Reliability of the senior fitness test in a modified environment

by

Tiffany Brooke VanGundy

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Exercise and Sport Science (Biological Basis of Physical Activity)

Program of Study Committee:
Warren D. Franke, Major Professor
Marian L. Kohut
Daniel W. Russell

Iowa State University

Ames, IA

2006

UMI Number: 1439906

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Table of Contents

Introduction

The growing number of people around the world living well into and beyond their seventies has generated new problems for gerontology researchers. For example, an estimated 12% of the U.S. population is 65 and older, and nearly 15% of the population within Iowa is comprised of members of this growing older adult set (U.S. Census Bureau, 2000). One of the most important issues facing older adults is maintaining a level of functional fitness (vs. physical) such that they can live independently and not be at an increased risk for falls. Functional fitness is defined as "having the physical capacity to perform normal everyday activities safely and independently without undue fatigue" (Rikli and Jones, 2001). Physical activity can be an important part of developing functional fitness. A sedentary lifestyle or suffering from disease states can generate substantial losses in functional fitness. In recognition of the importance of physical activity for older adults, the Healthy People 2010 objectives for physical activity and fitness recommend that adults 65 and older participate in at least 30 minutes of moderate physical activity, preferably daily (U.S. Department of Health and Human Services-Centers for Disease Control and Prevention, 2000).

Older adults who remain physically active can delay the onset, or even reverse the loss, of functional fitness and thus delay the age-related decline that leads to frailty (American College of Sports Medicine Position Stand, 1998; Gill, Williams, Richardson & Tinetti, 1996; Guralnik, Ferrucci, Simonsick, Salive & Wallace, 1995; Huang, Macera, Blair, Brill, Kohl & Kronenfeld, 1998; Jackson, Beard, Wier, & Blair, 1995; Lawrence and Jette, 1996; Morey, Pieper & Cornoni-Huntley, 1998). Physical frailty is defined as the loss of physiologic reserve that increases the risk of disability (Buchner and Wagner, 1992). The estimated annual cost for physical frailty is $54 to 80 billion, and it could grow to beyond $132 billion by 2030 unless disability rates are lowered (Select Committee on Aging, 1992).

Spirduso (1995) uses five hierarchical groupings to discriminate between common levels of functioning for older adults. They are as follows from lowest functional level to highest:

Physically Dependent – This group is comprised of those who cannot complete activities of daily living (ADL, e.g. dress, bathe, eat) and either live in nursing homes or require full-time care at home.

Physically Frail – Older adults in this group can complete all ADL's, but not all instrumental activities of daily living (IADL's, e.g. cook a meal, shopping, housework). Many still live at home, but they may require the aide of family members or services.

Physically Independent – Members of this group are capable of performing all IADL's. They live independently and function well within the community, but are generally sedentary. This group is comprised of ~70% of the older adult population, and they span a broad range of functional ability. The individuals in the bottom of this group may be near losing their independence.

Physically Fit – This group is comprised of older adults that often appear younger than their age. This is in part due to the fact that they exercise several times a week participating in activities such as walking, jogging, aerobic classes, tennis, or swimming.

Physically Elite – This is a very small group of older adults that consistently participate in aerobic and/or strength training, often times for competitions and races.

Clinicians and researchers have found it problematic to obtain a physiological assessment of older adults' immediate level of fitness and subsequent gains or declines as they age. Most protocols for determining an individuals' level of physical fitness are designed for younger adults, and are not appropriate for older adults (Rikli and Jones, 1999). Several researchers have generated functional fitness testing protocols for older adults. Two important tests in the U.S. are the American Alliance for Health, Physical Education, Recreation, and Dance (AAPHERD) Functional Fitness Assessment for Adults Over 60 Years of Age (Osness, Adrian, Clark, Hoeger, Rabb & Wiswell, 1990), and the Established Populations for Epidemiologic Studies of the Elderly (EPESE) (Guralnik, Simonsick, Ferrucci, Glynn, Berkman, Blazer, Scherr & Wallace, 1994). Other researchers have utilized physical-fitness performance based testing in their research worldwide. Tanaka, Kim, Yang, Shimamoto, Kokudo and Nishiji (1995) used a battery of physical tests to create an index of functional status for elderly Japanese men, and Van Heuvelen, Kempen, Ormel and Rispens (1998) used several performance based tests to relate physical fitness levels to age and physical activity in older adults in the Netherlands.

The Senior Fitness Test (SFT) developed by Rikli and Jones (1999) may be the best at assessing the six underlying functional fitness parameters for older adults: lower and upper body strength, aerobic endurance, lower and upper body flexibility, and agility/dynamic balance. Each test component of the SFT has been selected for its high content validity, criterion validity, construct validity, and reliability (Rikli and Jones, 1997;

1998). The SFT is usually performed in a fitness facility or large community facility. However, Rikli and Jones (2001) intentionally selected testing procedures that require very little equipment, and therefore theoretically could easily be adapted to other locations.

Nearly 20% (10,089,000) of the U.S. non-metropolitan residing population consists of those 60 and older, compared to only 15% of the metropolitan residing population (U.S. Department of Agriculture, 2002). There are important differences in these older adult populations in that rural elderly may suffer from higher morbidity and mortality from chronic diseases and are more sedentary than their urban counterparts (U.S. Department of Health and Human Services, 2000).

Most physical activity studies with older adults utilize subjects living in metropolitan areas and those residing in older adult facilities. Coming to a local laboratory facility to complete a sub-maximal exercise test is not a large logistical hurdle for participants living in metropolitan regions. However, in most rural areas there are often neither fitness centers nor research laboratories where one could perform the SFT. It seems reasonable to expect that a researcher or fitness leader should be able to travel to rural areas to conduct the SFT for an individual or group of older adults, instead of requiring the older adults (some of whom may not have a driver's license) to come to the nearest facility or lab. Unfortunately, the reliability of completing the SFT in a setting other than a fitness facility, especially the 6-Minute Walk test component, has yet to be tested.

