

ORIGINAL RESEARCH

INTRA- AND INTER-RATER RELIABILITY OF THE SELECTIVE FUNCTIONAL MOVEMENT ASSESSMENT (SFMA)

Kathryn R. Glaws, PT, DPT, CSCS^{1,2}
Christopher M. Juneau, PT, DPT, CSCS^{1,2}
Lindsay C. Becker, PT, DPT, SCS, CSCS^{1,2}
Stephanie L. Di Stasi, PT, PhD, OCS^{1,2,3,4,6}
Timothy E. Hewett, PhD^{1,2,3,4,5}

ABSTRACT

Purpose/Background: The Selective Functional Movement Assessment (SFMA) is a clinical assessment system designed to identify musculoskeletal dysfunction by evaluation of fundamental movements for limitations or symptom provocation. The purpose of this study was to determine the intra- and inter-rater reliability of the ten fundamental movement patterns of the SFMA in a healthy population using the SFMA categorical and criterion checklist scoring tools.

Methods: 35 healthy subjects (22.9 years +/- 1.9) were recorded with two digital video cameras (1-frontal view/1-sagittal view) while they performed the ten fundamental movements patterns that comprise the SFMA. Evaluators with varying experience with the SFMA (rater A, > 100 hours; rater B, 25 hours; and rater C, 16 hours) and not present at the initial data collection evaluated each video using categorical and criterion checklist scoring tools. Evaluators repeated this process at least one week later. The evaluators' composite results were compared between and within raters using the kappa coefficient and ICC's for categorical scoring and criterion checklist scoring, respectively.

Results: Substantial to almost perfect intra-rater reliability of the SFMA (kappa, % agreement) was observed for all raters using the categorical scoring tool (rater A: .83, .91; rater B: .78, .88; and rater C: .72, .85). The criterion checklist scoring tool yielded intra-rater ICCs (3,1; 95% confidence interval) ranging from good to poor with rater A demonstrating the highest reliability (ICC [SEM]) (.52 [2.36]) and rater C the lowest reliability (.26 [3.42]). Inter-rater reliability of the categorical scoring tool was slight to substantial (.41-.61, .69-.79) while the criterion checklist tool (ICC 2,1) demonstrated unacceptable inter-rater reliability when assessed in all raters together (.43 [2.7]).

Conclusions: As hypothesized, intra-and inter-rater reliability of categorical scoring and criterion checklist scoring of the ten fundamental movements of the SFMA was higher in raters with greater experience.

Key Words: Functional movement, reliability, Selective Functional Movement Assessment (SFMA)

¹ The Ohio State University Sports Medicine, Columbus, Ohio, USA

² The Ohio State University Wexner Medical Center, Columbus, Ohio, USA

³ Sports Health and Performance Institute

⁴ Department of Orthopedics, The Ohio State University, Columbus, Ohio, USA

⁵ Departments of Physiology & Cell Biology, Family Medicine, Orthopaedics and Biomedical Engineering, The Ohio State University, Columbus, Ohio, USA

⁶ The School of Health and Rehabilitation Science, The Ohio State University, Columbus, Ohio, USA

Acknowledgements:

The OSU Sports Medicine Clinical Outcomes Research Coordinator (CORC) Program provided funding for this study

CORRESPONDING AUTHOR

Kathryn Glaws
The Ohio State University Wexner Medical Center
2050 Kenny Road, Suite 3100
Columbus, Ohio 43221
Email: Kathryn.glaws@osumc.edu
Phone: (614)293-2385
Fax: (614)293-3066

INTRODUCTION

Musculoskeletal screening tools and movement based assessment models have become increasingly common in sports medicine practice as a means to identify injury risk and movement dysfunction, respectively.^{1,2,3} Impairments associated with musculoskeletal injuries are not often isolated to the injured joint and residual deficits can persist if these impairments are not addressed. These deficits may not be readily identifiable by traditional, joint-specific examination techniques.⁴ Screening and assessment tools that incorporate whole body functional movements may uncover important underlying impairments that allow for the development and implementation of targeted interventions to both maximize recovery after primary injury and prevent secondary injury.

The traditional medical model emphasizes identification of an anatomical source of pain via assessment of a tissue and/or task specific to the impaired joint. For example, manual muscle testing and goniometric measurements isolate a specific muscle or joint in a single plane of motion.^{1,4} This model focuses on diagnosis of pathology versus identification of impairments and functional limitations throughout the body; thus, rehabilitation guided by the traditional medical model prioritizes symptomatic treatment of the affected tissue. The traditional model may not be optimal for the management of musculoskeletal pathology, as it does not account for 'regional interdependence' or the concept that adjacent anatomical regions can contribute to or be the source of a patient's primary complaint.⁵ In contrast, observation of fundamental movement patterns allows for the assessment of dynamic neuromuscular control and the interaction of multiple joints and body regions. Because pain is known to alter motor control,⁶⁻¹¹ the assessment of painful patterns of movement may be unreliable for accurate identification of the cause of a patient's symptoms.⁶⁻¹¹ Therefore, the observation of multi-joint fundamental movement patterns can reveal non-painful impairments in adjacent regions that may contribute to the cause of the patient's symptoms.^{1,2,3}

