Effects of Footwear on Measurements of Balance and Gait in Women Between the Ages of 65 and 93 Years
Solveig A Arnadottir and Vicki S Mercer

The online version of this article, along with updated information and services, can be found online at: http://ptjournal.apta.org/content/80/1/17

Collections
This article, along with others on similar topics, appears in the following collection(s):
Geriatrics: Other
Tests and Measurements

e-Letters
To submit an e-Letter on this article, click here or click on "Submit a response" in the right-hand menu under "Responses" in the online version of this article.

E-mail alerts
Sign up here to receive free e-mail alerts
Effects of Footwear on Measurements of Balance and Gait in Women Between the Ages of 65 and 93 Years

Background and Purpose. Footwear is not consistently standardized in the administration of the Functional Reach Test (FRT), Timed Up & Go Test (TUG), and 10-Meter Walk Test (TMW). This study was conducted to determine whether footwear affected performance on these tests in older women. Subjects. Thirty-five women, aged 65 to 93 years, were recruited from assisted living facilities and retirement communities. Methods. Each subject performed the FRT, TUG, and TMW while wearing walking shoes, wearing dress shoes, and barefooted. Because of space constraints at the facilities where the testing was performed, 22 subjects performed the FRT and TUG on a linoleum floor and 13 subjects performed the tests on a firm, low-pile, carpeted floor. All 35 subjects completed the TMW on a firm, low-pile, carpeted floor. One-way repeated-measures analyses of variance (ANOVAs) and a Tukey Honestly Significant Difference test were used to compare the outcomes for the 3 footwear conditions, with separate ANOVAs conducted for the different floor surfaces for the FRT and TUG. Results. Subjects performed better on the FRT when barefooted or wearing walking shoes compared with when they wore dress shoes, regardless of floor surface. Differences were found among all footwear conditions for the TUG performed on the linoleum floor and for the TMW. For these tests, the women moved fastest in walking shoes, slower barefooted, and slowest wearing dress shoes. Conclusion and Discussion. Footwear should be documented and should remain constant from one test occasion to another when the FRT, TUG, and TMW are used in the clinic and in research. Footwear intervention may improve performance of balance and gait tasks in older women. [Arnadottir SA, Mercer VS. Effects of footwear on measurements of balance and gait in women between the ages of 65 and 93 years. Phys Ther. 2000;80:17–27.]

Key Words: Aging, Balance, Footwear, Gait, Measurement.

Solveig A Arnadottir
Vicki S Mercer
E valuation of physical impairments and functional limitations has become an essential part of clinical geriatrics as well as aging research. Physical function refers to the normal performance of an individual in managing daily routines and represents an important aspect of the individual’s overall health. Decline in physical mobility is a major concern for many older people. Even small improvements in the areas of mobility, balance, and gait may contribute valuable benefits in terms of quality of life. Therefore, measures of balance and gait performance are critical in the field of aging and essential to help health care professionals and researchers keep their focus on the real needs of the older population.

Multiple assessment instruments have been developed and validated, focusing on different aspects of physical performance. These instruments are designed to provide objective measurements of physical impairments or functional limitations for screening, evaluating status, monitoring changes, and predicting outcomes for individuals and populations. Many of these instruments, however, were developed for research purposes and are impractical for use in geriatric clinics because of their length, complexity, or equipment requirements or because they are not targeted toward older populations.

In order to be appropriate for any use, a measurement instrument should be examined in the population on which it will be used, and the psychometric properties of the test should be reported for that population. The instrument also should be safe, inexpensive, and easily incorporated into clinical practice, requiring minimal time and expertise to administer. Among the physical performance measures that fulfill these requirements are the Functional Reach Test (FRT), the Timed Up & Go Test (TUG), and measures of self-selected gait speed such as the 10-Meter Walk Test (TMW). All 3 of these scales are continuous measures and, therefore, theoretically more responsive to change than categorical scales. They can easily be administered in the subject’s own environment, which has advantages over testing in an artificial laboratory setting. The tests also were designed for use by different health care professionals, making them appropriate as multidisciplinary assessment instruments.

Because balance and gait have many different domains or components, no single measure of these abilities appears to be useful for all settings. The FRT captures the ability to control movement of the center of gravity over a fixed base of support, and the TMW and TUG include the ability to adjust the center of gravity continuously over a moving base of support. The FRT was developed by Duncan et al, who defined functional reach as “the maximal distance one can reach forward beyond arm’s length while maintaining a fixed base of support in the standing position.” In a sample of 128 community volunteers aged 20 to 87 years, FRT scores correlated with measurements of center-of-pressure excursion (Pearson $r = .71$) and had excellent test-retest reliability, with an intraclass correlation coefficient (ICC) of .92. Concurrent validity as a marker of physical frailty in community-dwelling elderly people (aged 66–104 years), predictive validity in identifying risk of falls in community-dwelling male veterans (aged 70–104 years), and sensitivity to change in balance in inpatient male veterans (aged 40–105 years) undergoing physical rehabilitation have been reported for the FRT. According to Light et al, a trained clinician should be capable of reading the FRT distance on a yardstick to the nearest 0.5 in (1.27 cm).