Therefore, the following was the background problem this research addressed: very little research has been done to determine factors related to increasing and maintaining functional and physical fitness in rural elderly or elderly who do not have either the desire to exercise or do not have access to fitness centers. This led to the main problem this research addressed: for any future research in this area to be completed, the fitness assessment tool (i.e., the SFT) must be deemed reliable outside a fitness or laboratory facility.

Consequently, the main purpose of this experiment was to examine the reliability of each component of the SFT when performed by older individuals in three different environments. Thus, the independent variable in this research was the testing environment or location, and the dependent variables were the outcome scores on each component of the SFT. The six components of the SFT are: the Chair Stand (lower body strength), the Arm Curl (upper body strength), the Chair Sit and Reach (lower body flexibility), the Back Scratch (upper body flexibility), the 8 ft. Up-and-Go (agility and dynamic balance), and the 6-Minute Walk (cardiopulmonary endurance). For the first trial, the five stationary components of the SFT were completed in a fitness facility gymnasium, with the 6-Minute Walk component completed by walking around a rectangular area on a gymnasium floor. This trial was repeated. The participants also subsequently completed the SFT in a partially modified setting and a modified setting. For the partially modified session, the stationary components were completed in the fitness facility gymnasium, but the 6-Minute Walk component was completed walking up and down an indoor hallway (smooth, tile flooring). The modified setting was each individual participant's home. The five stationary components were completed in their house, but the 6-Minute Walk component was completed walking outdoors (on sidewalks and smooth level surfaces whenever possible). It was hypothesized that each component of the SFT would have excellent test-retest reliability (Cronbach alpha $\geq .80$) in these three settings ($p<0.05$).

If the SFT was found reliable across the three environments, researchers could subsequently utilize the SFT with the under-studied, rural-residing older adults to obtain information about their physical fitness levels. In other words, the testing site would not constrain the use of the SFT. These findings may also motivate fitness leaders to begin rural community-based fitness programs for older adults and allow them to track functional fitness progress reliably. Testing like this is generally enjoyed by older adults, and it may motivate them to continue or to become more physically active after completing the SFT.

Methods

Participants

Participants were recruited and screened from university fitness programs, classes, and the surrounding community (n = 35). Participants were to be ambulatory, without the use of assistive devices, not suffering from joint pain, unstable cardiovascular disease, or any other medical condition that would be contraindicated for submaximal testing per the American College of Sports Medicine's guidelines (ACSM, 2006). Thirty-one participants were enrolled in the study. Four were excluded based on not meeting medical criteria. During the course of testing, seven participants did not complete all four weeks of testing, and thus twenty-four participants are included in data analysis. Participant characteristics are included in Table 1. Reasons for not completing all four weeks of testing included: one had aggravated arthritis after the first test session, two were unable to reschedule a missed session in a timely manner, two had illnesses for more than two weeks, two had medical conditions unrelated to the study that did not allow them to participate in any physical activity for longer than two weeks (surgery, fall at home). Drop outs were similar in gender, age, place of residence and activity status as those who completed the study. Each participant had the study fully explained to them, had their questions answered, and signed an informed consent document (approved by the Institutional Review Board of Iowa State University on Sept. 8, 2004).

Table 1. Participant Characteristics

Sex	**Age** (yrs)	**Current Exerciser**	**Residency**
Males (n=12)	Range = 65-85	Yes, n = 10	Metropolitan, n = 10
	Mean = 75 ± 6.0	No, n = 2	Rural, n = 2
Females (n=12)	Range = 66-79	Yes, n = 12	Metropolitan, n = 11
	Mean = 72 ± 4.3	No, n = 0	Rural, n = 1

Instruments

Simple devices were utilized that required no calibration such as the following: a digital stop watch for timing, a chair required for four of the component items, a measuring tape for the Chair Sit and Reach, and a hand weight (5 pounds for women, and 8 pounds for men) for the Arm Curl test. These tools have been recommended for use with the SFT (Rikli and Jones, 2001). The same equipment (including the chair) was used for each participant's home test session. A distance-measuring wheel (Model #07, Redington Counters Incorporated, Windsor, CT) was used to determine the distance walked in feet for participants in the building hallways and outdoors (partially modified and modified SFT). Accuracy of the distance-measuring wheel was assessed once a week by having it measure a known distance. A digital thermometer (Model #1449, Taylor Precision Products, Oak Brook, IL) was utilized to obtain the temperature during the outdoor 6-Minute Walk tests.

Procedures

Each participant completed the SFT on four separate occasions at least two days but not more than two weeks apart. The participants were tested in groups of no more than five at a time, and whenever possible when testing at the fitness facility gymnasium, in a group of at least two. Participants were tested individually for the 6-Minute Walk component that was completed up and down the hallways for the partially modified test and for the entire SFT that was completed at home. Each participant was asked not to engage in strenuous activity twenty-four hours prior to the testing and to avoid caffeine, smoking, and heavy eating within two hours prior to the test. They were instructed to eat a light meal beforehand and wear comfortable clothing and walking shoes on test days. Prior to each testing session, the participants were led through five to eight minutes of warm-up and stretching activities.

