

Validity and Reliability of Nintendo Wii Fit Balance Scores

Erik A. Wikstrom, PhD, ATC, FACSM

Department of Kinesiology and the Center for Biomedical Engineering Science, University of North Carolina at Charlotte

Context: Interactive gaming systems have the potential to help rehabilitate patients with musculoskeletal conditions. The Nintendo Wii Balance Board, which is part of the Wii Fit game, could be an effective tool to monitor progress during rehabilitation because the board and game can provide objective measures of balance. However, the validity and reliability of Wii Fit balance scores remain unknown.

Objective: To determine the concurrent validity of balance scores produced by the Wii Fit game and the intrasession and intersession reliability of Wii Fit balance scores.

Design: Descriptive laboratory study.

Setting: Sports medicine research laboratory.

Patients or Other Participants: Forty-five recreationally active participants (age = 27.0 ± 9.8 years, height = 170.9 ± 9.2 cm, mass = 72.4 ± 11.8 kg) with a heterogeneous history of lower extremity injury.

Intervention(s): Participants completed a single-limb-stance task on a force plate and the Star Excursion Balance Test (SEBT) during the first test session. Twelve Wii Fit balance activities were completed during 2 test sessions separated by 1 week.

Main Outcome Measure(s): Postural sway in the antero-posterior (AP) and mediolateral (ML) directions and the AP, ML,

and resultant center-of-pressure (COP) excursions were calculated from the single-limb stance. The normalized reach distance was recorded for the anterior, posteromedial, and posterolateral directions of the SEBT. Wii Fit balance scores that the game software generated also were recorded.

Results: All 96 of the calculated correlation coefficients among Wii Fit activity outcomes and established balance outcomes were interpreted as poor ($r < 0.50$). Intrasession reliability for Wii Fit balance activity scores ranged from good (intraclass correlation coefficient [ICC] = 0.80) to poor (ICC = 0.39), with 8 activities having poor intrasession reliability. Similarly, 11 of the 12 Wii Fit balance activity scores demonstrated poor intersession reliability, with scores ranging from fair (ICC = 0.74) to poor (ICC = 0.29).

Conclusions: Wii Fit balance activity scores had poor concurrent validity relative to COP outcomes and SEBT reach distances. In addition, the included Wii Fit balance activity scores generally had poor intrasession and intersession reliability.

Key Words: center of pressure, Star Excursion Balance Test, interactive gaming, rehabilitation

Key Points

- The Wii Fit balance activity scores had poor concurrent validity when compared with center-of-pressure outcomes from a 30-second single-limb stance and the anterior, posteromedial, and posterolateral reach distances from the Star Excursion Balance Test.
- Balance scores from the 12 tested Wii Fit activities generally had poor intrasession and intersession reliability and high minimal detectable change scores relative to the mean score of the activity.
- Wii Fit balance scores should not be used as an objective measure of progress during rehabilitation.

Athletic trainers and other allied health care professions have begun to use interactive gaming systems and technology to help rehabilitate patients with musculoskeletal and neurologic conditions.^{1–4} Interactive games, such as DanceDance Revolution (Konami Corp, Toyko, Japan), and gaming systems, such as PlayStation 2 (Sony Computer Entertainment America, LLC, Foster City, CA) and Wii (Nintendo of America Inc, Redmond, WA), can be seen in physical therapy and sports medicine facilities. The growing popularity of interactive gaming partially is due to the belief that playing interactive video games during a rehabilitation session can direct a patient's focus away from the repetitive and mundane nature of rehabilitation exercises and toward the fun and competitive aspects of the video game.^{5–7} Thus, interactive gaming technology subjectively improves adherence rates for

rehabilitation programs⁶ and presents allied health care providers with an adjunct to traditional rehabilitation exercises.

However, Nintendo recently released the Wii Balance Board (Nintendo of America Inc) as a controller for games such as Wii Fit (Nintendo of America Inc) and Wii Fit Plus (Nintendo of America Inc). The Wii Balance Board possesses characteristics similar to a research-grade force plate and contains transducers used to assess force distribution and the subsequent movements of an individual's center of pressure (COP). Thus, the Wii Balance Board combines current force-platform technology with the entertainment value of gaming while completing fitness activities, such as balance and yoga. Furthermore, the Wii Fit and Wii Fit Plus games include personalized settings to track an individual's progress over time. The completion of

these activities allegedly improves balance, and some health care providers have hypothesized that components of interactive games, such as balance scores and player ratings, might be related to sensorimotor function.⁶

Thus, using the Wii Balance Board while playing Wii Fit or Wii Fit Plus could be an effective assessment tool for athletic trainers and other allied health care providers because the board and games have the potential to improve balance and provide objective measures of balance over time to monitor rehabilitation progress. Clark et al⁸ compared COP data from a research-grade force plate with COP data from a Wii Balance Board. To do this, the authors interfaced the Wii Balance Board with a laptop computer using custom-written software, bypassing game software and scores. The results indicated that a Wii Balance Board produced raw COP means that were similar to the means that the force plate produced, and it had good to excellent test-retest reliability.⁸ However, the creation and use of custom-written software similar to that used by Clark et al⁸ is not practical in the clinical setting. Furthermore, doubt exists about the validity and reliability of the balance scores that arise from the software implemented in the Wii Fit and Wii Fit Plus games. Until this information is elucidated, clinicians using balance scores from the Wii Fit or Wii Fit Plus games to monitor progress during rehabilitation might be making clinical decisions based on inaccurate or unreliable data. Therefore, the purpose of this investigation was to determine the concurrent validity and the intrasession and intersession reliability of balance scores the Wii Fit game produces.

