, Nicole Lucchese, MS, Lenny D. Wiersma, PhD

ABSTRACT. Rose DJ, Lucchese N, Wiersma LD. Development of a multidimensional balance scale for use with functionally independent older adults. Arch Phys Med Rehabil 2006;87:1478-85.

Objective: To develop and evaluate the validity and reliability of a multidimensional balance scale—the Fullerton Advanced Balance (FAB) scale—suitable for use with functionally independent older adults.

Design: Psychometric evaluation of the scale's content and convergent validity, test-retest and intra- and interrater reliability, and internal rater consistency.

Setting: Urban community.

Participants: Forty-six community-residing older adults (mean $\pm$ standard deviation, 75 ± 6.2y), with (n=31) and without identified balance problems (n=15), participated in the study. Four physical therapists with expertise in the assessment and treatment of balance disorders in older adults also participated in the content validity and/or reliability phases of the study.

Interventions: Not applicable.

Main Outcome Measures: Spearman rank correlation coefficients for convergent validity, test-retest, intra- and interrater reliability, and homogeneity coefficient values for rater consistency.

Results: Test-retest reliability for the total balance scale score was high (ρ=.96). Interrater reliability for total score ranged from .94 to .97 whereas intrarater reliability coefficients ranged from .97 to 1.00. Homogeneity (H) coefficients were greater than .90 for 6 of the 10 individual test items and all 10 test items had H coefficients of greater than .75 for both rating sessions.

Conclusions: Preliminary results suggest that the FAB scale is a valid and reliable assessment tool that is suitable for use with functionally independent older adults residing in the community.

Key Words: Balance; Elderly; Outcome assessment (health care); Posture; Rehabilitation.

THE ADULT POPULATION 65 years and older has increased rapidly over the last century and is expected to reach 40 million by 2010 in the United States.[1] Heightened awareness of problems associated with aging is therefore warranted. One of the most common and potentially serious consequences associated with aging is falls. Approximately 35% of people over the age of 65 fall at least once per year, and 20% to 30% of falls result in moderate to severe injuries that adversely affect their mobility and independence.[2,3] The total cost associated with fall injuries in this population was \$20.2 billion in 1994 and is estimated to rise to \$32.4 billion by 2020.[4] These statistics indicate the need for immediate action to reduce falls and fall-related injuries among the older adult population.

Many factors contribute to increased fall risk among older adults, including impairments in balance and gait, decreased muscular strength, impaired vision, and increased cognitive impairment. Hazards in the home and community (eg, poor lighting, lack of stair railings, uneven sidewalks) also contribute to heightened fall risk, particularly among community-residing older adults. To understand why certain older adults are at a higher risk for falls, a number of clinical and laboratory measures of balance and/or gait have been developed.[5-13] Many of these same tests have also been used to show that fall risk and/or fall incidence rates can be appreciably lowered when exercise programs that specifically target identified impairments in balance, muscle strength, endurance, and gait are implemented.[14-17]

Whereas certain tests require the performance of a single task (eg, walking, tandem stance, single-leg stance, functional reach, sternal nudge), others require the performance of multiple tasks that are often similar to those performed during daily life (eg, transfers, stair climbing, reaching, picking up objects from the floor). Although single-item tests serve as useful screening tools, they provide little if any information about possible underlying balance impairments. In contrast, clinical tests comprised of multiple tasks have the potential to provide more useful information about the underlying system impairments as well as better treatment guidance for clinicians. Some examples of multi-item tests include the Performance Oriented Mobility Assessment (POMA),[12] comprising a balance and gait scale, the Gait Abnormality Rating Scale, which assesses 16 aspects of gait,[13] the Fast Evaluation of Mobility, Balance, and Fear (FEMBAF) that includes a risk factor questionnaire, 18 functional tasks, and a self-perception scale,[5,6] the Physiological Profile Assessment (PPA),[10] which has both a long (16 tests) and short (5 tests) version that includes 1 or more tests of vision, muscle force, peripheral sensation, reaction time, and postural sway, and the Berg Balance Scale (BBS) that measures performance on 14 functional tasks.[7,8] Of the aforementioned tests, the BBS is currently among the most widely used clinical tests to assess balance in community-residing older adults.

The reliability and validity of the BBS has been well established and its use in assessing functional abilities of older adults is widespread.[18-21] A limitation of the BBS becomes apparent, however, when assessing older adults who have

From the Center for Successful Aging (Rose, Lucchese) and Department of Kinesiology (Wiersma), California State University, Fullerton, CA.

Presented in part to the Joint Conference of the National Council on Aging and American Society on Aging, March 2003, Chicago, IL.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the authors or upon any organization with which the authors are associated.

Reprint requests to Debra J. Rose, PhD, Center for Successful Aging, KHS 245, California State University, 800 N State College Blvd, Fullerton, CA 92834, e-mail: *drose@fullerton.edu*.

0003-9993/06/8711-10938\$32.00/0
doi:10.1016/j.apmr.2006.07.263

balance problems but are also relatively high functioning. In fact, when first developing the scale, Berg et al acknowledged that "the lack of an item that requires a postural response to an external stimulus or uneven support surface is a limitation of the scale."[7(p310)] Berg further acknowledged that these omissions might limit the effectiveness of the scale when used with active older adults with less severe deficits. The scale's limitations have been confirmed by more recent studies. For example, Garland et al[22] studied postural responses to unilateral arm perturbation in young and elderly subjects. All participants received high scores on the BBS with minimal differences evident between the young and elderly groups. However, the older adult participants exhibited more varied responses to the perturbation when compared with younger adults. Garland suggested that the scale is limited in its ability to detect differences in balance that are subtler in nature or pose a higher level of balance challenge.