In the first session, each participant completed all components of the SFT in the fitness facility gymnasium and walked around an indoor rectangle for the 6-Minute Walk component. The purpose of this session was to familiarize them with the testing procedures. Test scores, especially on the 6-Minute Walk when completed in this rectangular format, are improved when a practice trial is completed once before a test trial (Rikli and Jones, 1998). Each subject then completed the SFT three more times. One trial was identical to the familiarization test in the fitness facility gymnasium. In the other two trials, the environment for testing was partially or totally modified. The partially modified SFT consisted of the first five component items being completed in the same fitness gymnasium facility as the initial SFT, and the 6-Minute Walk was performed individually in the hallways of the building. The participants walked up and down the length of a hallway, and were told to make as large a U-turn at each end as they needed to maintain their speed. Enright (2003) states that making abrupt turns during the 6-Minute Walk decreases walk distance. For the totally modified environment, each participant completed the first five components of the SFT in their home, and the 6-minute walk was done outdoors on level ground. While the first trial was always the practice SFT, the testing order of the next three trials was randomized and counter-balanced.

For each of the latter two walk tests, the participants were followed with a distance-measuring wheel to obtain the distance walked. When testing in each participant's home, the largest open area (e.g., the living room) was selected for the first five components of the SFT. For the first five component tests, twenty of the modified tests were completed on carpet and four on linoleum, tile, or hardwood. For the 6-Minute Walk, twenty-three were completed on paved surfaces, and one on gravel (outdoor temperature range 38.1 - 84.6°F, mean 63.9 ±15.7

˚F). Testing was completed in early summer and late fall seasons to ensure comfortable outdoor weather. This was also done to avoid the complication of normally increased physical activity levels that can occur in the early spring due to the weather warming up in the Midwest.

Each testing session took less than one hour to complete. Participants were encouraged to do their best on each test item, but not to push themselves to the point of overexertion or beyond what they thought was safe for them. During the timed upper and lower body strength tests, verbal encouragement was given half way through (i.e., "Keep going. Great job."). During the 6-minute walk, standardized verbal encouragement was given at each minute (i.e., minute 1 "You're off to a good start.", minute 2 "Keep up the good work.", minute 3 "You're half way through. Keep it up.", minute 4 "Keep up the good work.", minute 5 "One minute left. Give it your best."). Standardized encouragement protocols have been shown to be important in scoring accuracy on walking tests, specifically improving scores (Guyatt, Pugsley, Sullivan, Thompson, Berman, Jones, Fallen & Taylor, 1984; Steele, 1996).

In order to minimize variance introduced by multiple testers, only one technician supervised all the testing. The technician did not know participant scores from previous testing days to avoid any bias in reporting. To facilitate one-on-one communication with each participant in their homes, efforts were made to connect socially with each participant during every test session.

Statistical Design and Analysis

This is a reliability study. Intraclass correlation coefficients (ICC) were considered good if >.80. This method of assessment is sensitive to both measurement and systematic error (Thomas and Nelson, 2001). Each testing component of the SFT was studied independently of the others. Repeated measures ANOVA was also used to assess the extent to which test scores differed across trials. Data were assessed first in the chronological order tests were administered, ignoring the environments where tests were completed. This was done to assess whether there was a practice, or learning effect. Each test component was assessed for correlation between trials across time as well. Secondly, data were assessed by repeated measures ANOVA to compare the three environments where testing occurred (gym, partially modified-hallways, and modified-home). This assessment compared the effect of the environment where the test was performed on SFT performance. When differences were observed, effect sizes were determined. In addition, two-way repeated measures ANOVA was utilized to assess for gender by environment and order by environment interactions. A p value <0.05 was considered statistically significant. Data are expressed as mean ± standard deviation.

Results

Order Effects

Table 2 includes the SFT test scores in the order the tests were administered. In general, scores improved with repeated testing but the nature of this improvement varied by test component. The participants improved in the Chair Stand with trial 1 being less than trial 3 and 4 ($p<0.05$). Trials 2, 3, and 4 did not differ significantly from one another. For the Arm Curl, both trial 1 and 2 were significantly less than 3 and 4 ($p<0.05$). Trial 1 of the Chair Sit and Reach was less than trial 4, and the remaining comparisons between trials were not significantly different from one another ($p<0.05$). The Back Scratch yielded no significant differences between any of the trials ($p<0.05$). On the 8 ft. Up-and-Go participants were significantly faster on trial 4 than 1 and 2, and trial 3 differed from trial 1 ($p<0.05$). During the 6-Minute Walk, participants covered significantly less ground in trial 1 than in trial 3 and 4 ($p<0.05$). Correlations between performance on test components assessed in chronological order were all high ($r > .8$). The exception to this was the Chair Stand which had moderate correlations between trials ($r = .60$ to $.83$).

Effect of Environment

Mean participant performance on each SFT test component according to the environment where it was performed is depicted in Table 3. Except for the Chair Sit and Reach, performance on the three environmental conditional trials was better than the first, or practice, SFT trial ($p<0.05$). The Chair Sit and Reach did not differ significantly across trials. When the practice trial is removed from data analysis, none of the SFT test components were significantly different from one another in the three environmental conditions. The one exception was the Back Scratch test. Specifically when the Back

Scratch was done at home, participants scores significantly decreased from -3.0 ± 4.3 inches (gym) and -3.3 ± 4.3 inches (partially modified) to -3.7 ± 4.7 inches (modified-home) ($p<.05$). When assessing effect sizes, Cohen's D for the gym with the home Back Scratch score was .11, and Cohen's D was .16 for the partially modified with the home score. Intraclass correlation coefficients for all SFT test components across all three environmental conditions were high (Table 3). There was no gender by environment effect, but females consistently scored better than men in the Back Scratch (-1.2 ± 3.7 vs. -5.3 ± 4.1 inches, males vs. females, respectively). Finally, there was not a significant order of testing by environment interaction effect.

Table 2. Order Effects. Results are presented in chronological order of testing, independent of the environment in which the test was performed.