METHODS

Design and Participants

For this descriptive laboratory study, I used a test-retest study design to achieve the stated aims. A total of 45 participants (age = 27.0 ± 9.8 years, height = 170.9 ± 9.2 cm, mass = 72.4 ± 11.8 kg) completed 2 test sessions that were separated by 1 week. All participants were between 18 and 57 years of age and were *recreationally active*, which operationally was defined as participating in aerobic activities at least 3 times each week for at least a total of 90 minutes. In addition, participants had no known visual, vestibular, and neurologic impairments; had not had a concussion for at least 3 months before testing; and were not taking medication that might impair balance for at least 1 month before testing. However, no limitations were placed on the presence of acute (eg, lateral ankle sprain, anterior cruciate ligament [ACL] rupture) or chronic (eg, chronic ankle instability, ACL reconstruction) lower extremity conditions before enrollment in the study so that a more heterogeneous sample would be captured. A heterogeneous sample was used to ensure that reliability and validity would be established regardless of the type of participant tested.^{9,10} To quantify disability associated with acute and chronic lower extremity conditions, all participants completed 2 questionnaires: a basic lower extremity medical history questionnaire and the Lower Extremity Functional Scale (LEFS).¹¹ In total, 22 people had histories of at least 1 lower extremity musculoskeletal condition, with 6 participants reporting multiple conditions (ankle sprain = 6, ankle instability = 4, ACL rupture = 5,

meniscal injury = 2, lower extremity fracture = 5, other conditions = 6 [eg, low back pain, tendinitis]). People without histories of lower extremity injuries had an average LEFS score of 78.4 ± 2.3 , whereas people with histories of lower extremity injuries had an average LEFS score of 72.4 ± 8.7 out of 80 possible points. In research on the LEFS questionnaire, investigators^{11,12} have shown good clinometric properties across a wide range of musculoskeletal disorders of the lower extremity. All participants provided written informed consent before completing the test sessions, and the study was approved by the Institutional Review Board of the University of North Carolina at Charlotte.

Procedures

All participants reported to the Biodynamics Research Laboratory for both test sessions. Upon reporting for the first test session, they were exposed to and allowed to practice each of the balance tasks associated with the research protocol. This practice time was followed by a mandatory 10-minute rest period during which participants completed the medical history and LEFS questionnaire. They completed 2 test trials per limb, if applicable, of all balance tasks with 30-second rest periods between trials and 1-minute rest periods between tasks. Established balance tasks, which were completed during the first test session only, included a single-limb-stance task on a force plate and the Star Excursion Balance Test (SEBT).¹³ The established measures of static and dynamic balance used in our investigation were chosen for the following reasons: (1) they have been shown to be reliable,^{13,14} (2) they have been able to discriminate between groups with known balance impairments and controls,¹⁵⁻¹⁷ (3) they commonly have been reported in the literature,¹³⁻¹⁷ and (4) they were deemed to be functionally appropriate for the sample tested. Wii Fit balance activities were completed during the first and second test sessions. The order of test completion during each test session was randomized for each participant before the start of data collection to minimize potential effects of learning and fatigue. The second test session was completed 1 week after the first test session.

Established Balance Tasks. For the single-limb-stance task, participants stood barefoot on 1 limb on an instrumented force plate with hands on hips and eyes open.^{18,19} Participants completed two 30-second trials per limb on a triaxial force plate (model 4060NC; Bertec Corporation, Columbus, OH) mounted flush with the surrounding floor. Ground reaction force data were collected at 200 Hz and filtered using a fourth-order, zero-lag, low-pass filter with a cutoff frequency of 5 Hz. The chosen outcome measures were average sway in the anteroposterior (AP) and mediolateral (ML) directions and the AP, ML, and resultant COP excursions. Given that the single-limb-stance task trials were the same duration, the AP, ML, and resultant COP excursions are analogous to the average AP, ML, and resultant COP velocities (excursion divided by time). A 30-second time frame was chosen to keep the duration of the task similar to the duration of the Wii Fit balance activities.

Participants also completed 2 reach trials per limb in 3 SEBT reach directions (anterior, posteromedial, and posterolateral). Testing only these 3 directions is recommended because of the redundancy in information

