Reliability of Senior Fitness Test version in Spanish for older people in Tunja-Colombia

Elisa Andrea Cobo-Mejía, Marlene E. Ochoa González, Lida Y. Ruiz Castillo, Deisy M. Vargas Niño, Angélica M. Sáenz Pacheco, Carolina Sandoval-Cuellar

Universidad de Boyacá. Tunja-Colombia.

Received: 20.11.2015
Accepted: 14.03.2016

Summary

Introduction: *Senior Fitness Test*, test battery of functional capacity of the elderly, reliable test consists of seven original version of Rikli and Jones, evaluating the muscular strength, endurance, flexibility and agility. This text is the result of the research process "Validation of the *Senior Fitness Test*" which in its first stage, identified the face validity of the Spanish version, through a descriptive study, through the cross-cultural adaptation maintaining equivalence semantic, idiomatic and conceptual, with an index of global agreement of 0.9485. To determine the reliability of the Spanish version of the *Senior Fitness Test* from the internal consistency and reproducibility of the tests of the battery.
Methodology: *Senior Fitness Test* version 237 older adults between 60 and 90 years was applied in Spanish. Reliability was established by Cronbach Alpha and reproducibility through the intraclass correlation coefficient, the test applied 12 days apart. For registration and data analysis statistical program SPSS version 20.0 was used.
Results: Cronbach's alpha was 0.708, the reproducibility of evidence obtained an ICC between 0.851 and 0.960 with 95% and p <0.005 in the test - retest.
Conclusions: *Senior Fitness Test* test version in Spanish, have high psychometric properties, as evidenced by the level of agreement among experts within the validity of appearance, combined with the values obtained from the internal consistency and reproducibility, which supports the usefulness of Battery for assessing the functional capacity of older adults Colombians.

Key words:
Physical fitness.
Reproducibility of Results. Elder.
Validation Studies

Confiabilidad del *Senior Fitness Test* versión en español, para población adulta mayor en Tunja-Colombia

Resumen

Introducción: El *Senior Fitness Test*, batería de pruebas de la capacidad funcional del adulto mayor, compuesto por siete test confiables en versión original de Rikli y Jones, que evalúan la fuerza muscular, la resistencia aeróbica, la flexibilidad y la agilidad. Este texto, es resultado del proceso de investigación "Validación del *Senior Fitness Test*", que en su primera etapa, identificó la validez de apariencia de la versión en español, por medio de un estudio descriptivo, a través de la adaptación transcultural manteniendo la equivalencia semántica, idiomática y conceptual, con un índice de acuerdo global de 0,9485.
Objetivo: Determinar la confiabilidad de la versión en español del *Senior Fitness Test* a partir de la consistencia interna y la reproducibilidad de los tests de la batería.
Metodología: Se aplicó el *Senior Fitness Test* versión en español a 237 adultos mayores entre 60 y 90 años. La confiabilidad se estableció mediante el Alpha de Cronbach y la reproducibilidad a través del coeficiente de correlación intraclase, al aplicarse la prueba con 12 días de diferencia. Para el registro y análisis de datos se usó programa estadístico SPSS® versión 20.0.
Resultados: El Alfa de Cronbach fue de 0,708, la reproducibilidad de las pruebas obtuvo un CCI entre 0,851 y 0,960 con IC del 95% y un p<0,005 en el test – retest.
Conclusiones: Las pruebas del *Senior Fitness Test* versión en español, presentan altas propiedades psicométricas, evidenciado en el nivel de acuerdo de los expertos dentro de la validez de apariencia, sumado a los valores obtenidos en la consistencia interna y reproducibilidad, lo avala la utilidad de la batería para evaluar la capacidad funcional de los adultos mayores colombianos.

Palabras clave:
Aptitud física.
Reproducibilidad de resultados.
Anciano. Estudios de validación.

Financiación: artículo producto de un proyecto de investigación financiado por la Universidad de Boyacá y el Departamento Administrativo de Ciencia Tecnología e Innovación – Colciencias (Colombia), bajo la Beca Pasantía Jóvenes Investigadores e Innovadores 2013, entidades que no influyeron en el diseño del estudio.