Test Component	Mean Score	Standard Deviation
Chair Stand (# completed)		
First	14.8	3.3
Second	15.8	3.1
Third	16.4 *	2.9
Fourth	16.6 *	2.6
Arm Curl (# completed)		
First	18.0	3.8
Second	18.8	4.1
Third	20.0 * †	3.9
Fourth	20.9 * †	3.7
Chair Sit and Reach (in inches)		
First	1.8	5.2
Second	2.0	4.7
Third	2.1	5.0
Fourth	2.5 *	4.9
Back Scratch (in inches)		
First	-3.1	4.4
Second	-3.3	4.4
Third	-3.4	4.6
Fourth	-3.3	4.4
8 ft. Up-and-Go (in seconds)		
First	5.4	1.5
Second	5.2	1.5
Third	5.1 *	1.2
Fourth	4.9 * †	1.2
6-Minute Walk (in yards)		
First	622.9	83.8
Second	645.9	163.4
Third	674.0 *	93.6
Fourth	684.0 *	106.0

* $p<0.05$ vs. first trial; † $p<0.05$ vs. second trial

Table 3. Effect of Environment. Results are presented by testing environment, independent of the order of testing.

Test Component	Mean Score	St. Dev.	ICC
Chair Stand (# completed)			0.81
Gym (practice trial)	14.8	3.3	
Gym	16.4	2.7	
Partially Modified (hallway)	16.6	2.9	
Modified (home)	15.8	2.9	
Arm Curl (# completed)			0.79
Gym (practice trial)	18.0	3.8	
Gym	20.3	3.8	
Partially Modified (hallway)	19.8	3.4	
Modified (home)	19.5	4.5	
Chair Sit and Reach (in inches)			0.96
Gym (practice trial)	1.8	5.1	
Gym	2.2	5.1	
Partially Modified (hallway)	2.2	4.7	
Modified (home)	2.4	4.9	
Back Scratch (in inches)			0.97
Gym (practice trial)	-3.1	4.4	
Gym	-3.2	4.3	
Partially Modified (hallway)	-3.0	4.3	
Modified (home)	-3.7 *	4.7	
8 ft. Up-and-Go (in seconds)			0.92
Gym (practice trial)	5.4	1.5	
Gym	5.1	1.5	
Partially Modified (hallway)	5.0	1.2	
Modified (home)	5.1	1.3	
6-Minute Walk (in yards)			0.87
Gym (practice trial)	622.9	83.3	
Gym	670.2	110.5	
Partially Modified (hallway)	677.7	94.3	
Modified (home)	680.3	101.4	

* $p<0.05$ 4th trial differs significantly from the others

Discussion

The primary findings of the present study are two fold. First, there was a practice effect across the four SFT trials. Participants generally improved after a practice trial, but the improvement on all SFT test components with subsequent trials was relatively small. All trials correlated highly with one another, except the Chair Stand which had only moderate correlations. Therefore, a familiarization trial is important to perform, but only one practice trial is sufficient. Secondly, when comparing the environment where the SFT was performed (after completion of a practice SFT on a previous day), none of the SFT components differed by environment. The exception to this finding was the Back Scratch test, which varied significantly by trial environment. Moreover, each component of the SFT had high internal consistency. In other words, the location of the SFT test (i.e., gym, or "standard" environment, partially modified environment or in the gym with the 6-Minute Walk completed in the hallways, and the modified environment or at home with the 6-Minute Walk completed outside) does not significantly affect performance on the SFT.

Order Effects Analysis

There is an apparent learning effect on all but the Back Scratch test. All other test components were significantly different ($p<0.05$), and it is apparent that utilizing the first session as a practice test for familiarization is important. Although the Arm Curl and the 8 ft. Up-and Go both demonstrated significant differences until the third trial was performed, the magnitude of the improvement in scores that was obtained by performing an additional practice trial is not physiologically significant (on average 1 arm curl and .3 seconds faster respectively). Moreover, the effect sizes of these differences are small, further reinforcing the suggestion that only one practice trial is sufficient. The present findings support the previous findings from Rikli and Jones (1998) that participants performed better the second time when given the 6-Minute Walk test walking around a rectangle. However they did not report needing to have a practice trial for all the other test components.

The current findings are similar to other researchers' findings for the 6-Minute Walk with healthy older adults. Utilizing a same-day test-retest method, 86% of healthy subjects 20-80 years old (n=79) had their best score after the first trial, and the final fourth trial was the best score in nearly half of the participants (Gibbons, Fruchter, Sloan & Levy, 2001). However, they did not find any significant difference between the 3rd and 4th trials, and the best 6-Minute Walk score was significantly greater than the first trial ($p<0.05$). The average difference between first and last tests was only 30 meters, which is slightly less than that seen in the present study (mean 39 yard difference between first and last 6-Minute Walk scores). Trooster, Gosselink and Decramer (1999) reported finding significantly increased scores on the 6-Minute Walk tests when repeated two and a half hours after the first trial with 50-85 year old healthy participants. Steffen, Hacker and Mollinger (2002) studied ninety-six independent, community-dwelling older adults (ages 61-89) and found high same day test-retest reliability with both the 6-Minute Walk and the Up-and-Go tests (.95 and .97 respectively).

