Published in final edited form as:

Arch Phys Med Rehabil. 2016 April ; 97(4): 582–9.e2. doi:10.1016/j.apmr.2015.12.012.

Dimensionality and item-difficulty hierarchy of the lower-extremity Fugl-Meyer assessment in individuals with sub-acute and chronic stroke

Chitralakshmi K. Balasubramanian, PhD, PT¹, Chih-Ying Li, MS², Mark G. Bowden, PhD, PT^{2,3}, Pamela W. Duncan, PhD, PT, FAPTA, FAHA⁴, Steven A. Kautz, PhD^{2,3}, and Craig A. Velozo, PhD, OTR/L⁵

¹Department of Clinical and Applied Movement Sciences, University of North Florida, Jacksonville, FL

²Department of Health Science and Research and Division of Physical Therapy, College of Health Profession, Medical University of South Carolina. Charleston, SC

³Ralph H. Johnson VA Medical Center, Charleston, SC

⁴Department of Neurology, Wake Forest University, Winston-Salem, NC

⁵Division of Occupational Therapy, College of Health Profession, Medical University of South Carolina. Charleston, SC

Abstract

Objective—To investigate the dimensionality and item-difficulty hierarchy of the lower extremity section of the Fugl-Meyer assessment (FMA-LE).

Design—Secondary analyses of data pooled from four existing datasets: a Phase III randomized controlled trial investigating the effectiveness of body weight support and a treadmill for rehabilitation of walking post-stroke and three cross-sectional studies investigating the link between impaired motor performance post-stroke and walking.

Setting—University research centers and rehabilitation centers.

CORRESPONDING AUTHOR: Chitralakshmi K. Balasubramanian, PT, PhD, Associate Professor, Department of Clinical and Applied Movement Sciences, University of North Florida, 1 UNF Drive, Jacksonville, Florida 32224, Office no: (904) 620-3849, c.k-balasubramanian@unf.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

CONFLICT OF INTEREST

There is no conflict of interest associated with this study.

PRESENTATION

Partial results were presented in poster format at the 90th American Congress of Rehabilitation Medicine Annual Conference (Progress in Rehabilitation Research), Orlando, FL, November 2013 and the abstract was published (refer to citation below) Balasubramanian CK, Kautz SA, Velozo CA. Poster 44 Construct validity of the Fugl-Meyer assessment of lower extremity to evaluate motor impairment post stroke. *Archives of Physical Medicine and Rehabilitation*, Volume 94, Issue 10, Pages e27, October 2013.

Participants—A pooled sample of 535 individuals with a stroke (age = 61.91 ± 12.42 years; male = 313).

Interventions—Not applicable.

Main Outcome Measure(s)—Confirmatory factor analyses (CFA) and Rasch residual Principal component analysis (PCA) investigated the dimensionality of the FMA-LE. The Rasch analysis rating scale model (RSM) investigated item-difficulty hierarchy of the FMA-LE.

Results—The CFA showed adequate fit of a three-factor model with 2/3 indices (CFA = 0.95; TLI = 0.94; RMSEA = 0.124) showing good model fit. Rasch PCA showed that removal of the reflex and coordination items explained 90.8% of variance in the data, suggesting that the abnormal synergy items contributed to the measurement of a unidimensional construct. However, RSM results revealed deviations in the item-difficulty hierarchy of the unidimensional abnormal synergy items from the originally proposed stepwise sequence of motor recovery.

Conclusion(s)—Our findings suggest that the FMA-LE might represent a multidimensional construct challenging the use of a total score of the FMA-LE to predict lower extremity motor recovery. Removal of the misfit items resulted in creation of a unidimensional scale comprised of the abnormal synergy items. However, this unidimensional scale deviates from the originally proposed hierarchical ordering.

Keywords

Lower Extremity; Rehabilitation; Stroke

The Fugl-Meyer Assessment of the lower extremity (FMA-LE) is a widely used impairment-based outcome measure used to evaluate LE motor recovery and predict functional recovery in individuals with stroke.¹⁻⁴ The psychometric properties of the FMA-LE have been extensively tested using classical testing methods and the scale has shown good validity,⁴⁻⁸ excellent reliability^{9,10} and internal consistency.¹¹ However classical test methods investigate the assessment as a whole, whereas contemporary testing methods (e.g. item response theory, IRT) evaluate the item-level psychometric properties. The need for an item level analysis is apparent when examining the construct underlying the FMA-LE.

Twitchell and Brunnstrom established the conceptual framework that underlies the FMA-LE.^{12,13} Motor recovery of the LE was characterized to progress through a stepwise predictable sequence, from initial flaccid paralysis, to recovery of reflex activity, to the development of stereotypical synergistic voluntary movement and finally to the restoration of normal movement.^{12,13} The FMA-LE consists of 17 items arranged in a hierarchy (from reflex items, synergy items and coordination items), paralleling the hypothesized stepwise sequential recovery process. A total summed score of the 17 items is currently used in rehabilitation research trials to represent the level of LE motor recovery. However, the rationale for computation of a total score is lacking.