The TUG is typically used to evaluate basic mobility in elderly people. The test measures the time taken to stand up from a chair, walk 3 m at a comfortable and safe pace, turn around, walk back to the chair, and sit down. Podsiadlo and Richardson investigated reliability and validity for the TUG in a sample of 60 community-dwelling people, 60 to 90 years of age. Concurrent validity was based on correlations with log-transformed scores on the Berg Balance Scale ($r = -.81$), self-selected gait speed ($r = -.61$), and Barthel Index ($r = -.78$).

SA Arnadottir, PT, was a student at the University of North Carolina at Chapel Hill at the time this research was completed in partial fulfillment of the requirements for her Master of Science degree in physical therapy.

VS Mercer, PT, PhD, is Assistant Professor, Division of Physical Therapy, Department of Allied Health Sciences, CB# 7135, Medical School Wing E, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7135 (USA) (vmercer@css.unc.edu). Address all correspondence to Dr Mercer.

Concept and research design, writing, data analysis, facilities and equipment, and consultation were provided by Arnadottir and Mercer. Data collection, project management, fund procurement, subjects, and institutional liaisons were provided by Arnadottir. Jama Lynn Purser, Michael T Gross, and Cherie Rosemond also contributed to concept and research design and consultation (including review of manuscript before submission), and Cherie Rosemond contributed to provision of subjects.

This study was approved by the Committee on the Protection of the Rights of Human Subjects at the University of North Carolina at Chapel Hill.

This article was submitted August 20, 1998, and was accepted August 23, 1999.
Excellent intrarater reliability (ICC = .99) and interrater reliability (ICC = .99) were reported for a subgroup of 22 people.14

The TMW is a measure of self-selected walking speed,21,22 which, according to Cress et al.,23 is the best predictor of self-perceived function and overall physical performance. Test-retest reliability has been documented as ICC = .87 in older people (mean age = 74.5 years, SD = 5.7) with Parkinson disease (n = 14).22 Although gait speed slows with age, extreme slowness is an indication of frailty and is predictive of falls and nursing home placement.9

Standardization of test procedures is often critical for reliable generalization of results from one patient to another or from one facility to another.24 The type of footwear worn by the patient or subject is not consistently standardized in the administration of the FRT, TUG, and TMW. In the first article published on the FRT, Duncan et al.13 reported that subjects performed the test barefooted. Protocols for assessment of functional reach at the Center for the Study of Aging and Human Development at Duke University Medical Center provide for testing of subjects either barefooted or wearing shoes with heels of 1.27 cm (0.5 in) or less.21 The assumption being that these different footwear conditions produce essentially identical results. In previous studies in which the FRT was used as an outcome measure, the authors rarely mentioned footwear when describing the measurement procedures.25–28

Footwear also is not standardized for the TUG or TMW. Podsiadlo and Richardson14 simply described their subjects as wearing their “regular footwear” when performing the TUG. Other protocols for the TUG use similar statements of “regular” or “normal” footwear without further description.21,27 Footwear used during self-selected walking speed tests such as the TMW is commonly described as “normal walking shoes” or is not mentioned.15,22,29–31

Research on the direct effects of footwear on functional performance in the aging population is very limited. Briggs et al.32 found no effect of shoes versus no shoes on performance of the sharpened Romberg and one-legged stance tests among 71 female subjects with no known pathology between 60 and 86 years of age. For the “shoes on” condition, the outcomes of the tests were averaged across subjects, despite the wide variety of shoe styles the subjects wore. Lord and Bashford33 studied the effects of footwear on balance in 30 women aged 60 to 89 years. Outcome measures were postural sway, maximal balance range in the anterior-posterior direction, and coordinated stability using a “swaymeter.” The women performed better in flat shoes or barefoot than when they wore high-heeled shoes (all measurements were significant at P < .05). Postural sway was measured as the area traversed by the pen on the swaymeter and was, on average, 24 mm² less when performed barefoot than in low-heeled shoes and 50 mm² less when performed barefoot than in high-heeled shoes. Maximum balance range (anterior-posterior) was 1.4 cm larger in low-heeled shoes than in high-heeled shoes. On the coordinated stability test, the subjects received, on average, 6.8 fewer error points when wearing low-heeled shoes than when wearing high-heeled shoes and 7.6 fewer error points when barefoot than when wearing high-heeled shoes.

Several investigators34–45 have reported other effects of footwear on balance and gait characteristics. Few studies included older women, however, and no study encompassed the effect of footwear on FRT, TUG, and TMW scores. Older women are at high risk for functional decline, falls, and disablement as a consequence of poor balance and gait disorders.46–48 Research into these issues among elderly women is heavily dependent on assessment of physical performance and functional outcomes. At the same time, the characteristics of the “regular or normal footwear” worn by older women can vary greatly.92,33,49,50 Therefore, understanding the effects of footwear on balance and gait test scores in older women is essential.