Several models have been developed to evaluate functional movement and have demonstrated good intra- and inter-rater reliability.¹²⁻¹⁹ For example, the Functional Movement Screen (FMS), is utilized widely

by health professionals to screen individuals for risk of injury via identification of asymmetries and dysfunctional movements in a healthy population. The Selective Functional Movement Assessment (SFMA) is a system developed for clinicians to identify movement dysfunction in a population with known musculoskeletal injury.¹² Clinically, the SFMA begins with the evaluation of a set of fundamental movement patterns in order to identify limitations to the movements and/or symptom provocation.¹² The SFMA describes movements as "functional" based on meeting specific criterion for normalcy defined for each movement.¹² If a movement pattern is deemed functional and non-painful, then further investigation of that pattern is not recommended. Painful movement patterns are further assessed with caution, as pain is known to alter motor control⁶⁻¹¹ and continued active movement in painful patterns could exacerbate symptoms. Alternatively, these painful movements are addressed with pain modulating therapies/modalities and are ultimately used as a reassessment to determine effectiveness of treatment, as they should become non-painful with appropriate assessment and treatment of dysfunctional non-painful regions. Movement patterns that are deemed dysfunctional and non-painful are further examined using an algorithm of additional sequential tests to reveal the specific mobility or stability impairments causing the dysfunctional pattern.¹²

While the SFMA is becoming a common assessment tool used by trained clinicians to identify dysfunctional movement patterns and musculoskeletal impairments, to date there is no published study that has examined its reliability. The purpose of this study was to determine the intra- and inter-rater reliability of the fundamental movement patterns of the SFMA in a healthy population, using both the categorical scoring tool and criterion checklist scoring tools, in raters with various levels of experience. It was hypothesized that both the categorical and criterion checklist scoring tools would demonstrate higher intra-rater and inter-rater reliability in raters with more experience using the SFMA.

METHODS

Participants

Thirty-nine healthy subjects (N = 39, 27 males, 12 female; 22.9 years +/- 1.9) were enlisted via conve-

nience sampling and consisted primarily of physical therapy students and the Ohio State University Club Rugby team. Potential participants were excluded if they had undergone orthopedic surgery within the last 6 months, were currently pregnant, recorded any positive marks on the Physical Activity Readiness Questionnaire (PAR-Q) health assessment, or were under the age of 18. Neurologic conditions, specifically concussions, were not a part of the exclusion criteria; however, none of the subjects reported active neurologic or concussive symptoms at the time of testing. All subjects provided written informed consent prior to study participation.

Procedures

Each participant was video recorded while performing the fundamental movement patterns of the SFMA, which consisted of ten movements: 1) Cervical Flexion, 2) Cervical Extension, 3) Cervical Rotation, 4) Upper Extremity Pattern One (behind the back internal rotation), 5) Upper Extremity Pattern Two (behind the back external rotation), 6) Multi-segmental Flexion, 7) Multi-segmental Extension, 8) Multi-segmental Rotation, 9) Single Leg Balance, and 10) Overhead Deep Squat (Appendix A). All motions were performed bilaterally when applicable. The subjects were not informed of the grading criteria and were given only minimal cueing for the tasks to avoid influencing the subject's preferred movement pattern. The test administrator, a physical therapist with eight hours of didactic SFMA training, provided a demonstration of each movement prior to completion of the task. Subjects were instructed to raise a red index card if pain was present so that this was clear to the raters during evaluation. Raters were not present during this process.

Two-dimensional (frontal and sagittal view) video data was captured at 30 frames per second at 640x480 resolution using two webcams (LifeCam Studio 1080p, Microsoft Corp, Redmond, WA) while connected to a PC running Cortex motion analysis software (v4.0, Motion Analysis Corp, Santa Rosa, CA) to record and automatically name each trial as each subject performed the ten fundamental movements of the SFMA. The tasks were either performed at Position 1 (distance from sagittal plane camera = 2.1 meters and frontal plane camera = 1.8 m) or Posi-

tion 2 (distance from sagittal plane camera: 3.3 m, distance from frontal plane camera = 3.7 m). Two planes of video were required to document the quality of movements and positions were selected to ensure adequate viewing for the raters as it would be observed in clinical practice.

Evaluation

Raters: Three raters with varying levels of experience participated in this study. Rater A was a SFMA instructor, had attended over 100 hours of continuing education courses as an instructor and student, and had used the system extensively in the clinic for 3 years. Rater B had approximately 25 hours of education and 6 months of clinical use with the SFMA and rater C had 8 hours of education and was considered a novice with regard to clinical use of the system. Neither rater B nor rater C had taken a formal SFMA continuing education course. Education hours represented the amount of time each rater spent under supervision of a SFMA instructor learning the entire SFMA, including individual breakouts and interventions, and were not limited to the evaluation of the 10 fundamental movements that were assessed during this study.

Raters were not present during the data collection and later individually evaluated each video using two scoring sheets: (1) the categorical scoring sheet provided by the developers of the SFMA and (2) a 34 point criterion checklist scoring tool included in the SFMA manual (Appendices B and C). Scoring possibilities on the categorical score sheet included: Functional-Non-painful (FN), Functional-Painful (FP), Dysfunctional-Non-painful (DN), and Dysfunctional-Painful (DP). The raters were not given any instructions or restrictions regarding the number of times to watch each video nor were they prohibited from pausing the video at any time. Evaluators repeated the scoring process for intra-rater reliability between seven and fourteen days following the initial assessment.

