Further understanding of the Senior Fitness Test: Evidence from community-dwelling high function older adults in Hong Kong

Jing-Dong Liu[a,*], Binh Quach[b], Pak-Kwong Chung[b]

[a] Department of Physical Education, Sun Yat-Sen University, Guangzhou, China
[b] Department of Sport and Physical Education, Hong Kong Baptist University, Hong Kong SAR, China

ABSTRACT

Objectives: This study examined the factor structure of the Senior Fitness Test and examined its measurement and structural invariance across participants with normal body weight and participants who were overweight.
Methods: Eight hundred and forty-five community-dwelling older adults in Hong Kong participated in this study. Their functional fitness was assessed using the Senior Fitness Test. Confirmatory factor analysis was conducted to examine and compare two competing models: a one-factor model and a second-order model. Multiple-group confirmatory factor analysis was conducted to evaluate the measurement and structural invariance of the SFT measurement model. Multiple fit indices were used to evaluate the adequacy of the model fit to the data.
Results: The second-order model outperformed the one-factor model. For men participants, the second-order 4-factor model was indicated, while for women participants, the second-order 5-factor model was evidenced. The measurement and structural invariance of the SFT measurement model across normal weight and overweight participants were also supported.
Conclusion: This study provided initial evidence for the construct validity of the SFT. However, there were some differences in the factor structure of the SFT for men and women participants. Body composition did not necessarily reflect the functional fitness of men participants. All of the older adults responded to the testing items in a similar way regardless of body weight status in both men and women samples. The significance of each component for functional fitness was also similar for participants with different body weight statuses.

1. Introduction

Ageing is usually accompanied by a decline in physical and cognitive functioning, which creates economic and social challenges for the government and results in serious health problems for individuals. Physical fitness is a valid way to reflect physical functioning and plays a critical role in maintaining the everyday activities and independent living of older adults. Studies have consistently revealed that low levels of physical fitness are associated with a high risk of cardiovascular disease and all-cause mortality (Ross et al., 2016; Sui, Laditka, Hardin, & Blair, 2007), and high levels of physical fitness are associated with lower annual costs of medication (Laureano et al., 2014), higher quality of life (Brovold, Skelton, Sylliaas, Mowe, & Bergland, 2014; Chung, Zhao, Liu, & Quach, 2017; Munguía-Izquierdo, Santalla, & Lucia, 2015; Takata et al., 2010), and better cognitive functions such as learning and memory (Duzel, van Praag, & Sendtner, 2016; Kandola, Hendrikse, Lucassen, & Yücel, 2016).

According to the CDC's definition, physical fitness is the ability to carry out daily tasks with alertness and vigour, without undue fatigue and with enough energy reserves to meet emergencies or to enjoy leisure-time pursuits (Centers for Disease Control and Prevention, 2018). It can be further divided into two categories: health- and skill-related physical fitness (Liguori, 2017). However, for older people, the focus of physical fitness may shift from health promotion, disease prevention and performance enhancement to maintaining functional mobility (Rikli & Jones, 2001), which means being able to continue doing the things one wants and needs to do to stay strong, active and independent (Rikli & Jones, 1999). Functional fitness has been defined as "having the physical capacity to perform normal everyday activities safely and independently without undue fatigue" (Rikli & Jones, 1999). Rikli and Jones identified muscular strength (lower and upper body), aerobic endurance, flexibility (lower and upper body), agility/dynamic balance and body composition as the five distinct components of functional fitness (Rikli & Jones, 1999). To assess the functional fitness of older adults, the Senior Fitness Test (SFT) was developed based on the functional capacity framework to measure the five components (Rikli &

Jones, 1997). The SFT includes eight testing items, with the 30-Second Chair Stand Test and the 30-Second Arm Curl Test assessing muscular strength, the 6-Minute Walk Test or the 2-Minute Step Test assessing aerobic endurance, the Chair Sit-and-Reach Test and Back Scratch Test measuring flexibility, the 8-Foot Up-and-Go Test measuring agility/dynamic balance and the body mass index (BMI) measuring body composition. Either the 6-Minute Walk Test or the 2-Minute Step Test can be selected to measure aerobic endurance depends on space availability (Rikli & Jones, 2013).