Reflex, abnormal synergistic movements and coordination items of the FMA-LE measure more than one construct of motor control and recovery. Summing the responses from distinct constructs of a multidimensional assessment obscures what the total score might represent. Reflexes are simple and automatic responses to sensory stimulation that are physiologically

controlled at lower levels of the central nervous system.¹⁴ Contrary to automatic reflexive movements, voluntary movements after a stroke described as abnormal synergy patterns (“within synergy”, “combining synergy” and “out of synergy”) in the FMA-LE selectively represent the voluntary control of isolated movement likely to be controlled at the cortical level (e.g., motor cortex).¹⁴ Lastly, the process of recovery for LE gross coordination (“speed”, “dysmetria” and “tremor”) is likely to be controlled at both cortical and sub-cortical levels (e.g., cerebellum and basal ganglia),¹⁴ suggesting that the recovery process of coordination may physiologically differ compared to reflexive behaviors and abnormal synergy movements. Therefore, it can be argued that the construct underlying FMA-LE is multidimensional, capturing at least three different aspects of motor control and recovery, including (a) involuntary reflexive behaviors, (b) voluntary movements (represented as moving within, combining and out of synergy patterns), hereby referred to as ‘abnormal synergy’ movements and (c) gross coordination.

Rasch analysis, an IRT based methodology, offers the advantage of testing dimensionality and item-level psychometric properties of the scale as opposed to testing the instrument as a whole (using total scores).¹⁵ In addition, Rasch analysis allows an investigation of the item-difficulty hierarchy of the scale. Earlier studies have used IRT methodologies to evaluate the FMA.^{16,17} For instance, a Rasch based short form was developed from the total FMA comprising of upper and lower extremities.¹⁶ A computerized adaptive test (CAT) of the FMA has also been developed comprising of both upper and lower extremity items.¹⁷ Since the lower-extremity items cannot be selectively administered from the CAT, the application of the relevant FMA-LE items in rehabilitation trials focused on LE recovery is limited.

This study aimed to systematically investigate the dimensionality, item-difficulty hierarchy and item-level measurement properties of the FMA-LE using Rasch analysis. We hypothesized that the FMA-LE comprises three different constructs of reflex, abnormal synergy movements and gross coordination, and that the item-difficulty hierarchy will deviate from the originally proposed stepwise progression of motor recovery.

METHODS

Participants

Retrospective data pooled from four rehabilitation studies^{18–21} were analyzed (on-line only appendix). All participants gave their informed consent before participation and the study protocols were approved by the institutional review boards at the respective institutions.

Instrumentation

Trained physical therapists administered the FMA-LE using standardized testing procedures in all four studies. The FMA-LE assessment consists of 17 items and is the LE motor sub-section of the FMA, with 2 items measuring reflex activity, 11 items measuring synergistic movements and 3 items measuring coordination (Table 1). Scoring of each item is based on a 3-point ordinal scale (0 = *cannot perform*, 1 = *performs partially* and 2 = *performs fully*), except two reflex items (Table 1).

Data Analysis

Data Management—Item 1 ('Achilles reflex') and item 2 ('Patellar reflex') are measured on a 2-point ordinal scale as opposed to the rest of the items measured on a 3-point scale. Therefore, these items were rescored by assigning scores of '2' a score of '1' to be consistent with the rating of other items in the scale. The item 'normal reflexes' was deleted from all follow-up data analyses since it is not a dependent item and was scored based on the overall performances of other items. Sixteen items were used for confirmatory factor analysis (CFA) and a reduced set of 11 synergy items was used for principal components analysis (PCA) and Rasch analysis.

Psychometric Properties Analyses—Psychometric properties were analyzed in the following steps. First, a CFA (analyzed with Mplus 7.1),²² validated the hypothesized three-factor model. We chose weighted least squares means and variance adjusted (WLSMV) estimation as our model estimation method because the WLSMV is a robust estimator that does not assume normal distribution of variables, providing the best estimates when modelling categorical or ordered data.²³ The WLSMV provides model fit statistics such as comparative fit index (CFI), Tucker-Lewis index (TLI) and root mean square error of approximation (RMSEA) indices to identify the level of fit to the proposed model. We classified the items into three factors based on their descriptions: reflex (items 1–2), abnormal synergy (items 3–13) and coordination (items 15–17). The CFI, TLI and RMSEA indices explained the CFA results. The CFI/TLI compares the hypothesized model to a null model (with no factor at all); with benchmarks > 0.95 representing a good model fit, 0.90–0.95 representing a marginal model fit, and <0.90 representing a poor model fit.²⁴ The RMSEA compares expected values with actual values in a matrix after adjusting for sample size, a value less than 0.06 represents a good model fit.²⁴