If a patient population scores high initially on the BBS, then the scale's predictive ability is compromised.[20] It has also been argued that the BBS is more appropriate for use with frail populations due to the scale's emphasis on discrimination when balance is relatively poor. Support for this argument has been recently provided by Wee et al[21] who found the BBS to be a good predictor of length of stay and discharge destination among frail older patients admitted to a stroke rehabilitation unit. The fact that the BBS does not include any postural challenge measures, has unusually long tests of sitting balance (albeit an optional item), comfortable stance, and Romberg stance but a relatively short single-leg stance test item further suggests that the BBS is better suited for use with lower functioning persons. Finally, Newton[19] has reported ceiling effects when using the BBS to assess functionally independent older adults. She concluded that the tasks on the BBS may not be challenging enough to detect subtle balance deficits and recommended that more challenging tasks be included when testing active older adults.

When examining the relation between impairments in the systems controlling balance and the ability to perform activities that demand balance skills in community-residing older adults, Allison[23] identified further limitations with the BBS. Although performance on certain BBS test items was strongly associated with underlying impairments such as reduced limits of stability, lower-extremity joint range of motion, lower-extremity strength, and lower-extremity somatosensation, it could not be used to identify impairments evident in the visual and vestibular systems. Nor could any of the test items on the BBS be used to identify prolonged reaction times and automatic postural response latencies observed. These additional findings by Allison suggest the need for a clinical test that assesses more of the multiple dimensions of balance.

The purpose of this study was to develop a new balance assessment tool that could be used to identify balance problems of varying severity in functionally independent older adults and also evaluate more of the system(s) (eg, sensory, musculoskeletal, neuromuscular) that might be contributing to balance problems. More difficult static and dynamic balance tasks were included in the scale that would not only make it less prone to ceiling effects when used with more active older adults but also a more sensitive instrument when used to evaluate the effectiveness of an intervention conducted with this segment of the older adult population. Having a better understanding of the system(s) contributing to the identified balance problems will also provide better guidance for the clinician when developing a treatment plan for a patient. Given that balance activities are also becoming a more regular component of fitness programs designed for apparently healthy older adults, the development of a valid and reliable test of balance would also complement other measures of fitness currently used to assess an older adult client's functional abilities. To reflect the more challenging nature of the scale, it is called the Fullerton Advanced Balance (FAB) scale.

METHODS

The development of the FAB scale involved a multistage process. The following specific activities were completed: (1) a review of the conceptual frameworks used to identify the different dimensions of balance and/or the physiologic systems that contribute to postural control; (2) a review of the scientific literature describing the age-associated changes in postural control as well as their association with falls among older adults; (3) a review of previously published physical performance tests used to evaluate these age-associated changes and identify individuals at high risk for falls; (4) development of appropriate test items and clarification of test goals; (5) evaluation of the appropriateness of the individual test items, clarity of instructions, and scoring by a team of clinical experts; (6) pilot testing of the preliminary scale with older adults to establish appropriate test protocols, scoring procedures, and clarity of instructions; and (7) preliminary evaluation of the scale's reliability and content and convergent validity. All phases of the study involving the use of human subjects received prior approval by the local institutional review board.

Content Validity

We began the process of establishing content validity with a comprehensive review of the published literature to establish a conceptual and theoretical framework for the selection of individual test items. The conceptual framework that guided the selection of individual test items was the systems theory of postural control.[24] According to this theory, the neural control of posture and balance requires a complex interaction of neural (sensory and motor) and musculoskeletal systems. Shumway-Cook and Woollacott[24] identified 7 systems or mechanisms important for postural control. These included the sensory and musculoskeletal systems, sensory strategies, neuromuscular synergies, internal representations (cognition), and adaptive and anticipatory mechanisms. Lord et al[10] also selected their original PPA test items using a similar conceptual framework involving 5 systems (ie, vision, vestibular function, peripheral sensation, muscle force, reaction time).

Our next steps in deciding which test items to include in the FAB scale were to review the scientific literature that identified which systems or components of balance were most affected by the aging process and the extent to which the changes were associated with increased risk for falls. We also conducted a review of published laboratory and clinical tests that could successfully discriminate between different age groups[25-35] and fallers and nonfallers.[29,36-43] We then selected the preliminary test items for inclusion in the FAB scale that involved one or more of the 7 systems or mechanisms identified by Shumway-Cook and Woollacott[24] and/or that had been shown to discriminate between different age groups and faller status in the literature. The primary physiologic system(s) and/or dimensions of balance believed to be represented in 1 or more individual test items are presented in table 1. Evidence supporting the discriminatory power of certain test items is also presented in table 1. Finally, a review of the measurement literature and previously published clinical tests of balance assisted us in establishing the number of possible performance levels, the appropriate operational statements for each performance level, and overall task difficulty. The specific assess-

Table 1: Primary Systems and/or Mechanisms Evaluated on Each Individual Test Item