The SFT has been widely used to assess the functional fitness of older adults throughout the United States and in numerous other countries with materials reproduced in several languages, including Chinese, Danish, Korean, Japanese, Portuguese and Spanish (Chung, Zhao, Liu, & Quach, 2016; Gray, Paulson, & Powers, 2016; Moreno-Vecino et al., 2015; Nawrocka, Mynarski, & Cholewa, 2017; Rikli & Jones, 2013; Sardinha, Santos, Marques, & Mota, 2015). Numerous studies have demonstrated that all of the tests included in the SFT display good reliability (test-retest reliability) and excellent content validity, criterion validity and discriminant validity among older adults (Csuka & McCarty, 1985; Jones, Rikli, Max, & Noffal, 1998; Kline et al., 1987; Miotto, Chodzko-Zajko, Reich, & Supler, 1999; Podsiadlo & Richardson, 1991). Limited research has investigated the SFT's construct validity, which is one of the fundamental psychometric properties of any test or instrument. Construct validity pertains to the way items are organized and structured for a specific measure or test. According to the SFT, there are two potential measurement models: (1) all testing items load on one latent variable of functional fitness (the one-factor model) and (2) under the latent variable of functional fitness, there are five components (muscular strength, aerobic endurance, flexibility, agility/dynamic balance, body composition) with specific testing items (the second-order model). However, which model may best represent the latent variable of functional fitness has not been investigated. Although two studies investigated the factor structure of the SFT (Ho, Wu, Matthews, Chiang, & Lin, 2013; Konopack, Marquez, & Hu, 2008), they exhibited some limitations. First, although much work has been conducted to emphasize the importance of sex and gender-based analyses in health research and potential influences of sex and gender on health behaviours and outcomes (Bird & Rieker, 1999; Johnson, Greaves, & Repta, 2009), the two studies treated the sex as one of the predictors of selected fitness variables. They failed to evaluate the deeper question that whether women and men may react differently to the selected fitness tests or not. In other words, the question that whether the factor structure of the SFT would be different or not across women and men has not been addressed. Second, the two studies examined the factor structure of the SFT measurement model using selected items rather than the full set of items, meaning the answers to the research questions might have been based on limited information. Third, the two studies created a new variable (physical power) by combining the three components of muscular strength, aerobic endurance and agility/balance together, which is inconsistent with the assumption of the SFT. Forth, these studies used relatively small sample sizes, which limited the statistical power of their analyses. Fifth, these studies focused only on the factor structure and did not evaluate the performance of specific testing items, potentially limiting our understanding of the SFT testing items.

Therefore, the purpose of the present study was to examine the factor structure and measurement invariance of the SFT. To achieve the purpose, firstly, we examined whether there would be some differences on their performance in different physical fitness tests between women and men participants. Further, we examined and compared the two aforementioned models to determine which is better to represent the factor structure of the SFT in women and men samples respectively if sex differences were revealed. Furthermore, we examined whether the factor structure was invariant across two groups of people, those who were with normal body weight and those who were overweight (including obese people). The two groups of people were selected because body weight may be a confounder that influences individuals' reactions

to the testing items (Lloyd, Bishop, Walker, Kathryn, & Richardson, 2003; Pelletier, 1988; Sporiš, Jukić, Bok, Vuleta, & Harasin, 2011). Specifically, the measurement and structural invariance of the factor structure of the SFT were evaluated. Investigation on the factor structure of the SFT will reveal specific relationships of each component or test to the overall functional fitness of senior people and more importantly provide crucial evidence for the construct validity of the SFT.

2. Materials and methods

2.1. Participants

The data were taken from 944 older adults attending a cross-sectional study of the functional fitness of community-dwelling older adults from 22 community centres for older adults in different districts of Hong Kong (Kowloon = 8; New Territories = 10; Hong Kong Island = 4). District community centre is a type of community support services at district level to enable older adults to remain in the community and to lead a healthy, respectful and dignified life. All people who are 60 years older are qualified to register as members of community centres and join activities organized by the centres and seek help and support from the centres. The data were collected between March 2014 and April 2015 by 24 professional fitness assessors, who received a 3-h training workshop before data collection with the purpose of homogenizing and standardizing the assessment methods to reduce inconsistencies among the assessors. All tests were conducted in the older people community centres. Written informed consent was obtained from each participant before data collection. The study was approved by the Committee on the Use of Human and Animal Participants in Teaching and Research of the University. Data of 845 older adults out of the 944 participants were used in this study. Data of 99 participants were excluded given at least one missing variable was revealed.

2.2. Data collection

In this study, the 2-Minute Step Test was used to assess aerobic endurance. Therefore, the SFT testing items in this study includes the 30-Second Chair Stand Test, the 30-Second Arm Curl Test, the 2-Minute Step Test, the Chair Sit-and-Reach Test, the Back Scratch Test, the 8-Foot Up-and-Go Test, and BMI. In each testing session, seven testing stations were set up with each station tests one testing item. Each testing item was conducted strictly according to the SFT Manual (Rikli & Jones, 2013). At the start of each testing session, all participants were informed of the procedure and risks for each test. After a 10- to 15-minute rest, the participants received a series of measurements of height and weight from which BMI was calculated (weight/height^2). According to WHO expert consultation's recommendation, alternative cut-off values of BMI will be more appropriate for Asian populations (WHO expert consultation, 2004). Therefore, in this study, the cut-off value of 23 was used ($18.5 <$ BMI < 23.0 as normal weight and BMI ≥ 23.0 as overweight) (WHO expert consultation, 2004). A 10-minute warm up was then conducted, after which the participants were given pre-test instructions to ensure maximum safety and performance. The participants were motivationally encouraged, but not pushed to the point of overexertion or beyond what was easy for them. The tests were terminated whenever the participants showed signs of discomfort, such as pain, dizziness or fatigue.