We used the Rasch analysis rating scale model (RSM) to examine unidimensionality and item-difficulty hierarchy of the 11 abnormal synergy items. Rasch RSM was used because all the synergy items have the same 3-point rating scale structure. We also used Rasch residual PCA to examine unidimensionality of the scale. The PCA enables extraction of the primary (unidimensional) Rasch dimension enabling us to determine if the remaining residuals are meaningful or simply noise. The Rasch residual PCA can identify both the unidimensional Rasch dimension and the first contrast (that is, the first component in the correlation matrix of the residuals after extracting Rasch dimension).^{25–27}

Unidimensionality of a scale is suggested to be validated when the Rasch dimension explains 40% variance of the data, the first contrast of Rasch residual explains less than 5% variance of the data, and the eigenvalue of the first contrast is less than or equal to 2.0.^{25–27} Both CFA and the Rasch residual PCA were used to determine whether the latent trait underlying the data is unidimensional.²⁸ However, we recognize that the Rasch residual PCA method has the limitation of identifying only patterns in the data since the purpose of the PCA residuals method is to explain variance (primary Rasch measurement component and the remaining residual component) instead of constructing factors of the scale.

Following factor structure analysis (CFA and the Rasch residual PCA), we performed Rasch analysis to examine unidimensionality of the scale. In contrast to classical test theory, Rasch

analysis is an item-level analysis where the ability level of the person and the difficulty level of the items are placed on the same continuous variable providing statistics at the item and person level. Winsteps 3.75²⁹ was used to examine the item-level psychometric properties of the scale, including rating scale diagnosis statistics, person and item fit statistics, point-measure correlation, person separation reliability, person separation strata and ceiling/floor effects. The three main criteria for examining a rating scale included; a minimum of ten responses in each rating category, a monotonic pattern of category measure, and the outfit mean square for each rating scale category less than 2.0.³⁰ While the point-measure correlation is the Pearson correlation between the observations and the Rasch measures accounting for all observations, the expected point-measure correlation is the one between the expected observations and the Rasch measures. The person and item fit statistics were determined by the Infit mean square (Infit MNSQ) and Outfit mean square (Outfit MNSQ) value with a range of 0.6 to 1.4 and the value of standardized fit statistics (ZSTD) within ± 2 .³¹ The infit and outfit MNSQ values (fit statistics) are chi-square ratios that examine how well the empirical data fit the requirements of the ideal model.²⁸ Infit is an information-weighted fit statistics weighted by model variance, whereas outfit is an outlier-sensitive fit statistics.³² Person separation reliability, analogous to Cronbach's alpha, indicates the levels of person reliability across all items^{33,34} with the value of $\alpha = 0.90$ suggested for clinical application.³⁵ In addition, we calculated person strata using person separation index. Person separation index (Gp) is defined as true standard deviation (i.e., observed standard deviation² – average measurement error²) divided by averaged measurement error.³⁶ Person strata, calculated as $(4Gp + 1)/3$,³⁷ is the number of ability level with the center value three measurement errors apart, representing the level of person ability distinguished by all items. From a measurement perspective, instruments should be able to statistically divide a sample into at least three distinct person strata.³³ The ceiling/floor effects were determined by the item-person map,³⁸ using a standardized logit unit to place item-difficulty and person ability in the same continuum.

RESULTS

Pooled data from the four studies resulted in a total number of 535 individuals with stroke. The participant characteristics available from each study and of the pooled sample are summarized in the on-line only appendix.

3-factor Model Fit

The CFI/TLI values for the 3-factor model were 0.95/0.94, respectively. Whereas, the RMSEA value was 0.124. The reflex items showed high error variance (Patellar = 2584.41, Achilles = 176.84); followed by the coordination items (Tremor = 2.91; Dysmetria = 1.49; Speed = 9.41). The error variance for the 11 abnormal synergy items were low (0.06 – 0.31). Additionally, the reflex construct had only 2 items and the coordination construct had only 3 items; therefore, the remainder of the analyses focused on the measurement properties of the synergy construct (11 items).

Dimensionality, Item-difficulty hierarchy and Item-fit statistics of the 11 abnormal synergy items

The Rasch dimension of the 11 abnormal synergy items explained a total 90.8% of the variance in the data and no meaningful residuals were left after extracting this Rasch dimension. Additionally, the rating scale diagnostics met all three essential rating scale criteria. Point-measure correlations were also acceptable and the infit statistics for the 11 items fell within the guidelines (Table 2). However, 3 items had high outfit values, including item 7 'Within synergy – hip adduction' (Outfit MNSQ=2.28, ZSTD=4.6), item 3 'Within synergy – hip flexion' (Outfit MNSQ=1.92, ZSTD=2.8), and item 6 'Within synergy – hip extension' (Outfit MNSQ=1.90, ZSTD=2.7), Table 2. The difficulty level of the items ranged from -2.24 to 3.27 logits (5.51 logit range). Item 12 'out of synergy – knee flexion' (3.27 ± 0.09 logits) and item 13 'out of synergy – ankle dorsiflexion' (3.05 ± 0.09 logits) were the most difficult items. Conversely, item 8 'within synergy – knee extension' (-2.24 ± 0.14 logits) and item 6 'within synergy – hip extension' (-1.63 ± 0.12 logits) were the easiest items.