Test Item	Systems and/or Mechanisms Evaluated	Supporting Studies by Reference Number
1. Stand with feet together, eyes closed	Sensory systems and strategies (somatosensation, vision), internal representations, musculoskeletal components, neuromuscular synergies	25, 28, 42
2. Reaching forward to object	Sensory systems (vision), neuromuscular response synergies, musculoskeletal components, anticipatory mechanisms	30, 37, 41
3. Turn in full circle	Sensory systems and strategies (vestibular, vision), neuromuscular synergies, musculoskeletal components	7, 12, 41
4. Step up and over	Sensory systems and strategies (vision, somatosensation), anticipatory and adaptive mechanisms, neuromuscular synergies, musculoskeletal system	27, 38
5. Tandem walk	Sensory systems and strategies (vision, somatosensation), neuromuscular synergies, musculoskeletal components	35
6. Stand on one leg	Sensory systems (vision), anticipatory and adaptive mechanisms, musculoskeletal components	25, 40
7. Stand on foam, eyes closed	Sensory systems and strategies (vestibular), internal representations, neuromuscular synergies, musculoskeletal components	28, 42, 43
8. Two-footed jump	Neuromuscular synergies, musculoskeletal components, anticipatory and adaptive mechanisms	26, 34, 38
9. Walk with head turns	Sensory systems and strategies (vestibular, vision), neuromuscular synergies, adaptive mechanisms	41, 42
10. Reactive postural control	Neuromuscular synergies, adaptive mechanisms, musculoskeletal system	31, 32, 39

ment tools reviewed included the BBS, the FEMBAF, and the POMA.[5,7,8,12]

We administered the preliminary set of 10 test items to a group of 15 community-residing older adults (9 women, 6 men) who were physically active and had no identifiable impairments in balance. The purpose of this preliminary testing was to identify the best possible performance someone without observable balance deficits could expect to achieve on each test item. Only 4 of the 15 adults recorded less than a 4 on any individual test item, with the lowest score being 3. Total scores on the FAB scale ranged from 38 to 40. Each participant was also asked to rate the difficulty of each test item following its completion. The most difficult items identified were test items 5 (tandem walk), 7 (stand on foam, eyes closed), 8 (2-footed jump), and 9 (walk with head turns). Test item 10 (reactive postural control) was also perceived to be challenging because the participants were unaware that the test administrator's hand would be removed from the back as they leaned into the hand. We felt this was necessary to ensure that the movement response was indeed reactive. On the basis of this preliminary testing and the participants' feedback, the performance categories and scoring boundaries for each test item were established.

An expert panel of 4 experienced physical therapists that specialized in the assessment and treatment of older adults with balance disorders were then asked to review the 10 test items included on the scale. Specifically, each panel member had been a practicing physical therapist for more than 15 years with additional specialty certification in geriatrics and/or neurology. Each panel member was asked to review the appropriateness of the task and operational definitions delineating the level of performance on the 0 to 4 ordinal scale for each test item as well as the scale's overall appropriateness for the intended population. Each expert was also asked to evaluate the clarity of the instructions that accompanied each item. The panel's written feedback was then used to modify certain test items and/or operational statements on the original scale. Modifications were also made to the accompanying test administration instructions that were developed to standardize the manner in which each test item was administered and scored. Panel members unanimously agreed that the 10 test items originally identified for inclusion in the scale be retained. The majority of the changes recommended by the expert panel related to specific test item instructions, task difficulty, and the scaling of operational statements on certain test items. Once recommendations from the expert panel were reviewed and incorporated into the scale, we modified the scale and then returned it to the expert panel for a final evaluation. The final 10 items included in the FAB scale, along with each of the operational statements used to categorize performance on each item, are presented in appendix 1. Each test item was scored using a 0- to 4-point ordinal scale. The total test score possible was 40 points.

Convergent Validity

We obtained preliminary evidence for convergent validity by comparing scores on the FAB scale with the BBS in a group of 31 older adults with identified balance problems of varying severity. The total score on the BBS scale was correlated with the FAB scale total score using Spearman rank correlation analysis. A total of 24 women and 7 men, ranging in age from 63 to 84 years (mean ± standard deviation [SD], 75±6.2y) were recruited to participate in this phase of the scale's evaluation. All participants completed a health history and physical activity questionnaire that provided demographic and medical history information prior to the testing session. The 12-item Composite Physical Function (CPF) scale was also administered.[44] Participants were required to rate their ability to complete a range of basic, instrumental, and advanced activities of daily living. Specifically, they were asked to indicate whether they could do the activity (score of 2), could do it with difficulty or with help (score of 1), or could not perform the activity at all (score of 0). Respondents who scored between 22 and 24 points on the scale were categorized as high functioning, and respondents scoring below 13 points were categorized as low functioning. Respondents who scored between 13 and 21 were categorized at a moderate level of function. A broad range of functional abilities and health status was represented in the sample. This was evident by the number of medical diagnoses

Table 2: Intrarater Reliability Coefficients for Individual Test Items and Total Score on FAB Scale

FAB Item No.	Rater 1	Rater 2	Rater 3
1	No variance	No variance	No variance
2	1.00*	1.00*	1.00*
3	0.51	1.00*	0.99*
4	0.80*	1.00*	1.00*
5	0.96*	1.00*	0.98*
6	1.00*	1.00*	0.93*
7	0.99*	1.00*	0.97*
8	0.91*	1.00*	0.98*
9	0.79*	1.00*	0.95*
10	0.92*	1.00*	1.00*
Total score	0.99*	1.00*	0.93*

*Correlation is significant at .01 level (2-tailed).