2.3. Data analysis

Statistical analysis was performed using Statistics Package for the Social Sciences Version 24.0 (SPSS, Inc., Chicago, IL, USA). The descriptive statistics were reported as mean (M) values with standard deviation (SD). An independent *t*-test was conducted to examine sex differences in the variables of interest, and a chi-square test was

conducted to examine whether men and women participants were equally distributed into different age groups and the normal body weight and overweight groups. The factor structure of the SFT was evaluated by comparing the two measurement models using confirmatory factor analysis (CFA) (SPSS Amos, Inc., Chicago, IL, USA). The measurement invariance (factor loadings are constrained to be equivalent) and structural invariance (both factor loadings and path coefficients are constrained to be equivalent) of the SFT measurement model across the normal body weight and overweight groups were examined by using multiple-group CFA. Specifically, a baseline model (configural model) was established first, and then two increasingly constrained models were specified to examine the invariance of measurement (measurement weights) and structural parameters (structural weights) across the two groups.

Multiple fit indices were used to evaluate the adequacy of the model fit to the data, including the chi-square value, the Comparative Fit Index (CFI), the root mean square error of approximation (RMSEA) accompanied by its 90% confidence interval (CI) and the standardized root mean square residual (SRMR). A CFI value > .90 indicates adequate model fit and a value > .95 is the benchmark for models with excellent fit. Values ≤ .08 and .06 (respectively) for RMSEA and SRMR advocate a model with acceptable fit to the data (Hu & Bentler, 1999). Chi-square difference test and information criteria of Akaike's Information Criteria (AIC) and the Bayesian Information Criterion (BIC) were used in model comparisons. It should be noted that AIC and BIC are "badness of fit" index, which means that a model with lower information criteria values would be considered to fit the data better than the one with higher values. Because the likelihood ratio (the chi-square difference test) was sensitive to non-normality and influenced by the sample size, the conclusion of the invariance analysis in this study was reached based on the change in the CFI test, with a change in CFI smaller than .01 from the less constrained model to the more constrained model, indicating invariance (Cheung & Rensvold, 2002).

3. Results

3.1. Descriptive statistics

The demographic information and the descriptive statistics of the variables for overall population, women and men participants are presented in Tables 1–3, respectively. As shown in Table 1 that significant sex differences were observed in the majority of the variables of interest, the following data analyses were conducted separately for men

and women participants. The correlations between all testing items in women and men samples can be found in Table 4.

3.2. Factor structure and measurement invariance

For men participants, Table 5 shows that the one-factor model demonstrates a poor fit to the data and that all of the testing items except for BMI significantly loaded on one latent variable of functional fitness. The results suggest that BMI failed to significantly contribute to the latent variable of function fitness. After removing the BMI item, the fit to the data was much improved. Therefore, two second-order models (5-factor model and 4-factor model with BMI removed) were sequentially evaluated. The second-order 4-factor model without the BMI item displayed an excellent fit to the data, which was much better than other three models (Table 5). Fig. 1 shows the relationships between the testing items and latent variables. These results suggest that the latent variable of functional fitness was more reasonably represented by four components (muscle strength, aerobic endurance, flexibility, and agility and balance) with specific testing item(s) in men participants. The results of the invariance analysis of the second-order 4-factor model across the normal and overweight groups reveal that the Chi-square difference test was not significant and the changes in CFI between the configural model and the measurement and structural weight models were also less than .01, indicating that both the measurement and structural weights of the model were invariant across the two groups of participants (see Table 6). The results suggest that men participants, regardless of their body weight status, responded to the SFT testing items in a similar way and that the contribution of specific components to the latent variable of functional fitness was also similar across the two groups.

For women participants, as shown in Table 5, the one-factor model displayed a poor fit to the data, but all of the testing items significantly loaded on the latent variable of functional fitness. Further examination of the second-order 5-factor model revealed that the model displayed an acceptable and better fit to the data than the one-factor model. Fig. 2 shows the relationships between the testing items and latent variables. These results suggest that the latent variable of functional fitness was more reasonably represented by the five components (body composition, muscle strength, aerobic endurance, flexibility, and agility and balance) with specific testing item(s). The results of the invariance analysis reveal that both the measurement and structural invariance of the second-order model across normal and overweight groups were supported (see Table 6). The results suggest that women participants

Table 1
Participants characteristics, descriptive statistics, and group and sex differences.

	Total (n = 845)	Women (n = 471)	Men (n = 374)	p-value
Age (years), mean (SD)	74.44 (6.20)	74.41 (6.37)	74.49 (5.99)	.836
Age group (years), n (%)				.103
65–69	231 (27.3)	132 (28)	99 (26.5)	–
70–74	196 (23.2)	109 (23.1)	87 (23.3)	–
75–79	207 (24.5)	115 (24.4)	92 (24.6)	–
80–84	175 (20.7)	88 (18.7)	87 (23.3)	–
85+	36 (4.3)	27 (5.7)	9 (2.4)	–
Height (cm), mean (SD)	155.86 (8.32)	150.85 (5.60)	162.23 (6.44)	< . 001
Weight (kg), mean (SD)	58.15 (10.77)	54.06 (9.33)	63.31 (10.24)	< . 001
BMI (kg/m2), mean (SD)	24.02 (3.55)	24.00 (3.88)	24.05 (3.09)	.877
Body weight group (kg/m2) n (%)				.208
Overweight (BMI > = 23 kg/m2)	503 (59.5)	271 (57.54)	232 (62.03)	–
Normal (18 kg/m2 < BMI < 23 kg/m2)	341 (40.5)	200 (42.46)	142 (37.97)	–
30 s Arm curl test (no. of reps), mean (SD)	12.56 (4.44)	12.49 (4.43)	12.67 (4.45)	.557
30 s Chair stand test (no. of stands), mean (SD)	15.01 (4.14)	14.99 (4.10)	15.04 (4.20)	.873
Back scratch test (cm), mean (SD)	−6.14 (11.99)	−3.25 (9.55)	−9.77 (13.66)	< .001
Chair sit and reach test (cm), mean (SD)	0.84 (12.63)	4.51 (11.48)	−3.77 (12.52)	< .001
8-foot up and go test (seconds), mean (SD)	6.72 (1.99)	6.96 (2.01)	6.41 (1.93)	< .001
2-min step test (no. of steps), mean (SD)	82.37 (20.84)	80.45 (19.94)	84.80 (21.7)	< .001

