Pérez-Vigo, C; Sanchez-Lastra M.A.; Martínez-de-Quel O.; Ayan, C. (2022) Reliability and Validity of the V-Sit-and-Reach and Toe-Touch Tests in Preschoolers. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 22 (88) pp. 969-984 Http://cdeporte.rediris.es/revista/revista88/artfiabilidad1418.htm
DOI: https://doi.org/10.15366/rimcafd2022.88.015

ORIGINAL

RELIABILITY AND VALIDITY OF THE V-SIT-AND-REACH AND TOE-TOUCH TESTS IN PRESCHOOLERS

FIABILIDAD Y VALIDEZ DE LAS PRUEBAS V-SIT-AND-REACH Y TOE-TOUCH EN PREESCOLARES

Pérez-Vigo, C[1]; Sanchez-Lastra M.A.[2]; Martínez-de-Quel O.[3]; Ayan, C.[4]

[1] Doctor. Departamento de Didácticas Especiales, Facultad de Ciencias de la Educación y el Deporte, Universidad de Vigo, Pontevedra (Spain) Criseducayprepara@hotmail.es

[2] Doctorando en Educación, Deporte y Salud. Facultad de Ciencias de la Educación y el Deporte, Universidad de Vigo, Pontevedra (Spain) misanchez@uvigo.es

[3] Departamento de Didáctica de las Lenguas, Artes y Educación Física. Facultad de Educación, Universidad Complutense de Madrid (Spain) odequel@edu.ucm.es

[4] Profesor titular en el Departamento de Didácticas Especiales, Facultad de Ciencias de la Educación y el Deporte, Universidad de Vigo, Pontevedra. Well-Move Research Group. IIS Galicia Sur (Spain) cayan@uvigo.es

Spanish-English translator: Ricardo Collazo Carballa, ricardo@bonza.es

Código UNESCO / UNESCO code: 3212 Salud Pública / Public Health
Clasificación Consejo de Europa / Council of Europe classification: 17. Otras: Actividad Física y Salud / Others: Physical Activity and Health

Recibido 20 de julio de 2020 **Received** July 20, 2020
Aceptado 9 de abril de 2021 **Accepted** April 9, 2021

ABSTRACT

This psychometric study aimed at assessing the test-retest reliability and criterion-related validity of the V-Sit and Reach (VS) and Toe-Touch (TT) tests when performed by preschoolers. Out of an initial sample of 158 participants 132 performed the tests on three occasions with a week interval. Construct validity was assessed by means of the passive straight leg raise (PSLR) test. Both tests showed excellent reliability (ICC=0.91-0.93 for the VS; ICC=0.93-0.97 for the TT). Statistically significant correlations ($p<0.05$) were observed between both tests and the PSLR test for all of the groups that were analyzed, except for the TT in the case of three-year old group. In conclusion, these findings indicate that both the VS and the TT are reliable tests that show

moderate criterion-related validity for assessing lower-body flexibility in preschoolers. The VS may be a preferable test over the TT.

KEY WORDS: Psychometric properties; Flexibility; Children; Assessment

RESUMEN

Este estudio de carácter psicométrico tuvo como objetivo evaluar la fiabilidad y validez de las pruebas V-Sit and Reach (VS) y Toe –Touch (TT) en población preescolar. De un total inicial de 158 participantes, 132 realizaron las pruebas en tres ocasiones. La prueba de elevación pasiva de la pierna recta (PSLR) se usó como indicador de validez. Ambas pruebas mostraron una fiabilidad excelente (CCI=0,91-0,93 para el VS; CCI=0,93-0,97 para el TT). Se observaron correlaciones estadísticamente significativas (p<0,05) entre ambas herramientas y la prueba de PSLR para todos los grupos analizados, excepto en el TT en el grupo de 3 años de edad. En conclusión, estos hallazgos indican que tanto el VS como el TT son pruebas con una validez moderada para evaluar la flexibilidad de la parte inferior del cuerpo en preescolares. La prueba VS puede ser preferible sobre la prueba TT.