(mean, 2.2±1.5; range, 0−6) and number of prescription medications reported (mean, 3.5±3.0; range, 0−13) on the health history portion of the questionnaire. The most commonly reported medical diagnoses included osteo- or rheumatoid arthritis (n=7), cardiovascular disease (n=4), high blood pressure (n=10), and macular degeneration or other vision problems (n=3). According to the cutoff scores associated with the CPF scale, 3 of the 31 participants were defined as low functioning, 21 at a moderate level of function, and 7 were classified as high functioning.

Test-Retest Reliability

The same group of 31 older adults who participated in the convergent validity phase of the study also volunteered for the test-retest reliability phase. Each participant was tested on the FAB scale on 2 separate occasions, 2 to 4 days apart. The same tester administered the scale on both occasions. The total FAB score and each of the individual item scores obtained on the first and second test dates were then compared to determine the scale's level of test-retest reliability.

Intra- and Interrater Reliability

To determine intrarater and interrater reliability, we asked 3 of the 4 physical therapists who served on the expert panel to serve as raters. Each rater was provided with a videotape showing 10 older adults (age range, 65−81y) being tested on the FAB scale. The older adults (6 women, 4 men) shown in the video were selected from the larger group who participated in the test-retest reliability phase of the study and who represented a range of functional abilities. On the basis of their CPF scores, 2 were classified as low functioning, 6 were moderate functioning, and 2 were high functioning. Detailed test administration instructions and the scoring form associated with the scale were also sent to each rater. Each rater was instructed to watch the first videotape and score the participants after a single viewing of each test item. This instruction was intended to make the scoring situation similar to a typical "live" testing situation. After the first viewing was completed, each rater mailed the video and 10-scored balance scales back to the test developers. One week later they were asked to watch a second videotape showing the same group of older adults (in a different test order) and once again record and send the second video and their scores back to the test developers for further analysis.

We determined test-retest reliability and intrarater and interrater reliability using Spearman rank correlation coefficients (ρ). The ρ constitutes the nonparametric version of the Pearson r, and is the appropriate statistical technique to use when correlating 2 test scores that are based on ordinal rankings (ie, 0−4).[18,45] Homogeneity (H) coefficients[46] were also calculated to determine the extent to which each test item on the balance scale was rated similarly by the raters across the 2 rating sessions. This analysis provided a measure of the level of internal consistency across raters.

RESULTS

The results of the Spearman rank correlation analysis indicating convergent validity produced a significant ($P<.01$) but moderate correlation of .75 when total scores using the BBS were compared with the total scores on the FAB scale. The moderate correlation obtained suggests that the 2 tests were measuring a similar construct, but it was not so high as to suggest that the 2 scales were necessarily measuring the same dimensions of balance.

Test-retest reliability was also established using Spearman rank correlation coefficients. The calculated ρ for the total FAB score was .96, demonstrating high reliability. All correlations between individual test items were also significant at the .01 level and ranged from .52 to .82. The lowest coefficients were associated with items 3 (.52) and 7 (.64). To determine intrarater reliability, the first and second set of scores for each of the 3 raters was compared using Spearman rank-order coefficient analysis. Total score as well as individual item correlations for each rater are presented in table 2. For all of these analyses, item number 1 on the balance scale produced no variance; each rater awarded the highest mark of 4 to all participants at each review. With the exception of a nonsignificant correlation obtained for rater 1 on test item number 3 (turn in a full circle), each rater showed good to excellent consistency in the way each test item was scored across the 2 rating sessions.

Interrater reliability was established by comparing the scores among the 3 raters for each of the 10 participants who were videotaped during the first rating session. Because it is only possible to compare 2 sets of scores using a Spearman rank-order correlation analysis, we compared each possible combination of 2 raters individually. The correlation coefficients obtained for the total scores were all high for each pair of raters, although the correlation coefficients ranged from low to very high on certain individual test items (table 3). Although the correlations exceeded .80 for 6 of the 10 test items across each pair of raters, lower correlations were obtained across 1 or more pairs of raters for items 3 (turn in a full circle), 4 (step up

Table 3: Interrater Reliability Coefficients for Individual Test Items and Total Score on FAB Scale

Test Item No.	Rater 1 vs Rater 2	Rater 2 vs Rater 3	Rater 3 vs Rater 1
1	No variance	No variance	No variance
2	0.82*	1.00*	0.82*
3	0.39	0.75†	0.57
4	0.22	0.61	0.80*
5	0.83*	0.93*	0.94*
6	1.00*	0.93*	0.93*
7	0.73†	0.95*	0.66†
8	0.96*	0.97*	0.96*
9	0.88*	0.75†	0.91*
10	0.95*	0.91*	0.94*
Total Score	0.96*	0.97*	0.94*

*Correlation significant at .01 level (2-tailed).
†Correlation significant at .05 level (2-tailed).

Table 4: Homogeneity Coefficients for the 10-Item FAB Scale

Item	Time 1	Time 2
1: Feet together, eyes closed	1.00±0.00	1.00±0.00
2: Reach forward for object	0.98±0.08	0.95±0.11
3: Turn in full circle	0.86±0.19	0.78±0.21
4: Step up and over	0.93±0.12	0.98±0.08
5: Tandem walk	0.81±0.18	0.86±0.15
6: Stand on one leg	0.98±0.08	1.00±0.00
7 Stand on foam, eyes closed	0.90±0.13	0.95±0.11
8: Two-footed jump	0.85±0.13	0.86±0.15
9: Walk with head turns	0.78±0.22	0.75±0.12
10: Reactive postural control	0.93±0.12	0.93±0.12

NOTE. Values are mean ± SD. Right-table probability of *H* (.01) = 1.00 (as provided by Aiken[46]).

and over), 7 (stand on foam, eyes closed), and 9 (walk with head turns).