The purpose of our study was to determine the effects of footwear on FRT, TUG, and TMW scores in older women. We limited our focus to older women because they are at higher risk for disablement than are men.17,48 Women aged 65 years and older are 50% more likely to report a fall in the previous year than are elderly men.51 In addition, women wear high-heeled dress shoes that may have a great impact on balance and gait performance. Consideration of footwear effects may improve the reliability for measurement tools used to assess the status and progress of older women at high risk for functional decline and disablement.

Two separate hypotheses were put forward. The first hypothesis was that older women would demonstrate the longest FRT distances when barefooted and the shortest distances when wearing dress shoes (high-heeled), with intermediate distances in the walking shoe (low-heeled) condition. This hypothesis was based on potential positive effects of increased proprioceptive input and negative effects of increased heel height on postural stability. More precise foot position awareness has been associated with barefoot standing as compared with shoed conditions.52 Better awareness of foot position might result in improved FRT performance in the barefoot condition. Increases in heel height, however, may result in decreases in the overall base of support.25,38,40,53

Excellent intrarater reliability (ICC = .99) and interrater reliability (ICC = .99) were reported for a subgroup of 22 people.14

The TMW is a measure of self-selected walking speed,21,22 which, according to Cress et al.,23 is the best predictor of self-perceived function and overall physical performance. Test-retest reliability has been documented as ICC = .87 in older people (mean age = 74.5 years, SD = 5.7) with Parkinson disease (n = 14).22 Although gait speed slows with age, extreme slowness is an indication of frailty and is predictive of falls and nursing home placement.9

Standardization of test procedures is often critical for reliable generalization of results from one patient to another or from one facility to another.24 The type of footwear worn by the patient or subject is not consistently standardized in the administration of the FRT, TUG, and TMW. In the first article published on the FRT, Duncan et al.13 reported that subjects performed the test barefooted. Protocols for assessment of functional reach at the Center for the Study of Aging and Human Development at Duke University Medical Center provide for testing of subjects either barefooted or wearing shoes with heels of 1.27 cm (0.5 in) or less.21 The assumption being that these different footwear conditions produce essentially identical results. In previous studies in which the FRT was used as an outcome measure, the authors rarely mentioned footwear when describing the measurement procedures.25–28

Footwear also is not standardized for the TUG or TMW. Podsiadlo and Richardson14 simply described their subjects as wearing their “regular footwear” when performing the TUG. Other protocols for the TUG use similar statements of “regular” or “normal” footwear without further description.21,27 Footwear used during self-selected walking speed tests such as the TMW is commonly described as “normal walking shoes” or is not mentioned.15,22,29–31

Research on the direct effects of footwear on functional performance in the aging population is very limited. Briggs et al.32 found no effect of shoes versus no shoes on performance of the sharpened Romberg and one-legged stance tests among 71 female subjects with no known pathology between 60 and 86 years of age. For the “shoes on” condition, the outcomes of the tests were averaged across subjects, despite the wide variety of shoe styles the subjects wore. Lord and Bashford33 studied the effects of footwear on balance in 30 women aged 60 to 89 years. Outcome measures were postural sway, maximal balance range in the anterior-posterior direction, and coordinated stability using a “swaymeter.” The women performed better in flat shoes or barefoot than when they wore high-heeled shoes (all measurements were significant at P < .05). Postural sway was measured as the area traversed by the pen on the swaymeter and was, on average, 24 mm² less when performed barefoot than in low-heeled shoes and 50 mm² less when performed barefoot than in high-heeled shoes. Maximum balance range (anterior-posterior) was 1.4 cm larger in low-heeled shoes than in high-heeled shoes. On the coordinated stability test, the subjects received, on average, 6.8 fewer error points when wearing low-heeled shoes than when wearing high-heeled shoes and 7.6 fewer error points when barefoot than when wearing high-heeled shoes.

Several investigators34–45 have reported other effects of footwear on balance and gait characteristics. Few studies included older women, however, and no study encompassed the effect of footwear on FRT, TUG, and TMW scores. Older women are at high risk for functional decline, falls, and disablement as a consequence of poor balance and gait disorders.46–48 Research into these issues among elderly women is heavily dependent on assessment of physical performance and functional outcomes. At the same time, the characteristics of the “regular or normal footwear” worn by older women can vary greatly.92,33,49,50 Therefore, understanding the effects of footwear on balance and gait test scores in older women is essential.

The purpose of our study was to determine the effects of footwear on FRT, TUG, and TMW scores in older women. We limited our focus to older women because they are at higher risk for disablement than are men.17,48 Women aged 65 years and older are 50% more likely to report a fall in the previous year than are elderly men.51 In addition, women wear high-heeled dress shoes that may have a great impact on balance and gait performance. Consideration of footwear effects may improve the reliability for measurement tools used to assess the status and progress of older women at high risk for functional decline and disablement.

Two separate hypotheses were put forward. The first hypothesis was that older women would demonstrate the longest FRT distances when barefooted and the shortest distances when wearing dress shoes (high-heeled), with intermediate distances in the walking shoe (low-heeled) condition. This hypothesis was based on potential positive effects of increased proprioceptive input and negative effects of increased heel height on postural stability. More precise foot position awareness has been associated with barefoot standing as compared with shoed conditions.52 Better awareness of foot position might result in improved FRT performance in the barefoot condition. Increases in heel height, however, may result in decreases in the overall base of support.25,38,40,53
superior movement of the center of gravity, and displacement of the line of gravity closer to the anterior margin of the base of support. These factors could contribute to poorer performance on FRT in the dress shoe condition.