Rater Scoring

Categorical Scoring: Raters scored each fundamental movement as FN, FP, DN or DP. Criteria for each of these scores are not included on the categorical scoring tool and therefore were not available for use during scoring. (Appendix B)

Criterion Checklist Scoring: The raters viewed each trial and assigned an ordinal scale rating to each fundamental movement task. A total score of zero was considered perfect technique without compensation while a score of 34 represented the failure of all criteria. Cumulative scores (0-34) from each rater for each subject were used for all statistical analysis (Appendix C). Prior to statistical analysis, the data were compiled and coded by a researcher not involved in rating.

It is important to note that the two scoring tools used nearly identical rating methods, which are taught as a part of the SFMA continuing education course. The major difference between the tools is that the criterion-checklist requires raters to identify specific criteria that are not met and provides a numerical score. Additionally, the criterion-checklist lacks items regarding quality of the movement and instead focuses on quantitative aspects.

Statistical Analysis

Categorical Scoring: The evaluators' composite results were compared within raters and between raters (A-B, A-C, B-C) using both absolute agreement and the kappa coefficient based upon the categorical classification of each movement. These analyses are standard in similar studies.³ Kappa coefficients were used to quantify strength of agreement. Interpretation of the Kappa coefficient has been described as: ≤ 0 = poor, .01-.2 = slight, .21-.4 = fair, .41-.6 = moderate, .61-.8 = substantial, .81-1.0 = almost perfect.^{20,21} Kappa coefficients were calculated using Microsoft Excel 2010.

Criterion Checklist Scoring: A mean composite score for trial one and trial two was calculated for each rater's data. Intra- and inter-rater reliability were assessed using Intra-class Correlation Coefficients (ICCs) with 95% confidence intervals (CI) and Standard Error of Measurement (SEM) calculated on the composite scores of each subject. Data were compared between raters (ICC [2,1]); and within each rater over the two scoring sessions (ICC [3,1]). Interpretation of reliability results were based on the following criterion: ICC > 0.75 = good, $0.50 - 0.74$ = moderate, and < 0.50 = poor.²² Inter-rater reliability, using ICC (2,1), was calculated during trial one only between all three raters collectively and between each rater (A-B, A-C, B-C).

Reliability analyses with ICCs were performed using SPSS 17.0 (SPSS Inc., Chicago, IL).

RESULTS

A total of thirty-nine (39) subjects met the inclusion criteria; however, data from 35 (24 males, 11 females) of these subjects were used in the final statistical analyses as a result of data errors in four subjects due to equipment malfunction or video file corruption. None of the subjects experienced pain when completing the fundamental movements. When using the categorical tool, subjects were scored as FN most frequently when performing (1) cervical rotation to the left, (2) upper extremity pattern two – left and (3) upper extremity pattern two – right (Table 1). Evaluators scored subjects as DN most often when completing the overhead deep squat and multi-segmental rotation right and left, respectively. Rater B reported the highest total number of dysfunctional patterns, followed by rater A and rater C when using the categorical scoring tool. Rater B also demonstrated the most DN ratings per subject (mode: rater A = 7, rater B = 9, rater C = 7). When using the criterion checklist, rater A reported the highest average score using the criterion-based scoring tool in both trials, followed by rater B and rater C, respectively (Table 2).

Categorical Scoring Reliability: Substantial to almost perfect intra-rater reliability of the composite scoring of the SFMA (kappa, % agreement) was observed for all raters. Rater A demonstrated the highest intra-rater reliability (.83, .91), followed by rater B (.78, .88). Rater C demonstrated the lowest intra-rater reliability (.72, .85) (Table 3). Inter-rater reliability of composite scoring was interpreted as slight to substantial. The two raters with the most experience, rater A and B, demonstrated the highest inter-rater reliability (.76, .88). Rater A and rater C demonstrated fair inter-rater reliability (.30, .74) and raters B and C, the least experienced raters, demonstrated the lowest inter-rater reliability (.20, .62). (Table 4)

Intra-rater reliability of each of the component fundamental movements varied between raters. Rater A showed the highest range of kappa values followed by rater B and rater C, respectively (.41-.94, .34-.89). (Table 5) The most experienced rater demonstrated the highest intra-rater reliability with single

Table 1. Frequency of Functional/Non-painful vs Dysfunction/Non-Painful Ratings, Categorical Scoring