PALABRAS CLAVE: Propiedades psicométricas; Flexibilidad; Niños; Evaluación.

INTRODUCTION

Flexibility is defined as the intrinsic property of body tissues, including muscles and connective tissues, that determines the range of motion that can be achieved without injury in a joint or group of joints (Pate et al., 2012). Significant loss of flexibility is considered an aggravating factor in health and quality of life, aspects that become especially evident during aging (Matos-Duarte & Berlanga, 2020). Therefore, this physical capacity is considered as one of the components of health-related physical fitness (Delgado-Marín et al., 2020), even from preschool age, being inversely related to pain (especially of the lumbar area of the back) and to the risk of injuries (Pate et al., 2012). In fact, in childhood age, it has been observed that reduced levels of hamstring flexibility negatively affect motor skills (Lopes et al., 2017) and posture adopted in activities of daily living (Coelho et al., 2014). Therefore, it has been suggested that its assessment is important in this population (Ayán Pérez et al., 2018; Sánchez-Lastra et al., 2020). Thus, being able to detect whether there is reduced flexibility of the lower extremities could be considered a preventive health strategy in the Early Childhood Education stage.

At the preschool stage, field tests are recommended to assess health-related components of physical fitness, such as flexibility. These are easy to use, the equipment and cost is minimal, and a large number of participants can be assessed in a relatively short period of time. However, before a field test can be considered a useful measurement tool, its psychometric properties, especially levels of reliability and validity, must first be analyzed.

There are two tests that are commonly employed to assess lower extremity flexibility and are easy to use in the school setting: the V sit-and-reach (VS) and the Toe-Touch (TT). The VS was proposed by the North American association "President's Challenge", as a test to measure the flexibility of the lower back and hamstring muscles (Pate et al., 2012). It is a modified version of the "Sit-and-Reach," the advantage of which is that a box is not needed for its development, required in the original test (Cooper Institute for Aerobics Research, 1992). The TT, originally known as the "Fingertip-to-floor" was designed as part of a battery of tests aimed at muscle assessment (Kraus & Eisenmenger-Weber, 1945) and has been used to assess hamstring extensibility (Kippers & Parker, 1987).

The psychometric properties of the VS and TT have been previously analyzed, albeit in young adults and patients with low back pain (Ayala et al., 2012; Mayorga-Vega, Merino-Marban, et al., 2014; Mayorga-Vega, Viciana, et al., 2014). However, research on the psychometric properties of flexibility tests when performed by preschoolers are scarce. Thus, after reviewing the literature, it seems that only one study has been conducted in this regard, focused on the Sit-and-reach test (Ayán Pérez et al., 2018). Therefore, its psychometric properties in the preschool context are unknown, showing the need to expand scientific knowledge in this regard.

AIM

The aim of this research was to evaluate the reliability and validity of the VS and TT tests when performed by preschool children.

MATHERIALS AND METHODS

DESIGN

This research had a psychometric design in which two flexibility tests were performed twice, after a previous familiarization period, with an interval of one week, in order to identify their test-retest reliability. To determine their validity, their results were contrasted with those obtained in a test considered as Gold Standard. The analysis of the psychometric properties of the tests was carried out with data registered during the third school term of the 18/19 academic year and the second term of the 19/20 academic year.

SUBJECTS

The participants were healthy Spanish children who were recruited from three kindergartens in urban areas of northern Spain. The inclusion criteria for participation were: a) being between 3 and 5 years old and b) having no medical condition that could hinder performance of the field-based tests proposed in the research. Children who showed intellectual or physical disabilities that prevented them from understanding the test protocol or performing the tests correctly, were excluded from the research. Written

informed consent was received from the kindergarten principal and parents or guardians of all the children who took part in the research study. A total of 158 participants (49.3% girls), volunteered for the investigation. The study design was approved by the Ethics Committee of the Faculty of Education and Sports Science (University of Vigo)

MEASUREMENTS

Antropometry. Weight (kg) and height (cm) were measured without shoes and with light clothing. Each child's body mass index (BMI) was calculated using the following formula: body mass/height2 (kg/m2).