We derived additional information on the interrater reliability of the FAB scale through use of the Aiken homogeneity (*H*) coefficient.[46] Coefficients were calculated across the raters for each of the 10 participants and compared to the right-tailed probability table developed by Aiken.[46] Table 4 includes the mean *H* coefficient values for each item for the 10 subjects at both testing times. Table 5 includes frequency counts for each item across raters. The homogeneity coefficient cutoff value for 3 raters on a 5-category response scale was 1.00, which would only be attained on exact agreement by all 3 raters. For example, if 2 of the clinicians rated a participant as a 4 on a balance task but the third rated the participant as a 3, the homogeneity coefficient would be a .75, which would be below the 1.00 cutoff value at the .01 α level.

As indicated in table 4, 6 of the 10 items had mean *H* coefficients .90 or greater and all 10 items had *H* coefficients greater than .75 for both testing times. Frequency counts in table 5 further indicate that for most items, the majority of the clinicians rated the participants with a high degree of consistency. Items with the highest *H* values included item 1 (feet together, eyes closed), item 2 (reach forward for object), item 6 (stand on 1 leg), and item 7 (stand on foam, eyes closed). The 3 items with the lowest rater agreement included item 3 (turn in full circle), item 5 (tandem walk), and item 9 (walk with head turns).

The relatively low *H* coefficients for items 3, 5, and 9 were reasonably consistent with low ρ coefficients obtained in test-retest analyses, but not all of the inter- or intrarater coefficients corresponded with obtained *H* coefficients. Individual or between-rater comparisons are dependent on consistency over time, but homogeneity coefficients are indicative of agreement between the raters on test scores, not on consistency over time. Thus, *H* coefficients serve as a measure of internal consistency of rater scores and provide a different form of reliability inferences than do traditional examinations of consistency over time.

DISCUSSION

Our primary purpose in developing the FAB scale was to produce a functional assessment tool that could assist in the identification of balance problems among functionally independent older adults. Although other balance assessment tools exist and are routinely used to assess balance in the clinical setting, we felt there was a need for a test that could more comprehensively address the multiple dimensions of balance and also identify emerging balance problems in higher functioning older adults. The systems theory of postural control was the conceptual framework that we used to identify appropriate test items for this new balance scale. For inclusion in the new balance scale, we selected test items that evaluated the contribution to balance of each of the sensory systems as well as the motor (voluntary and involuntary) and musculoskeletal systems. Because the BBS does not adequately evaluate the contribution of the 3 sensory systems to balance, or include a measure of reactive postural control, we contend that its ability is limited to detect functional limitations that are sensory in origin or to identify more subtle balance deficits emerging in higher functioning older adults. The need for a balance test comprised of more challenging test items was also considered important based on ceiling effects reported by other researchers and our own experiences when administering the BBS to community-residing older adults. To be practical, the scale also needed to be relatively quick and easy to administer, require little equipment, and, most important, demonstrate that it is valid and reliable.

In addition to establishing the content validity of the FAB scale, high test-retest reliability and interrater and intrarater reliability was also established when the total FAB score was used. A high level of internal consistency across raters was also evident, based on the high *H* coefficients obtained. It must be acknowledged, however, that the physical therapists who served as raters were geriatric specialists who had considerable experience assessing older adults with balance impairments using this type of clinical scale.

Good to excellent intrarater reliability was also established for 2 of the raters on all of the individual test items, and for 8 test items for the third rater. This finding suggests that if the same experienced clinician evaluates a patient's performance

Table 5: Frequency Counts of Homogeneity Coefficients for the FAB Scale by Individual Test Item

	Time 1			Time 2		
Item	*H*	Frequency	%	*H*	Frequency	%
1	1.00	10	100	1.00	10	100
2	0.75	1	10	0.75	2	20
	1.00	9	90	1.00	8	80
3	0.50	1	10	0.50	2	20
	0.63	1	10	0.63	2	20
	0.75	2	20	0.75	2	20
	1.00	6	60	1.00	4	40
4	0.75	3	30	0.75	1	10
	1.00	7	70	1.00	9	90
5	0.50	1	10			
	0.63	1	10	0.63	1	10
	0.75	4	40	0.75	4	40
	1.00	4	40	1.00	5	50
6	0.75	1	10			
	1.00	9	90	1.00	10	100
7	0.75	4	40	0.75	2	20
	1.00	6	60	1.00	8	80
8				0.63	1	10
	0.75	6	60	0.75	4	40
	1.00	4	40	1.00	5	50
9	0.25	1	10	0.50	1	10
	0.75	6	60	0.75	8	80
	1.00	3	30	1.00	1	10
10	0.75	3	30	0.75	3	30
	1.00	7	70	1.00	7	70

NOTE. Right-table probability of *H* (.01) = 1.00 (as provided by Aiken[46]). Frequency represents the number of participants (out of 10) rated at each homogeneity coefficient across the 3 raters.