Correspondence: Elisa A Cobo-Mejía
E-mail: eacobo@uniboyaca.edu.co

Introduction

In acknowledgement of the need for an instrument to assess the functional capacity of older people - understood as the physical condition required to perform daily activities safely, independently and without excessive fatigue - researchers from the State University of California, Fullerton, have developed and assessed a battery of tests for functional capacity, which includes performance standards, called the Senior Fitness Test (SFT). 7,183 North Americans aged between 60 and 94 years participated in the study[1].

Among the physical condition variables included in the test were: muscle strength (upper and lower limbs), aerobic resistance, flexibility (upper and lower limbs) and agility (1), where each item of the SFT was compared to the respective gold standard, to discover the psychometric properties, apart from height and weight. Some properties reported in the literature are:

- Sitting and standing up from a chair test: Reliability test – retest in contrast with a maximum repetition (MR) on leg press of r = 0.78 for men and 0.71 for women[2,3].
- Elbow bend test: Validity with a r = 0.82 of correlation with Cybex machine arm curl performance[4].
- 6-Minute walking test: with correlation r = 0.82 for men and r = 0.71 for women between this test and the Treadmill test[5].
- The 2-minute step test: correlation with the Rockport test r = 0.73[6] and strong reliability evidence with a CIC of 0.8 and of 0.93, specifically in frail adults. For this group, moderate evidence was reported in intra and inter assessment reliability and in construct validity[7].
- Sitting and touching the toes test using a chair (measures lower-body flexibility): reports r = 0.83 in correlation with the goniometry[1,3].
- Joining the hands behind the back test: given that there is no gold standard, it should be considered a "consensus element", an example left to expert criteria[8].
- Standing up, walking 8 feet and sitting again test: Correlation of r = 0.81 with the Berg Balance Scale and with the Barthel ADL Index of r = 0.789.

In general the SFT reports reliability according to the co-efficient of intraclass correlation (CIC) using the ANOVA variance analysis method, between 0.80 to 0.98[10]. Likewise, Langhammer and Stanghelle describe the association between the gold standard of a maximum repetition vs sitting and standing up from the chair in the same way, with the other SFT tests, with a CIC of 0.73 to 0.838. Likewise, Levy *et al.* report values of convergent validity of 0.70 and test re-test reliability, CIC = 0.80 (95% CI)[11].

On the other hand, Boneth *et al.* determine the test reproducibility - re-test via CIC, with a confidence interval (CI = 95%), for the elbow flex test in 30 seconds a CIC = 0.88 was obtained and for the standing up and sitting in a chair test in 30 seconds a CIC = 0.78 was obtained, indicating that these tests have good reproducibility to assess the muscle strength of upper and lower limbs in functionally independent older adults[12].

Likewise, Pedrero-Chamizo *et al.* measure the physical conditions of non-institutionalised and independent adults to relate them to gender and age, for which they use the SFT tests alongside anthropometric measurements and additional tests, proposed in accordance with protocols that assess motor components. This way they obtain a reliability of the physical aptitude tests of over 90%, apart for the 2.5 metre walk - returning - sitting again, the intraclass reliability values for all were of 0.80 - 0.90, with the majority of the values positioned at 0.90 or above, which indicates that the tests have a good relative reliability in the pilots carried out[13].

In accordance with the aforementioned, the aim of this study is to assess the reliability from the internal consistence and the reproducibility of the SFT version in Spanish for the older adult population in Tunja, thus this is this second stage of the research process "Assessment of the Senior Fitness Test", a study that initially determined the surface validity obtaining a globally approved index of 0.9485 and a comprehensibility of 85.2%, referred by subjects from diverse levels of schooling and ages[14]. These studies are of interest to the assessment and intervention processes of the functional capacities of older adults.

Material and method

A descriptive study, which established the confidence (reliability), understood as the parameter that assesses the consistence and precision of a measurement, enables the discovery of the degree to which the results - obtained based on the application of an instrument - are affected by the error that has been committed in the measurement; if it is slight, the instrument result can be trusted[15].