The second hypothesis was that older women would have larger TUG scores and slower self-selected gait speeds when wearing dress shoes than when walking shoes, with intermediate values when barefooted. This hypothesis was based on previous reports that increases in heel height produce decreased ankle joint stability and shock absorption capacity at the feet and ankles, as well as greater energy costs of walking. These factors could result in decreased speed of movement during sit-to-stand activities as well as during gait. The women were expected to perform better in the walking shoe condition than when barefooted because of the shock absorption afforded by the walking shoes.

Method

Subjects

A convenience sample of 35 women, aged 65 to 93 years (X=80, SD=6.48), was recruited from 2 assisted living facilities and 2 retirement communities in North Carolina. After receiving permission from each facility’s administrator, the principal investigator (SAA) invited female residents to participate in the study. Interested subjects were contacted by telephone for an initial screening. The first 35 volunteers were included in the study, provided that they met the following inclusion criteria: (1) were 65 years of age or older, (2) owned at least one pair of walking shoes and at least one pair of dress shoes, (3) wore these shoes at least occasionally, (4) had at least 90 degrees of shoulder flexion, (5) transferred independently, (6) stood unsupported for 30 seconds or more, (7) could walk independently at least 20 m and turn 180 degrees, with or without an ambulatory aid, and (8) did not wear a lower-extremity brace or orthosis. Subjects were excluded based on performance criteria but not on the basis of medical diagnosis. Additional exclusion criteria included any inability to follow standardized test instructions or to perform the FRT, TUG, and TMW under all footwear conditions. Subjects were required to bring their own pairs of walking shoes and dress shoes to the data collection session.

The sample consisted of 5 women who lived in assisted living facilities and 30 women who lived independently in retirement communities. Only 2 women, both in assisted living facilities, used walking aids. The subjects reported a variety of medical diagnoses, and 23 of the 35 subjects had at least one foot abnormality. Subject characteristics are reported in Table 1.

<table>
<thead>
<tr>
<th>Subject characteristics (N=35)</th>
<th>X</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>80</td>
<td>6.48</td>
<td>65–93</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.61</td>
<td>0.06</td>
<td>1.52–1.75</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>64.3</td>
<td>9.1</td>
<td>46–87</td>
</tr>
<tr>
<td>Medical diagnoses a</td>
<td>2.1</td>
<td>1.3</td>
<td>0–5</td>
</tr>
<tr>
<td>Prescribed medications b</td>
<td>2.7</td>
<td>2.4</td>
<td>0–12</td>
</tr>
<tr>
<td>Foot abnormalities c</td>
<td>1.3</td>
<td>1.5</td>
<td>0–7</td>
</tr>
<tr>
<td>Ankle dorsiflexion (°)</td>
<td>11</td>
<td>6.8</td>
<td>−5–25</td>
</tr>
<tr>
<td>Ankle plantar flexion (°)</td>
<td>51.7</td>
<td>7.9</td>
<td>30–80</td>
</tr>
<tr>
<td>Heel height (cm)–walking shoes</td>
<td>1.0</td>
<td>0.3</td>
<td>0.5–1.7</td>
</tr>
<tr>
<td>Heel height (cm)–dress shoes</td>
<td>5.3</td>
<td>1.2</td>
<td>4.0–8.8</td>
</tr>
<tr>
<td>Area under the heel (cm²)–dress shoes</td>
<td>8.1</td>
<td>8.3</td>
<td>0.7–33.4</td>
</tr>
</tbody>
</table>

* Most frequently reported types of medical diagnoses were orthopedic (n=21), cardiovascular (n=19), metabolic (n=14), and respiratory (n=11).
* Most frequently reported prescribed medications were antiplatelet agents (n=12) and antihypertensives (n=11).
* Most commonly observed foot abnormalities were hallux valgus (n=12) and hammertoes/claw toes (n=12).
* Negative flair of the sole of the shoe was observed in 7 pairs of walking shoes and 35 pairs of dress shoes.

Instruments

Footwear was a nominal variable consisting of 3 categories: barefoot, walking shoes, and dress shoes. Shoes were categorized as walking shoes or dress shoes based on heel height, measured to the closest millimeter, and whether the shoe was a lace-up, buckled, Velcro-fastened, or slip-on type. A dress shoe was defined as a firm-soled, slip-on shoe with a heel height of at least 4 cm (1.6 in). A walking shoe was defined as a laced-up, buckled, or Velcro-fastened shoe, with a heel height of 0 to 2 cm (0–0.8 in), including athletic shoes and oxford-type shoes. The heel height was established by measuring the vertical distance from the floor to the insole at the front of the heel and subtracting from that measurement the sole thickness beneath the first metatarsal head. Other shoe characteristics such as the flare of the sole, firmness of the sole, and area under the high heel were qualitatively judged and documented. In an effort to maximize ecological validity for making inferences about the effects of customary footwear (worn in everyday life) and minimize the confounding effect of novelty of footwear, subjects were tested wearing their own shoes. See Table 1 for a summary of shoe characteristics.

