Archives of Rehabilitation Research and Clinical Translation (2023) 5, 100296

ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Rehabilitation Research and Clinical Translation

Archives of Rehabilitation Research and Clinical Translation 2023;5:100296

Available online at www.sciencedirect.com

OPEN ACCESS
ARCHIVES of Rehabilitation Research & Clinical Translation
An OPEN ACCESS JOURNAL serving the rehabilitation community

Original Research

The Functional Balance Ability Measure: A Measure of Balance Across the Spectrum of Functional Mobility in Persons Post-Stroke

Check for updates

Jasmine J. Cash, MS [a,b], Craig A. Velozo, PhD, OTR/L [b,c,*], Mark G. Bowden, PT, PhD [d], Bryant A. Seamon, PT, DPT, PhD [b,e]

[a] *Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, Charleston, SC*
[b] *Ralph H Johnson VA Health Care System, Charleston, SC*
[c] *Division of Occupational Therapy, Department of Rehabilitation Sciences, College of Health Professions, Medical University of South Carolina, Charleston, SC*
[d] *Department of Clinical Integration and Research, Brooks Rehabilitation, Jacksonville, Florida*
[e] *Division of Physical Therapy, Department of Rehabilitation Sciences, College of Health Professions, Medical University of South Carolina, Charleston, SC*

KEYWORDS
Postural balance;
Psychometrics;
Rehabilitation;
Stroke

Abstract *Objective:* To determine whether the measurement properties of an instrument that combines items from the Berg Balance Scale (BBS) and the Functional Gait Assessment (FGA) called the *Functional Balance Ability Measure (FBAM)* supports measuring balance across the functional mobility spectrum.
Design: Retrospective cohort.
Setting: Item-level data were from an archival research database.
Participants: Ambulatory individuals (N=93, BBS=50 [29-56], FGA=16 [0-30], Fugl-Meyer Assessment of Lower Extremities=27 [14-34], self-selected walking speed=0.4±0.2 m/s, mean age ± SD, 61.7±11.3y; 30.1% female) with chronic stroke (≥6 months).

List of abbreviations: BBS, Berg Balance Scale; FBAM, Functional Balance Ability Measure; FGA, Functional Gait Assessment; PCA, principal component analysis; RM, Rasch Measurement; RSM, Andrich Rating Scale model.

Preliminary findings of this work were presented at the 2023 American Physical Therapy Association Combined Sections Meeting in San Diego California.
Supported by VA Office of Research and Development (ORD) Rehabilitation R&D Service [Merit Review A6365-R, 1I01RX002665-01A2], National Institutes of Health [NIH P20 GM109040], Foundation for Physical Therapy Research Promotion of Doctoral Studies (PODS) Scholarship – Level I, SCTR TL1 (T32) Predoctoral Clinical & Translational Research Award [TL1 TR001451 & UL1 TR001450]
Disclosures: The investigators have no financial or nonfinancial disclosures to make in relation to this project.
Cite this article as: Arch Rehabil Res Clin Transl. 2023;5:100296
* Professor emeritus

https://doi.org/10.1016/j.arrct.2023.100296

Interventions: Not applicable.
Main Outcome Measures: Unidimensionality was evaluated with a principal components analysis (PCA) of residuals. FBAM rating-scale characteristics, item hierarchy, item and person fit, and person separation were investigated using the Andrich Rating Scale Model.
Results: PCA findings indicate the FBAM is sufficiently unidimensional. Rating scale structure was appropriate without modifying the original BBS and FGA scoring systems. Item hierarchy aligned with clinical and theoretical predictions (hardest item: FGA-gait with narrow base of support, easiest item: BBS-sitting unsupported). One item (BBS-standing on 1 foot) misfit, however, removal marginally affected person measures and model statistics. The FBAM demonstrated high person reliability (0.9) and 6 people (~6%) misfit the expected response pattern. The FBAM separated participants into 4 statistically distinct strata, without a floor or ceiling effect.
Conclusions: The FBAM is a unidimensional measure for balance ability across a continuum of functional tasks. Rating-scale characteristics, item hierarchy, item and person fit, and person separation support the FBAM's measurement properties in persons with chronic stroke. Future work should investigate measurement with fewer items and whether the FBAM addresses barriers to adoption of standardized balance measures in clinical practice.