Table 1. Wii Fit^a Activities

Assessment	Description	Stance Type	Balance Type	Outcome Measure
Basic balance	Evaluates ability to control frontal-plane weight distribution.	Dual limb	Dynamic	Number of targets hit
Agility	Evaluates ability to move center of pressure to targeted areas.	Dual limb	Dynamic	Number of targets hit
Walking	Evaluates frontal-plane center-of-pressure displacement during 20 stationary walking steps.	Dual limb	Dynamic	Balance percentage
Steadiness	Evaluates resultant center-of-pressure displacement with visual, sway reference, and no visual feedback.	Dual limb	Static	Stillness percentage
Single-leg stance	Evaluates frontal-plane center-of-pressure displacement during a 30-s single-limb stance.	Single limb	Static	Left and right stability
Deep breathing	Evaluates resultant center-of-pressure movement during a deep-breathing exercise.	Dual limb	Static	Points 0–100
Tree	Evaluates frontal-plane center-of-pressure movement during a tree yoga pose.	Single limb	Static	Points 0–50 per limb
Standing knee	Evaluates frontal-plane center-of-pressure movement while repeatedly pulling a knee to a chest.	Single limb	Static	Points 0–50 per limb
Palm tree	Evaluates resultant center-of-pressure movement while performing a sustained heel raise.	Dual limb	Static	Points 0–100
Single-leg extension	Evaluates frontal-plane center-of-pressure movement while performing 10 repetitions of the hip extension exercise.	Single limb	Dynamic	Points 0–50 per limb
Single-leg twist	Evaluates frontal-plane center-of-pressure movement while performing 10 repetitions of the hip-flexion and adduction exercise.	Single limb	Dynamic	Points 0–50 per limb
Sideways leg lift	Evaluates frontal-plane center-of-pressure movement while performing 10 repetitions of the hip-abduction exercise.	Single limb	Dynamic	Points 0–50 per limb

^a Nintendo of America Inc, Redmond, WA.

obtained when all 8 directions are tested.¹⁶ To conduct the reach trials, each participant placed the stance limb in the center of the SEBT grid and placed his or her hands on the hips. Next, he or she reached as far as possible in the 3 reach directions with the contralateral limb. *Reach distance* was defined as the farthest point that an individual could touch without accepting weight and while maintaining balance through the return to a bilateral stance.^{13,16,20,21} All reach distances were normalized to the participant's leg length (distance from the anterosuperior iliac spine to the ipsilateral medial malleolus). The SEBT is a reliable measure of dynamic balance (intraclass correlation coefficient [ICC] [2,1] = 0.85–0.96).¹³

Wii Fit Balance Activities. Twelve Wii Fit activities were chosen to validate as many potential assessment tools as possible (Table 1). These activities consisted of a variety of dual-limb and single-limb support tasks that were both static and dynamic. All tasks were conducted using a standard Wii game console, Wii Balance Board, and Wii Fit software and were completed in accordance with the game software. The game console was connected to a projector (model XD490U DLP; Mitsubishi Electric Visual Solutions America, Inc, Irvine, CA) so that the gaming screen was 2 m in front of the participants and the center of the screen was at eye level. During each test session, 7 Wii Fit activities were completed (deep breathing, tree, standing knee, palm tree, single-leg extension, single-leg twist, sideways leg lift) twice per limb. The remaining activities (basic balance, agility, walking, steadiness, and single-leg stance) are part of the Wii Fit's body test, and the game software randomly selects 2 of the 5 activities to be completed per body test. To collect as much data as possible, 2 body tests were completed for each test session. Thus, each participant completed 5 body test activities (basic balance, agility, walking, steadiness, and single-leg stance) per test session. The Figure illustrates the testing sequence for the Wii Fit activities. Two trial averages were

not always available for the body test activities because of the random selection made by the game software. In such cases, the single trial score was used for further analysis.

Statistical Analysis

Wii Fit balance activity scores initially were submitted to separate independent-samples *t* tests. These analyses determined that no differences existed between the left and right limbs of individuals who reported no history of lower extremity conditions when applicable ($P > .05$) and no differences existed between the uninjured limbs and the limbs of the participants who self-reported a history of lower extremity injury ($P > .05$). Therefore, data from both the left and right limbs of all participants were used to calculate concurrent validity and intrasession and intersession reliability.

To determine concurrent validity, Pearson product moment correlation calculations were used to assess the strength of the relationships between the Wii Fit balance activity scores and the COP and SEBT outcome measures. The quality of the correlation was interpreted as follows: less than 0.50 indicated *poor validity*; 0.50 to 0.75, *moderate to good validity*; and more than 0.75, *excellent validity*.²² Intrasession and intersession (1-week) reliability was quantified using dependent-samples *t* tests and ICCs (2,1). The precision of each Wii Fit activity outcome measure was calculated using the standard error of measurement (SEM). In addition, minimal detectable change (MDC) scores were used to quantify how much a Wii Fit balance score would need to change after a rehabilitation intervention for that change to be considered confidently a true change (ie, improvement or degradation) due to the intervention and not due to measurement error.²³ Reliability coefficients were interpreted as follows: less than 0.69 was considered *poor*; 0.70 to 0.79, *fair*; 0.80 to 0.89, *good*; and 0.90 to 1.00, *excellent*.²² The α level was