To identify it, in the SFT, firstly the internal consistency was discovered, for which 237 voluntary adults aged over 60 years from the city of Tunja were studied, who gave their informed prior consent; a sample obtained from a population of 1,330 people with an Alfa of 0.85[16] and a reliability of 95%, that attend 4 integral centres for physical activity. Furthermore a proportional stratified probabilistic sample was used in each centre, obtaining a sample fraction to discover the number of older adults to assess in each one, as displayed in Table 1. Once the stratified sample had been obtained, the participants were chosen using a simple random sample, using Epidat® software.

Secondly, to establish the intraclass relationship in related groups, the simple correlation formula was used in one group:

$$n = 3 + \frac{K}{C^2} \qquad K = (Z\alpha + Z\beta)^2$$

$$C = 0{,}5 \ln \frac{(1 + r)}{(1 - r)}$$

In which: R = expected co-efficient of correlation.

With an expected correlation of 90%, a confidence interval of 95% and a power of 90%, obtaining a sample size of 38 individuals[17].

For reliability, firstly, based on the analysis of the internal consistency, defined as the degree to which all the elements of an instrument measure the same attribute or dimension[18], the Cronbach Alfa is established. Secondly, to identify the test re-test, which reflects the

Table 1. Stratified sample with proportional allocation.

Physical activity groups	No. older adults	Wi	Proportional allocation
Comfaboy	70	0.052631579	12.4737
Irdet	700	0.526315789	124.7368
Indeportes	440	0.330827068	78.4060
Club Nueva Vida	120	0.090225564	21.3835

Source: own creation.

Table 2. Central trend measurements of the subject population (n: 237).

	Average	*SD	**CI 95%	Median	***Min	****Max
Age - Years	68.73	6.653	67.88 - 69.58	68.30	60	90
Weight - Kg.	64.15	10.41	62.81 - 65.48	64.40	22.40	98.80
Height - cm	152.52	6.773	151.66 - 153.39	152.00	135	193
Body Mass Index (BMI) - kg/m^2	27.67	3.94	27.16 - 28.17	27.54	18.64	41.68
Sitting and standing up from a chair - rep.	14.14	3.54	13.68 – 14.59	14.00	5	24
Elbow flex - rep.	15.63	3.71	15.15 - 16.10	16.00	6	27
6-Minute walk - mts	524.55	116.03	509.70 - 539.40	530.70	138.00	787.00
2-Minute walk - No steps	72.71	19.57	70.21 - 75.22	74.00	19	124
Sitting and touching feet using a chair - cm.	-7.81	9.95	-9.08 - -6.53	-6.00	-33.0	18.5
Joining hands behind the back - cm.	-12.05	-12.48	-13,54- -10.45	-12.00	-64.0	81.0
Standing, walking 8 feet. turning and sitting again - follow-up.	6.06	1.41	5,88 - 6.25	5.760	3.50	12.41

*SD: Standard deviation **CI 95%: Confidence interval at 95%. ***Min: minimum value. ****Max: maximum value.
Source: Own creation.

degree to which similar results are obtained when a scale is applied on different occasions separated by a short interval of time[19], in which for this study the time between the first and the second assessment was 12 days, the CCI was identified.

This research was approved by the Boyaca University Ethics Committee, which also requires the application of informed consent from each of the participants, considering that established in the Colombian Motion 008430 of 1993, by which the minimum risk is classified when contemplating the application of physical tests.

Finally, for the recording and analysis of data, the statistics programme SPSS® version 20.0 was used, in which the Cronbach Alfa was established based on the percentiles of the 6 SFT tests, in the first instance, with a 6-minute walking test, and the second with the 2-minute step test, which is considered an alternative for measuring aerobic capacity. The body mass index (BMI) was excluded given that its values follow an inversely proportional distribution to the other tests, because the higher the BMI, the greater the body weight of the individual.

Theoretically it was considered that a 0.0 value indicates that there is no correlation between the scores and that 1.0, represents perfect correlation[20]. The minimum accepted value is 0.70 and the maximum expected is 0.90; values between 0.80 and 0.90 are preferred, however, if a better instrument is not available, lower values are accepted[21].

For the CIC the scores are interpreted as high when they are equal to or greater than 0.70, moderate between 0.5 and 0.69 and low if below 0.50[22]. These measurements express the stability and repeatability of the SFT when the test is used on two different occasions, in the same conditions and with the same population.