Approximately 30 measures are available for the clinical assessment of balance.[1] The plethora of measures can be attributed to test development using classical test theory and attempts to measure balance with respect to specific mobility tasks. The variety allows clinicians to select measures specific to patient-reported goals and mobility complaints, however, it can make it difficult to implement a standardized approach to quantify balance ability. Lack of standardized measures introduces unwanted variability in practice and can prevent the use of shared language and data to progress innovation in care.[2] Additionally, clinicians and researchers must select multiple measures since few instruments measure balance ability across a functional mobility spectrum.

This problem can be addressed by combining multiple existing measures. A previously created "bedside to community" unified balance measure using Rasch Measurement (RM) Theory methodology in a neurologically diverse sample demonstrated that less balance ability is needed to perform easier tasks (eg, transfers and standing), and more balance ability is required for more difficult tasks (eg, walking).[3] However, the measure did not include items that would theoretically extend the measure to capture lower (eg, sitting) and higher (eg, stair navigation) functioning mobility tasks. Additionally, the scoring structure was modified to fit RM Theory model expectations. Maintaining the scoring structure of commonly used instruments would reduce relearning of how to score each instrument in the context of a combined instrument.

We propose that maintaining scoring fidelity and capturing functional mobility across a spectrum can be accomplished by combining items from the Berg Balance Scale (BBS) and Functional Gait Assessment (FGA). Clinical practice guidelines strongly support the use of the BBS and the FGA[4] in neurologic rehabilitation. The BBS is a 14-item measure with excellent test-retest and interrater reliability[5,6] and excellent construct validity[5] in stroke. However, the BBS does not have difficult mobility tasks like walking or stair navigation. The FGA is a 10-item measure derived from the Dynamic Gait Index using RM Theory, that has excellent inter-rater, intra-rater, and test-retest reliability and construct validity[7] in stroke. However, the FGA does not include sitting and standing tasks. Both measures fit RM Theory models,[5,6,8,9] which use probability to examine a measurement tool's ability to quantify abstract constructs. The Andrich Rating Scale model (RSM), an extension of the Rasch model for polytomous data, assumes the probability of successfully performing an item is dependent on the relation between a person's ability and item difficulty. Results order a measure's items with respect to the construct (ie, least to most ability), which allows linear interpretation of scores with set interval distances.[10] Given the prior evidence that the BBS and FGA fit RM Theory models and that their rating scales for item performance are similar (ie, ratings increase with better performance) RM Theory analyses should be suitable to combine the BBS and FGA into 1 measure of balance ability.

The purpose of this research was to construct a comprehensive measure of balance ability across the functional mobility spectrum by combining items from the BBS and FGA called the Functional Balance Ability Measure (FBAM). We hypothesize that a combined measure would be unidimensional and fit the RSM, indicating sound measurement properties and maintain traditional scoring for BBS and FGA items. We also hypothesize that the items will indicate a hierarchy consistent with clinical expectations, or balance ability increases with functional task difficulty.

Methods

Data Source

Study methodology was informed by the RULER statement.[11] This study is a retrospective, cross-sectional analysis of 93 individuals with chronic (>6 months) stroke who were previously participated in research. Research trials sponsored by the NIH Center of Biomedical Research Excellence in Stroke Recovery (COBRE) include cross-sectional and randomized designs to investigate mechanisms of functional recovery. Informed consent was obtained for participation in all studies. Deidentified data were extracted from an archival database approved by the Institutional Review Board (IRB). IRB approval of this secondary analysis was not required.[12] Individuals were included if they had a BBS and FGA completed within 90-days of each other. Item-level data were taken from the first visit date of a study to prevent any confounding influence from individual research projects. Summary demographic data were analyzed using R Statistical Software.[a]

RM Theory Analysis

RM Theory analysis was performed with the RSM in Winsteps.[b]

Sample Size

Sample size recommendations for evaluating RSM measurement properties are based on the desired confidence interval for stable item calibrations or person measures. A range of 64 to 144 persons is acceptable for stable calibrations within a 0.5 logit with a 95% confidence interval.[13] A sample size of 50 persons is commonly recommended, with a maximum of 61 persons, for calibrations within 1 logit with a 99% confidence interval for polytomous items like in the FGA.[13] Our sample size of 93 persons is adequate for calibrations between 0.5 to 1 logit, which is consistent with recommended useful stability for measurement.[14]