	Functional Non-painful Total (Male/Female)			Dysfunctional Non-painful Total (Male/Female)		
	Rater A	Rater B	Rater C	Rater A	Rater B	Rater C
Cervical Flexion	23 (18/5)	19 (16/3)	32 (22/10)	12 (6/6)	16 (8/8)	3 (2/1)
Cervical Extension	24 (16/8)	22 (14/8)	29 (18/11)	11 (8/3)	13 (10/3)	6 (6/0)
Cervical Rotation Left	29 (19/10)	29 (19/10)	30 (20/10)	6 (5/1)	6 (5/1)	5 (4/1)
Cervical Rotation Right	24 (15/9)	30 (20/10)	29 (19/10)	11 (9/2)	5 (4/1)	6 (5/1)
UE Pattern 1 Right	23 (13/10)	9 (6/3)	13 (5/8)	12 (11/1)	26 (18/8)	22 (19/3)
UE Pattern 1 Left	28 (18/10)	9 (5/4)	18 (9/9)	7 (6/1)	26 (19/7)	17 (15/2)
UE Pattern 2 Right	28 (18/10)	29 (19/10)	27 (17/10)	7 (6/1)	6 (5/1)	8 (7/1)
UE Pattern 2 Left	29 (18/11)	30 (19/11)	28 (17/11)	6 (6/0)	5 (5/0)	7 (7/0)
Multi-segmental Flexion	9 (7/2)	9 (7/2)	18 (13/5)	26 (17/9)	26 (17/9)	17 (11/6)
Multi-segmental Extension	4 (2/2)	10 (7/3)	16 (9/7)	31 (22/9)	25 (17/8)	19 (15/4)
Multi-segmental Rotation Right	3 (2/1)	2 (1/1)	12 (8/4)	32 (12/10)	33 (13/10)	23 (16/7)
Multi-segmental Rotation Left	6 (5/1)	3 (1/2)	18 (11/7)	29 (19/10)	32 (23/9)	17 (13/4)
Single Leg Stance Right	11 (7/4)	4 (3/1)	18 (11/7)	24 (17/7)	31 (21/10)	17 (13/4)
Single Leg Stance Left	15 (11/4)	5 (4/1)	20 (13/7)	20 (13/7)	30 (20/10)	15 (11/4)
Overhead Deep Squat	2 (2/0)	4 (4/0)	6 (4/2)	33 (22/11)	31 (20/11)	29 (20/9)

Table 2. Mean of Total Score, Criterion Checklist Scoring

	Rater A	Rater B	Rater C
Trial 1	11.6 (3.0)	11.0 (3.1)	7.9 (3.6)
Trial 2	11.9 (3.6)	10.0 (3.6)	7.5 (3.3)

(SD; Standard Deviation)

Table 3. Intra-rater Reliability, Categorical Scoring

	Kappa Coefficient	Absolute Agreement
Rater A	.83	.91
Rater B	.78	.88
Rater C	.72	.85

Table 4. Inter-rater Reliability, Categorical Scoring (Trial 1)

	Kappa Coefficient	Absolute Agreement
Rater A/Rater B	.76	.88
Rater A/Rater C	.30	.74
Rater B/Rater C	.20	.62

leg stance on the left (kappa = .94) and the lowest with left cervical rotation (.41). Rater B showed the highest intra-rater reliability with upper extremity pattern two on the right (.89) and the lowest with

upper extremity pattern one on the right (.34). Rater C demonstrated the highest and lowest intra-rater reliability when scoring the single leg stance on the left (.76) and multi-segmental extension (.25), respectively.

Inter-rater reliability of the component fundamental movements was lower than intra-rater reliability (Table 6). Raters A and B demonstrated the highest range of kappa values (.07-.79) followed by raters A and C (.18-.62) and raters B and C (-.31-.68), respectively. Raters A and B demonstrated the highest inter-rater reliability with multi-segmental rotation to the right (.79). The movement with the lowest inter-rater reliability was upper extremity pattern one on the right evaluated by the least experienced raters (-.31).

Criterion Scoring Reliability: Intra-rater ICCs (3, 1) ranged from good to poor, with rater A having the highest reliability with the smallest SEM (ICC, SEM) (.86, 1.2), followed by rater B (.71, 1.7) and C (.59, 2.2) respectively (Table 7). The two most experienced raters (raters A and B) had the highest inter-rater reliability within both trials (.68 and .52, respectively), interpreted as moderate reliability. The least experienced rater, rater C, demonstrated poor inter-rater reliability with the other two raters in trials one (.40 and .31) and two (.26 and .34) (Tables 8 and 9), while

Table 5. Intra-Rater Reliability, Categorical Scoring: Component Fundamental Movements

	Rater A		Rater B		Rater C	
	Kappa	% Agrmt	Kappa	% Agrmt	Kappa	% Agrmt
Cflex	0.72	0.89	0.70	0.86	0.44	0.89
Cext	0.75	0.89	0.75	0.89	0.80	0.94
C rot L	0.41	0.80	0.68	0.91	0.60	0.89
C rot R	0.66	0.88	0.44	0.89	0.53	0.86
UE 1 L	0.68	0.91	0.43	0.74	0.66	0.83
UE 1 R	0.88	0.94	0.70	0.89	0.61	0.83
UE 2 L	0.89	0.97	0.77	0.94	0.60	0.89
UE 2 R	0.68	0.91	0.89	0.97	0.41	0.83
MSF	0.46	0.80	0.85	0.94	0.72	0.86
MSE	0.87	0.97	0.34	0.71	0.25	0.63
MSR L	0.73	0.97	0.52	0.91	0.77	0.89
MSR R	0.89	0.94	0.52	0.91	0.62	0.83
SLS L	0.94	0.97	0.60	0.89	0.76	0.89
SLS R	0.79	0.91	0.68	0.91	0.83	0.91
OHDS	0.65	0.97	0.52	0.91	0.72	0.91

(**Cflex**:Cervical flexion; **C Rot**: cervical rotation; **UE 1**: upper extremity pattern 1; **UE 2**: upper extremity pattern 2; **MSF**: multi-segmental flexion; **MSE**: multi-segmental extension; **MSR**: multi-segmental rotation; **SLS**:single leg stance; **OHD**: overhead deep squat)

Table 6. Inter-Rater Reliability, Categorical Scoring: Component Fundamental Movements