on the FAB scale on 2 separate occasions, high reliability can be expected. Conversely, the reliability between the 3 pairs of raters varied considerably across the 10 test items. Good to high correlation values were obtained across all pairs of raters on test items 1, 2, 5, 6, 8, and 10 whereas lower correlation values were found for test items 3, 4, 7, and 9 across 1 or more pairs of raters. Possible explanations for the lower correlation values associated with certain individual test items include less than optimal videotape viewing angles for items 7 and 9 and the lack of clear test administration instructions in the case of test item 3. Two of the 3 raters reported difficulty in determining whether the head turned sufficiently on each head turn in item 9 because of the sagittal viewing angle provided. When a test administrator conducts this test item, he/she is usually positioned immediately behind the person being tested and therefore better able to determine how far the head turns in each direction. Space limitations precluded us from filming from the rear on this test item. Raters also reported difficulty in determining if and when the eyes opened during test item 7 for some participants because the camera was providing a wide-angle versus close-up view of the individual being tested. Although intrarater reliability was not adversely affected by the use of a videotape-based scoring method, good reliability was more difficult to obtain when certain test item scores were compared across the 3 pairs of raters. Finally, the low correlations associated with item 3 were likely due to the lack of clear test administration instructions initially provided for that test item. Two raters were unsure as to when they should discontinue counting steps after each full circle was completed. One rater indicated that she counted any additional preparatory steps taken prior to the turn being started in the opposite direction. Test administration instructions were further clarified on the basis of this feedback.

One of the advantages of the FAB scale is that it is quick to administer, requiring approximately 10 to 12 minutes. In contrast to the BBS, which is comprised of 14 test items, the FAB scale has only 10 test items. The average time for a trained assessor to administer the FAB scale is approximately 5 to 10 minutes shorter than the time required to administer the BBS and 45 to 50 minutes shorter than for the PPA developed by Lord et al.[10] The PPA also requires more equipment and set-up time prior to testing. Although more equipment is required to administer the FAB scale when compared with the BBS, it is relatively inexpensive and easily accessible (required equipment is listed in appendix 1). In addition, the test can be administered in a relatively small area, making it an easy test to administer in the home.

Additional studies will be needed to establish additional psychometric properties of this new balance scale. For example, it will be important to evaluate the relative strengths of this scale in comparison to other scales or tests currently used to assess balance as well as its discriminant and predictive validity across different levels of disability or fall risk. The relative responsiveness, or ability of the FAB scale to detect changes in balance over time is another form of test validity that will be important to investigate in future studies. This form of validity is particularly important when assessing the efficacy of an exercise intervention designed to improve an individual's balance abilities. Although the focus of this first study was on the test administrator's ability to correctly score each test item, future research should also address the degree to which different groups of professionals reliably administer the scale to older adults representing different levels of function. It will also be important to carefully evaluate whether all of the 10 test items are needed on the scale. The fact that no variance was observed for the first test item (stand with feet together) in the current sample may suggest that this item could be eliminated from the scale. Although the inclusion of this test item can be justified conceptually and demonstrates good discriminatory power across age groups, other test items on the scale also evaluate similar dimensions of balance (ie, contribution of somatosensory and visual systems to balance).

CONCLUSIONS

The results of this study indicate that the FAB scale is a very promising new balance assessment tool that is both valid and reliable when used to assess multiple dimensions of balance in community-residing older adults by experienced clinicians on multiple occasions.

APPENDIX 1: INDIVIDUAL TEST ITEMS AND SCORING CATEGORIES ASSOCIATED WITH FAB SCALE

Test Item and Verbal Instructions	Scoring Categories
1. Standing with feet together and eyes closed. "Bring your feet together, fold your arms across your chest, close your eyes when you are ready, and remain as steady as possible until I instruct you to open your eyes."	0 Unable to obtain the correct standing position independently. 1 Able to obtain the correct standing position independently but unable to maintain the position or keep the eyes closed for at least 10 seconds. 2 Able to maintain the correct standing position with eyes closed for more than 10 seconds but less than 30 seconds. 3 Able to maintain the correct standing position with eyes closed for 30 seconds but requires close supervision. 4 Able to maintain the correct standing position safely with eyes closed for 30 seconds.
2. Reaching forward to an object. "Try to lean forward to take the pencil from my hand and return to your starting position without moving your feet from their present position." Equipment: 12-inch ruler and pencil	0 Unable to reach the pencil without taking >2 steps. 1 Able to reach the pencil but needs to take 2 steps. 2 Able to reach the pencil but needs to take 1 step. 3 Can reach the pencil without moving the feet but requires supervision. 4 Can reach the pencil safely and independently without moving the feet.
3. Turn in full circle. "Turn around in a full circle, pause, and then turn in a second full circle in the opposite direction."	0 Needs manual assistance while turning. 1 Needs close supervision or verbal cueing while turning. 2 Able to turn 360° but takes more than 4 steps in both directions. 3 Able to turn 360° but unable to complete in ≤4 steps in 1 direction. 4 Able to turn 360° safely and takes ≤4 steps in both directions.

APPENDIX 1: INDIVIDUAL TEST ITEMS AND SCORING CATEGORIES ASSOCIATED WITH FAB SCALE (cont'd)