Table 2
Participants characteristics, descriptive statistics, and group differences of women sample.

	Women (n = 471)	Normal (n = 200)	Overweight (n = 271)	p-value
Age (years), mean (SD)	74.41 (6.37)	74.74 (6.74)	74.16 (6.07)	.335
Age group (years), n (%)				.283
65–69	132 (28)	55 (27.5)	77 (28.4)	
70–74	109 (23.1)	45 (22.5)	64 (23.6)	
75–79	115 (24.4)	43 (21.5)	72 (26.6)	
80–84	88 (18.7)	41 (20.5)	47 (17.3)	
85 +	27 (5.7)	16 (8)	11 (4.1)	
Height (cm), mean (SD)	150.85 (5.60)	150.75 (5.77)	150.91 (5.48)	.758
Weight (kg), mean (SD)	54.06 (9.33)	47.01 (5.51)	59.26 (8.08)	< .001
BMI (kg/m2), mean (SD)	24.00 (3.88)	20.65 (1.66)	26.48 (3.10)	< .001
30 s Arm curl test (no. of reps), mean (SD)	12.49 (4.43)	11.96 (4.69)	12.87 (4.19)	.026
30 s Chair stand test (no. of stands), mean (SD)	14.99 (4.10)	15.44 (4.35)	14.66 (3.88)	.044
Back scratch test (cm), mean (SD)	−3.25 (9.55)	−.837 (8.98)	−5.04 (9.59)	< .001
Chair sit and reach test (cm), mean (SD)	4.51 (11.48)	4.12 (11.85)	4.80 (11.21)	.525
8-foot up and go test (seconds), mean (SD)	6.96 (2.01)	6.74 (1.98)	7.12 (2.01)	.045
2-min step test (no. of steps), mean (SD)	80.45 (19.94)	82.03 (19.82)	79.29 (19.98)	.140

with normal body weight and overweight participants responded to the SFT testing items similarly. Furthermore, the contributions of specific components to the latent variable of functional fitness were similar for normal and overweight women participants.

4. Discussion

This study examined the factor structure of the SFT and evaluated whether participants with different body weight statuses (normal weight versus overweight) responded to the SFT testing items similarly (measurement weight invariance) and whether the contribution of the SFT testing items to the latent variable of functional fitness (structural weights invariance) was similar across the two groups of people. The use of CFA allowed researchers to evaluate the relationships between testing items (e.g., the 30-Second Arm Curl Test) and latent factors (e.g., functional fitness or muscle strength), providing a rationale for including or excluding certain items or components to identify a theoretically reasonable factor structure of a psychometrically sound tool. More importantly, the results of multiple-group CFA provide evidence for the stability and invariance of the factor structure across different population groups.

This study provides the first evidence for the construct validity of the SFT among community-dwelling older adults in Hong Kong. The results suggest that there are some sex differences in the factor structure of the SFT. For men participants, the second-order model with four components (muscle strength, aerobic endurance, flexibility, and

agility/balance) better represented the factor structure of the SFT. In contrast, for women participants, the second-order model with five components (muscle strength, aerobic endurance, flexibility, agility/balance, and body composition) better represented the factor structure of the SFT. The results of this study suggest that body composition (BMI) may not necessarily reflect the functional fitness of men participants. Although body composition may contribute negatively to women participants' functional fitness, the strength is relative low (−.13). No research has thoroughly investigated the factor structure of the SFT. Konopack et al. found that the correlations between six selected testing items of the SFT were insignificant or relatively low and concluded that the one-factor model may not be supported (Konopack et al., 2008). Therefore, they explored the latent structures of physical power and flexibility instead of functional fitness. Ho et al. (2013) replicated the two latent structures by including different correlates (physical activity) in the model in a small group of older Chinese adults in Taiwan. The two studies provided initial insight into the factor structure and contributed to our understanding of the SFT. However, these two studies analysed the data of men and women together, potentially influencing the generalizability of the conclusions. Unlike those of previous studies, our results reveal that most of the SFT testing items were significantly associated with each other (low to moderate) across men and women participants. These results provide initial evidence for examining the two potential measurement models. Results of this study reveal that the second-order model outperformed the one-factor model for both men and women participants, providing empirical

Table 3
Participants characteristics, descriptive statistics, and group differences of men sample.