Unidimensionality and local dependence

Unidimensionality of the FBAM was evaluated with a principal component analysis (PCA) of residuals using Winsteps.[b] PCA findings were compared against threshold guidelines[15] to evaluate a potential second dimension. The FBAM was considered unidimensional: 1) the model explains >50% of the variance, 2) eigenvalue ratios of <3, and 3) the observed variance of the first contrast was <4%. Our eigenvalue criterion was modified from the often used >2 to >3, as it was determined that eigenvalues of less than 1 item over the threshold of 2 (eg, 2.5), are within the acceptable noise of the RSM.[15]

If all unidimensionalty criteria were not met, we examined additional dimensions to determine whether they degraded measurement quality with disattenuated correlations between person measures derived from item clusters on the proposed primary and secondary dimensions from the PCA. Correlations >0.82 indicate that clusters of items are dependent, that is, are representative of the same latent construct,[15] and therefore not multidimensional.

We additionally tested item dependence with correlations between item standardized residuals from the PCA. Items were considered locally dependent if their correlation was >0.7 indicating items shared more than 50% of their variance.[11,16]

Rating-scale structure

We tested whether rating-scale structure for the BBS and FGA items could be maintained against the 3 essential criteria outlined by Linacre: 1) each rating scale category should have ≥10 observations; 2) the observed average measure should advance monotonically for each category; and 3) outfit mean-squares should be <2.[17]

Item and person statistics

Items and people were considered to fit the RSM if their mean-squared infit and outfit statistics were <1.4 and their z scores were <2.[17,18] Misfitting items were removed serially and additively, and model characteristics were compared with and without items. Misfitting persons were only removed if >10% were found to misfit the model's item response pattern expectations. Person reliability, measures, and strata were compared between each iteration of removing misfitting items or persons to examine the effect of item removal on the FBAM's measurement properties. A ceiling or floor effect was identified if 15% of persons had a maximum or minimum measure, respectively.[19]

Item difficulty hierarchy

The item hierarchy was used to determine if the FBAM had theoretical and clinical validity using the RSM. Items were considered overlapping, or to have a similar difficulty level, if the item's measure estimate was within 2 standard errors (SE) of another item.

Separation index

Person separation was calculated to determine the FBAM's ability to differentiate individuals into distinct strata[20]:

$$Strata = \frac{4 * (person\ separation\ index) + 1}{3}$$

RSM fit statistics

RSM fit statistics were calculated to determine the strength of the model fit to the sample.

Measurement scoring

RM Theory models provide measurement values in a logit scale anchored at 0, meaning measurement values can be negative or positive and include decimals. We converted the FBAM's logit measurement scale to a more easily interpretable 100-point scale (ie, 0-no balance, 100-full balance) by anchoring the mean of the scale to 50.[10] We generated a linear regression formula to convert traditional scores on the BBS and FGA to the 100-point FBAM interval measure scale.

Table 1 Participant demographics

N=93	
BBS Score	50 [29-56]
FGA Score	16 [0-30]
FMA-LE Score	27 [14-34]
Self-selected Walk Speed (m/s)*	0.4 (0.2)
Time Between Assessments (d)	23.9 (26.4)
Time Since Stroke (months)†	59.0 (62.2)
Age (years)	61.7 (11.3)
Sex (n)	
Female	n=28
Male	n=65
Stroke Location (n)	
Both	n=3
Left	n=38
Right	n=52

NOTES: Categorical data presented as median [min-max]; Continuous variables presented as mean ± SD; BBS-Berg Balance Scale; FGA-Functional Gait Assessment; FGA-LE-Functional Gait Assessment of the Lower Extremities;
* 11 missing data points.
† 3 missing data points.

Results

Participant demographics

Item-level data from 93 individuals (age=61.7±11.3 years; 28F/65M) were included (table 1). Individuals had a median score of 50 (range 29-56) on the BBS, and 16 (range 0-30) on the FGA. Median Fugl Meyer Assessment of the Lower Extremities score was 27 (range 14-34) and mean treadmill self-selected walking speed was 0.4±0.2 m/s. Mean time between the BBS and FGA measures was 23.89±26.35 days. Mean time since stroke was 59.0±62.2 months, and most participants had a stroke on the right hemisphere (n=52). All participants were ambulatory and community dwelling. The studies did not record gait aid or assistive device use for ambulation.