V-Sit and Reach. For the VS, the children were instructed to sit on the floor with their knees straight and feet separated by about 12 inches to form a V-shape leg position. A ruler was placed between the legs with the 9-inch mark located at the heel line. The children, palms downs and placing one hand on top of the other, were told to reach forward slowly as far as possible sliding their hands along the ruler and to hold the position for two seconds (Hui, Yuen, Morrow, & Jackson, 1999).

Toe-Touch. The TT was performed with a "sit-and-reach" box (height 12.5 inches) with a 6-inch mark representing the point at which the children's fingertips were in line with their toes. The 6-inch mark was chosen instead of the 0-inches mark to avoid negative scores, which have been observed in pediatric populations performing this test (Kippers & Parker, 1987). Thus, following these procedures, the scoring allowed for test scores to be reported as a positive number for both tests (Ayala, Sainz de Baranda, De Ste Croix, & Santonja, 2012). The children were placed in the standing position on the box, with their knees extended and feet hip-width apart. They were instructed to bend forward as far as possible, while maintaining their knees, arms, and fingers fully extended and to hold the position for approximately five seconds (Muyor, Vaquero-Cristóbal, Alacid, & López-Miñarro, 2014).

For both the VS and TT tests, the final position that the children reached was considered to be the test score and the furthest point reached with the finger tips was recorded (to the nearest tenth of an inch). The best of two trials for each test was retained for analysis. One tester supervised the correct execution of both tests, while other evaluator registered the scores obtained (Ayán Pérez et al., 2018).

Passive straight leg raise test. The passive straight leg raise test (PSLR), was performed while the children were in a supine position on a mat (Hui et al., 1999). The axis of a digital goniometer was aligned with the axis of the hip joint, while the stationary arm was placed in line with the trunk, with the movable arm positioned in line with the femur. The children's preferred leg was passively moved into hip flexion until tightness was felt by the children or the tester, who assisted with moving the leg through flexion and made sure that the knee remained straight during the leg raise, while the other evaluator recorded the

maximum angle (degree) read from the goniometer at the point of maximum hip flexion.

PROCEDURES

To identify the test-retest reliability of the VS and the TT, both tests were performed on three occasions, following previous procedures in this regard (Ayán Pérez et al., 2018). All tests were carried out on groups of 15 children during the daily break time in the kindergarten's gymnasium. The evaluation sessions took place twice a week, on alternate days, over a period of four weeks. During the first week anthropometric measurements were carried out, and the protocol of the tests was explained. The children were encouraged to perform the tests after observing their execution by their teacher and they were also allowed to perform several attempts in order to become familiar with them and avoid a learning effect. During the following week, the VS (first session) and the TT (second session) were performed. This period was considered as the "test" phase. The third week was devoted to carry out the PSLR. The obtained results were used to identify the validity of the VS and the TT, as this test is considered the "Gold Standard" in this type of research.

The sample size calculation showed that a minimum of 108 participants were necessary if a statistical power level of 0.90 was desired, with an anticipated effect size (Cohen's d) of 0.4 and an alpha level (p-value) of 0.05.

STATISTICAL ANALYSIS

Mean and standard deviation were calculated in order to describe the children's characteristics. A Student's t test for independent samples was used to determine the differences between sex groups, in conjunction with an evaluation of normality using the Kolmogorov-Smirnov and Shapiro-Wilk tests and homoscedasticity through the Levenne test (Atkinson & Nevill, 1998). Effect sizes of sex differences were estimated by Cohen's d, interpreted following the classic values (Lakens, 2013): 0.2 small effect, 0.5 medium effect and 0.8 large effect. Comparison between age groups was performed by one-way ANOVA and Games-Howell post hoc.