The person ability ranged from -4.54 to 5.01 , with a 9.55 logit range (Figure 1). Person ability was one standardized deviation higher than item-difficulty. The mean person ability was 1.73 logits higher than the mean item-difficulty. Ten percent of the sample ($51/535=9.5\%$) had the maximum raw score (highest ability), indicating a ceiling effect (Figure 1). Additionally, the person separation reliability index was 0.91, person separation index was 2.37 and person separation strata were 3.49 for the 11 abnormal synergy items.

DISCUSSION

The purpose of this study was to investigate the dimensionality and item-difficulty hierarchy of the FMA-LE. Application of IRT-based methodologies supported our hypothesis that FMA-LE represents a multidimensional construct, reflecting three distinct constructs: reflex, abnormal synergy and coordination. After removing the reflex and coordination items, the 11 abnormal synergy items of the FMA-LE scale contributed to the measurement of a single construct. However, the difficulty order of the abnormal synergy items deviated from the stepwise recovery process proposed by others. Our findings have implications for using the total score of the FMA-LE assessment to predict lower extremity motor recovery.

Results from the CFA and Rasch PCA supported our hypothesis of multidimensionality. First, the CFA showed adequate to good fit of a 3-factor model with 2/3 indices showing good model fit. Specifically, the CFI/TLI values of 0.95/0.94 showed good model fit whereas the RMSEA value of 0.124 (>0.06) did not meet the model fit criteria. Second, the Rasch PCA revealed that after removal of the reflex and coordination factors, the abnormal synergy factor explained a substantial (90.8%) variance in the data, implying that the reflex and coordination items measure a different construct when compared to the abnormal synergy items. Therefore, based on our findings of multidimensionality, we suggest discontinuing the use of the total score (34) of the FMA-LE to measure motor recovery. However, our findings of multidimensionality should be considered preliminary since not all the model fit criteria were met for the hypothesized 3-factor model. Future studies are warranted to confirm our findings.

Multidimensionality of the FMA is also implicitly suggested in other reports.^{16,39} When evaluating the measurement characteristics of the upper extremity section of the FMA, Woodbury and colleagues showed that the upper extremity reflex items misfit the construct measured by the rest of the items suggesting a multidimensional assessment.³⁹ Involuntary reflexive behaviors are also physiologically different from voluntary movements (such as those movements occurring within, combining and out of synergies).¹⁴ While developing a short form of the FMA, Hsieh and colleagues reported that not only the reflex items but 2/3 of the coordination items ('Tremor' and 'Dysmetria') had high fit statistics and subsequently removed them from the short form.¹⁶ Notably, coordination as captured by the FMA-LE is also considerably different from the intra- and inter-limb coordination required for lower extremity functions, such as walking.^{40–42} Therefore, the multidimensionality of FMA-LE has some literature support.

While our results suggested that the FMA-LE was multidimensional, the 11 abnormal synergy items seemed unidimensional. However, the item-difficulty hierarchy of the unidimensional abnormal synergy items revealed several deviations from the originally proposed stepwise sequence of motor recovery.¹² For instance, Item 12 'Out of synergy – knee flexion' was calibrated as the most challenging item contrary to the original hierarchy. This item requires flexing the affected knee in standing and its high difficulty level can be explained by the additional balance demands to support weight on one leg. Another deviation from the original hierarchy was that of the ankle items. All ankle items irrespective of whether movement occurred within, combining or outside of synergies clustered towards the more difficult end of the hierarchy. Contemporary motor control literature provides some rationale for why the ankle items, irrespective of the synergy pattern, maybe more challenging than the proximal hip and knee movements. Ankle dorsiflexor and plantarflexor muscle groups receive direct corticospinal innervation during voluntary activities.⁴³ Therefore, damage to the corticospinal tract may be expected to cause greater impairment in ankle control compared to proximal hip and knee joints. Our finding of distal ankle control being harder than proximal hip and knee control also conflicts with the proximal-to-distal sequence suggested in the earlier motor recovery works.⁴⁴