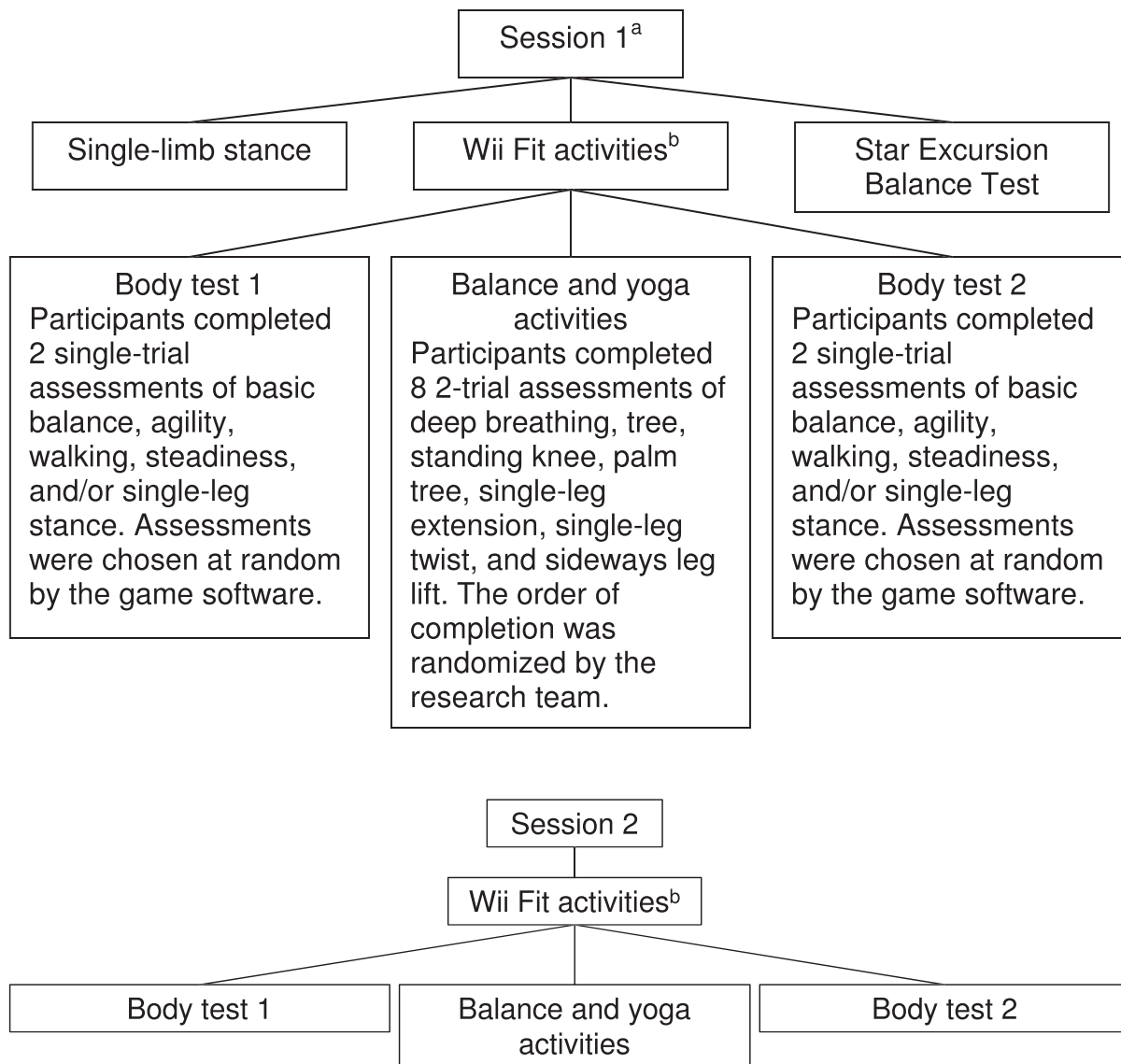


Figure. Flow of the data collection sequence for each test session. ^a Indicates completion of the single-leg stance and Wii Fit (Nintendo of America Inc, Redmond, WA) test and Star Excursion Balance Test were randomized for each participant during test session 1. ^b Indicates that during each test session, the Wii Fit activities, which included 2 body tests and 8 other activities, also were randomized for each participant.

set a priori at $\leq .05$ to determine if (1) the relationships were correlated, (2) the within-sessions trials were different, and (3) the between-sessions trial averages were different.

RESULTS

The group means, standard deviations, and number of limbs that completed a task during the first test session are presented in Table 2. The Pearson product moment correlations indicated that 4 relationships were correlated (Table 3). However, all 96 of the calculated relationships between Wii Fit balance activity scores and established balance task outcomes, even those identified as correlated, were interpreted as poor (< 0.50). Thus, concurrent validity between Wii Fit balance scores and established balance tasks outcome was not established. Indeed, the highest correlation coefficient was between the Wii Fit walking

activity score and the ML ($r = 0.32$, $P = .045$) and resultant ($r = 0.32$, $P = .05$) COP excursion outcomes.