Unidimensionality and local dependence

The PCA showed that the RSM explained 68.1% of the variance, surpassing the 50% threshold. The first contrast made up 3.9% of the unexplained variance, meeting our criteria (<4%). While the corresponding eigenvalue of 2.78 was less than the desired threshold (<3), the disattenuated correlation between the primary and secondary dimensions was 0.96, meeting our criteria (>0.82). Based on these results, we determined the FBAM adequately met the assumption of unidimensionality. No local item dependence was found (table 2).

Rating scale structure

Rating scale structures for BBS and FGA items met all 3 essential criteria. Each category had >10 observations, the observed average measure increased as categories increased along the rating scale, and infit and outfit mean-squares were <2 (table 3).

Item and person statistics

"Stand on one foot" (BBS) had an infit mean-squared of 1.74 and a *z* score of 4.3, and an outfit mean-square of 1.76 with a *z* score of 4.1 (table 3), indicating the item did not meet criteria. Six individuals (~6%) demonstrated misfit to the RSM. Removal of "stand on one foot" resulted in 2 misfitting items (BBS-"Turing to Look Behind", and BBS-"Standing with Feet Together"). There was no difference in person measures between the 2 models.[21] Therefore, we determined that keeping the "stand on one foot" item did not degrade FBAM measurement properties, and its inclusion maintains fidelity with the original BBS. The FBAM had high person reliability (0.9 and Cronbach's alpha = 0.92) and high item reliability (0.96), with no ceiling or floor effect.

Separation index

The FBAM separated individuals into 4.44 statistically distinct strata, with a person separation index of 3.08.

Item difficulty hierarchy and score conversion

Figure 1 presents the person-item map of the FBAM with person ability (left side) in relation to item difficulty (right side) with converted FBAM scores, that is, from logits to a 100-point scale. Most FGA items (orange) had higher difficulty or required more balance ability than BBS items (blue). The hardest FBAM item was from the FGA; "gait with narrow base of support" (difficulty=76.61 SE 1.15). The easiest FGA item was "gait and pivot turn" (difficulty=52.67 SE 1.14). The hardest BBS item was "stand on one foot" (difficulty=61.68 SE

Table 2 Local item dependence

Item Pairs		Correlation
Standing unsupported (BBS)	Standing with eyes closed (BBS)	0.60
Standing with one foot in front (BBS)	Stand on one foot (BBS)	0.38
Gait with horizontal head turns (FGA)	Gait with vertical head turns (FGA)	0.32
Standing to sitting (BBS)	Transfers (BBS)	0.31
Standing to sitting (BBS)	Sitting to standing (BBS)	0.25
Gait with vertical head turns (FGA)	Gait and pivot turn (FGA)	0.25

NOTES: Only item pairs with standardized residual correlations >0.2 are presented.

Table 3 Rating scale structure of the BBS and the FGA

Score	Frequency Counts (%)	Observed Average	Infit Mean-Square	Outfit Mean-Square
BBS				
0	22 (2)	-6.24	1.10	1.06
1	58 (4)	-0.28	1.33	1.71
2	112 (9)	3.04	0.78	0.62
3	222 (17)	13.34	1.02	0.96
4	888 (68)	25.05	1.33	1.18
FGA				
0	116 (12)	-13.87	0.75	0.77
1	257 (28)	-2.88	0.90	0.83
2	366 (39)	5.71	0.92	0.96
3	191 (21)	14.20	1.02	1.02

0.81), and easiest FBAM item was from the BBS; "sitting unsupported" (difficulty=11.06 SE 11.90). The mean ability level of the sample was 63.75 SE 2.41. Twelve items were within 2 SE of the next easiest item (table 4), indicating similar levels of difficulty.

RSM fit statistics

Global fit statistics demonstrated RSM fit to the sample (table 5).