The previous analysis is framed within the classic theory of the tests, according to which the "reliability of a test is linked to the random measurement errors present in the scores obtained based on their application. Therefore, a test will be more reliable the fewer the measurement errors contained within the scores obtained from the subjects to whom they were applied"[15].

Results

The SFT Spanish version was applied to 237 older adults aged between 60 and 90 years, with an average age of 68.7 years, with a SD of 6.65 years, of which 209 were women (88.2%) and 28 men (11.8%). The predominant social stratum was 2, with 48.1%, in terms of education over half (52.7%) had primary studies, 22.8% had secondary education, 13.5% had university studies and 11% had none (Table 2).

Firstly the inter-rater reliability was identified using the CIC, coefficient that measures the degree of agreement between two or more observers. The results are displayed in Table 3, where there is a high inter-rater reliability (greater than 0.7) an average of 0.91434.

For the case of the SFT, the global Cronbach Alfa was 0.708 for the 6 dimensions, in turn, the Alfa was discovered with the 2-minute step

Table 3. Inter-rater reliability.

OBS	Total
OB2	0.92155556
OB3	0.96288889
OB4	0.822
OB5	0.96533333
OB6	0.90011111

Source: Own creation.

Table 4. Internal consistency of the SFT version in Spanish - Cronbach Alfa.

Cronbach Alfa	Number of elements
0.708	6
0.704	6*

*SFT with 2-minute step test.
Source: Own creation.

Table 5. Test re-test - CIC.

Test name	Measurements	Intraclass correlation[b]	Confidence interval 95%			F test with true value 0		
			Lower limit	Upper limit	Value	gl1	gl2	Sig.
Sitting and standing up from a chair test	Individual	0.760[a]	0.557	0.872	8.451	39	39	0.000
	average	0.864[c]	0.716	0.932	8.451	39	39	0.000
Elbow flex test	Individual	0.740[a]	0.561	0.854	6.664	39	39	0.000
	average	0.851[c]	0.719	0.921	6.664	39	39	0.000
Two-minute step test	Individual	0.923[a]	0.859	0.958	24.675	39	39	0.000
	average	0.960[c]	0.924	0.979	24.675	39	39	0.000
Six-minute walk test	Individual	0.862[a]	0.756	0.925	13.49	39	39	0.000
	average	0.926[c]	0.861	0.961	13.49	39	39	0.000
Sitting and touching feet using a chair test	Individual	0.860[a]	0.752	0.923	13.242	39	39	0.000
	average	0.925[c]	0.858	0.96	13.242	39	39	0.000
Joining hands behind the back test	Individual	0.860[a]	0.739	0.925	14.584	39	39	0.000
	average	0.925[c]	0.85	0.961	14.584	39	39	0.000
Standing, walking 8 feet, turning and sitting again test	Individual	0.896[a]	0.813	0.944	18.558	39	39	0.000
	average	0.945[c]	0.897	0.971	18.558	39	39	0.000

[a]: The estimator is the same, whether or not the interaction effect is present; b: Co-efficient of intraclass correlation type A, using an absolute agreement definition; c: This estimation is calculated assuming that the interaction effect is not present, as otherwise it cannot be estimated.
Source: Own creation.

test, obtaining 0.704, considered as acceptable, to ensure the internal consistency of the scale seen from its homogeneity upon assessing the same phenomenon-construct or theoretical dimension[23], as displayed in Table 4. The Cronbach Alfa was similar to any of the tests that measure aerobic capacity. In general, this measure indicates that the SFT is composed of tests that measure the same "functional capacity" attribute, seen in greatest location in the percentile of the subject in each of the best physical condition tests.

Likewise, the CIC is identified, revealing that the average measurements are around 1 in each of the tests and the minimum value is 0.851. In relation to the elbow flex test, all the measurements were obtained with CI of 95% and a $p<0.005$ as displayed in Table 5, remembering that values over 0.8 indicate suitable reliability. The average for the seven dimensions of the CIC is 0.913, thus demonstrating a significant correlation and reproducibility of the tests[23].