The 6-Minute Walk test was first utilized to assess cardiopulmonary fitness in patient populations. Other researchers assessing diseased populations have found that repeated testing improves scores for the 6-Minute Walk. One study with patients with chronic heart failure found that they could achieve reproducible results on a 6-Minute-Walk test after 2 tests, with one week intervals between walks (Guyatt, Sullivan, Thompson, Fallen, Pugsley, Taylor & Berman, 1985). Patients with chronic obstructive airway disease tested twelve times over four weeks showed a learning effect over the first nine tests, and the researchers recommended performing four to five walk tests to remove the effect of learning (Knox, Morrison & Muers, 1988). When performing two 6-Minute Walk tests with emphysema patients (n=761), researchers found significantly improved 6-Minute Walk scores on the second test whether same-day or different day testing (Sciurba, Criner, Lee, Mohsenifar, Shade, Slivka & Wise, 2003). Even when encouragement was standardized during testing, a learning effect was still seen on 6-Minute-Walk tests in patients with CHF and chronic cardiopulmonary illness (Guyatt, et al., 1984).

In healthier populations, such an extensive practice effect has not been observed. It appears that when testing patient populations two or more practice trials may be necessary, but when testing healthy elderly populations, one practice trial is all that is needed whether doing same day test-retest or testing on different days. Steele (1996) recommends performing an additional walk if there was more than a 10% improvement in scores on subsequent walks. This may be a good rule to follow for both healthy and patient populations.

Along with Steffen, et al. (2002), others have studied the test-retest reliability of the timed Up-and-Go test. One utilized cognitively impaired and unimpaired elderly participants with tests separated by 112 ± 73 days. The researchers found that using two sessions for this test achieved only moderate (.56) test-retest reliability in a clinical and screening setting (Rockwood, Awalt, Carver & MacKnight, 2000). However, they did note that the tests were completed in different environments and often times utilizing different testers. Another study found very reliable intra-rater (.99) and inter-rater (.99) Up-and-Go test scores with twenty frail elderly participants (Podsiadlo and Richardson, 1991).

Our findings indicate that a practice SFT is important for all test components and not just the 6-Minute Walk. On tests where a genuine effort can affect the score, this makes intuitive sense. On the other hand, ICCs were consistently high (.79 to .97) for all SFT test components and strongly suggest that all participants' scores change in the same direction and amount with practice. Only the Chair Stand had moderate correlation between the practice and the other trials (r = .60-.72). After the first Chair Stand trial, the correlation was high (r > .8), and all other test components displayed high correlation between trials (even with the first trial). A practice trial may only be necessary when making comparisons between participants or comparisons to norms. If simply making within participant comparisons, a practice trial may not be necessary for healthy participants.

Effect of Environment Analysis

When studying the three environmental SFT test sessions (and ignoring the practice trial), it was found that participants can be tested in their homes and achieve reliable scores when appropriate equipment is utilized and the outdoor walking ground is level. No randomized order by environment or gender by environment effects were noted. Thus the order that the SFT was completed in did not have any effect on performance, and males and females did not perform differently in the three environments.

Interestingly, there was not a significant difference between 6-Minute Walk scores in the three environments, even though this test component had the most variability across environments. In a pilot study completed at Iowa State University comparing walking around the standard SFT rectangle to walking around an oval track, participants scored significantly consistently worse when walking around the rectangle. It was believed that this was due to the participants having to slow down to round the corners of the rectangle as tightly as possible. Sciurba, et al. (2003) found that participants walking on a continuous course had a significantly longer walk distance than those who walked on a straight (i.e., out and back) course with turns. It is possible that this difference was not observed in the present study because when walking around the rectangle, the participants had the additional motivation of competition since they could see the other participants. Furthermore, in the present study when walking up and down the hallway participants were allowed to make large turns so as to be able to keep up their speed.

The Back Scratch test component appeared to have slightly lower scores when tested in each participant's home. However, the intraclass correlation coefficient for the Back Scratch was .97. Therefore, all subjects tended to respond to the environment similarly and simply did not perform as well when tested at home. It is possible that this lower score was due to the fact that the Back Scratch was the test component that participants seemed to dislike. Older adults tend to not be very flexible. If they experienced discomfort from their maximal efforts on previous test sessions, they may not have wanted to stretch their shoulders as far. Peer pressure was also absent when this test was performed in their homes. There is also the possibility that they may not have warmed up their upper bodies as well when at home because they are not in the mind set of exercising when at home. These lower scores were not seen in the lower body flexibility test or the Chair Sit and Reach. Their legs were relatively more warmed up due to other SFT components. However, there is no component testing that inadvertently warms up the shoulders other than the actual warm up. The mean difference across Back Scratch scores was 0.5 to 0.7 inches, and it can be argued that this is not physiologically significant for this test. The standard deviations for all three environmental Back Scratch test means were quite large, and thus the test may not be a good representation of upper body flexibility (Thomas and Nelson, 2001). Effect sizes of home with the gym and home with the partially modified trials were very small (Cohen's D .11 and .16, respectively). Thus, the distributions of scores have a high amount of overlap and are not really very different. Parenthetically, the women tended to do better than the men. Yet when the scores were expressed as normative data and compared (Rikli and Jones, 2001), they no longer differed significantly. In other words, these subjects were similar to their peers.

Limitations of Study

First, most of the participants were physically active and knew how hard they could push themselves, because they participated in physical activity on a regular basis. They all appeared to be highly motivated, and Enright (2003) notes that high motivation results in greater 6-Minute Walk scores. Furthermore, Seele (1996) reports that when participants are involved in a regular walking program, one walk test may be adequate to achieve reliable scores.

Second, the goal of this study was to be able to recruit a broad spectrum of older adults--physically active and inactive--as well as those residing in the rural and metropolitan areas. When recruiting inactive participants, it seemed that the older adults who do not habitually exercise have many health issues or utilize a cane or walker. Recruiting rural participants (i.e. those who lived far from the university) proved difficult in the absence of an incentive for their participation. For future research studying the rural older adult population, it is recommended that there be some form of compensation or benefit to the subjects for their participation.

Although the participant population appeared fairly homogenous the standard deviations associated with mean scores were all large. When assessing scores by percentile norms, there were no differences between males and females and there was a broad range of ability on every component of the SFT (typically from 0 to 100th percentile). Overall, the participants were active, but they were not exceptionally fit.