Administration of the FRT, TUG, and TMW followed protocols used at the Center for the Study of Aging and Human Development at Duke University Medical Center. Functional reach was measured with the subject in a standing position, with her dominant upper extremity.

Table 1. Subject Characteristics and Shoe Characteristics
ity next to a wall. The dominant arm was used for consistency with procedures described by Duncan et al.\textsuperscript{13} for development of the FRT. The selection of dominant arm was based on the subject’s self-report of the hand used for writing. The subject was asked to attain a comfortable standing position, and the position of her feet was marked on the floor for each footwear condition. A measuring stick with a built-in level was placed on the wall at acromion height, leveled, and secured to the wall with marking tape. The subject made a fist and raised her dominant arm to approximately 90 degrees of shoulder flexion. In this position, the placement of the end of the third metacarpal bone along the measuring stick was recorded to the closest centimeter as position 1. The subject was then instructed to reach as far forward as possible without taking a step or losing balance, and the location of the end of the third metacarpal was recorded to the closest centimeter as position 2. Functional reach was defined as the difference between the 2 positions. After 2 practice trials, 3 measurements of functional reach were recorded and averaged to establish the FRT measure.\textsuperscript{9,21}

The walking distance for the TUG\textsuperscript{14} was measured with a metal tape measure and marked with tape. The same armchair, with hardwood arms and cushioned back and seat, was used for all subjects. The armchair’s seat height was 44 cm (17.3 in), its seat depth was 44 cm (17.3 in), and its arm height (measured from the floor) was 63 cm (24.8 in). Each subject began the test in a sitting position with her back against the chair back, hands on the chair arms, and customary walking aid (if required) in front of her. She was instructed to perform at “a comfortable and safe pace for you.” On the word “go,” the subject stood up, walked at a self-selected pace to a line 3 m away, turned around, returned to the chair, and sat. The tester timed the TUG to the nearest hundredth of a second using a digital stopwatch. Timing began on the word “go” and ended when the subject returned to the start position. The subject performed 1 practice trial and 2 trials for data collection. The test trials were averaged to give a TUG score (in seconds).\textsuperscript{9,14,21}

The walking distance for the TMW\textsuperscript{9,21,22} was measured with a metal tape measure and marked with tape on the floor. Five extra meters were measured and marked, both ahead and at the end of the 10-m distance, to allow the subject enough distance to accelerate and decelerate. The subject was instructed to walk “at your normal, comfortable pace” and was allowed to use her customary walking aid, if needed. The tester timed the TMW to the nearest hundredth of a second using a digital stopwatch. Timing began when the subject’s leading foot crossed the start line and ended when the leading foot crossed the finish line. The subject performed 1 practice trial and 2 test trials. The test trials were averaged, and the outcome was used to calculate the self-selected gait speed over 10 m.\textsuperscript{9,14,21}

Testing Procedure

To describe the population under study, demographic data on age, information on residential settings, and self-reports of medical history and medications were collected during a telephone interview. The number of chronic diseases was computed as the sum of self-reported diagnoses, and the number of medications was computed as the sum of currently prescribed medications.

Following the telephone interview, data collection was conducted at the assisted living facility or retirement community where the participant resided. Prior to testing, test procedures were described to each subject, and informed consent was obtained using a form approved by the Committee on the Protection of the Rights of Human Subjects at the University of North Carolina at Chapel Hill. The tests were administered in a common room in each facility. Each testing session took approximately 1 hour to complete. Because of space constraints at the facilities where the testing was performed, 22 subjects performed the FRT and TUG on a linoleum floor, and 13 subjects performed the tests on a firm, low-pile, carpeted floor. All subjects completed the TMW on firm, low-pile, carpeted floor. The principal investigator, a physical therapist with 6 years of clinical experience, administered all tests. To minimize effects of experimenter bias, the investigator followed standardized test instructions and did not examine any of the data until all subjects had been tested.

In order to describe the sample, data related to subject and footwear characteristics were collected before testing and during rest periods between performance measures. Subjects’ height and weight were obtained through self-report. Ankle range of motion in dorsiflexion and plantar flexion was measured using standard goniometric techniques.\textsuperscript{38} Subjects were assessed for the presence of foot abnormalities, including hallux valgus, hallux rigidus, hammertoes, claw toes, overlapping or underriding toes, painful corns, and ulcers.\textsuperscript{35,55,59–61} Joint position sense was estimated as present or impaired as described by Schenkman et al.\textsuperscript{62} Any foot pain experienced during each of the tests was documented. Shoe characteristics such as heel height, area under the heels, flair of the heel, and firmness of the sole were documented.