Discussion

Considering that the psychometric properties of the test are suitable as they display a higher Cronbach Alfa, and good reproducibility, (data that agrees with that found in different studies such as that by Ávila et al., in which a battery of physical capacities is applied to older adults in Quebec, and which includes SFT tests, such as standing, walking 2.54 metres and sitting back down. (This latter test was modified as it should be performed five times in the least time possible)); the first test applied reported a CIC = 0.99 and in the second a CIC of 0.67 to 0.73; and a CIC of 0.97 for 50 older adults, with an application interval of 48 hours. With relation to the estimated internal consistence with the Cronbach Alfa, 0.74 is obtained, which represents a satisfactory and suitable result for a measurement instrument with research purposes[24].

In turn, Boneth *et al.,* assess the test-re-test reproducibility and the agreement level between test measurements 30 sec (30-s) Arm

Curl and 30-s Chair Stand, for which two opportunities were applied by the same assessor with an interval of 4 to 8 days, determining the reproducibility via the CIC with a CI =95%, reporting for the elbow flex test in 30 seconds 0.88 and for the standing and sitting on the chair test in 30 seconds of 0.78, highlighting that these tests have a good level of agreement for assessing muscle strength of both the upper and lower body in functionally independent older adults[12].

Continuing, Ries *et al.*, in the assessment of the test-re-test reliability for the 6-minute walking test and standing, walking 2.5 metres, turning and sitting (Timed-Up-and-Go/TUG), in a population of older adults diagnosed with Alzheimer, stable and previously classified according to the degree of severity, they have high CIC; for TUG, CIC = 0.985-0.988 (p <0.001) and for the 6 minute walk CIC = 0.982-0.987 (p <0.001), checking that the reliability data is not influenced by the degree of dementia[25].

Next, Olivares *et al.* apply the SFT tests with the aim of associating them to quality of life and perspective of health, among other aspects. Their results regarding the joining hands behind the back test report a CIC = 0.96 in lower body flexibility, in the Sit and Reach test, a CIC = 0.95 and in the 6-minute walk, a CIC = 0.95[26]. Furthermore, Hilgenkamp et al. determine the reliability test of a battery of physical aptitude tests applied to older adults with different degrees of intellectual and mobility disability, in which they included the sitting and standing from the chair test and the sitting and touching the toes using a chair test, revealing a CIC between 0.63 and 0.96 for all the tests[27].

Finally, this SFT validation process established the surface validity, the internal consistency and the reproducibility, for their application in the Spanish version, which shows that it may be used in the practice given that it reports suitable validity and reliability. Likewise, it is reported that not considering the co-morbidities of older adults could be a limitation in the study, an aspect that could vary the test results. Therefore, these results open the possibility of performing new studies in which other psychometric properties of the test are addressed, and aspects such as the aforementioned are considered. However, these results are of interest to the prescription of exercise in older adults, as the battery tests are suitable for the language and for this population.

Acknowledgements

To the Department of Languages at the Boyaca University and the physical activity comprehensive centres: Indeportes, IRDET, Club Nueva Vida and Comfaboy.