In addition to the deviations in item-difficulty hierarchy of the abnormal synergy items, several of these items were calibrated at similar difficulty levels suggesting measurement redundancy. Additionally, the abnormal synergy items revealed moderate internal consistency, had substantial ceiling effect and the discriminatory ability seemed low (person separation index = 2.36). Therefore, while the abnormal synergy items were unidimensional, measurement could be improved by revising current items along with potentially addition of new items. Nonetheless, it is notable that the investigation of the staged hierarchical progression in the FMA-LE is inherently limited by the structure and the procedures of the assessment. There are an unequal number of items across the progressive stages (i.e. only 2 items each represent the combined and out-of-synergy movements whereas 7 items represent the within synergy movements). Moreover, the postural transitions from supine to sitting and ultimately to standing that are required for assessing the successive stages of motor recovery additionally create balance demands. Nevertheless, until an improved version is developed, the abnormal synergy score may remain of value for assessing LE motor control. Therefore,

we suggest discontinuing the use of the total score of the FMA-LE and use the abnormal synergy score to predict lower extremity motor recovery.^{19,40,45}

Our suggestions are different from others. Crow and colleagues investigated the hierarchical properties of the motor sections of both upper and lower extremity FMA in individuals with acute and chronic stroke and supported the use of a total score of the FMA.^{46,47} The inconsistent results may primarily be due to different study methodologies Crow and colleagues used the Guttman scale analysis as opposed to CFA and Rasch analysis.^{46,47} They propose that the Guttman scale analysis is suitable to investigate the hierarchical properties of the FMA due to the underlying nature of the construct, i.e. staged progression of motor recovery. However, Crow et al.^{46,47} do not take into account the concept of a linear continuum (which they refute) and staged progressions (which they support) are separate. Linearity reflects whether all items measure a linear construct whereas staged progression refers to where individuals lie on that linear construct.

Despite the differences in methodology between our study and Crow et al.'s work,^{46,47} there are many similarities in our results. Item 12 'out of synergy - knee flexion' was the most challenging item in both our and Crow et al.'s study. Furthermore, similar to our findings, Crow et al. reported that the ankle items were the next most challenging items clustering towards the harder end of the hierarchy. Lastly, both our findings and Crow et al.'s study found that the within-synergy hip and knee items were the least challenging. Individual hip and knee items deviated from their strict ordering within-synergy in both studies. Only the ordering of one item (Item 10 'combined synergy – knee flexion') differed between our results. In our results Item 10 was one of the easier items, whereas in Crow et al.'s study it was one of the more challenging items. Despite their findings of deviations in the rank ordering similar to our results, Crow et al. suggest the continued use of a total score. Contrarily, our findings of misfit items conflict with the use a total score of the FMA-LE.

Study limitations

Note that, since this is a retrospective study, the generalizability of the study is limited. Additionally, several assessors were involved in the four studies which could attribute to instrumentation error. We also recognized that our proposed 3-factor model included two factors (i.e. reflex and coordination) with low number of items (i.e. two items for reflex factor and three items for coordination factor). Thus, the items for these two factors may not be able to completely capture the latent trait of each.

CONCLUSIONS

Our results are suggestive of multidimensionality of the FMA-LE, conflicting the use of a total FMA-LE score to measure motor recovery. Particularly, use of a total score from a multidimensional assessment may decrease the construct validity of the FMA-LE, contributing to lower responsiveness of the scale to interventions aimed at improving lower extremity motor recovery. Removal of the misfit items of reflex and coordination resulted in 11 items (assessing abnormal synergy) that contribute to the measurement of a single construct. While these 11 abnormal synergy items were unidimensional, the hierarchical ordering of item-difficulty deviated from the 'stepwise' sequence of motor recovery

originally proposed by the instrument developers suggesting improvement in these items. In the interim, the abnormal synergy score may prove to be useful to measure lower-extremity motor recovery.

Acknowledgments

FINANCIAL SUPPORT FOR INCLUDED STUDIES

Study # 1: This material was based on work supported by the National Institute of Neurologic Diseases and Stroke and the National Center for Medical Rehabilitation Research (grant RO1 NS050506).

Study # 2: Data collection was supported by HD46820.

Study # 3: Data collection was supported by the Rehabilitation Research & Development Service of the Veteran's Affairs.

Study # 4: Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20-GM109040.

Abbreviations

FMA-LE	Fugl-Meyer assessment of lower extremity
IRT	Item-response theory
CFA	Confirmatory factor analysis
PCA	Principal component analysis
CFI	Comparative fit index
TLI	Tucker-Lewis index
RMSEA	Root mean square error of approximation
RSM	Rating scale model
MNSQ	Mean square error
ZSTD	Standardized fit statistics
Gp	Person separation index

References

1. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scand J Rehabil Med.* 1975; 7:13–31. [PubMed: 1135616]
2. Gladstone DJ, Danells CJ, Black SE. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabil Neural Repair.* 2002; 16:232–240. [PubMed: 12234086]
3. Shelton FD, Volpe BT, Reding M. Motor impairment as a predictor of functional recovery and guide to rehabilitation treatment after stroke. *Neurorehabil Neural Repair.* 2001; 15:229–237. [PubMed: 11944745]