Subjects performed the FRT, TUG, and TMW in a randomized order. The tests were performed in the sequence in which the subjects drew their names from a hat at the beginning of the testing period. The order of the footwear conditions was counterbalanced among the
subjects so that all possible sequences of footwear conditions were equally represented. To avoid undue fatigue, subjects rested 3 minutes between footwear conditions and 1 minute between different functional measurements. For the first footwear condition, the tester explained each test and demonstrated it in a standardized manner. The TMW was not demonstrated. For the 2 following footwear conditions, the tester repeated the explanation but did not demonstrate the test. Subjects received the same number of practice and test trials for all footwear conditions.

To assess test-retest reliability of these measurements, a subgroup of subjects was asked to repeat the tests 7 days after the original testing date. The first 12 subjects who agreed to repeat the testing participated in the reliability part of the study. The ICC (3,1) was calculated for the 3 tests performed under each category of footwear. The 9 ICCs ranged between .94 and .99, indicating good test-retest reliability for the measurements under all footwear conditions. The difference in scores between test occasions was not significant for any of the tests ($P > .05$), with mean absolute differences of 0.7 cm for the FRT (all 3 conditions), 1.34 to 1.48 seconds for the TUG, and 0.01 to 0.03 m/s for the TMW.

### Results

The subjects completed all tests without difficulty, except for 1 subject who reported chronic foot pain that remained constant throughout the testing session. The descriptive statistics for the FRT, TUG, and TMW scores under the 3 footwear conditions are documented in Table 2. Huynh and Feldt’s estimator ($\varepsilon$) ranged from .994 to 1.000 for all 5 repeated-measures ANOVAs, indicating that the sphericity assumption was not violated. The ANOVAs revealed an overall footwear condition effect for FRT scores for subjects tested on either carpet ($F = 29.57; df = 2.24; P < .0001; \varepsilon = .999$) or lino-

### Data Analysis

Descriptive statistics were obtained and data were screened for outliers and to determine whether assumptions for repeated-measures analysis of variance (ANOVA) were met. This was done by visual inspection initially, using box plots and stem-and-leaf plots of the dependent variables. We also analyzed the residuals to determine whether statistical assumptions were met. Residuals were standardized by dividing them by their standard deviation. A residual that was $\geq 3$ of these units in absolute size was considered an outlier. Logarithmic transformation was performed on the TUG scores (TUG_{log}) to correct for a positively skewed data set. The TUG_{log} scores were used in all ANOVA procedures and post hoc comparisons. The raw TUG scores (in seconds) were used when describing the subjects.

A one-way repeated-measures ANOVA was used for each test to compare the outcomes on the FRT, TUG, and TMW for the 3 different footwear conditions. Because subjects were tested on different floor surfaces for the FRT and TUG, separate repeated-measures ANOVAs were performed for the 2 surfaces for these 2 tests. For the omnibus $F$ tests, a Bonferroni adjustment was used to control Type I error probability at the .05 level. Because 5 repeated-measures ANOVAs were performed, the level of significance was set at .01 (.05/5) for each test. Following a significant overall $F$ test, post hoc comparisons among footwear conditions were performed using the Tukey Honestly Significant Difference (Tukey HSD) test with a significance level of $P < .05$. The 95% confidence interval (95% CI) also was calculated for each point estimate.

### Table 2.

Means, Standard Deviations, Ranges, and 95% Confidence Intervals (CI) for Each Category of Footwear for Functional Reach Test (FRT), Timed Up & Go Test (TUG), and 10-Meter Walk Test (TMW)

<table>
<thead>
<tr>
<th>Category</th>
<th>$\bar{X}$</th>
<th>SD</th>
<th>Range</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRT (cm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barefoot</td>
<td>35.2</td>
<td>6.6</td>
<td>18–45.3</td>
<td>32.9–37.5</td>
</tr>
<tr>
<td>Walking shoes</td>
<td>34.2</td>
<td>6.0</td>
<td>21.7–43.7</td>
<td>32.1–36.3</td>
</tr>
<tr>
<td>Dress shoes</td>
<td>30.2$^a$</td>
<td>5.7</td>
<td>21–41.7</td>
<td>28.3–32.2</td>
</tr>
<tr>
<td>TUG (s)$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barefoot</td>
<td>13.53$^c$</td>
<td>7.3</td>
<td>7.32–48.35</td>
<td>11.03–16.04</td>
</tr>
<tr>
<td>Walking shoes</td>
<td>12.82$^c$</td>
<td>6.45</td>
<td>6.8–42.55</td>
<td>10.61–15.04</td>
</tr>
<tr>
<td>Dress shoes</td>
<td>14.02$^c$</td>
<td>6.73</td>
<td>7.41–45.26</td>
<td>11.71–16.34</td>
</tr>
<tr>
<td>TMW (m/s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barefoot</td>
<td>1.11$^c$</td>
<td>0.28</td>
<td>0.39–1.62</td>
<td>1.02–1.21</td>
</tr>
<tr>
<td>Walking shoes</td>
<td>1.19$^c$</td>
<td>0.29</td>
<td>0.48–1.75</td>
<td>1.09–1.29</td>
</tr>
<tr>
<td>Dress shoes</td>
<td>1.06$^c$</td>
<td>0.26</td>
<td>0.39–1.62</td>
<td>0.98–1.16</td>
</tr>
</tbody>
</table>

$^a$ Significantly different ($P \leq .05$) from means for both walking shoe and barefoot conditions.