Intrasession and intersession reliability estimates, including t and P values, precision (SEM), and MDC scores, for the Wii Fit balance scores are presented in Tables 4 and 5, respectively. None of the Wii Fit balance activity scores differed between trials 1 and 2 during the first test session (Table 4). One score from the Wii Fit balance activity (single-leg stance) received an intrasession ICC point estimate that was interpreted as good. Three balance scores (steadiness, deep breathing, and single-leg twist) had intrasession ICC point estimates that were interpreted as fair, whereas the remaining 8 activities had intrasession ICC point estimates of poor reliability. Similarly, 11 of the 12 Wii Fit balance scores demonstrated poor intersession reliability, whereas deep breathing demonstrated a fair reliability point estimate. Four other balance scores

Table 2. Performance on Established and Wii Fit Balance Tasks (Mean ± SD)

Activity	No. Completed ^a	Mean ± SD
Sway, cm		
Anteroposterior	90	7.28 ± 2.50
Mediolateral	90	6.21 ± 7.26
Center-of-pressure excursion, m		
Anteroposterior	90	2.26 ± 1.71
Mediolateral	90	3.88 ± 6.07
Resultant	90	5.05 ± 6.34
Star Excursion Balance Test, %		
Anterior	90	84.0 ± 8.9
Posterolateral	90	86.0 ± 11.2
Posteromedial	90	91.9 ± 10.2
Wii Fit		
Basic balance, levels passed	34	4.0 ± 0.6
Agility, boxes hit	34	13.4 ± 1.6
Walking, %	39	62.3 ± 16.2
Steadiness, %	34	61.5 ± 13.4
Single-leg stance, points out of 50	39	68.4 ± 8.0
Deep breathing, points out of 50	45	86.1 ± 7.1
Tree, points out of 100	90	32.3 ± 7.7
Standing knee, points out of 100	90	41.1 ± 7.3
Palm tree, points out of 100	90	63.9 ± 15.2
Single-leg extension, points out of 100	90	37.2 ± 6.0
Single-leg twist, points out of 100	90	40.4 ± 4.1
Sideways leg lift, points out of 100	90	36.3 ± 7.2

^a Indicates the total number of participants who completed at least 1 trial of the listed activity during the first test session.

(agility, single-leg extension, single-leg twist, sideways leg lift) differed between the first and second test sessions, with scores improving during the second test session in all 4 activities. High MDC scores, which represent in units of the outcome measure the minimal change that must occur in the Wii Fit activity before the clinician can be truly confident that the change is not due to error, also indicated poor stability of Wii Fit balance scores.

DISCUSSION

The purpose of this investigation was to determine the suitability of Wii Fit balance activity scores as objective assessment tools to monitor rehabilitation progression. Indeed, an inexpensive (relative to traditional force platforms), portable, valid, and reliable way to objectively assess balance would be valuable to a wide range of allied health care providers.⁸ The most important findings of this investigation were that Wii Fit balance scores possessed poor levels of concurrent validity and generally poor intrasession and intersession reliability estimates, as well as high MDC scores. These findings are contrary to those of Clark et al,⁸ who demonstrated that the raw COP data collected by a Wii Balance Board had reliability estimates comparable with those collected by a traditional research laboratory force platform. However, they bypassed the Wii software by interfacing the Wii Balance Board with a laptop computer using custom-written software, which limits the direct comparisons between this investigation and their work.⁸

Poor levels of concurrent validity (ie, Wii Fit balance scores were correlated poorly with balance scores from

Table 3. Relationships Between Wii Fit^a Activities and Established Balance Tasks

Wii Fit Activity (n ^b)	Sway						Center-of-Pressure Excursion						Star Excursion Balance Test							
	Anteroposterior		Mediolateral		P		Anteroposterior		Mediolateral		P		Resultant		Anterior		Posterolateral		Posteromedial	
	r	P	r	P	r	P	r	P	r	P	r	P	r	P	r	P	r	P	r	P
Basic balance (34)	0.15	.41	0.14	.46	0.24	.19	0.30	.10	0.26	.15	-0.19	.29	0.26	.15	-0.19	.29	0.26	.15	-0.06	.74
Agility (34)	-0.26	.15	-0.17	.35	0.06	.72	-0.07	.69	-0.13	.94	<-0.01	.99	-0.13	.94	<-0.01	.99	-0.13	.94	0.25	.15
Walking (39)	0.09	.57	0.13	.43	0.30	.06	0.32 ^c	.045	0.32 ^c	.05	0.21	.20	0.32 ^c	.05	0.21	.20	0.32 ^c	.05	0.21	.20
Steadiness (34)	-0.08	.67	-0.04	.85	0.024	.18	0.27	.133	0.25	.15	<0.01	.99	0.25	.15	<0.01	.99	0.25	.15	0.12	.51
Single-leg stance (39)	-0.09	.59	-0.21	.21	-0.13	.42	-0.07	.66	-0.09	.57	-0.19	.25	-0.09	.57	-0.19	.25	-0.09	.57	-0.04	.79
Deep breathing (45)	0.17	.25	0.10	.51	0.14	.36	0.13	.40	0.14	.38	0.20	.19	0.14	.38	0.20	.19	0.14	.38	0.17	.28
Tree (90)	-0.05	.63	0.02	.86	0.09	.39	0.06	.56	0.07	.51	0.07	.53	0.07	.51	0.07	.53	0.07	.51	0.06	.58
Standing knee (90)	-0.02	.88	0.03	.81	0.07	.53	0.05	.64	0.06	.60	0.02	.88	0.06	.60	0.02	.88	0.06	.60	-0.06	.57
Palm tree (90)	-0.16	.31	0.08	.62	0.02	.90	0.01	.94	0.03	.85	0.10	.53	0.03	.85	0.10	.53	0.03	.85	0.08	.60
Single-leg extension (90)	-0.07	.51	-0.11	.63	0.02	.85	-0.02	.85	-0.01	.93	0.07	.49	-0.01	.93	0.07	.49	-0.01	.93	-0.02	.84
Single-leg twist (90)	0.03	.76	-0.06	.55	0.04	.70	-0.02	.89	-0.01	.98	0.21 ^c	.048	-0.01	.98	0.21 ^c	.048	-0.01	.98	0.20	.06
Sideways leg lift (90)	0.14	.19	-0.03	.77	0.06	.57	-0.01	.96	0.01	.93	-0.03	.80	0.01	.93	-0.03	.80	0.01	.93	0.01	.96

^a Nintendo of America Inc, Redmond, WA.