Score conversion formula

The following regression formula predicts FBAM interval measures on the 100-point scale from traditional BBS and FGA scores:

$$\textit{FBAM measure} = (\textit{BBS score} + \textit{FGA score}) * 0.76 + 16.55$$

For example, an individual with BBS score of 35 and FGA score of 9 will score 50 on the FBAM 100-point scale, indicating a "moderate" balance ability:

$$50 = (35 + 9) * 0.76 + 16.55$$

Discussion

The purpose of this investigation was to generate a single comprehensive measure of balance ability across the continuum of functional mobility. We combined the BBS and FGA using RM Theory to examine the tool's measurement properties with the RSM in a chronic stroke population. We found that the RSM supported the measurement properties of the FBAM, and the FBAM itself demonstrated sufficient rating scale characteristics, adequate item and person fit, and satisfactory item hierarchy in line with clinical expectations.

Previous work assessing the BBS using RM Theory collapsed the rating-scale into 3 categories and found that the psychometric properties were nearly identical to that of the five-category BBS,[22] however, this group did observe an increase in a ceiling effect. The psychometric properties of the FGA have only been evaluated in older, community-dwelling, neurologically healthy adults.[23] There was no ceiling effect and all items, except "walking on a level surface", met clinical expectations of item difficulty. We hypothesized that the original rating scales for the BBS and FGA could be analyzed together since the rating scales are reflective of task performance for each item with 0 (BBS and FGA) indicating poor performance and 4 (BBS) or 3 (FGA) indicating higher task performance. We did not find any concerns with the original BBS and FGA rating scale structures confirming our hypothesis. This allowed us to maintain fidelity with traditional performance scoring which we propose will improve clinical adoption because of clinician familiarity with BBS and FGA scoring systems. We found that the previously observed BBS ceiling effect can be eliminated by combining the 2 instruments. We also found that the item difficulty hierarchy supported clinical expectations of how balance ability demands increase with increasing functional mobility task difficulty.

Previous work demonstrated the possibility of constructing a single measure of balance ability from multiple measures.[3] We built on their work by combining the BBS and FGA to measure balance from "sitting" to "stair" performance. In contrast, we selected the BBS and FGA because these measures are commonly used in practice and clinicians are likely to be familiar with item scoring, thus facilitating clinical adoption. PCA results supported our hypothesis, indicating that the FBAM measures 1 latent construct: balance ability across the functional mobility spectrum. These results dispelled concerns that the BBS and FGA were measuring 2 separate constructs that is, sitting and standing balance (BBS), and walking and stair balance (FGA). Importantly, the original BBS and FGA scoring system and items were maintained, setting the stage for facilitating clinical translation.

Person and item fit

Our results indicate that "stand on one foot" misfit the RSM. BBS items "stand on one foot" and "standing with feet together" for post-stroke individuals,[8] and "sitting unsupported" and "standing unsupported" for a heterogenous neurologic population,[3] and the FGA item "gait with a narrow base of support" for community-dwelling older adults[7] have all been previously removed from analyses due to