	Men (n = 374)	Normal (n = 142)	Overweight (n = 232)	p-value
Age (years), mean (SD)	74.49 (5.99)	75.20 (6.4)	74.05 (5.70)	.076
Age group (years), n (%)				< .001
65–69	99 (26.5)	39 (27.5)	60 (25.9)	
70–74	87 (23.3)	26 (18.3)	61 (26.3)	
75–79	92 (24.6)	28 (19.7)	64 (27.6)	
80–84	87 (23.3)	43 (30.3)	44 (19.0)	
85 +	9 (2.4)	6 (4.2)	3 (1.3)	
Height (cm), mean (SD)	162.23 (6.44)	162.62 (5.87)	161.99 (6.77)	.365
Weight (kg), mean (SD)	63.31 (10.24)	55.63 (5.57)	68.00 (9.59)	< .001
BMI (kg/m2), mean (SD)	24.05 (3.09)	21.01 (1.46)	25.90 (2.23)	< .001
30 s Arm curl test (no. of reps), mean (SD)	12.67 (4.45)	12.22 (4.29)	12.94 (4.54)	.129
30 s Chair stand test (no. of stands), mean (SD)	15.04 (4.20)	15.31 (4.20)	14.87 (4.20)	.328
Back scratch test (cm), mean (SD)	−9.77 (13.66)	−6.14 (12.68)	−11.99 (13.79)	< .001
Chair sit and reach test (cm), mean (SD)	−3.77 (12.52)	−4.76 (11.93)	−3.16 (12.85)	.229
8-foot up and go test (seconds), mean (SD)	6.41 (1.93)	6.35 (2.10)	6.45 (1.82)	.664
2-min step test (no. of steps), mean (SD)	84.80 (21.7)	83.67 (22.04)	85.48 (21.51)	.436

Table 4
Correlations among variables.

		2.	3.	4.	5.	6.	7.
1. BMI	1.00	−.304**	−.021	.093*	−.066	.161**	−.100*
2. Back scratch test	−.273**	1.00	.259**	−.290**	.251**	.093*	.235**
3. Chair sit and reach test	.018	.220**	1.00	−.277**	.299**	.208**	.305**
4. 8-foot up and go test	.062	−.278**	−.240**	1.00	−.555**	−.234**	−.460**
5. 30 s Chair stand test	−.115*	.152**	.162**	−.556**	1.00	.373**	.506**
6. 30 s Arm curl test	.055	.124*	.149**	−.364**	.395**	1.00	.320**
7. 2-min step test	.009	.206**	.163**	−.515**	.470**	.383**	1.00

  ** Correlation is significant at the 0.01 level (2-tailed).

  * Correlation is significant at the 0.05 level (2-tailed). Correlations in lower and upper diagonal matrix are for women and men participants respectively.

Table 5
Model fit indices for different models.

	Model	χ^2	df	CFI	RMSEA (90% CI)	SRMR	AIC	BIC	Model Comparison	Δ χ^2 (Δdf)
Men										
N = 374	M1: One factor model	60.58	14	0.898	0.094 (0.071–0.119)	0.063	88.580	143.519		
	M2: One factor model#	19.42	9	0.975	0.056 (0.020–0.090)	0.039	43.421	124.469	M2 vs M1	41.46 (5)**
	M3: 2nd order-5 factor model	50.87	12	0.915	0.093 (0.068–0.120)	0.057	82.866	145.654	M3 vs M1	9.87 (2)**
	M4: 2nd order-4 factor model#	9.66	7	0.994	0.032 (0.000–0.076)	0.023	41.107	107.458	M4 vs M3	41.21 (5)**
Women										
N = 471	M1: One factor model	80.20	14	0.887	0.100 (0.080–0.122)	0.066	108.203	166.371		
	M3: 2nd order-5 factor model	66.54	12	0.907	0.098 (0.076–0.122)	0.061	98.542	165.02	M1 vs M3	13.66 (2)**

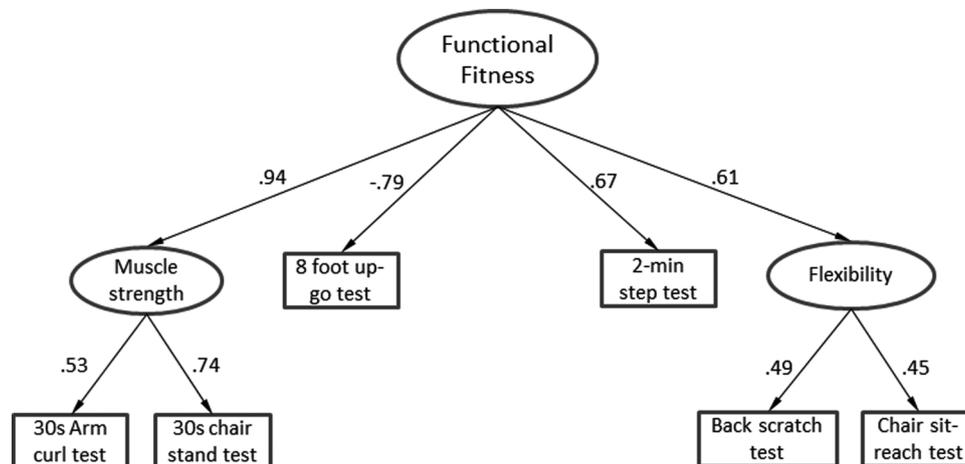*#Note* = for these models, BMI test was removed. **p < .01.