^b Indicates the total number of participants who completed at least 1 trial of the listed Wii Fit activity during the first test session.

^c Indicates a difference ($P \leq .05$).

Table 4. Intrasection Reliability Estimates for Wii Fit Balance^a Activity Scores

Wii Fit Activity (n ^b)	Trial 1 Mean ± SD	Trial 2 Mean ± SD	<i>t</i> Value	<i>P</i> Value	Intrasection Intraclass Correlation Coefficient	Standard Error of Measurement	Minimal Detectable Change
Basic balance (5)	4.0 ± 0.7	4.0 ± 0.8	1.7	.17	0.57	0.5	1.4
Agility (7)	14.4 ± 2.1	14.6 ± 1.1	-0.2	.84	0.64	1.6	4.4
Walking (7)	56.9 ± 14.8	71.7 ± 19.8	-1.9	.11	0.43	13.0	36.0
Steadiness (7)	64.1 ± 6.7	63.1 ± 8.8	0.4	.72	0.75	3.9	10.8
Single-leg stance (6)	72.83 ± 4.7	69.0 ± 7.2	1.9	.12	0.80	2.7	7.5
Deep breathing (45)	86.3 ± 7.9	86.3 ± 8.3	0.041	.97	0.72	4.3	11.9
Tree (90)	32.6 ± 7.4	32.6 ± 9.3	0.035	.97	0.56	5.6	15.5
Standing knee (90)	40.5 ± 8.8	42.5 ± 8.0	-1.7	.09	0.39	6.5	18.0
Palm tree (45)	63.6 ± 18.1	62.9 ± 19.5	0.2	.85	0.40	14.6	40.5
Single-leg extension (90)	37.1 ± 7.5	37.1 ± 6.7	-0.031	.98	0.67	4.1	11.4
Single-leg twist (90)	40.1 ± 4.5	40.7 ± 5.0	-1.3	.19	0.72	2.5	6.9
Sideways leg lift (90)	36.7 ± 7.9	35.9 ± 9.9	0.8	.45	0.47	6.5	18.0

^a Nintendo of America Inc, Redmond, WA.

^b Indicates the total number of participants who completed 2 trials of the listed Wii Fit activity during the first test session.

established balance tasks) mean that clinicians cannot predict accurately how well an individual will perform on a Wii Fit activity based on COP outcomes during a 30-second single-leg stance or his or her reach distance during an SEBT or vice versa. The established measures of balance that I used (single-leg stance on a force plate and SEBT) are not considered the criterion standard for static and dynamic balance. However, a criterion standard for static or dynamic balance has not been established in the literature. In addition, several of the Wii Fit activities are performed during dual-limb support, but both of the established measures used are completed during single-limb support. Thus, the current established measures of balance might not have been the most appropriate tasks from which to establish concurrent validity for some Wii Fit activities, which was a limitation of the investigation. However, a static dual-limb support task was not chosen because static dual-limb support tasks are not considered challenging enough to evaluate postural control or monitor the rehabilitation progress of physically active individuals with lower extremity musculoskeletal conditions.

Another potential reason for the poor validity estimates observed in this investigation is the algorithm used by the Wii Fit software to calculate the balance scores. Currently, the exact algorithms used to calculate Wii Fit balance

scores are unknown to the general public, but regardless of the calculation details, the software would apply the algorithm the same way for each trial, thereby giving standard outcomes for each balance board activity. Thus, the Wii Fit software might explain the poor concurrent validity, but the gaming software is not likely responsible for the poor reliability estimates observed. In the future, researchers should compare the Wii Fit dual-limb support tasks with an established dual-limb support task and attempt to determine concurrent validity with other established static and dynamic balance tasks (eg, Berg Balance Scale, Balance Error Scoring System, hop to stabilization) and outcomes (eg, time to boundary, time to stabilization). In addition, concurrent validity should be confirmed in various age groups and clinical populations through mutually exclusive investigations (ie, separately determine concurrent validity in people with acute ankle sprains, chronic low back pain, and other lower extremity conditions). Using this systematic approach in the future will provide for each injured group more statistical power than and expand on the clinical applicability of the results obtained from the current heterogeneous sample.