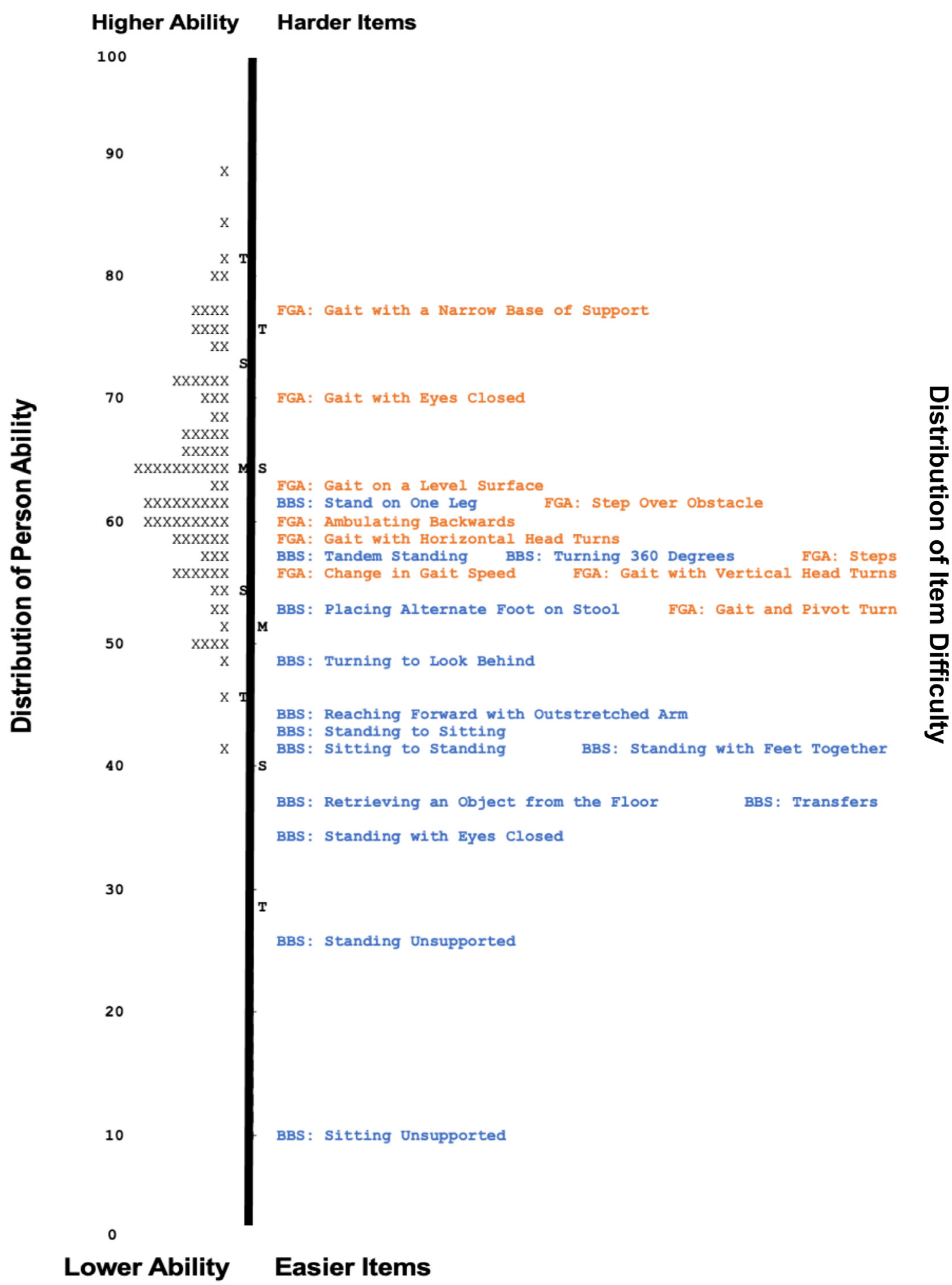

Fig 1 Key: M = mean, S = 1 standard deviation, T = 2 standard deviations, X = 1 individual; Note: Items on the same line share a similar difficulty level.

misfit. However, compared to these previous analyses, we investigated the effect of misfit on measurement quality (ie, person measures) before removing items. We found that removing the item "stand on one foot" did not degrade measurement properties, and person measures did not significantly change outside of a 95% confidence interval. Therefore, we opted to keep the item because its inclusion helps maintain similarity with current practice and it does not degrade measurement.

Item difficulty hierarchy

Item maps and measures reflected our hypothesized hierarchy difficulty, which was in line with clinical and theoretical reasoning; that balance ability would increase with functional task difficulty. Certain items in the FBAM overlapped based on each item's model SE (ie, within 2 SE of another item), and may therefore have a similar difficulty level (eg, BBS-"stand on one foot", BBS-"turning 360 degrees", and