Fig. 1. The 2nd order model and the standardized relationships among variables for men.

Table 6
Model fit indices for different models in invariance analysis.

	Model	χ^2	df	CFI	ΔCFI	RMSEA (90% CI)	SRMR	Model Comparison	Δ χ^2 (Δdf)
Men									
N = 374	Normal (N = 142)	6.36	7	1.000		0.000 (0.000–0.098)	0.031		
	Overweight (N = 232)	9.84	7	0.989		0.042 (0.000–0.097)	0.029		
	M1: Configural model	16.19	14	0.995	–	0.021 (0.000–0.056)	0.031		
	M2: Measurement weights	18.52	17	0.996	0.001	0.015 (0.000–0.051)	0.038	M2 vs M1	2.23 (3)
	M3: Structural weights	20.27	19	0.997	0.001	0.013 (0.000–0.048)	0.040	M3 vs M2	1.75 (2)
Women									
N = 471	Normal (N = 200)	17.31	12	0.978		0.047 (0.000–0.093)	0.044		
	Overweight (N = 271)	34.57	12	0.928		0.083 (0.052–0.117)	0.061		
	M1: Configural model	55.05	24	0.944		0.053 (0.034–0.071)	0.047		
	M2: measurement weights	64.61	28	0.936	0.008	0.053 (0.036–0.070)	0.058	M2 vs M1	8.56 (4)
	M3: Structural weights	64.91	30	0.939	0.003	0.05 (0.033–0.066)	0.059	M3 vs M2	0.3 (2)

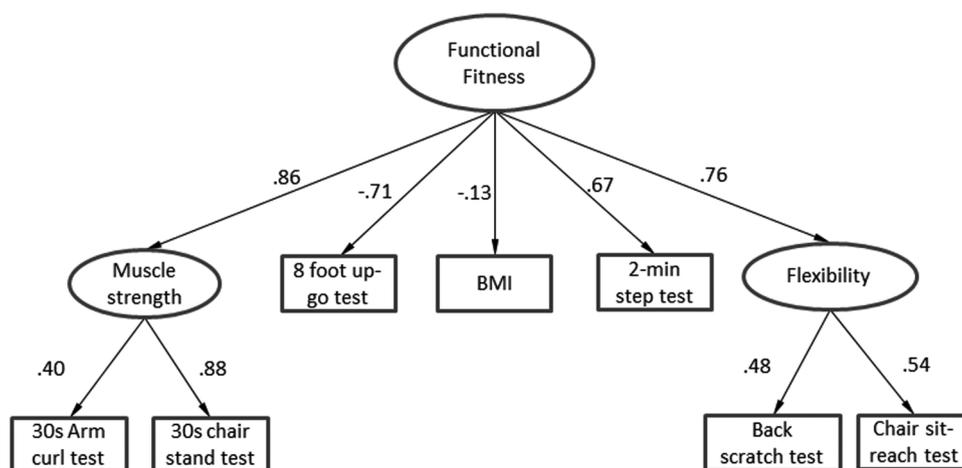*#Note* = for these models, BMI test was removed.

Fig. 2. The 2nd order model and the standardized relationships among variables for women.

evidence for the factor structure of the SFT.

In practical setting, physical fitness instructors and practitioners are always mindful of the selection of particular tests for their clients, especially for those special populations such as older adults. Body weight and body composition have been important indicators considered in the selection of physiological tests for these populations. However, no study has investigated whether body composition or body weight influences older adults' reactions to the selected testing items of the SFT. This study answered the research question by using multiple-group CFA to evaluate the measurement and structural invariance of the SFT. It was found that both measurement and structural invariance were supported across participants with normal body weight and participants who were overweight. This result implies that among both men and women participants, the participants in the two weight groups responded to the testing items of the SFT in a similar way. This is the first study that provided empirical support for the application of the SFT among older adults with different body weight statuses. This finding provided practitioners practical reference for their selection of physical fitness tests. Furthermore, the contribution of specific components to the latent variable of functional fitness was similar across both groups of participants. In other words, the significance of each component for functional fitness was similar across the two groups. This result provides further evidence for practitioners to apply the SFT battery among older adults with different body weight statuses.

It was found that some correlations between testing items, components and functional fitness differed between the men and women samples. For example, the correlation between flexibility and functional fitness in the women sample was much higher than that in the men sample. A similar difference was also observed in the relationship between muscle strength and the 30-Second Chair Stand Test. However, given that the factor structure of the SFT is different for the men and women populations, no statistical test was conducted to evaluate whether the differences were statistically significant. If the differences were statistically significant, they would have suggested that the weighting of the component to functional fitness or the testing item to the component differed across the men and women participants. The results of this study reminded practitioners that in their practice they might need to interpret the contributions and significance of different components or tests to the overall physical fitness cautiously as there may be some differences between men and women older adults. Future research may further investigate this matter.