These results show that testing active older adults in their homes yield outcomes no different than testing them at a fitness facility. While these results are promising, generalizing these results to include the inactive, rural older adult population and those suffering from major medical illnesses and diseases should be done with caution. These populations were either not represented well or not at all in the participant population for this study.

It is also important to note that there was one tester for all subjects, so it is possible that utilizing different testers would not achieve the same internal consistency. Therefore, it is recommended that if there are several testers in future research, they should test the same participants across time. This is especially recommended if SFT testing is utilized for determining gains or losses in fitness across time.

Conclusion

There are no significant differences in all test component scores between utilizing the SFT in the modified environments and the fitness facility when a practice trial is performed first at the fitness facility. Although correlations between trials tested in chronological order are high (except the Chair Stand which has moderate correlation of the last three with the first trial) and there is a slight practice effect on all tests. Therefore, this research emphasizes the necessity for performing a practice SFT trial first on a separate day. However, this may only be necessary when making between participant comparisons. The ICCs were quite good (.79 to .97) on all components of the SFT for the three environments (unmodified (gym), partially modified (hallway), modified (home)). The applications of this research should aid in future research studying functional fitness in the older adult rural population via the use of the SFT to specifically measure their functional fitness.

Appendix: Literature Review

Rationale for Use of the Senior Fitness Test

The SFT was designed by Rikli and Jones (1997) to specifically address the weaknesses of other tests currently used to assess physical or functional performance in older adults. The Activities of Daily Living (ADL's, e.g. dressing, bathing, eating) (Katz, Ford & Moskowitz, 1963) and Instrumental Activities of Daily Living (e.g. prepare meals, shopping, housework) (Lawton and Brady, 1969) scales are good at assessing functional limitations in the dependent and frail elderly. Rikli and Jones (1997) argued that few tests are aimed at the 70% of independently living older adults, and they do not encompass specific parameters such as muscle function, aerobic endurance, and flexibility that support those older individuals in their ability to live independently.

While the American Alliance for Health, Physical Education, Recreation, and Dance (AAPHERD) Functional Fitness Assessment for Adults Over 60 Years of Age (Osness, Adrian, Clark, Hoeger, Rabb & Wiswell, 1990) made important strides by publishing research measuring underlying physical parameters, Rikli and Jones (1997) pointed out that it has several flaws. It uses a half-mile walk test, and those who cannot complete a half-mile simply receive no score. Furthermore, there is no assessment for lower body strength, a major predictor for subsequent disability (Guralnik et al., 1995) and a contributor to incidence of falling (de Rekeneire, Visser, Peila, Nevitt, Cauley, Tylavsky, Simonsick & Harris, 2003). The SFT utilizes a 6-minute walk test, granting even those who can take a few steps a score, and the SFT has a timed chair stand test to measure lower body strength.

The Established Populations for Epidemiologic Studies of the Elderly (EPESE) is another short test battery designed to test mobility in older adults by measuring strength, balance, and gait speed (Guralnik, Simonsick, Ferrucci, Glynn, Berkman, Blazer, Scherr & Wallace, 1994). However, Rikli and Jones (1999) noted that while it was successful at classifying large populations of community-residing older adults into broad categories of functional status, it did not measure on a continuum. For example, 22% of participants could not perform the 5-times chair-stand test (i.e. couldn't reach the "floor" of the test) and almost 50% received a perfect score on the tandem balance task (i.e. reached the "ceiling").

The aim of Rikli and Jones (1999) in creating the SFT was to generate a testing battery (see Table 4) that included the following aspects: measured performance on a continuum across all levels of abilities (e.g. eliminates floor and ceiling effects) in 60+ year old subjects, was socially acceptable and meaningful to the subjects, was adaptable to other testing environments (e.g. did not require major, specific equipment or space), and was reliable and valid. Each item of the SFT was selected for content validity, criterion validity, or construct validity, and reliability by an advisory panel made up of researchers at the Ruby Gerontology Center at California State University and nationally noted gerontology researchers and exercise physiologists (Rikli and Jones, 1997; 1998). Additionally, the SFT has age group percentile norms (5 year groups, gender separated), based on data from a nationwide study involving over 7,000 independent-living older adults from 267 test sites and 21 different states (Rikli and Jones, 2001).

Table 4. Components of the Senior Fitness Test

Test Item	Purpose	Scoring
Chair Stand	Lower-body strength	Total number of stands completed in 30 seconds
Arm Curl	Upper-body strength	Total number of arm curls completed in 30 seconds
Chair Sit and Reach	Lower-body flexibility	Distance from tips of middle fingers to toe of shoe in inches
Back Scratch	Upper-body flexibility	Distance of overlap or distance between tips of middle fingers in inches
8 ft. Up-and-Go	Agility and dynamic balance	Time to stand, round a cone 8 feet away, and return to seat
6-Minute Walk	Aerobic endurance	Distance walked in yards

Need for Study of Rural Residing Older Adults

Much of the research regarding the physical activity of older adults has been completed with physically independent, urban dwellers or with community-residing, assisted-living dwellers that may have disease states that have affected their functional fitness. There is a moderate population of independent (or living with family), rural-residing older individuals in the Midwest that has largely not been studied because logistically there are not fitness centers or university laboratories nearby.

Nearly 20% (10,089,000) of the U.S. non-metropolitan population consists of those 60 years and older, compared to only 15% of the metro population (U.S. Department of Agriculture, 2002). Rural residing elderly may suffer from higher morbidity and mortality from chronic diseases and are more sedentary than their urban counterparts (U.S. Department of Health and Human Services, 2000). A desired outcome of the present study was to be able to utilize the SFT in a home or community (e.g. church, senior center, shopping mall) setting to study this population. Therefore, to ensure validity of future research, the reliability of the SFT in a setting outside a fitness center or laboratory setting needs to be established.