The general poor reliability estimates and high MDC scores strongly suggest that clinicians should not use the balance scores produced by Wii Fit activities as an

Table 5. Intersession Reliability Estimates for Wii Fit Balance^a Activity Scores

Wii Fit Activity (n ^b)	Day 1 Mean ± SD	Day 2 Mean ± SD	<i>t</i> Value	<i>P</i> Value	Intersession Intraclass Correlation Coefficient	Standard Error of Measurement	Minimal Detectable Change
Basic balance (10)	4.0 ± 0.5	4.0 ± 0.7	0.8	.44	0.67	0.3	0.8
Agility (12)	12.7 ± 2.4	14.0 ± 2.1	-1.9	.09	0.57	1.5	4.2
Walking (15)	66.1 ± 25.9	61.4 ± 18.3	0.5	.60	0.45	16.3	45.2
Steadiness (5)	59.4 ± 18.0	67.2 ± 7.3	-1.2	.31	0.58	8.2	22.7
Single-leg stance (6)	71.7 ± 3.0	67.5 ± 5.2	2.2	.08	0.57	2.7	7.5
Deep breathing (45)	82.1 ± 13.0	86.5 ± 7.2	-3.0	.005	0.74	5.1	14.1
Tree (90)	30.2 ± 10.0	32.6 ± 7.1	-2.3	.02	0.57	5.6	15.5
Standing knee (90)	40.7 ± 8.9	41.4 ± 6.7	-0.6	.54	0.37	7.0	19.4
Palm tree (45)	65.1 ± 15.4	64.2 ± 13.8	0.3	.75	0.29	12.2	33.8
Single-leg extension (90)	34.5 ± 7.6	37.2 ± 6.2	-2.8	.007	0.37	5.5	15.2
Single-leg twist (90)	37.5 ± 7.4	40.3 ± 4.2	-3.6	.001	0.46	4.1	11.4
Sideways leg lift (90)	31.2 ± 10.8	36.1 ± 7.3	-4.1	<.001	0.47	6.5	18.0

^a Nintendo of America Inc, Redmond, WA.

^b Indicates the total number of participants who completed at least 1 trial of the listed Wii Fit activity during both the first and second test session.

objective tool to monitor progress during rehabilitation. Indeed, the generally high MDC scores indicate that a substantial and impractical change in Wii Fit balance scores is needed to exceed the error of the measurement. For example, a change equal to roughly 50% of the recorded mean for the tree activity is needed to be confident that intersession change (Table 5) is due to the intervention delivered and not to the measurement error.

A plausible cause of the generally poor intrasession and intersession reliability involves the novelty of the Wii Fit activities to the participants. In this investigation, participants were allowed to practice each Wii Fit activity once (per test session) before completing the test trials. One practice trial per activity creates potentially insufficient self-confidence and does not allow participants to feel comfortable with the tasks, and it should be considered a limitation of the investigation. The balance score improvements between the first and second test sessions in the agility, single-leg extension, single-leg twist, and sideways leg lift tasks provide evidence that supports this hypothesis. Increasing the number of trials used to calculate the Wii Fit activity averages possibly could influence positively both intrasession and intersession reliability estimates. For example, a standard protocol with multiple practice trials and at least 3 test trials used to calculate average scores might improve the reliability of the Wii Fit activities, as previously reported with the SEBT.¹³ Thus, the reliability and subsequently the concurrent validity of Wii Fit balance scores might improve as a user's experience with the game activities increases or as a greater number of trials serve as the basis for the averages used for analysis. However, future investigations are needed to confirm or refute this hypothesis. Research also is needed to confirm both the intrasession and intersession reliability of the Wii Fit activities associated with the body test activities (ie, basic balance, agility, walking, steadiness, and single-leg stance). Indeed, during this investigation, only a few participants (as few as 5) completed these activities, which limits the power available to quantify reliability estimates and the ability to draw solid conclusions from the data.

I undertook this investigation because of the growing use of interactive gaming technology, in particular the Wii and Wii Fit, by allied health care providers and especially athletic trainers. The results suggest that Wii Fit balance scores, as produced by the game software, should not be used as tools to assess or monitor progress during rehabilitation. If future researchers confirm these findings, several emerging techniques might enable allied health care providers to take advantage of the portability and cost advantages of a Wii and Wii Balance Board while overcoming the apparent poor reliability and validity observed in this investigation.^{8,24–27} In general, these techniques involve bypassing the scores produced by the game software and simply using the raw data from the Wii Balance Board or other peripheral devices. However, without specific skill sets, use of these techniques likely will require collaboration with units or individuals familiar with computer programming. Most importantly, the findings do not condemn the use of a Wii, Wii Balance Board, or other interactive gaming system by allied health care providers. Indeed, the use of interactive gaming technologies as tools or exercise in the rehabilitation process is recommended because of the large number of purported benefits (eg, improving balance,⁶ improving

compliance⁶) associated with using interactive gaming technology.

CONCLUSIONS

Wii Fit balance scores had poor concurrent validity compared with the COP outcomes from a 30-second single-limb stance and the anterior, posteromedial, and posterolateral reach distances from the SEBT. In addition, balance scores associated with the 12 Wii Fit activities tested in this investigation generally had poor intrasession and intersession reliability and generally high MDC scores relative to the mean score of the activity. However, future research is needed to confirm the reliability of those activities associated with the Wii Fit body test. These findings indicate the Wii Fit balance scores should not be used as an objective measure of rehabilitation progress, but using a Wii or Wii Fit in a clinical setting is still recommended based on the other purported benefits.

ACKNOWLEDGMENTS

This study was funded by a research grant from the Mid Atlantic Athletic Trainers' Association.