Test Item and Verbal Instructions	Scoring Categories
4. Step up and over. "Step up onto the bench with your right leg, swing your left leg directly up and over the bench, and step off on the other side. Repeat the movement in the opposite direction with your left leg as your leading leg." Equipment: 6-inch high × 14–18-inch wide bench	0 Unable to step onto the bench without loss of balance or manual assistance. 1 Able to step up onto the bench with leading leg, but trailing leg contacts bench or leg swings around bench during the swing-through phase in both directions. 2 Able to step up onto the bench with leading leg, but trailing leg contacts bench or swings around the bench during the swing-through phase in 1 direction. 3 Able to complete the step up and over in both directions but requires close supervision in 1 or both directions. 4 Able to complete the step up and over in both directions safely and independently.
5. Tandem walk. "Walk along the line, placing one foot directly in front of the other such that the heel and toe are in contact on each step forward. I will tell you when to stop." Equipment: Masking tape (2 inches wide)	0 Unable to complete 10 steps independently. 1 Able to complete the 10 steps with >5 interruptions. 2 Able to complete the 10 steps with <5 but more than 2 interruptions. 3 Able to complete the 10 steps with 2 or fewer interruptions. 4 Able to complete the 10 steps independently and with no interruptions.
6. Stand on one leg. "Fold your arms across your chest, lift your preferred leg off the floor, without touching your other leg, and stand with your eyes open as long as you can."	0 Unable to try or needs assistance to prevent falling. 1 Able to lift leg independently but unable to maintain position for >5 seconds. 2 Able to lift leg independently and maintain position for >5 but ≤12 seconds. 3 Able to lift leg independently and maintain position for >12 but <20 seconds. 4 Able to lift leg independently and maintain position for the full 20 seconds.
7. Stand on foam, eyes closed. "Step up onto the foam and stand with your feet shoulder-width apart. Fold your arms over your chest, and close your eyes when you are ready. I will tell you when to open your eyes." Equipment: Two Airex[a] balance pads with 18×18-inch sheet of nonslip material	0 Unable to step onto foam and/or maintain standing position independently with eyes open. 1 Able to step onto foam independently and maintain standing position but unable or unwilling to close eyes. 2 Able to step onto foam independently and maintain standing position with eyes closed for ≤10 seconds. 3 Able to step onto foam independently and maintain standing position with eyes closed for >10 seconds but <20 seconds. 4 Able to step onto foam independently and maintain standing position with eyes closed for 20 seconds.
8. Two-footed jump. "Try to jump as far but as safely as you can with both feet." Equipment: Yard stick and masking tape.	0 Unwilling or unable to attempt or attempts to initiate 2-footed jump but 1 or both feet do not leave the floor. 1 Able to initiate 2-footed jump but one foot leaves the floor or lands before the other. 2 Able to perform 2-footed jump but unable to jump further than the length of their own feet. 3 Able to perform 2-footed jump and achieve a distance greater than the length of their own feet. 4 Able to perform 2-footed jump and achieve a distance greater than twice the length of their own feet.
9. Walk with head turns. "Walk forward while turning your head from left to right with each beat of the metronome. I will tell you when to stop." Equipment: Metronome set to 100 beats per minute	0 Unable to walk 10 steps independently while performing 30° head turns at an established pace. 1 Able to walk 10 steps independently but unable to perform 30° head turns at an established pace. 2 Able to walk 10 steps but veers from a straight line while performing 30° head turns at an established pace. 3 Able to walk 10 steps in a straight line while performing head turns at an established pace but head turns <30° in one or both directions. 4 Able to walk 10 steps in a straight line while performing 30° head turns at established pacing.
10. Reactive postural control. "Slowly lean back into my hand until I ask you to stop."	0 Unable to maintain upright balance; no observable attempt to step; requires manual assistance to restore balance. 1 Unable to maintain upright balance; takes more than 2 steps and requires manual assistance to restore balance. 2 Unable to maintain upright balance; takes more than 2 steps but is able to restore balance independently. 3 Unable to maintain upright balance; takes 2 steps but is able to restore balance independently. 4 Unable to maintain upright balance but able to restore balance independently with only 1 step.

Adapted from Rose.[47] Reprinted with permission. (Additional instructions related to testing procedures provided on pages 66 to 68.)