	Raters A vs B		Rater A vs C		Rater B vs C	
	Kappa	% Agrmt	Kappa	% Agrmt	Kappa	% Agrmt
Cflex	0.77	0.89	0.30	0.74	0.20	0.63
Cext	0.62	0.83	0.62	0.86	0.38	0.74
C rot L	0.71	0.83	0.30	0.83	0.25	0.80
C rot R	0.38	0.77	0.32	0.74	0.46	0.86
UE 1 L	0.07	0.40	0.19	0.60	0.22	0.60
UE 1 R	0.20	0.54	0.28	0.63	-0.31	0.51
UE 2 L	0.68	0.91	0.53	0.86	0.60	0.89
UE 2 R	0.72	0.91	0.58	0.86	0.64	0.89
MSF	0.64	0.86	0.44	0.71	0.38	0.69
MSE	0.15	0.71	0.27	0.66	0.05	0.54
MSR L	0.37	0.86	0.33	0.66	0.16	0.57
MSR R	0.79	0.97	0.34	0.77	0.21	0.71
SLS L	0.36	0.71	0.61	0.80	0.22	0.57
SLS R	0.48	0.80	0.60	0.80	0.22	0.60
OHDS	0.64	0.94	0.18	0.83	0.54	0.89

(**Cflex**:Cervical flexion; **C Rot**: cervical rotation; **UE 1**: upper extremity pattern 1; **UE 2**: upper extremity pattern 2; **MSF**: multi-segmental flexion; **MSE**: multi-segmental extension; **MSR**: multi-segmental rotation; **SLS**: single leg stance; **OHD**: overhead deep squat)

also demonstrating the lowest intra-rater reliability (.59). Inter-rater reliability of trial one was found to be poor when assessed in all raters together (.43) (Table 8).

DISCUSSION

The purpose of this study was to investigate the intra-rater and inter-rater reliabilities of the SFMA in raters with varying experience levels. The current findings show that intra-and inter-rater reliability of categorical scoring and criterion checklist scoring of the fundamental movements of the SFMA were reflective of level of experience, which supported the *a priori* hypothesis. Specifically, fundamental SFMA movement scoring demonstrated substantial to excellent intra-rater reliability when using the categorical scoring tool but poor to good intra-rater reliability when using criterion-based scoring tool. Substantial and moderate inter-rater reliability was found for more experienced users with categorical and criterion scoring, respectively, while poor reliability was found for least experienced rater.

Table 7. Intra-rater Reliability, Criterion Checklist Scoring

Rater (ICC 3,1)	ICC	95% CI	SEM	MDC
Rater A	0.86	0.74-0.93	1.2	3.3
Rater B	0.71	0.46-0.85	1.7	4.8
Rater C	0.59	0.32-0.77	2.2	6.0

(**ICC 3,1**:Intra-rater Intra-class Correlation Coefficients; **CI**: Confidence Intervals; **SEM**:standard error of the measurement; **MDC**: minimal detectable change)

Table 8. Inter-rater Reliability, Criterion Checklist Scoring (Trial 1)

Group (ICC 2,1)	ICC	95% CI	SEM	MDC
All Raters	0.43	0.12-0.67	2.7	7.5
Rater A vs. B	0.68	0.45-0.82	1.7	4.76
Rater B vs. C	0.31	-0.50-0.60	3.04	8.42
Rater A vs. C	0.40	-0.10-0.73	2.92	8.09

(**ICC 2,1**:Inter-rater Intraclass Correlation Coefficients; **CI**: Confidence Intervals; **SEM**: standard error of the measurement; **MDC**: minimal detectable change)

Table 9. Inter-rater Reliability, Criterion Checklist Scoring (Trial 2)

Group (ICC 2,1)	ICC	95% CI	SEM	MDC
Rater A vs. B	0.52	0.15-0.74	2.36	6.55
Rater B vs. C	0.34	-0.01-0.64	2.77	7.67
Rater A vs. C	0.26	-0.10-0.58	3.42	9.48

(ICC 2,1: Inter-rater Intraclass Correlation Coefficients; CI: Confidence Intervals; SEM: standard error of the measurement; MDC: minimal detectable change)

The reliability data presented here are similar to that observed in reliability studies examining other movement-based screening and assessment systems. Rater A, who had significantly more experience compared to the other raters and is also a SFMA instructor, demonstrated the best intra- and inter-rater reliabilities using both scoring tools, followed by raters B and C, respectively. Novice raters of the Functional Movement Screen™ show poor to good reliability while experienced raters show good to excellent reliability.^{13,14,15} In addition, the inter-rater reliability between novice and expert raters of the Landing Error Scoring System (LESS) and the drop jump landing for ACL risk were shown to range from moderate to excellent and good to excellent, respectively.^{23,24} In sum, these data indicate that clinical screening tools are most reliability implemented by experienced users; however, the amount of training necessary to maximize intersession reliability of the SFMA is unknown.

While total categorical and criterion scoring methods indicated moderate intra- and inter-rater reliability, some patterns were less reliable than others. Multi-segmental extension and upper extremity pattern one demonstrated slight to fair inter-rater reliability between all raters when categorical scoring was used. Intra-rater reliability of multi-segmental extension was also fair for the two least experienced raters, while the most experienced rater demonstrated almost perfect reliability. Evaluation of both multi-segmental extension and upper extremity pattern one require assessment of qualitative criterion including the presence or absence of a uniform spinal curve and scapular dyskinesia, respectively. These criteria may be more difficult to assess compared to criterion associated with specific landmarks. Experienced raters likely have increased familiarity viewing the fundamental movements and with the rating

criteria, which may allow them to more reliably score each movement. In addition, results from this study may indicate a trend for experienced raters to be more stringent using the criterion checklist scoring tool. For example, rater A averaged four more errors per subject than rater C. As a result of viewing more fundamental movements, experienced raters may have a higher benchmark for scoring functional particularly when evaluating qualitative aspects.