References

1. Rikli R, Jones J. *Senior Fitness Test Manual.* California. Editorial Human Kinetics; 2001. p. 1-160.
2. Burger H, Marincek C. Functional testing of elderly subjects after lower limb amputation. *Prosthet Orthot Int.* 2001;25:102-7.
3. Miotto JM, Chodzko-Zajko WJ, Reich JL, Supler MM. Reliability and validity of the Fullerton Functional Fitness Test: An independet replication study. *J Aging Phys Activ.* 1999;7:339-53.
4. Osness WH, Adrian M, Hoeger W, Rabb D, Wiswell R. *Functional fitness assessment for adults over 60 years: A field Based Assessment.* Reston: AAHPERD; 1990. p. 1-24.
5. Rikli R, Jones J. The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. *J Aging Phys Activ.* 1988;6:363-75.
6. Dugas JC, White KN, Gunter KB. The effect of a novel ADL-based training program on performance of activities of daily living and physical fitness. *J Aging Phys Activ.* 2007; 15:13-25.
7. Pin TW. Psychometric Properties of 2-Minute Walk Test: A Systematic Review. *Arch Phys Med Rehabil.* 2014;95:1759-75.
8. Langhammer B Stanghelle JK. Functional fitness in elderly Norwegians measured with the Senior Fitness Test. *Rev Advances in Physiotherapy.* 2011; 13:137-44.
9. Podsiadlo D, Richardson S. The timed "up and go": A test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* 1991;39:142-48.
10. Rikli R, Jones J. *Senior Fitness Test.* Segunda ed. Champaign: Editorial Human Kinetics; 2013.p. 1-200.
11. Levy SS, Macera CA, Hootman JM, Coleman KJ, Lopez R, Nichols JF, *et al.* Evaluation of a multi-component group excercise program for adults with arthritis: Fitness and excercise for people with arthritis (FEPA). *Disabil Health J.* 2012;5:305-11.
12. Boneth M, Ariza CL, Angarita A, Parra J, Monsalve A, Gomez E. Reproducibilidad de las pruebas Arm Curl y Chair Stand para evaluar resistencia muscular en población adulta mayor. *Revista Ciencias de la Salud.* 2012;10:15-29.
13. Pedrero-Chamizo R, Gómez-Cabello A, Delgado S, Ramirez-Llarena S, Rodríguez-Marroyo JA, Cabanillas E, *et al*. Physical fitness levels among independent non-institutionalized spanish elderly: The elderly EXERNET multi-center study. *Arch Gerontol Geriatr*. 2012;55:406-16.
14. Ochoa ME, Cobo EA, Ruiz LY, Vargas DM, Sandoval C. Cross-cultural adaptation of the English version of the Senior Fitness Test to Spanish. *Rev Fac Med*. 2015; 62: 387-97.
15. Meneses J, Barrios M, Bonillo A, Cosculluela A, Lozano LM, Turbany J, Valero S. *Psicometría.* Barcelona. Editorial UOC; 2014.p. 1-273.
16. Swearinyes J, Brach J, Working Geriatric Assesment work; selecting useful measures. *Phys Ther.* 2014;81:1233-53.
17. Velasco Rodríguez VM, Martínez Ordaz VA, Roiz Hernández J, Huazano García F, Nieves Rentería A. Muestreo y Tamaño de Muestra. Una guía práctica para personal de salud que realiza investigación. Buenos Aires. Editorial: e-libro.net; 2003.p.57. (Consultado 13/11/2015). Disponible en: http://es.scribd.com/doc/239620151/Muestreo-y-Tamano-de-Muestra-2003#scribd.
18. Kottner J, Streiner DL. Internal consistency and cronbach´s alpha: A comment on Beckman *et al. Int J Nurs Stud.* 2010;47:926-28.
19. McCrae RR, Kurtz JE, Yamagata S, Terracciano A. Internal consistency, retest reliability and their implications for personality scale validity. *Pers Soc Psychol Rev.* 2011; 15:28-50.
20. Kottner J, Streiner DL. Internal consistency and Cronbach´s alpha: A comment on Beckman *et al. Int J Nurs Stud.* 2010;47:926-28.
21. Celina H, Campo Arias A. Aproximación al uso del coeficiente alpha de Cronbach. *Rev Colomb Psiquiatr.* 2005;34:572-80.
22. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehabil.* 2005;84:719-23.
23. Ramada JM, Serra-Pujadas C, Delclós-Clanchet GL. Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. *Salud Pública de Méx.* 2013;55:57-66.
24. Ávila-Funes JA, Gray-Donald K, Payette H. Medición de las capacidades físicas de adultos mayores de Quebec: un análisis secundario del estudio NuAge. *Rev Salud Pública Mex.* 2006;48:446-54.
25. Ries JD, Echternach JL, Nof L, Gagnon Blodgett M. Test-Retest Reliability and minimal detectable change scores for the timed "Up & Go" Test, the Six-Minute Walk Test, and Gait Speed in people with Alzheimer disease. *Rev Phys Ther.* 2009;89:569-79.
26. Olivares PR, Gusi N, Prieto J, Hernandez-Mocholi MA. Fitness and health-related quality of life dimensions in community-dwelling middle aged and older adults. *Health Qual Life outcomes.* 2011;9:117.
27. Hilgenkamp TI, Van Wijck R, Evenhuis HM. Feasibility and reliability of physical fitness tests in older adults with intellectual disability: a pilot study. *J Intellect Dev Disabil.* 2012;37:158-62.