4. Hsieh Y, Wu C, Lin K, Chang Y, Chen C, Liu J. Responsiveness and validity of three outcome measures of motor function after stroke rehabilitation. *Stroke*. 2009; 40:1386–1391. [PubMed: 19228851]
5. Malouin F, Pichard L, Bonneau C, Durand A, Corriveau D. Evaluating motor recovery early after stroke: comparison of the Fugl-Meyer Assessment and the Motor Assessment Scale. *Arch Phys Med Rehabil*. 1994; 75:1206–1212. [PubMed: 7979930]
6. Nadeau S, Arsenault AB, Gravel D, Bourbonnais D. Analysis of the clinical factors determining natural and maximal gait speeds in adults with a stroke. *Am J Phys Med Rehabil*. 1999; 78:123–130. [PubMed: 10088586]
7. Dettmann MA, Linder MT, Sepic SB. Relationships among walking performance, postural stability, and functional assessments of the hemiplegic patient. *Am J Phys Med*. 1987; 66:77–90. [PubMed: 3578493]
8. Bernspång B, Asplund K, Eriksson S, Fugl-Meyer AR. Motor and perceptual impairments in acute stroke patients: effects on self-care ability. *Stroke*. 1987; 18:1081–1086. [PubMed: 3686581]
9. Duncan PW, Propst M, Nelson SG. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Phys Ther*. 1983; 63:1606–1610. [PubMed: 6622535]
10. Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke. *Phys Ther*. 1993; 73:447–454. [PubMed: 8316578]
11. Wood-Dauphinee SL, Williams JI, Shapiro SH. Examining outcome measures in a clinical study of stroke. *Stroke*. 1990; 21:731–739. [PubMed: 2339453]
12. Twitchell TE. The restoration of motor function following hemiplegia in man. *Brain*. 1951; 74:443–480. [PubMed: 14895765]
13. Brunnstrom S. Motor testing procedures in hemiplegia: based on sequential recovery stages. *Phys Ther*. 1966; 46:357–375. [PubMed: 5907254]
14. Wolpert, DM.; Pearson, J.; Ghez, CP. *Principles of Neural Science*. 5th. McGraw-Hill Professional; 2012. The organization and planning of movement.
15. Hambleton RK, Jones RW. An NCME Instructional Module on: Comparison of classical test theory and item response theory and their applications to test development. *Educ Meas Issues Pract*. 2005; 12:38–47.
16. Hsieh Y-W, Hsueh I-P, Chou Y-T, Sheu C-F, Hsieh C-L, Kwakkel G. Development and validation of a short form of the Fugl-Meyer motor scale in patients with stroke. *Stroke*. 2007; 38:3052–3054. [PubMed: 17916763]
17. Hou W-H, Shih C-L, Chou Y-T, et al. Development of a computerized adaptive testing system of the Fugl-Meyer motor scale in stroke patients. *Arch Phys Med Rehabil*. 2012; 93:1014–1020. [PubMed: 22440481]
18. Duncan PW, Sullivan KJ, Behrman AL, et al. Body-weight-supported treadmill rehabilitation after stroke. *N Engl J Med*. 2011; 364:2026–2036. [PubMed: 21612471]
19. Clark DJ, Ting LH, Zajac FE, Neptune RR, Kautz SA. Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke. *J Neurophysiol*. 2010; 103:844–857. [PubMed: 20007501]
20. Worthen LC, Kim CM, Kautz SA, Lew HL, Kiratli BJ, Beaupre GS. Key characteristics of walking correlate with bone density in individuals with chronic stroke. *J Rehabil Res Dev*. 2005; 42:761–768. [PubMed: 16680613]
21. Routson RL, Kautz SA, Neptune RR. Modular organization across changing task demands in healthy and poststroke gait. *Physiol Rep*. 2014; 2:e12055. [PubMed: 24963035]
22. Muthen, BO.; Muthen, LK. *MPlus User's Guide*. 6th. Los Angeles, CA: 2006.
23. Brown, TA. *Confirmatory Factor Analysis for Applied Research*. New York: Guildford; 2006.
24. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model A Multidiscip J*. 1999; 6:1–55.
25. Linacre JM. Rasch model estimation: further topics. *J Appl Meas*. 2004; 5:95–110. [PubMed: 14757994]