$^b$ For the TUG, statistical analyses were performed on log-transformed data.

$^c$ Means for all three footwear conditions were significantly different ($P \leq .05$) from each other.
leum (F=24.27; df=2,42; P<.0001; ε=1.000), for TUGlog scores for subjects tested on linoleum (F=18.64; df=2,42; P<.0001; ε=1.000), and for TMW scores (F=47.29; df=2,68; P<.0001; ε=1.000). For the 13 subjects who performed the TUG on carpeted floor, the main effect for footwear conditions was not significant (F=3.26; df=2,24; P=.061; ε=.994).

Tukey HSD post hoc pair-wise comparisons revealed that the subjects performed better on the FRT when they were barefoot or wore walking shoes compared with when they wore dress shoes, regardless of whether they performed the test on carpet (HSD.05=2.12) or linoleum (HSD.05=1.56). The mean absolute difference in FRT scores between barefoot and dress shoe conditions was 5.1 cm. There was no difference between the barefoot and walking shoe conditions on either floor surface. Measurements taken during all footwear conditions were different from each other for the TUGlog scores obtained from testing on the linoleum floor (HSD.05=.023) and for the TMW scores (HSD.05=.032). Subjects performed best in walking shoes and worst in dress shoes, with intermediate scores in the barefoot condition. The mean absolute difference between walking shoe and dress shoe conditions was 1.4 seconds on the TUG and 0.13 m/s on the TMW.

Figure 1 shows individual differences in FRT scores based on footwear. Thirty-four of the 35 subjects reached farther while barefooted than in their dress shoes (Fig. 1b). Thirty-two subjects reached farther in walking shoes than in dress shoes (Fig. 1c). Individual differences in FRT scores varied greatly between the walking shoe and barefoot conditions (Fig. 1a). Twelve subjects reached farther in walking shoes, 21 subjects reached farther when barefooted, and 2 subjects had equal FRT scores under barefoot and walking shoe conditions.

Figures 2 and 3 present individual differences based on footwear conditions for the TUG and TMW, respectively. Twenty-seven subjects had a better score on the TUG when wearing walking shoes compared with barefoot (Fig. 2a), and 31 subjects walked faster wearing walking shoes than barefoot on the TMW (Fig. 3a). Twenty-seven subjects moved faster barefoot than in dress shoes on the TUG (Fig. 2b), and 28 subjects walked faster barefoot than wearing dress shoes on the TMW (Fig. 3b). Thirty subjects performed better in walking shoes than in dress shoes on the TUG (Fig. 2c), and all except 1 subject performed better in walking shoes than in dress shoes on the TMW (Fig. 3c).

**Discussion**

The results of this study show that type of footwear affects the measurements obtained with the FRT, TUG, and TMW in older women. These results complement previous research evidence for the effects of high heels, walking shoes, and a barefoot condition on kinematic and kinetic movement characteristics. In view of the magnitude of the effects, clinicians and researchers should view type of footwear as an important factor when using these common clinical tests. A change from the barefoot condition to the dress shoe condition produced an average 15% decline in FRT scores (Tab. 2). The mean absolute difference of 5.1 cm between barefoot and dress shoe conditions was of the same magnitude as the mean change score (1.99 in [5.05 cm]) reported by Weiner et al for subjects undergoing inpatient physical
rehabilitation. Given the reliability of FRT scores, a change of the magnitude probably would be interpreted as reflecting an actual change in performance abilities if footwear effects were not taken into account. Footwear effects on TUG and TMW scores were smaller, with a change from the walking shoe condition to the dress shoe condition producing an average 10% to 12% change in scores. We believe that the footwear effects are important, however, as potential sources of error variability that can easily be controlled for.

The detrimental effects of high heels on FRT scores are consistent with the findings of Lord and Bashford, who used a “swaymeter” to examine balance in older women under the same 3 footwear conditions used in our study. The magnitude of the effect in our study varied considerably among subjects, but several subjects reached 5 to 10 cm farther when barefoot or wearing walking shoes than when they wore dress shoes (Figs. 1b and 1c).

The lack of a difference in FRT scores between barefoot and walking shoe conditions agrees with the observations of Briggs and colleagues. They found no effect of wearing shoes versus not wearing shoes on performance in sharpened Romberg and one-legged stance tests among older women with no known pathology. One explanation for this outcome may be that the walking shoes had a variety of different characteristics that could affect the amount of postural stability they provided. Many of the walking shoes worn by the women in our study had a positive flare of the sole (ie, angulation of the midsole and outer sole material away from the midline of the shoe), affording an increased base of support compared with the barefoot condition. The increased base of support may have counteracted the negative effects that shoes can have on joint position sense and balance. Some of the walking shoes had relatively firm, thin soles, which may facilitate joint position sense, as suggested by Robbins et al. The walking shoes also had a range of heel heights. Increased heel height, up to a certain point, may provide an advantage in forward functional reach to individuals with limited dorsiflexion.