References

1. Tideiksaar R. Falling in old age. 2nd ed. New York: Springer; 1997.
2. Hausdorff JM, Rios DA, Edelber HK. Gait variability and fall risk in community-living older adults: a 1-year prospective study. Arch Phys Med Rehabil 2004;82:1050-6.
3. Alexander BH, Rivara FP, Wolf ME. The cost and frequency of hospitalization for fall-related injuries in older adults. Am J Public Health 1992;82:1020-3.
4. Englander F, Hodson TJ, Terregrossa RA. Economic dimensions of slip and fall injuries. J Forensic Sci 1996;41:733-46.
5. Arroyo JF, Herrmann F, Saber H, et al. Fast evaluation test for mobility, balance, and fear: a new strategy for the screening of elderly fallers [abstract]. Arthritis Rheum 1994;37:S416.
6. Di Fabio RP, Seay R. Use of the "Fast evaluation of mobility, balance, and fear" in elderly community dwellers: validity and reliability. Phys Ther 1997;77:904-15.
7. Berg KO, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiother Can 1989;41:304-11.
8. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83(Suppl 2):S7-11.
9. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil 1992;73:1073-80.
10. Lord SR, Menz HB, Tiedemann A. A physiological profile approach to falls risk assessment and prevention. Phys Ther 2003;83:237-52.
11. Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients: the physical performance test. J Am Geriatr Soc 1990;38:1105-12.
12. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 1986;34:119-26.
13. Wolfson L, Whipple R, Amerman P, Tobin JN. Gait assessment in the elderly: a gait abnormality rating scale and its relation to falls. J Gerontol A Biol Sci Med Sci 1986;45:M12-9.
14. Rose DJ, Clark S. Can the control of bodily orientation be significantly improved in older adults with a history of falls? J Am Geriatr Soc 2000;48:275-82.
15. Rubenstein LZ, Josephson KR, Trueblood PR. Effects of a group exercise program on strength, mobility, and falls among fall-prone elderly men. J Gerontol A Biol Sci Med Sci 2000;55:M317-21.
16. Barnett A, Smith B, Lord SR, Williams M, Baumand A. Community-based exercise improves balance and reduces falls in at-risk older people: a randomized controlled trial. Age Ageing 2003;32:407-14.
17. Chang JT, Morton SC, Rubenstein LZ, et al. Interventions for the prevention of falls in older adults: systematic review and meta-analysis of randomized clinical trials. BMJ 2004;328:680-3.
18. Bogle Thorbahn LD, Newton RA. Use of the berg balance test to predict falls in elderly persons. Phys Ther 1996;76:576-85.
19. Newton RA. Balance screening of an inner city older adult population. Arch Phys Med Rehabil 1997;78:587-91.
20. Wee JY, Bagg SD, Palepu A. The Berg balance scale as a predictor of length of stay and discharge destination in an acute stroke rehabilitation setting. Arch Phys Med Rehabil 1999;80:448-52.
21. Wee JY, Wong H, Palepu A. Validation of the Berg Balance Scale as a predictor of length of stay and discharge destination in stroke rehabilitation. Arch Phys Med Rehabil 2003;84:731-5.
22. Garland SJ, Stevenson TJ, Ivanova T. Postural responses to unilateral arm perturbation in young, elderly, and hemiplegic subjects. Arch Phys Med Rehabil 1997;78:1072-7.
23. Allison L. Relationships between postural control system impairments and disabilities [master's thesis]. Corvallis: Oregon State Univ; 1997.
24. Shumway-Cook A, Woollacott MH. Motor control: theory and practical applications. Philadelphia: Lippincott Williams & Wilkins; 2001.
25. Bohannon RW, Larkin PA, Cook AC. Decrease in timed balance scores with aging. Phys Ther 1984;64:1067-70.
26. Bosco C, Komi PV. Influence of aging on the mechanical behaviour of leg extensor muscles. Eur J Appl Physiol 1980;45:209-19.
27. Chen HC, Ashton-Miller JA, Alexander NB, Schultz AB. Effects of age and available response time on ability to step over an obstacle. J Gerontol 1994;49:M227-33.
28. Cohen H, Heaton LG, Congdon SL, Jenkins HA. Changes in sensory organization test scores with age. Age Ageing 1996;25:39-44.
29. Chandler JM, Duncan PW, Studenski SA. Balance performance on the postural stress test: comparison of young adults, healthy elderly, and fallers. Phys Ther 1990;70:410-5.
30. Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. J Gerontol 1990;45:192-7.
31. Luchies CW, Alexander NB, Schultz AB, Ashton-Miller J. Stepping responses of young and old adults to postural disturbances: kinematics. J Am Geriatr Soc 1994;42:506-12.
32. McIlroy WE, Maki BE. Age-related changes in compensatory stepping in response to unpredictable perturbations. J Gerontol A Biol Sci Med Sci 1996;51:M289-96.
33. Pearson MB, Bassey EJ, Bendall MJ. Muscle strength and anthropometric indices in elderly men and women. Age Ageing 1985;14:49-54.
34. Skelton DA, Greig CA, Davies JM, Young A. Strength, power and related functional ability of healthy people aged 65-89 years. Age Ageing 1994;23:371-7.
35. Speers RA, Ashton-Miller JA, Schultz AB, Alexander NB. Age differences in abilities to perform tandem stand and walk tasks of graded difficulty. Gait Posture 1998;7:207-13.
36. Campbell AJ, Borrie MJ, Spears GF, Jackson SL, Brown JS, Fitzgerald JL. Circumstances and consequences of falls experienced by a community population 70 years and over during a prospective study [published erratum appears in: Age Ageing 1990;19:345-6]. Age Ageing 1990;19:136-41.
37. Duncan PW, Studenski S, Chandler J, Prescott B. Functional reach: predictive validity in a sample of elderly male veterans. J Gerontology 1992;47:M93-8.
38. Lord SR, McLean D, Stathers G. Physiologic factors associated with injurious falls in older people living in the community. Gerontology 1992;38:338-46.
39. Studenski S, Duncan PW, Chandler J. Postural responses and effector factors in persons with unexplained falls: results and methodologic issues. J Am Geriatr Soc 1991;39:229-34.
40. Vellas BJ, Wayne SJ, Romero L, Baumgartner RN, Rubenstein LZ, Garry PJ. One-leg balance is an important predictor of injurious falls in older persons. J Am Geriatr Soc 1997;45:735-8.
41. Allison L, Fuller K. Balance and vestibular disorders. In: Umphred DA, editor. Neurological rehabilitation. 4th ed. St. Louis: Mosby; 2001. p 616-60.
42. Shephard NT, Telian SA. Programmatic vestibular rehabilitation. Otolaryngol Head Neck Surg 1995;112:173-82.
43. Shumway-Cook A, Horak F. Assessing the influence of sensory interaction on balance. Phys Ther 1986;66:1548-50.
44. Rikli RE, Jones CJ. The reliability and validity of the 6-minute walk test as a measure of physical endurance in older adults. J Aging Phys Act 1998;6:363-75.
45. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River: Prentice Hall; 2000.
46. Aiken LR. Three coefficients for analyzing the reliability and validity of ratings. Educ Psychol Meas 1985;45:131-42.
47. Rose DJ. Fallproof: a comprehensive balance and mobility program. Champaign: Human Kinetics; 2003.

Supplier

a. Alcan Airex AG, Specialty Foams, CH-5643 Sins, Switerland.