The intraclass correlation coefficient (ICC) between test and re-test including the Cohen's d effect size of the test-retest differences, was used to determine relative reliability. Results of ICC were interpreted using these ranges: 0.90-1 excellent, 0.80-0.89 good, 0.70-0.79 fair and <0.69 poor reliability. Standard Error Measurement (SEM) and Minimal Detectable Change (MDC) were used to evaluate absolute reliability (Shrout & Fleiss, 1979). They were calculated as follows: SEM = [SD·√(1-ICC)] and MDC = [1.96·SEM·√2(1-ICC)] where SD represents the standard deviation of the test-retest differences and 1.96 is the z-score corresponding to 95% CI. SEM and MDC were also expressed as a percentage [SEM% = (SEM/meantest-retest)·100] and [MDC% = (MDC/meantest-retest)·100]. These values were interpreted as follows: MDC% > 30% was considered poor, from 10% to 30% was considered acceptable, and <10% was considered excellent (Ayán Pérez et al., 2018). SEM% values below

10% were interpreted as excellent absolute reliability (Flansbjer, Holmbäck, Downham, Patten, & Lexell, 2005). Lastly, to evaluate test-retest variability and possible systematic bias, Bland-Altman plots were used, showing test-retest differences against test-retest mean and representing the 95% limits of agreement (LOA).

Criterion-related validity was assessed by means of the Pearson's Correlation Coefficient (PCC), which was interpreted as follows: >0.8 high degree of correlation, 0.79-0.60 moderate degree of correlation, 0.59-0.30 fair degree of correlation, and <0.29 poor degree of correlation (Chan, 2003). All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS Inc. Version 25.0, Chicago, IL) and Microsoft Excel 2011.

RESULTS

Twenty-six of participants initially recruited, did not complete all the assessment tests, since they had not attended class at one of the measurement sessions. Therefore, the final study sample consisted of 132 preschoolers (68 boys and 64 girls, with a mean age of 4.76±0.85 years), who completed all the measurements and provided valid data for the analysis (Table 1). No difficulties were observed in the children's understanding or execution of the tests, and no adverse effects were observed as a result of their performance.

Table 1. Descriptive values (mean ± SD) and comparison between sex groups (Student t-test and Cohen's d effect size).

Variables	All (n =132)				Boys (n =68)		Girls (n =64)		*t* de Student, p	Cohen's *d*
	Mean	SD	Max.	Min.	Mean	SD	Mean	SD		
Age (years)	4.76	0.85	2.89	5.87	4.76	0.81	4.75	0.89	0.952	0.011
Body mass (kg)	18.66	3.19	12.80	36.30	18.73	2.75	18.50	3.57	0.681	0.071
Height (cm)	106.36	6.68	92	121	107.32	6.86	105.20	6.31	0.068	0.316
BMI ($kg*m^{-2}$)	16.43	1.73	13.66	25.63	16.21	1.28	16.65	2.10	0.147	-0.253
VS test (cm)	22.24	6.67	5.00	39.00	19.95	6.31	24.59	6.24†	0.000	-0.695
VS retest (cm)	22.65	7.35	4.00	42.00	19.90	7.00	25.48	6.66†	0.000	-0.760
TT test (cm)	16.15	6.19	1.00	30.00	14.23	5.78	18.09	5.78†	0.000	-0.625
TT retest (cm)	16.60	6.59	1.00	31.00	14.99	5.99	18.16	6.02†	0.005	-0.482
Goniometry (PSLRT) (º)	91.85	13.26	61.50	122.00	87.88	13.44	95.97	11.88†	0.000	-0.611

*BMI = body mass index; VS = V sit and reach; TT = Toe Touch. †Significantly greater values than the other sex group ($p < 0.05$).

Statistically significant differences in flexibility levels were found in the girl subgroup. The scores obtained in the VS and the PSLRT indicated an inverse relationship between age and flexibility, a trend that not was observed when the TT was used as reference (table 2).

Table 2. Descriptive values (mean ± SD) and comparison between age groups (ANOVA and Games-Howell post hoc).