It will be important for researchers to be able to continue to study the effects of physical fitness training programs for the elderly. Latham, Bennett, Stretton and Anderson (2004) recently published a systematic review of studies completed with elderly subjects participating in progressive resistance strength training programs to reduce the onset of disability. Out of the sixty-two studies reviewed, only seven utilized home-based exercise programs, and seven others utilized a fitness center along with the subject's home. Clearly, the rural elderly or those only capable of utilizing home-based exercise programs have largely been ignored. Furthermore, their findings were that muscular strength and the functional parameter of gait speed were improved with progressive resistance strength training with this population. However, they concluded that the effects of this type of activity on physical disability for older adults remained unclear. This conclusion surely warrants continued research utilizing resistance training programs and the elderly population.

It is also important to assess differences between the rural population and their metropolitan residing counterparts. Henry, Webster-Gandy and Varakamin (2001) highlight the differences in activity levels within participants of a rural Thailand community, and differences between those residents and other participants residing in a residential facility. Elderly adults who remained with family in the rural setting scored higher on the physical activity level scale and were leaner. This is a developing country and cannot be compared with the Midwest of the United States. Nevertheless, this study shows the importance of looking at the effects of differences in the residences of our elderly on their physical activity levels, functional fitness, and subsequent quality of life.

Pomeroy (2003) completed a three-month stage-matched, theory based nursing educational intervention study in Missouri utilizing community-dwelling rural residing older adults. They were compared to a control group who received a delayed modified intervention. The intervention group increased their accumulated steps as measured by pedometers (and thus presumably their level of physical activity) by the end of the intervention period. This study highlights that educational strategies alone motivate older adults to increase their activity levels. Furthermore, it also establishes the need for future longitudinal studies directed at the generation and maintenance of physically active behaviors for rural elderly.

Home-based physical activity programs could be utilized by rural elderly. Nelson, Layne, Bemstein, Nuemberger, Castaneda, Kaliton, Hausdorff, Judge, Buchner, Roubenoff and Singh (2004) utilized a multi-dimensional home based exercise program to determine its effects on functional performance in older adults. This six month study utilized progressive strength, balance, and general physical activity intervention with thirty-six functionally impaired elderly. The researchers found that increased scores on the Physical Performance Test (assessing activities of daily living), on the EPESE, on the specific physical dimension dynamic balance. However, there were no changes in measures of physiologic capacity such as maximal gait speed, strength, or overall endurance. The researchers suggested that unsupervised exercise sessions for older adults may not be ideal. Future research in this area should yield a clearer picture of whether minimal supervision during exercise sessions can significantly improve strength and cardiopulmonary endurance for near-frail elderly.

Overall, there is a lack of published physical activity research using elderly who do not have access to an established fitness center with its equipment, or older adults who wish only to participate in home-based fitness programs. Determining the reliability of the SFT when used outside a lab or fitness facility is an important step towards researching exercise prescription and follow-up for these individuals.

References

American College of Sports Medicine (ACSM). *ACSM's Guidelines for Exercise Testing and Prescription, 7th Ed.* (2006). Ed. M.H. Whaley. Baltimore, Maryland: Lippincott Williams & Wilkins.

American College of Sports Medicine (ACSM). American College of Sports Medicine position stand. Exercise and physical activity for older adults. (1998). *Medicine and Science in Sports and Exercise, 30,* 992-1008.

Buchner, D.M., & Wagner, E.H. (1992). Preventing frail health. *Health Promotion and Disease Prevention, 8,* 585-591.

de Rekeneire, N., Visser, M., Peila, R., Nevitt, M.C., Cauley, J.A., Tylavsky, F.A., Simonsick, E.M., & Harris, T.B. (2003). Is a fall just a fall: Correlates of falling in healthy older persons. The health, aging, and body composition study. *Journal of the American Geriatrics Society, 51,* 841-846.

Enright, P.L. (2003). The six-minute walk test. *Respiratory Care, 48,* 783-785.

Gibbons, W.J., Fruchter, N., Sloan, S. & Levy, R.D. (2001). Reference values for a multiple repetition 6-minute walk test in healthy adults older than 20 years. *Journal of Cardiopulmonary Rehabilitation, 21,* 87-93.

Gill, T.M., Williams, C.S., Richardson, E.D., & Tinetti, M.E. (1996). Impairments in physical performance and cognitive status as predisposing factors for functional dependence among nondisabled older persons. *Journal of Gerontology: Medical Sciences, 51A,* M283-M288.

Guralnik, J.M., Simonsick, E.M., Ferrucci, L., Glynn, R.J., Berkman, L.F., Blazer, D.G., Scherr, P.A., & Wallace, R.B. (1994). A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. *Journal of Gerontology, 49,* M85-M94.

Guralnik, J.M., Ferrucci, L., Simonsick, E.M., Salive, M.E., & Wallace, R.B. (1995). Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *The New England Journal of Medicine, 332,* 556-561.

Guyatt, G.H, Sullivan, M.J., Thompson, P.J., Fallen, E.L., Pugsley, S.O., Taylor, D.W. & Berman, L.B. (1985). The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. *Canadian Medical Association Journal, 132,* 919-923.

Guyatt, G.H., Pugsley, S.O., Sullivan, M.J., Thompson, P.J., Berman, L., Jones, N.L., Fallen, E.L., & Taylor, D.W. (1984). Effect of encouragement on walking test performance. *Thorax, 39,* 818-822.