Table 4 Items in measure order

Item	Measure	Model Standardized Error	Infit		Outfit	
			Mean-Square	z-Score	Mean-Square	z-Score
Gait with narrow base of support (FGA)	76.61	1.15	1.23	1.47	1.03	0.22
Gait with eyes closed (FGA)	69.72	1.05	1.12	0.89	1.10	0.70
Gait level surface (FGA)*	63.04	1.03	0.63	-3.11	0.65	-2.86
Step over obstacle (FGA)*	61.90	1.04	1.10	0.73	1.07	0.57
Stand on one foot (BBS)*	61.68	0.81	1.74	4.31	1.76	4.08
Ambulating backwards (FGA)	60.41	1.04	0.59	-3.39	0.59	-3.35
Gait with horizontal head turns (FGA)*	58.03	1.07	0.77	-1.72	0.92	-0.55
Steps (FGA)*	57.50	1.07	0.65	-2.76	0.66	-2.63
Turning 360 degrees (BBS)*	57.18	0.84	0.77	-1.63	1.03	0.24
Standing with one foot in front (BBS)*	56.86	0.84	1.00	0.02	1.06	0.39
Gait with vertical head turns (FGA)*	55.71	1.09	0.91	-0.60	0.88	-0.74
Change in gait speed (FGA)	55.53	1.10	0.98	-0.06	0.99	0.00
Placing foot on alternate stool (BBS)*	52.79	0.91	1.24	1.44	0.97	-0.09
Gait and pivot turn (FGA)	52.67	1.14	1.15	1.03	1.13	0.82
Turning to look behind (BBS)	48.43	1.05	1.06	0.42	1.38	1.31
Reaching forward with outstretch arm (BBS)	44.71	1.23	1.10	0.52	1.58	1.50
Standing to sitting (BBS)*	42.39	1.37	0.96	-0.12	0.80	-0.36
Sitting to standing (BBS)*	41.16	1.47	1.00	0.10	1.83	1.61
Standing with feet together (BBS)	40.82	1.49	1.42	1.62	1.02	0.19
Retrieving object from floor (BBS)*	37.55	1.80	0.62	-1.40	0.46	-1.01
Transfers (BBS)	37.04	1.85	0.74	-0.88	0.43	-1.04
Standing with eyes closed (BBS)	33.84	2.27	0.73	-0.69	0.30	-1.17
Standing unsupported (BBS)	26.28	3.82	0.80	-0.16	0.45	-0.41
Sitting unsupported (BBS)	11.06	11.90	Minimum measure†			

* Items are within 2 standard errors of the next easiest item on the hierarchy (ie, items directly below).
† In-fit and Out-fit cannot be calculated for items with minimum or maximum measure values.

FGA-"steps"). While it can be debated if overlapping items share the same difficulty level, what is most important is that there are items that cover the spread of person ability, which is the case in our results. The person mean of 63.75 SE 2.41 indicated that the sample performed higher than the mean difficulty of the combined instrument (FBAM item mean was centered at 50), however there was no reported ceiling effect, indicating that the spread of the items for this population was sufficient.

Separation

We calculated 4 statistically distinct strata, or 4 balance ability levels, for the FBAM (via the calculated value of 4.44). This distribution would be useful to identify patient subgroups and observe progression through each strata.

Future practice and research

Investigators should consider testing the measurement properties of a shorter FBAM because of the overlap of items observed. If overlapping items represent a similar difficulty level, it may be that a person's ability level can be measured using only 1 of the overlapping items, thus reducing patient and therapist assessment burden. An instrument with fewer items could take the form of either a shorter instrument, or computerized adaptive test. Based on the scoring of each item on the FBAM, a computerized adaptive test would adapt to an individual's performance.

Additionally, barriers and facilitators to the implementation of an instrument derived from current clinical measures (eg, FBAM), as opposed to a completely new instrument, should be investigated. The inclusion of a score conversion equation is intended to facilitate interpretability of scores

Table 5 Andrich Rating Scale Model fit indicators

Indicator	Statistic	Degrees of Freedom	*P* value
Log-Likelihood chi-squared	3274.02	2111	<0.001
Pearson Global chi-squared	2053.57	2111	<0.001
Akaike Information Criterion (AIC)	3330.02		
Schwarz Bayesian Information Criterion (BIC)	3488.73		
Global Root-Mean-Square Residual	0.59		

and can be implemented in such a way that will create a common scoring system.

Study Limitations

First, our study findings have limited generalizability to persons who are older and more functionally limited. This is due to the younger mean age of our cohort compared to other cohorts,[24] and that our sample was community dwelling and ambulatory. Secondly, although traditional scores and interval measures are correlated there is error associated when using linear regression to convert scores to interval measures. Score conversion error is lower near the center of the measure and higher at the extreme ends. Thirdly, data were retrospective, and only included persons consented in previous studies causing potential selection bias. Lastly, BBS and FGA scores were not always collected within the same study limiting our ability to control for potential changes in functional ability between test points. We minimized this concern by using individuals in the chronic phase of stroke, limiting time between test dates, and selecting data from the initial study date.

Conclusion

The FBAM demonstrated sufficient rating scale characteristics, item and person fit, and satisfactory item hierarchy in line with clinical expectations. The FBAM additionally separated our sample into 4 distinct strata, or levels of balance ability. We recommend investigation into the measurement properties the FBAM with fewer items, and the FBAM's implementation in stroke rehabilitation.