While the criterion checklist model demonstrated poor to moderate inter-rater reliability between examiners of different experience, intra-rater reliability of the experienced rater was good. The format of this tool, specifically the inclusion of listed criteria and a numerical score, lends itself to potential use as a clinical outcome tool. This scoring method produces ordinal data, in contrast to the categorical scoring method, which allows for increased objectivity and ease of tracking. Data from this preliminary study highlights the reliability of the SFMA in the hands of an experienced user; however validation and responsiveness to change in a clinical population must be determined prior to use of the SFMA as an outcome tool.

Evidence-based clinical implementation of the SFMA categorical scoring tool requires validation for its use in identifying dysfunction. Plich et al performed a correlational study to determine whether the SFMA related to the self-reported outcomes in patients treated for neck or low back conditions.²⁵ The reported results showed a significant relationship between the scores of the Neck Disability Index and the criterion scoring tool, thus establishing a preliminary positive relationship between patients' self-reported outcomes and the SFMA.²⁵

Residual musculoskeletal impairments following injury and supervised rehabilitation are common and may increase an individual's risk for a second injury.²⁶⁻³⁰ These impairments often are not isolated to the injured joint and can be missed by traditional, joint-specific examination techniques. Decreased hip mobility in golfers with low back pain³¹ and decreased core and proximal hip control in overhead athletes with shoulder pain^{30,32} are just a few examples of these regionally related impairments associated with primary injury. Similar regional impairments have been associated with secondary injury. Deficits

in postural stability and proximal hip control in the transverse plane were among four factors found to be predictive of secondary ACL injury.²⁷ Use of the SFMA may provide a more comprehensive understanding of the mobility and stability impairments throughout the body than the traditional medical model, and subsequently may aid the development of targeted interventions to maximize recovery after primary injury. Ultimately, resolution of total body impairments following primary injury may help to decrease risk of a re-injury or a secondary injury.

There are limitations to this study. Video recording does not mimic the clinical setting, where the patient can be viewed from multiple angles, receive verbal clarification, or be palpated to determine if grading criteria have been met. However, the use of video assessment has been used in similar reliability studies and limits external variability.¹³ Second, while the criterion checklist scoring tool is included in the SFMA handbook this tool is not commonly utilized and none of the raters had prior experience with the tool. This is the first study to evaluate the reliability of the SFMA, but its validity has not yet been examined. The raters in this study retrospectively reported that the criteria on the checklist were not fully inclusive of movement impairments they would have scored as an impairment necessitating further evaluation. For example, failing criteria listed on the criterion checklist for the Overhead Deep Squat movement did not include lifting of the heels during the action, but this would be deemed compensatory or dysfunctional per the categorical scoring criteria. In addition, criteria did not fully account for qualitative aspects of the movements that may be observed during UE patterns (such as scapular dyskinesis) or differences in perceived exertion during bilateral movements, further highlighting that refinement of the tool may also be necessary. Finally, the SFMA was developed for use on a clinical population, whereas this study examined healthy participants, thus limiting its external validity.

Future research should focus on determination of the reliability of raters with similar experience and rating experience, which may help determine the minimum education and experience required to establish adequate agreement of the scoring for the

SFMA during clinical use. Additional research is also necessary to determine the reliability of the categorical and criterion checklist scoring of the SFMA in a clinical population, as this is the target population for which this screening system was designed. Finally, validation of both the categorical scoring tool for clinical use and the criterion checklist scoring tool for potential use as an outcome tool would support the implementation of the SFMA in clinical populations.

CONCLUSION



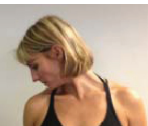


Scoring of the ten fundamental movements of the SFMA, using the categorical and criterion checklist model in a healthy population, is most reliable between sessions when performed by a single experienced rater. Future research should focus on the determination of the reliability and validity of the SFMA in a population with known musculoskeletal injury, as this is the population for which this assessment system is intended.

REFERENCES

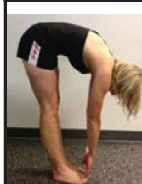
1. Cook G, Burton L, Hoogenboom B. Pre-participation screening: The use of fundamental movements as an assessment of function- Part 1. *N Am J Sports Phys Ther.* 2006;1(2):62-72.
2. Kiesel K, Plisky PJ, Voight ML. Can serious injury in professional football be predicted by a preseason functional movement screen? *N Am J Sports Phys Ther.* 2007;2(3):147-158.
3. Schneiders A, Davidsson A, Horman E, Sullivan SJ. Functional Movement Screen Normative Values in a Young, Active Population. *Int J Sport Phys Ther.* 2011;6(2):75-82.
4. Holzbaur R, Murray W, Delp S. A Model of the Upper Extremity for Simulating Musculoskeletal Surgery and Analyzing Neuromuscular Control. *Ann Biomed Eng.* 2005;33(6):829-840.
5. Wainer R, Whitman J, Cleland J, Flynn T. Regional Interdependence: A Musculoskeletal Examination Model Whose Time Has Come. *J Orthop Sports Phys Ther.* 2007;37(11):658-660.
6. Dickx et al. "The effect of unilateral muscle pain on recruitment of the multifidus during autonomic contraction. An experimental pain study." *Manual Therapy.* 2010; 15(4):364-369.
7. Ferreira PH, Ferreira ML, Hodges PW. "Changes in recruitment of the abdominal muscles in people with low back pain: ultrasound measurement of muscle activity." *Spine.* 2004;29(22):2560-2566.