26. Linacre JM. Predicting responses from Rasch measures. *J Appl Meas.* 2010; 11:1–10. [PubMed: 20351444]
27. Linacre, JM. A User's Guide to Winsteps Ministep 3.70.0: Rasch Model Computer Programs. Chicago, IL: Winsteps; 2012.
28. Bond, TG.; Fox, CM. Applying the Rasch Model : Fundamental Measurement in the Human Sciences. 2nd. Lawrence Erlbaum Associates Publishers; 2007. Chapter 12: Model fit and unidimensionality.
29. Linacre, JM. Winsteps 3.75 Rasch Measurement Computer Program. 2006.
30. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas.* 2002; 3:85–106. [PubMed: 11997586]
31. Bond, TG.; Fox, CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, N.J.: Lawrence Erlbaum Associates Publishers; 2001. Model Fit and Unidimensionality; p. 235-260.
32. Linacre JM. What do Infit and Outfit, Mean-square and Standardized mean? *Rasch Meas Trans.* 2002; 16:878.
33. Fisher WJ. Reliability, separation, strata statistics. *Rasch Meas Trans.* 1992; 6:328.
34. Linacre JM. Prioritizing misfit indicators. *Rasch Meas Trans.* 1995; 9:422–423.
35. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ.* 1997; 314:572. [PubMed: 9055718]
36. Wright BD. Number of Person or Item Strata: $(4 \times \text{Separation} + 1)/3$. *Rasch Meas Trans.* 2002; 16:888.
37. Wright, BD.; Masters, GN. Rating Scale Analysis Rasch Measurement. MESA Press; 1982.
38. Velozo CA, Wang Y, Lehman L, Wang J-H. Utilizing Rasch measurement models to develop a computer adaptive self-report of walking, climbing, and running. *Disabil Rehabil.* 2008; 30:458–467. [PubMed: 18297500]
39. Woodbury ML, Velozo CA, Richards LG, Duncan PW, Studenski S, Lai S-M. Dimensionality and construct validity of the Fugl-Meyer Assessment of the upper extremity. *Arch Phys Med Rehabil.* 2007; 88:715–723. [PubMed: 17532892]
40. Bowden MG, Clark DJ, Kautz SA. Evaluation of abnormal synergy patterns poststroke: relationship of the Fugl-Meyer Assessment to hemiparetic locomotion. *Neurorehabil Neural Repair.* 2010; 24:328–337. [PubMed: 19794132]
41. Kautz SA, Patten C. Interlimb influences on paretic leg function in poststroke hemiparesis. *J Neurophysiol.* 2005; 93:2460–2473. [PubMed: 15590727]
42. Thompson AK. Interlimb coordination during locomotion: finding available neural pathways and using them for gait recovery. *Clin Neurophysiol.* 2012; 123:635–637. [PubMed: 21930427]
43. Capaday C, Lavoie BA, Barbeau H, Schneider C, Bonnard M. Studies on the Corticospinal Control of Human Walking. I. Responses to Focal Transcranial Magnetic Stimulation of the Motor Cortex. *J Neurophysiol.* 1999; 81:129–139. [PubMed: 9914274]
44. Gordon, C. Assumptions underlying physical therapy interventions, theoretical and historical perspectives. In: Carr, JH.; Shepherd, RB., editors. *Movement Science: Foundations for Physical Therapy in Rehabilitation.* 2nd. Gaithersburg: Aspen; 2000.
45. Brown DA, Kautz SA. Increased workload enhances force output during pedaling exercise in persons with poststroke hemiplegia. *Stroke.* 1998; 29:598–606. [PubMed: 9506599]
46. Crow JL, Harmeling-van der Wel BC. Hierarchical properties of the motor function sections of the Fugl-Meyer assessment scale for people after stroke: a retrospective study. *Phys Ther.* 2008; 88:1554–1567. [PubMed: 18927197]
47. Crow JL, Kwakkel G, Bussmann JBJ, Goos JAG, Harmeling-van der Wel BC. Are the hierarchical properties of the Fugl-Meyer assessment scale the same in acute stroke and chronic stroke? *Phys Ther.* 2014; 94:977–986. [PubMed: 24677254]

- The lower-extremity fugl-meyer assessment represents a multi-dimensional construct.
- Our results challenge use of the total score to predict lower-extremity motor recovery.
- The abnormal synergy items of the lower-extremity fugl-meyer are unidimensional.
- The abnormal synergy items deviate from the originally proposed hierarchical order.

Software Suppliers List

Winsteps 3.75

Winsteps software was developed and maintained by John Michael Linacre, Ph.D. He is the Research Director of [Winsteps.com](http://winsteps.com) software and an adjunct professor at the University of the Sydney. He can be reached via Rasch Measurement Forum (<http://raschforum.boards.net/>). Winsteps company is located at 14525 SW Millikan Way # 27615, Beaverton, Oregon 97005, United States

Mplus 7.1

Mplus software was developed and maintained by Dr. Linda Muthén and Dr. Bengt Muthén. They can be reached via Mplus Discussion Platform (<https://www.statmodel.com/cgi-bin/discus/discus.cgi>). More detailed information of the Mplus software can be found at the Muthén & Muthén Home page (<http://www.statmodel.com/>).