	3 years old (*n* =26)	4 years old (*n* =38)	5 years old (*n* =68)	ANOVA		Cohen's d effect size		
Variables	M ± SD	M ± SD	M ± SD	F	*p*	3-4 years	4-5 years	3-5 years
Age (years)	3.43 ± 0.29†‡	4.37 ± 0.24‡	5.49 ± 0.24	841.92	0.000	3.57	4.67	7.84
Body mass (kg)	17.00 ± 2.27‡	18.07 ± 2.39‡	19.61 ± 3.55	8.00	0.001	0.46	0.52	0.90
Height (cm)	99.21 ± 4.94†‡	105.80 ± 5.07‡	109.37 ± 5.90	32.50	0.000	1.32	0.65	1.87
BMI (kg*m^{-2})	17.25 ± 1.75†	16.09 ± 1.30	16.31 ± 1.86	4.00	0.021	-0.76	0.14	-0.52
VS test (cm)	25.96 ± 5.76‡	24.01 ± 5.02‡	19.87 ± 6.90	11.28	0.000	-0.36	-0.69	-0.96
VS retest (cm)	26.58 ± 5.99‡	24.27 ± 6.20‡	20.27 ± 7.59	9.31	0.000	-0.38	-0.58	-0.93
TT test (cm)	0.17 ± 6.02	3.05 ± 7.19	0.48 ± 5.47	2.59	0.079	0.44	-0.41	0.05
TT retest (cm)	0.67 ± 6.70	3.46 ± 7.77	0.92 ± 5.69	2.17	0.118	0.39	-0.38	0.04
Goniometry (PSLRT) (º)	96.63 ± 12.77‡	96.14 ± 15.16‡	87.68 ± 10.93	7.82	0.001	-0.04	-0.65	-0.76

* BMI = body mass index; VS = V sit and reach; TT = Toe Touch. †Significant differences between age groups 3 and 4 years old ($p < 0.05$). ‡Significant differences between age groups 3 and 5 years or 4 and 5 years old ($p < 0.05$).

The VS showed an excellent reliability (ICC = 0.91-0.93), except for the five-year old group (ICC = 0.88), while the reliability of the TT was excellent in all cases (ICC = 0.93-0.97). Absolute reliability was excellent, with SEM and MDC percentage values ranging from 2.7 to 5.1% and from 1.9 to 4.9% for the VS and from 2.4 to 5.1% and from 1.1 to 3.5% for the TT (Table 3).

Table 3. Reliability of V-Sit & Reach and Toe-Touch test, according to sex and age groups.

		95% CI of ICC		Test-retest differences	t Student			SEM		MDC	
	ICC	Lower	Higher	M ± SD	t	p	Effect size	score	%	score	%
all (n = 132)											
VS	0.936	0.911	0.955	0.41 ± 3.42	-1.371	0.173	0.058	0.862	3.842	0.603	2.684
TT	0.951	0.930	0.965	0.44 ± 2.75	-1.862	0.065	0.066	0.610	3.723	0.375	2.290
boys (n = 68)											
VS	0.934	0.893	0.959	-0.05 ± 3.34	0.124	0.902	-0.008	0.857	4.300	0.610	3.060
TT	0.948	0.913	0.968	0.76 ± 2.77	-2.258	0.027	0.129	0.632	4.330	0.401	2.744
girls (n = 64)											
VS	0.917	0.862	0.950	0.90 ± 3.49	-2.058	0.044	0.139	1.005	4.013	0.801	3.198
TT	0.945	0.910	0.967	0.07 ± 2.71	-0.208	0.836	0.012	0.633	3.491	0.410	2.263
3 years old (n = 26)											
VS	0.937	0.860	0.971	0.62 ± 2.86	-1.111	0.277	0.106	0.720	2.740	0.502	1.912
TT	0.937	0.860	0.971	0.50 ± 3.13	-0.815	0.423	0.079	0.787	5.103	0.549	3.559
4 years old (n = 38)											
VS	0.883	0.776	0.939	0.27 ± 3.68	-0.446	0.658	0.047	1.255	5.199	1.188	4.920
TT	0.970	0.943	0.984	0.41 ± 2.56	-0.983	0.332	0.055	0.443	2.426	0.213	1.165
5 years old (n = 68)											
VS	0.938	0.900	0.961	0.40 ± 3.51	-0.953	0.344	0.056	0.875	4.361	0.605	3.013
TT	0.935	0.896	0.960	0.44 ± 2.74	-1.341	0.184	0.079	0.697	4.437	0.491	3.129

Data regarding the reliability of both the VS and TT (test-retest) are shown in Figure 1.