Suppliers

a. R v.4.1.3; R Core Team.
b. Winsteps v.5.2.4; Linacre, J.M.

Corresponding author

Jasmine Cash, MS, Department of Health Sciences and Research, Medical University of South Carolina, 77 President St., Charleston, SC, USA 29425. *E-mail address:* cashja@musc.edu.

Acknowledgments

We thank Heather Knight and the NIH COBRE Stroke Recovery research staff for data archival and management.

References

1. Tyson SF, Connell LA. How to measure balance in clinical practice. A systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions. Clin Rehabil 2009;23:824-40.
2. Liao Z, Quintana Y. Challenges to global standardization of outcome measures. AMIA Jt Summits Transl Sci Proc 2021: 404-9.
3. La Porta F, Caselli S, Susassi S, Cavallini P, Tennant A, Franceschini M. Is the Berg Balance Scale an internally valid and reliable measure of balance across different etiologies in neurorehabilitation? A revisited Rasch analysis study. Arch Phys Med Rehabil 2012;93:1209-16.
4. Moore JL, Potter K, Blankshain K, Kaplan SL, O'Dwyer LC, Sullivan JE. A core set of outcome measures for adults with neurologic conditions undergoing rehabilitation: a clinical practice guideline. J Neurol Phys Ther 2012;42:174-220.
5. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. Phys Ther 2008;88:559-66.
6. Alghadir AH, Al-Eisa ES, Anwer S, Sarkar B. Reliability, validity, and responsiveness of three scales for measuring balance in patients with chronic stroke. BMC Neurol 2018;18:141.
7. Van Bloemendaal M, Bout W, Bus SA, Nollet F, Geurts AC, Beelen A. Validity and reproducibility of the Functional Gait Assessment in persons after stroke. Clin Rehabil 2019;33:94-103.
8. Straube D, Moore J, Leech K, Hornby TG. Item analysis of the berg balance scale in individuals with subacute and chronic stroke. Top Stroke Rehabil 2013;20:241-9.
9. Miyata K, Tamura S, Kobayashi S, Takeda R, Iwamoto H. Berg Balance Scale is a valid measure for plan interventions and for assessing changes in postural balance in patients with stroke. J Rehabil Med 2022;54:jrm00359.
10. Bond T. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 3rd ed. Routledge; 2015 Third Edition.
11. Mallinson T, Kozlowski AJ, Johnston MV, et al. Rasch Reporting Guideline for Rehabilitation Research (RULER): the RULER Statement. Arch Phys Med Rehabil 2022;103:1477-86.
12. HHS.gov. Revised Common Rule. 2017. Available at: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html. Accessed September 12, 2022.
13. Linacre JM. Sample size and item calibration stability. Rasch Meas Trans 1994;7.
14. Lee OK. Variance in mathematics and reading across grades: grade equivalents and logits. Rasch Meas Trans 1992;6.
15. Linacre JM. Winsteps® Rasch measurement computer program User's Guide. Beaverton, Oregon: Winsteps.com
16. Linacre JM. A User's Guide to WINSTEPS® MINISTEP Rash-Model Computer Programs. Program Manual 5.1.5
17. Linacre JM. Rasch power analysis: size vs. significance: standardized chi-square fit statistic. Rasch Meas Trans 2003;17:918.
18. Wright BD, Linacre JM. Reasonable mean-square fit values. Rasch Measurement Transactions 1994;8:370.
19. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293-307.
20. Wright BD, Masters GN. Number of person or item strata: (4*separation + 1)/3. Rasch Meas Trans 2002;16:888.
21. Linacre JM. When to stop removing items and persons in Rasch misfit analysis? Rasch Meas Trans 2010;23:1241.
22. Wang CH, Hsueh IP, Sheu CF, Yao G, Hsieh CL. Psychometric properties of 2 simplified 3-level balance scales used for patients with stroke. Phys Ther 2004;84:430-8.
23. Beninato M, Ludlow LH. The Functional Gait Assessment in older adults: validation through Rasch Modeling. Phys Ther 2016;96:456-68.
24. Kissela BM, Khoury JC, Alwell K, et al. Age at stroke. Neurology 2012;79:1781-7.