8. Hodges PW, Richardson CA. "Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds." *Arch Phys Med Rehabil.* 1999;80(9):1005-1012.
9. Hodges PW, Tucker K. "Moving differently in pain: a new theory to explain the adaptation to pain." *Pain.* 2011;152:S90-98.
10. Kiesel et al. "Experimentally induced pain alters the EMG activity of the lumbar multifidus in asymptomatic subjects." *Manual Therapy.* 2012;17(3):236-240.
11. Kiesel KB, Uhl T, Underwood FB, Nitz AJ. "Rehabilitative ultrasound measurement of select trunk muscle activation during induced pain." *Manual Therapy.* 2008;13(2):132-138.
12. Cook, G. *Movement: Functional Movement Systems: Screening, Assessment, and Corrective Strategies.* On Target Publications. Santa Cruz, CA. 2010.
13. Gribble P, Brigle J, Pietrosimone B, Pfile K, Webster K. Intrarater reliability of the Functional Movement Screen. *J Strength Cond Res.* 2013;27(4):978-81.
14. Minick K, et al. Interrater reliability of the functional movement screen. *J Strength Cond Res.* 2010;24(2):479-86.
15. Smith CA, Chimera NJ, Wright NJ, Warren M. Interrater and intrarater reliability of the functional movement screen. *J Strength Cond Res.* 2013;27(4):982-987.
16. Teyhen D, et al. The Functional Movement Screen: A Reliability Study. *J Orthop Sports Phys Ther.* 2012;42(6). 530-540.
17. Harris-Hayes M, Van Dillen L. "The inter-tester reliability of physical therapists classifying low back pain problems based on the movement system impairment classification system." *PM&R.* 2009; 117-126.
18. Sahrman S. *Movement system impairment syndromes of the extremities, cervical and thoracic spines.* Elsevier Health Sciences, 2010.
19. Trudelle-Jackson E, Sarvaiya-Sjaj S, Wang S. Interrater reliability of a movement impairment-based classification system for lumbar spine syndromes in patients with chronic low back pain. *J Orthop Sports Phys Ther.* 2008;38(6):371-376.
20. Landis, J.R., Kock, G.G. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.
21. Sim J, Wright C. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005;85:257-268.
22. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 2nd ed. Upper Saddle River, NJ: Prentice Hall;2000.
23. Padua et al. Reliability of the Landing Error Scoring System-real time, a clinical assessment tool of jump-landing biomechanics. *J Sport Rehabil.* 2011;20(2): 145-156.
24. Ekegren et al. Reliability and validity of observational risk screening in evaluating dynamic knee valgus. *J Orthop Sports Phys Ther.* 2012;39(9):665-674.
25. Plich SB, et al. Relationship Among Performance on the Selective Functional Movement Assessment and NDI and ODI Scores in Patients with Spine Pain. *Orthopaedic Practice.* 2013;24(1):12.
26. Greene HS, Cholewicki J, Galloway MT, Nguyen CV, Radebold A. A history of low back injury is a risk factor for recurrent back injuries in varsity athletes. *Am J Sports Med.* 2001;29(6):795-800.
27. Hewett TE, Di Stasi SL, Meyer GD. Current concepts for injury prevention in athletes after anterior cruciate ligament reconstruction. *Am J Sports Med.* 2013;41(1):216-24.
28. Paterno MV, Schmitt LC, Ford KR, Rauh MH, Myer GD, Huang B, Hewett TE. Biomechanical measures during landing and postural stability predict second anterior cruciate ligament injury after anterior cruciate ligament reconstructions and return to sport. *Am J Sports Med.* 2010;38(10):1968-1978.
29. Orchard JW. Intrinsic and extrinsic risk factors for muscles strains in Australian football. *Am J Sports Med.* 2001;29(3):300-303.
30. Zazulak BT, Hewett TE, Reeves NP, Goldberg B, Cholewicki J. Deficits in neuromuscular control of the trunk predict knee injury risk: a prospective biomechanical-epidemiologic study. *Am J Sports Med.* 2007;(7):1123-30.
31. Vad VB et al. Low back pain in professional golfers: the role of associated hip and low back range-of-motion deficits. *Am. J. Sports Med.* 2004;32(2):494-497.
32. Burkhart S, Morgan C, Kibler B. Shoulder Injuries in Overhead Athletes. *Clin Sport Med.* 2003;19(1): 125-158.