Despite the lack of a difference in FRT scores between barefoot and walking shoe conditions, clinicians and researchers should not assume that these 2 conditions produce identical outcomes on the FRT. Based on the individual differences in FRT performance when barefoot versus in walking shoes, lack of consistency between these 2 conditions could be a source of measurement error when testing an individual over time. The mean absolute difference in FRT scores between the barefoot and walking shoe conditions (2.45 cm) was more than 3 times that obtained when footwear was kept constant during reliability testing (0.7 cm). Eight subjects demonstrated differences of 4 cm or greater (Fig. 1a).

The results for the TUG, performed on a linoleum floor, and the TMW, performed on carpet, were as hypothesized. The subjects had larger TUG scores and slower gait speeds when wearing dress shoes than when wearing walking shoes, with intermediate values when barefoot. The better performance in walking shoes compared with barefoot is consistent with the results from Dobbs et al, who reported faster self-selected walking speeds in shoes as compared with barefoot in subjects up to 89 years of age.

![Figure 2](http://ptjournal.apta.org/)

**Figure 2.** Individual differences in Timed Up & Go Test (TUG) measurements among footwear conditions: (a) barefoot minus walking shoes, (b) barefoot minus dress shoes, (c) walking shoes minus dress shoes.
The shock absorption provided by walking shoes may allow people to walk faster without increasing the impact loading of the body. The magnitude of the effect on individual TUG scores was generally less than 2 seconds, although 1 subject required almost 6 seconds longer to complete the TUG when barefoot than when wearing walking shoes (Fig. 2a). For individual TMW scores, differences between walking shoe and barefoot conditions were less than 0.2 m/s for all but 3 subjects (Fig. 3a).

The poor performance on the TUG and TMW in the dress shoe condition is consistent with studies that demonstrated slower gait in high-heeled shoes compared with low-heeled shoes. The magnitude of the effects on individual scores was greater than for the walking shoe versus barefoot condition, with several subjects demonstrating differences greater than 2 seconds on the TUG (Figs. 2b and 2c) and greater than 0.2 m/s on the TMW (Figs. 3b and 3c). High-heeled dress shoes, therefore, appear to have a particularly pronounced influence on TUG and TMW scores in older women. The lack of a difference between footwear conditions when the TUG was performed on carpet was most likely the result of low statistical power (n = 13), but may be related to the softer surface. Conclusions, however, will need to await research with greater statistical power.

Some researchers, focusing on the biomechanical effects of footwear on human movement, provided their subjects with standardized shoes to control for specific footwear characteristics. Testing subjects in new shoes, however, may influence postural responses to footwear. Additionally, subjects usually wear their own shoes during testing in clinical settings. Therefore, the fact that subjects in our study were tested in their own shoes decreases information about shoe characteristics but should improve generalizability of the results. Test performance on different floor surfaces was not examined, but the results of our study appear to be valid for the FRT both on linoleum and carpeted floor, for the TUG on linoleum floor, and for the TMW on carpeted floor.

The results of our study have several implications for research and clinical practice. A clinician or researcher may want to assess individuals under many of the footwear conditions that they typically encounter in their daily lives. An older woman who wears high-heeled shoes to church every Sunday, for example, should be tested while wearing high heels as well as her other customary footwear. When assessing changes in an individual’s balance or gait abilities over time, however, we believe that comparisons should be made only between very similar footwear conditions. In addition, footwear and the testing surface should be described when reporting test results for research or clinical purposes.

Results of our study also provide information about the effects of footwear on general stability and gait in older women. Results focusing on mean and individual differences in performance for the various footwear conditions suggest that footwear intervention may improve performance of gait and balance tasks by older women. For example, the FRT distances for subject 26 were 34 cm (13.4 in), 27.3 cm (10.7 in), and 24.3 cm (9.6 in) in the barefoot, walking shoe, and dress shoe conditions, respectively. These results indicate dramatically decreased balance control in dress shoes compared with the barefoot condition. This type of information may
assist clinicians in making recommendations to their clients about safer footwear.

Further research is needed to identify the specific shoe characteristics that provide the greatest benefits for given physical and environmental conditions. Additional research is also needed in order to use outcomes on the FRT, TUG, and TMW to recommend footwear selection for older individuals. The smallest clinically meaningful difference for each of these measures must be determined. Only then will it be possible to describe whether a difference in performance, associated with a change in footwear, is important.

Conclusions
This study indicates that the type of footwear an older woman is wearing can have an effect on her scores on the FRT, TUG, and TMW. Performance of the FRT on linoleum or carpeted floor was worse when wearing dress shoes than for either barefoot or walking shoe conditions. Scores on the TUG performed on linoleum floor and the TMW performed on carpeted floor were best when subjects wore walking shoes and worst when they wore dress shoes, with intermediate values for the barefoot condition. Footwear should be carefully documented and should remain constant from one test occasion to another when the FRT, TUG, and TMW are used in the clinic and in research. In addition, footwear intervention should be considered as a way of improving gait and balance in older women.

References
Effects of Footwear on Measurements of Balance and Gait in Women Between the Ages of 65 and 93 Years
Solveig A Arnadottir and Vicki S Mercer