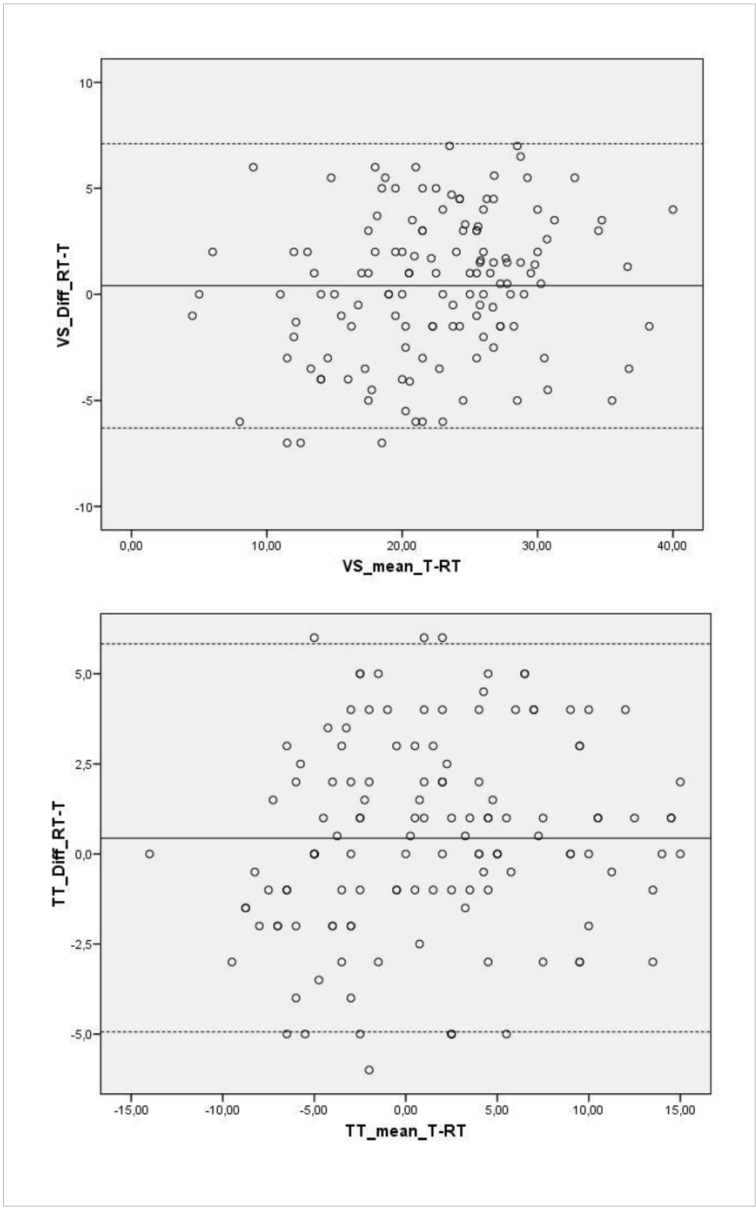

Figure 1. Bland-Altman plots for VS and TT. Continuous line represents the mean value with the upper and lower discontinuous lines representing the limits of agreement (1.96 SD).

Table 4 shows the data on the criterion-related validity of the VS and TT tests.
4 se muestran los datos relativos a la validez de las pruebas VS y TT.

Table 4. Degree of association between goniometry (passive straight leg raise test) and VS and TT by sex and age.