Muthén & Muthén company is located at 3463 Stoner Avenue, Los Angeles, CA 90066, United States. Below are the company contact information:

Tel: (310) 391-9971; Fax: (310) 391-8971

Email Contacts:

For questions about sales: sales@statmodel.com

For technical support questions: support@statmodel.com

For Mplus registration: registration@statmodel.com

Figure 1. Item person map of the 11 abnormal synergy items

This figure presents a map of person abilities and the 11 abnormal synergy item difficulties plotted on a logit scale. Person abilities ("."- representing 1–4 persons, "##" representing 5 persons) range from a low of –5 logits at the bottom of the figure to a high of 6 logits at the top of the figure. Items are plotted at their mean difficulty level with items that are easier to endorse at the bottom of the map and items that are difficult to endorse at the top of the map. Letters to the left of the vertical dashed line are means and standard deviations for person abilities and to the right of the vertical line are mean and standard deviations for the item

difficulties. Abbreviations: M = mean, S = 1 standard deviation and T = 2 standard deviations.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 1

The Lower Extremity section of the Fugl-Meyer Assessment Scale

Test	Item #	Item	Scoring Criteria
I. Reflex Activity	1	Achilles	0- No reflex activity can be elicited 2- Reflex activity can be elicited
	2	Patellar	
II. Flexor Synergy (in supine)	3	Hip Flexion	0- Cannot be performed at all 1- Partial motion 2- Full motion
	4	Knee Flexion	
	5	Ankle Dorsiflexion	
III. Extensor Synergy (in side lying)	6	Hip Extension	0- Cannot be performed at all 1- Partial motion 2- Full motion
	7	Adduction	
	8	Knee Extension	
	9	Ankle Plantar Flexion	
IV. Movement combining synergies (sitting: knees free of chair)	10	Knee Flexion beyond 90°	0- No active motion 1- From slightly extended position, knee can be flexed but not beyond 90° 2- Knee flexion beyond 90°
	11	Ankle Dorsiflexion	0- No active flexion 1- Incomplete active flexion 2- Normal Dorsiflexion
V. Movement out of synergy (Standing hip at 0°)	12	Knee Flexion	0- Knee cannot flex without hip flexion 1- Knee begins flexion without hip flexion, but does not reach to 90° or hip flexes during motion 2- Full motion as described
	13	Ankle Dorsiflexion	0- No active motion 1- Partial motion 2- Full motion
VI. Normal Reflexes (sitting)	14	Knee flexors Patellar Achilles	0- At least 2 of 3 phasic reflexes are markedly hyperactive 1- One reflex is markedly hyperactive or at least 2 reflexes are lively 2- No more than one reflex is lively and none are hyperactive
VII. Coordination/Speed-Sitting: Heel to opposite knee (5 repetitions in rapid succession)	15	Tremor	0- Marked tremor 1- Slight tremor 2- No tremor
	16	Dysmetria	0- Pronounced or unsystematic dysmetria 2- No dysmetria
	17	Speed	0- Activity is more than 6 seconds longer than unaffected side 1- (2-5-9) seconds longer than unaffected side 2- Less than 2 seconds difference

TABLE 2

Item Measure Table of the 11 Abnormal Synergy Items

Item No.	Item Name	Score		Infit				Outfit		Pt. Measure Correlation	
		Raw	Mean Measure	Model S.E.	MNSQ	ZSTD	MNSQ	ZSTD	Observed	Expected	
12	Out of synergy – knee flexion	301	3.27	0.09	0.97	-0.4	0.93	-0.4	0.77	0.77	
13	Out of synergy – ankle dorsiflexion	326	3.05	0.09	0.76	-3.9	0.68	-2.4	0.82	0.78	
11	Combined synergy – ankle dorsiflexion	529	1.36	0.09	0.8	-3.4	0.7	-3.7	0.82	0.78	
5	Within synergy – ankle dorsiflexion	631	0.48	0.10	0.86	-2.1	0.71	-3.1	0.79	0.76	
9	Within synergy – ankle plantarflexion	673	0.08	0.10	1.29	3.9	1.29	2.2	0.71	0.74	
10	Combined synergy – knee flexion	721	-0.41	0.10	0.85	-2.1	0.67	-2.4	0.75	0.72	
7	Within synergy – hip adduction	777	1.05	0.11	1.3	3.5	2.28	4.6	0.61	0.68	
4	Within synergy – knee flexion	803	-1.4	0.12	0.85	-1.8	0.61	-1.7	0.7	0.66	
3	Within synergy – hip flexion	812	-1.52	0.12	0.93	-0.8	1.92	2.8	0.65	0.66	
6	Within synergy – hip extension	819	-1.63	0.12	1.16	1.8	1.9	2.7	0.6	0.65	
8	Within synergy – knee extension	856	-2.24	0.14	1.04	0.4	1.5	1.3	0.58	0.61	