Appendix A. FUNDAMENTAL MOVEMENTS: *“Please remove your shoes for all activities. I will perform each exercise and give you instructions. Please let me know if you have any pain with any of the activities.” (If the subject performs any movements incorrectly, please provide cueing at your discretion for instructional aid.)*

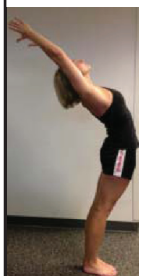
	<p>Active Cervical Flexion: “Stand in a tall position with feet together and toes pointing forward.”</p> <p>-“Please bring your chin down to your chest with your body upright.”</p>
	<p>Active Cervical Extension: “Stand in a tall position with feet together and toes pointing forward.”</p> <p>-“Please look up to the ceiling.”</p>
	<p>Cervical Rotation Side-Bend: “Stand in a tall position with feet together and toes pointing forward.”</p> <p>-“Please turn your head as far as you can to the right/left and then bend down and touch your chin towards your collarbone.” (Note: Avoid shoulder elevation)</p>
	<p>Upper Extremity Pattern 1-Medial Rotation Extension: “Stand in a tall position with feet together and toes pointing forward.”</p> <p>-“Take your Right/Left arm and reach behind your back and aiming to touch the bottom of your opposite shoulder blade.”(Note: scapular dyskinesis)</p>
	<p>Upper Extremity Pattern 2-Lateral Rotation Flexion: “Stand in a tall position with feet together and toes pointing forward.”</p> <p>-“Take your Right/Left arm and reach up and behind your head towards the top of your opposite shoulder blade.” (Note: scapular dyskinesis)</p>

Appendix A. (continued)



Multi-segmental Flexion: “Stand in a tall position with feet together and toes pointing forward.”

-“Please bend down and try to touch your toes.” (Note: knees straight)



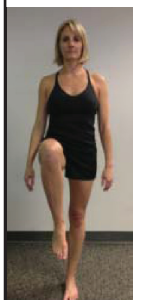
Multi-segmental Extension: “Stand in a tall position with feet together and toes pointing forward.”

-“Please keep your hands above your head and reach back as far as you can.”
(Note: knees straight)



Multi-segmental Rotation: “Stand in a tall position with feet together and toes pointing forward.”

-“Please place your hands by your side and rotate your entire body Right/Left trying to look behind you while keeping your feet still.”



Single Leg Stance (eyes open, then closed): “Stand in a tall position with feet together and toes pointing forward.”

-“1. Lift your Right/Left leg so that your hip and knee make 90 degree angles. Please hold this for 10 seconds. 2. Now lift your leg to the same 90 degree position and then close your eyes. Hold this position for 10 seconds.”



Overhead Deep Squat: “Please start by placing feet shoulder width and toes pointing forward.”

-“Please place your arms over your head and move them slightly outside your shoulders with your elbows straight. Hold this position and squat down as far as you can towards the floor.” (Note: subject should squat back and down, not over toes.)

Appendix B.

THE SELECTIVE FUNCTIONAL MOVEMENT ASSESSMENT

SFMA SCORING		FN	FP	DP	DN
Active Cervical Flexion					
Active Cervical Extension					
Cervical Rotation-Lateral Bend		L			
		R			
Upper Extremity Pattern 1 (MRE)		L			
		R			
Upper Extremity Pattern 2 (LRF)		L			
		R			
Multi-Segmental Flexion					
Multi-Segmental Extension					
Multi-Segmental Rotation		L			
		R			
Single Leg Stance		L			
		R			
Overhead Deep Squat					

THE SELECTIVE FUNCTIONAL MOVEMENT ASSESSMENT

Name	Date	Total Score
Cervical Flexion	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary	
<input type="checkbox"/> Can't Touch Sternum_____		
Cervical Extension		
<input type="checkbox"/> Greater than 10 Degrees of Parallel_____		
Cervical Rotation		
<input type="checkbox"/> <i>RIGHT</i> – Can't Touch Chin to Mid-Clavicle_____		
<input type="checkbox"/> <i>LEFT</i> - Can't Touch Chin to Mid-Clavicle_____		
Upper Extremity	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i>	
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Can't Touch Inferior Angle of the Contralateral Scapula		
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Can't Touch Spine of the Contralateral Scapula		
Multi-Segmental Flexion	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary	
<input type="checkbox"/> Can't Touch Toes and Return to Standing Position_____		
<input type="checkbox"/> < 70 Degrees Sacral Angle_____		
<input type="checkbox"/> No Posterior Weight Shift (T-L Junction over foot) _____		
<input type="checkbox"/> Non-Uniform Spinal Curves_____		
Multi-Segmental Extension	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary	
<input type="checkbox"/> ASIS Doesn't Clear the Toes_____		
<input type="checkbox"/> Can't Maintain Normal (≥ 170 degrees) Shoulder Flexion_____		
<input type="checkbox"/> Spine of Scapula Doesn't Clear the Heels_____		
<input type="checkbox"/> Non-Uniform Spinal Curves_____		
Multi-Segmental Rotation	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i>	
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Pelvis Rotation < 50 degrees_____		
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Trunk/shoulder < 50 degrees more than pelvis_____		
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Spinal/Pelvic Deviation_____		
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Excessive Knee Flexion_____		
Single Leg Stance	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i>	
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Eyes Open Standing < 10 seconds_____		
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Eyes Closed Standing < 10 seconds_____		
<input type="checkbox"/> <i>RIGHT</i> <input type="checkbox"/> <i>LEFT</i> Loss of Height_____		
Overhead Squating	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary	
<input type="checkbox"/> Loss of Shoulder Flexion_____		
<input type="checkbox"/> Thoracic Flexes_____		
<input type="checkbox"/> Hips Don't Break Parallel_____		
<input type="checkbox"/> Sagittal Plane Deviation of Lower Extremity Rt. _____ Lt, _____		