PCC goniometry	VS				TT			
	Test		Retest		Test		Retest	
Grupos	PCC	*p*	PCC	*p*	PCC	*p*	PCC	*p*
All (n = 132)	r = 0.620†	0.000	r = 0.647†	0.000	r = 0.618†	0.000	r = 0.609†	0.000
Boys (n = 68)	r = 0.581†	0.000	r = 0.628†	0.000	r = 0.617†	0.000	r = 0.619†	0.000
Girls (n = 64)	r = 0.568†	0.000	r = 0.571†	0.000	r = 0.529†	0.000	r = 0.526†	0.000
3 years old (n = 26)	r = 0.538†	0.005	r = 0.398†	0.044	r = 0.372	0.061	r = 0.320	0.111
4 years old (n = 38)	r = 0.572†	0.000	r = 0.618†	0.000	r = 0.727†	0.000	r = 0.727†	0.000
5 years old (n = 68)	r = 0.624†	0.000	r = 0.705†	0.000	r = 0.655†	0.000	r = 0.661†	0.000

PCC= Pearson's correlation coefficient †= significant PCC value ($p < 0.05$)

Statistically significant correlations were observed between both tests and the PSLR test for all of the groups that were analyzed, except for the TT in the case of three-year old group. The correlation was moderate for the whole sample. Associations between the VS and the PSLR test were fair for the three-year old group (r = 0.398-0.538). The VS and TT showed a moderate validity for the four-year (r = 0.572-0.727) and five-year old groups (r = 0.624-0.705).

DISCUSSION

There is a lack of scientific evidence regarding the psychometric properties of flexibility field-based tests when performed by preschoolers. This study aimed to bridge this gap by providing information on the absolute and relative reliability, as well as criterion-related validity of two widely used tests.

Evidence indicates that flexibility levels are influenced by sex and age, even at younger ages (Ayán Pérez et al., 2018). The results obtained from both tests confirmed that girls were more flexible than boys, and the VS scores indicated an inverse relationship between age and flexibility. However, this trend was not observed when the TT scores were analyzed. These findings could imply that the VS is a more sensitive test for measuring flexibility at an early age.

Both the VS and the TT showed excellent relative reliability. Although there is no previous scientific evidence on reliability values for both tests in preschoolers, these data can be compared with those reported for adult population. In the case of the VS, reliability values have generally been good or excellent in adults (ICC = 0.86-0.99) (Ayala & de Baranda, 2012). On the TT, Vrbik reported an internal consistency ranging between 0.96 and 0.98, when taken by children in Primary Education (Vrbik et al., 2017). The relative reliability found in this study for both tests was slightly higher than that demonstrated by the original Sit-and-Reach in preschoolers (ICC = 0.74-0.94) (Ayán Pérez et al., 2018; Ortega et al., 2015).

It should be noted that, from a statistical point of view, the exclusive use of the ICC to assess reliability can be considered insufficiently accurate, as it is affected by the heterogeneity of the sample and does not indicate the presence of a possible systematic bias (Hopkins, 2000). Therefore, the amount of random error that occurs when a single individual is tested multiple times (SEM), as well as the minimum amount of change needed to overcome measurement error (MDC), were included in the statistical analysis as an additional strategy to identify the absolute reliability of both tests. In this regard, it was observed that changes of around 2.68% (about 0.60 cm) for the VS and 2.29% (about 0.38 cm) for the TT would indicate the existence of changes of considerable magnitude in ischiosural extensibility between both measurements (test-retest). The results obtained point to the VS and TT presenting excellent absolute reliability, a result that has been previously observed for the Sit-and-Reach when performed by preschool children with an adapted drawer (Ayán Pérez et al., 2018).

Several studies have analyzed the validity of the VS (Mayorga-Vega, Merino-Marban, et al., 2014) and the TT (Mayorga-Vega, Viciana, et al., 2014) using the PSLR test as a measure of reliability, reporting moderate to acceptable degrees of correlations (VS, r = 0.44-0.65; TT, r = 0.54-0.79). In this study, both tests showed moderate correlations, which were even higher than those found in the Sit-and-Reach in preschool population (r=0.352-0.51)2. These findings are in agreement with those of Hui et al. (1999), who found that the VS shows higher levels of validity than the Sit-and-Reach in young adults. In contrast, other authors have found that the Sit-and-Reach showed higher validity than the VS in a similar population (López-Miñarro et al., 2008). Additionally, and in contrast to these ideas, Ayala and de Baranda (2012) reported that the Sit-and-Reach has a higher validity coefficient than the TT in active young adults.