This study has certain limitations. First, although a large sample of older adults was recruited to participate in this study, the convenience sampling method was used. Future studies may consider using a random sampling approach to recruit more representative samples. Second, all of the participants in this study were healthy older adults;

therefore, the conclusion may not be generalizable to the unhealthy older people (e.g., older adults diagnosed with sarcopenia). The SFT has an adapted version for special populations, and researchers may extend the investigation of the factor structure of the adapted version of the SFT to special populations. Third, BMI has been used as reference for grouping of body weight status. Waist circumference has been considered as one more appropriate reference marker of obesity and cardiovascular risk (Shen et al., 2006; Torres-Castillo et al., 2018). Researchers may consider using waist circumference as grouping marker in future study. Forth, age has been widely reported to be associated with the functional fitness level of older adults. Given that the sample size in each age group was relatively small in this study, an invariance analysis across age groups was not conducted. Future research is encouraged to further investigate the question with a larger sample.

5. Conclusion

The present study reveals that different factor structures of the SFT were observed among women and men older adults. The factor structure of the SFT was better represented by a second-order 4-factor model for men sample and a second-order 5-factor model for women sample. Body composition (BMI) was found not to be significantly associated with the latent variable of functional fitness in the men sample. In both the men and women samples, regardless of body weight status (normal vs. overweight), the participants responded to the SFT testing items in similar ways. The contributions of the SFT components to the latent variable of functional fitness in the normal weight and overweight samples were also similar.

Authors' contributions

JDL conceived of the study, developed the research design, performed statistical analyses, interpreted the data and drafted the initial manuscript; PKC completed data collection, interpreted the data, and revised the manuscript; BQ collected the data, interpreted the data and revised the manuscript. All authors have read and approved the final version of the manuscript, and agree with the order of presentation of the authors.

Conflict of interest statement

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by the Faculty Research Grant (FRG) from

References

Bird, C. E., & Rieker, P. P. (1999). Gender matters: An integrated model for understanding men's and women's health. *Social Science & Medicine, 48*(6), 745–755. https://doi.org/10.1016/S0277-9536(98)00402-X.

Brovold, T., Skelton, D. A., Sylliaas, H., Mowe, M., & Bergland, A. (2014). Association between health-related quality of life, physical fitness, and physical activity in older adults recently discharged from hospital. *Journal of Aging and Physical Activity, 22*(3), 405–413. https://doi.org/10.1123/JAPA.2012-0315.

Centers for Disease Control and Prevention (2018). *Physical fitness. CDC Glossary of Terms* https://www.cdc.gov/physicalactivity/basics/glossary/index.htm Accessed August 30, 2018.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

Chung, P. K., Zhao, Y., Liu, J. D., & Quach, B. (2016). Functional fitness norms for community-dwelling older adults in Hong Kong. *Archives of Gerontology and Geriatrics, 65*, 54–62. https://doi.org/10.1016/j.archger.2016.03.006.

Chung, P. K., Zhao, Y., Liu, J. D., & Quach, B. (2017). A canonical correlation analysis on the relationship between functional fitness and health-related quality of life in older adults. *Archives of Gerontology and Geriatrics, 68*, 44–48. https://doi.org/10.1016/j.archger.2016.08.007.

Csuka, M., & McCarty, D. J. (1985). Simple method for measurement of lower extremity muscle strength. *The American Journal of Medicine, 78*, 77–81.

Duzel, E., van Praag, H., & Sendtner, M. (2016). Can physical exercise in old age improve memory and hippocampal function? *Brain, 139*, 662–673. https://doi.org/10.1093/brain/awv407.

Gray, M., Paulson, S., & Powers, M. (2016). Maximal, not habitual, walking velocity is more highly correlated to functional fitness of community-dwelling older adults. *Journal of Aging & Physical Activity, 24*(2), 305–310. https://doi.org/10.1123/japa.2015-0078.

Ho, L. A., Wu, H. J., Matthews, T. D., Chiang, J., & Lin, Y. J. (2013). Factor structure and correlates of functional fitness of older adults. *International Journal of Gerontology, 7*, 158–161. https://doi.org/10.1016/j.ijge.2012.11.009.

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. https://doi.org/10.1080/10705519909540118.

Johnson, J. L., Greaves, L., & Repta, R. (2009). Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health, 8*, 14. https://doi.org/10.1186/1475-9276-8-14.

Jones, C. J., Rikli, R. E., Max, J., & Noffal, G. (1998). The reliability and validity of a chair sit-and-reach test as a measure of hamstring flexibility in older adults. *Research Quarterly for Exercise and Sport, 69*, 338–343. https://doi.org/10.1080/02701367.1998.10607708.

Kandola, A., Hendrikse, J., Lucassen, P. J., & Yücel, M. (2016). Aerobic exercise as a tool to improve hippocampal plasticity and function in humans: Practical implications for mental health treatment. *Frontiers in Human Neuroscience*, 373. https://doi.org/10.3389/fnhum.2016.00373.

Kline, G. M., Porcari, J. P., Hintermeister, R., Freedson, P. S., Ward, A., McCarron, R. F., et al. (1987). Estimation of VO2 max from a one-mile track walk, gender, age and body weight. *Medicine and Science in Sports and Exercise, 19*, 253–259.

Konopack, J. F., Marquez, D. X., Hu, L., et al. (2008). Correlates of functional fitness in older adults. *International Journal of Behavioral Medicine, 15*, 311–318. https://doi.org/10.1080/10705500802365557.