Henry, C.J.K., Webster-Gandy, J., & Varakamin, C. (2001). A comparison of physical activity levels in two contrasting elderly populations in Thailand. *American Journal of Human Biology, 13,* 310-315.

Huang, Y., Macera, C.A., Blair, S.N., Brill, P.A., Kohl, H.W. 3rd & Kronenfeld, J.J. (1998). Physical fitness, physical activity, and functional limitation in adults aged 40 and over. *Medicine and Science in Sports and Exercise, 30,* 1430-1435.

Jackson, A.S., Beard, E.F., Wier, L.T., & Blair, S.N. (1995). Changes in aerobic power of men, ages 25-70 years. *Medicine and Science in Sports and Exercise, 27,* 113-120.

Katz, S., Ford, A.B., & Moskowitz, R.W. (1963). Studies of illness in the aged—The index of ADL: A standardized measure of biological and psychosocial function. *Journal of the American Medical Association, 185,* 914-918.

Latham, N.K., Bennett, D.A., Stretton, C.M., & Anderson, C.S. (2004). Systematic review of progressive resistance strength training in older adults. *Journal of Gerontology: Medical Sciences, 59A*, 48-61.

Lawrence, R. & Jette, A.M. (1996). Disentangling the disablement process. *Journal of Gerontology: Social Sciences, 51B,* 5173-5182.

Lawton, M.P. & Brody, E.M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist, 9,* 179-186.

Morey, M.C., Pieper, C.F., & Cornoni-Huntley, J. (1998). Physical fitness and functional limitations in community-dwelling older adults. *Medicine and Science in Sports and Exercise, 30,* 715-723.

Nelson, M.E., Layne, J.E., Bernstein, M.J., Nuemberger, A., Castaneda, C., Kaliton, D., Hausdorff, J., Judge, J.O., Buchner, D.M., Roubenoff, R., & Finatarone Singh, M.A. (2004). The effects of multidimensional home-based exercise on functional performance in elderly people. *Journal of Gerontology: Medical Sciences, 59A,* 154-160.

Osness, W.H., Adrian, M., Clark, B., Hoeger, W., Rabb, D., & Wiswell, R. (1990). *Functional fitness assessment for adults over 60 years.* Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.

Pomeroy, Sherry Lynn Hobgood. (2003). Increasing physical activity in rural elderly. Dissertation, University of Missouri, Columbia.

Podsiadlo, D. & Richardson, S. (1991). The timed "up & go": a test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society, 39,* 142-148.

Rikli, R.E. & Jones, C.J. (1997). Assessing physical performance in independent older adults: issues and guidelines. *Journal of Aging and Physical Activity, 5,* 244-261.

Rikli, R.E. & Jones, C.J. (1998). The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. *Journal of Aging and Physical Activity, 6,* 363-375.

Rikli, R.E. & Jones, C.J. (1999). Development and validation of a functional fitness test for community-residing older adults. *Journal of Aging and Physical Activity, 7*, 129-161.

Rikli, R.E. & Jones, C.J. (2001). *Senior Fitness Test Manual.* Champaign, IL: Human Kinetics.

Rockwood, K., Awalt, E., Carver, D. & MacKinght, C. (2000). Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian Study on Health and Aging. *Journal of Gerontology: Biological and Sciences and Medical Sciences, 55A,* M70-M73.

Sciurba, F., Criner, G.J., Lee, S.M., Mohsenifar, Z., Shade, D., Slivka, W. & Wise R.A. (2003). Six-minute walk distance in chronic obstructive pulmonary disease. Reproducibility and effect of walking course layout and length. *American Journal of Respiratory and Critical Care Medicine, 167,* 1522-1527.

Select Committee on Aging, U.S. House of Representatives. (1992). Aging research: benefits outweigh the costs (Publ. No. 102-871). Washington, D.C: U.S. Government Printing Office.

Spirduso, W.W. (1995). *Physical dimensions of aging.* Champaign, IL: Human Kinetics, 338-355.

Steele, B. Timed walking tests of exercise capacity in chronic cardiopulmonary illness. *Journal of Cardiopulmonary Rehabilitation, 16,* 25-33.

Steffen, T.M., Hacker, T.A. & Mollinger, L. (2002). Age- and gender-related test performance in community-dwelling elderly people: six-minute walk test, berg balance scale, timed up & go test, and gait speeds. *Physical Therapy, 82,* 128-137.

Tanaka, K., Kim, H.S., Yang, J.H., Shimamoto, H., Kokudo, S. & Nishijima, T. (1995). Index of assessing functional status in elderly Japanese men. *Applied Human Science: Journal of Physiological Anthropology, 14,* 65-71.

Thomas, J.R. & Nelson, J.K. (2001). *Research methods in physical activity, 4th Ed.* Champaign, IL: Human Kinetics, 181-200.

Troosters, T., Gosselink, M. & Decramer, M. (1999). Six minute walking distance in healthy elderly subjects. *European Respiratory Journal, 14,* 270-274.

U.S. Census Bureau Iowa Quick Facts. (2000). Retrieved elderly population estimates for U.S. and Iowa on April 5, 2005. http://quickfacts.census.gov/qfd/states/19000.html

U.S. Department of Agriculture, Economic Research Service. (2002). Rural population and migration: rural elderly. Retrieved elderly population estimates for place of residence on April 5, 2005. http://www.ers.usda.gov/Briefing/Population/elderly/

U.S. Department of Health and Human Services-Centers for Disease Control and Prevention. (2000). *Healthy People 2010.* Washington, DC: U.S. Government Printing Office.

Van Huelven, M.J., Kempen, G.I., Ormel, J., & Rispens, P. (1998). Physical fitness related to age and physical activity in older persons. *Medicine and Science in Sports and Exercise, 30,* 434-431.