REFERENCES

1. Brumels KA, Young AM. Use of DDR in rehabilitation and prevention of athletic injuries. *Clin Kinesiol*. Summer 2006. http://findarticles.com/p/articles/mi_6810/is_2_60/ai_n28416099. Accessed November 18, 2011.
2. Deutsch JE, Borbely M, Filler J, Huhn K, Guarrera-Bowly P. Use of a low-cost, commercially available gaming console (Wii) for rehabilitation of an adolescent with cerebral palsy. *Phys Ther*. 2008;88(10):1196–1207.
3. Dixon T. A Wii spot of fun. *Australian Ageing Agenda*. March/April 2008:28–32.
4. Wuang YP, Chiang CS, Su CY, Wang CC. Effectiveness of virtual reality using Wii gaming technology in children with Down syndrome. *Res Dev Disabil*. 2011;32(1):312–321.
5. González-Fernández M, Gil-Gómez JA, Alcañiz M, Noé E, Colomer C. eBaViR, easy balance virtual rehabilitation system: a study with patients. *Ann Rev Cyberther Telemed*. 2010;8:49–53.
6. Middlemas DA, Basilicato J, Prybicien MA, Savoia J, Biodoglio J. Incorporating gaming technology into athletic injury rehabilitation. *Athl Train Sports Health Care*. 2009;1(2):79–84.
7. Saposnik G, Teasell R, Mamdani M, et al. Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle. *Stroke*. 2010;41(7):1477–1484.
8. Clark RA, Bryant AL, Pua Y, McCrory P, Bennell K, Hunt M. Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. *Gait Posture*. 2010;31(3):307–310.
9. Docherty CL, Rybak-Webb K. Reliability of the anterior drawer and talar tilt tests using the LigMaster joint arthrometer. *J Sport Rehabil*. 2009;18(3):389–397.
10. Wikstrom EA, Tillman MD, Chmielewski TL, Cauraugh JH, Naugle KE, Borsa PA. Self-assessed disability and functional performance in individuals with and without ankle instability: a case control study. *J Orthop Sports Phys Ther*. 2009;39(6):458–467.
11. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. *Phys Ther*. 1999;79(4):371–383.
12. Yeung TS, Wessel J, Stratford PW, Macdermid J. Reliability, validity, and responsiveness of the lower extremity functional scale for inpatients of an orthopaedic rehabilitation ward. *J Orthop Sports Phys Ther*. 2009;39(6):468–477.
13. Hertel J, Miller SJ, Denegar CR. Intratester and intertester reliability during the Star Excursion Balance Tests. *J Sport Rehabil*. 2000;9(2):104–116.

14. Hertel J, Olmsted-Kramer LC, Challis JH. Time-to-boundary measures of postural control during single leg quiet standing. *J Appl Biomech.* 2006;22(1):67–73.
15. Arnold BL, De La Motte S, Linens S, Ross SE. Ankle instability is associated with balance impairments: a meta-analysis. *Med Sci Sports Exer.* 2009;41(5):1048–1062.
16. Hertel J, Braham RA, Hale SA, Olmsted-Kramer LC. Simplifying the star excursion balance test: analyses of subjects with and without chronic ankle instability. *J Orthop Sports Phys Ther.* 2006;36(3):131–137.
17. Wikstrom EA, Fournier KA, McKeon PO. Postural control differs between those with and without chronic ankle instability. *Gait Posture.* 2010;32(1):82–86.
18. Tropp H, Odenrick P, Gillquist J. Stabilometry recordings in functional and mechanical instability of the ankle joint. *Int J Sports Med.* 1985;6(3):180–182.
19. Tropp H, Odenrick P. Postural control in single-limb stance. *J Orthop Res.* 1988;6(6):833–839.
20. Gribble PA, Hertel J, Denegar CR, Buckley WE. The effects of fatigue and chronic ankle instability on dynamic postural control. *J Athl Train.* 2004;39(4):321–329.
21. Olmsted LC, Carcia CR, Hertel J, Shultz SJ. Efficacy of the Star Excursion Balance Tests in detecting reach deficits in subjects with chronic ankle instability. *J Athl Train.* 2002;37(4):501–506.
22. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates Inc; 1988.
23. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. *J Clin Epidemiol.* 2001;54(12):1204–1217.
24. Anderson F, Annett M, Bischof WF. Lean on Wii: physical rehabilitation with virtual reality Wii peripherals. *Stud Health Technol Inform.* 2010;154:229–234.
25. Bokhari R, Bollman-McGregor J, Kahoi K, Smith M, Feinstein A, Ferrara J. Design, development, and validation of a take-home simulator for fundamental laparoscopic skills: using Nintendo Wii for surgical training. *Am Surg.* 2010;76(6):583–586.
26. González-Fernández M, Gil-Gómez JA, Alcañiz M, Noé E, Colomer C. eBaViR, easy balance virtual rehabilitation system: a study with patients. *Stud Health Technol Inform.* 2010;154:61–66.
27. Lange B, Flynn S, Proffitt R, Chang CY, Rizzo AS. Development of an interactive game-based rehabilitation tool for dynamic balance training. *Top Stroke Rehabil.* 2010;17(5):345–352.

Address correspondence to Erik A. Wikstrom, PhD, ATC, FACSM, Department of Kinesiology, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223. Address e-mail to ewikstrom@uncc.edu.