Laureano, M. L., Martins, R. A., Sousa, N. M., Machado-Rodrigues, A. M., Valente-Santos, J., & Coelho-E-Silva, M. J. (2014). Relationship between functional fitness, medication costs and mood in elderly people. *Revista da Associação Médica Brasileira, 60*(3), 200–207. https://doi.org/10.4025/jphyseduc.v27i1.2725.

Liguori, G. (2017). *ACSM's Health-related physical fitness assessment manual* (5th ed.). Wolters Kluwer.

Lloyd, L. K., Bishop, P. A., Walker, J. L., Kathryn, R. S., & Richardson, M. T. (2003). The influence of body size and composition on FITNESSGRAM(r) test performance and the adjustment of FITNESSGRAM(r) test scores for skinfold thickness in youth. *Measurement in Physical Education and Exercise Science, 7*(4), 205–4226. https://doi.org/10.1207/S15327841MPEE0704_1.

Miotto, J. M., Chodzko-Zajko, W. J., Reich, J. L., & Supler, M. M. (1999). Reliability and validity of the Fullerton Functional Fitness Test: An independent replication study. *Journal of Aging and Physical Activity, 7*, 339–353. https://doi.org/10.1123/japa.7.4.339.

Moreno-Vecino, B., Arija-Blázquez, A., Pedrero-Chamizo, R., Alcázar, A., Gómez-Cabelloc, A., Pérez-López, F. R., et al. (2015). Associations between obesity, physical fitness, and urinary incontinence in non-institutionalized postmenopausal women: The elderly EXERNET multi-center study. *Maturitas, 82*(2), 208–214. https://doi.org/10.1016/j.maturitas.2015.07.008.

Munguía-Izquierdo, D. S., Santalla, A., & Lucia, A. A. (2015). Cardiorespiratory fitness, physical activity and quality of life in McArdle patients. *Medicine and Science in Sports and Exercise, 47*, 799–808. https://doi.org/10.1249/mss.0000000000000458.

Nawrocka, A., Mynarski, W., & Cholewa, J. (2017). Adherence to physical activity guidelines and functional fitness of elderly women, using objective measurement. *Annals of Agricultural and Environmental Medicine, 24*(4), 632–635. https://doi.org/10.5604/12321966.1231388.

Pelletier, D. L. (1988). The effects of occupation, leisure activities, and body composition on aerobic fitness in Western Samoan men. *Human Biology, 60*(6), 889–899.

Podsiadlo, D., & Richardson, S. (1991). The timed "up and 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. (2001). *Senior fitness test manual.* Champaign, IL: Human Kinetics.

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. https://doi.org/10.1123/japa.7.2.129.

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. https://doi.org/10.1123/japa.5.3.244.

Rikli, R. E., & Jones, C. J. (2013). Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. *Gerontologist, 53*(2), 255–267. https://doi.org/10.1093/geront/gns071.

Ross, R., Blair, S. N., Arena, R., Church, T. S., Després, J. P., Franklin, B. A., et al. (2016). Importance of assessing cardiorespiratory fitness in clinical practice: A case for fitness as a clinical vital sign: A scientific statement from the american heart association. *Circulation, 134*(24), e653–e699. https://doi.org/10.1161/CIR.0000000000000461.

Sardinha, L. B., Santos, D. A., Marques, E. A., & Mota, J. (2015). Criterion-referenced fitness standards for predicting physical independence into later life. *Experimental Gerontology, 61*, 142–146. https://doi.org/10.1016/j.exger.2014.12.012.

Shen, W., Punyanitya, M., Chen, J., Gallagher, D., Albu, J., P-Sunyer, X., et al. (2006). Waist circumference correlates with metabolic syndrome indicator better than percentage fat. *Obesity, 14*(4), 727–736. https://doi.org/10.1038/oby.2006.83.

Sporiš, G., Jukić, I., Bok, D., Vuleta, D., & Harasin, D. (2011). Impact of body composition on performance in fitness tests among personnel of the Croatian navy. *Collegium Antropologicum, 35*(2), 335–339.

Sui, X., Laditka, J. N., Hardin, J. W., & Blair, S. N. (2007). Estimated functional capacity predicts mortality in older adults. *Journal of the American Geriatrics Society, 55*(12), 1940–1947. https://doi.org/10.1111/j.1532-5415.2007.01455.x.

Takata, Y., Ansai, T., Soh, I., Awano, S., Yoshitake, Y., Kimura, Y., et al. (2010). Quality of life and physical fitness in an 85-year-old population. *Archives of Gerontology and Geriatrics, 50*(3), 272–276. https://doi.org/10.1016/j.archger.2009.04.005.

Torres-Castillo, N., Campos-Perez, W., Gonzalez-Becerra, K., Hernandez-Cañaveral, I., Vizmanos, B., Muñoz-Valle, J., Martinez-Lopez, E., et al. (2018). Waist circumference is an anthropometric parameter that identifies women with metabolically unhealthy phenotypes. *Nutrients, 10*(4), 447. https://doi.org/10.3390/nu10040447.

WHO expert consultation (2004). Appropriate body mass index for Asian populations and its implications for policy and intervention strategies. *Lancet, 363*(9403), 157–163. https://doi.org/10.1016/S0140-6736(03)15268-3.