In the present investigation, despite finding moderate validity for the TT, it should be noted that this test did not show a significant association with the PSLR test in the 3-year-old group, indicating a lack of criterion validity in this group. In this regard, it has been observed that toe-to-ground distance may be influenced by anthropometric characteristics such as arm size or spinal range of motion (Carregaro et al., 2007). Based on this, it could be hypothesized that, by modifying the drawer dimensions, it would be possible for shorter children to achieve better results, as has been observed for the Sit-and-Reach (Ayán Pérez et al., 2018). In this regard, further research is needed.

The results of this study provide two data of particular interest. First, they indicate that the absolute and relative reliability of the VV and TT tests is high for children aged 3 to 5 years. This is relevant not only from a psychometric point of view, but also because it provides information on the level of physical fitness in 3-year-old children, an aspect that has been relatively little studied in the literature (Arufe-Giráldez, 2020). Secondly, after analyzing the degree of association between the VS and the TT with the PSLR test, it appears that both tests demonstrate moderate criterion validity, except in the case of the TT when performed by 3-year-old children. Considering all this and taking into account that the performance of the VS is simpler, it might be advisable to perform this test to estimate hamstring extensibility in preschoolers.

Despite the originality of this study, there are a number of limitations that deserve to be highlighted. First, the TT and VS are indirect measures of flexibility, since they involve movement of the whole body. This aspect favors that their validity may be influenced by anthropometric factors that were not evaluated. Because of this, the use of another methodological approach to estimate the relative validity criteria, such as measuring hip range of motion by means of a radiograph or using an inclinometer, would have allowed a better identification of this psychometric property. Second, it should be noted that the subgroup analysis was limited by the small sample size.

CONCLUSIONS

The results of this research indicate that the VS and the TT are reliable lower extremity flexibility assessment tests for use in preschool education, showing moderate validity. The VS may be a preferable test over the TT, considering the

inverse relationship between its results and the age of the participants, as well as the moderate validity presented for all age groups.

REFERENCES

Arufe-Giráldez, V. (2020). ¿Cómo debe ser el trabajo de Educación Física en Educación Infantil? (¿How should Physical Education work in Early Childhood Education be?). *Retos*, (37), 588-596. https://doi.org/10.47197/retos.v37i37.74177

Atkinson, G., & Nevill, A. M. (1998). Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sports Medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Ayala, F., & de Baranda, P. (2012). Reliability and validity of sit-and-reach tests, systematic review. *Revista Andaluza de Medicina Del Deporte*, *5*, 57–66. https://doi.org/10.1016/S1888-7546(12)70010-2

Ayala, F., Sainz de Baranda, P., De Ste Croix, M., & Santonja, F. (2012). Reproducibility and criterion-related validity of the sit and reach test and toe touch test for estimating hamstring flexibility in recreationally active young adults. *Physical Therapy in Sport*, *13*(4), 219–226. https://doi.org/10.1016/j.ptsp.2011.11.001

Ayán Pérez, C., Álvarez Pérez, S., González Baamonde, S., & Martínez de Quel, Ó. (2018). Influence of the Box Dimensions on the Reliability and Validity of the Sit and Reach in Preschoolers. *Journal of Strength and Conditioning Research*. https://doi.org/10.1519/JSC.0000000000002886

Carregaro, R., Silva, L., & Gil Coury, H. (2007). Comparação entre dois testes clínicos para avaliar a flexibilidade dos músculos posteriores da coxa. *Revista Brasileira de Fisioterapia*, *11*(2), 139–145. https://doi.org/10.1590/S1413-35552007000200009

Chan, Y. H. (2003). Biostatistics 104: correlational analysis. *Singapore Medical Journal*, *44*(12), 614–619. http://www.ncbi.nlm.nih.gov/pubmed/14770254

Coelho, J. J., Graciosa, M. D., de Medeiros, D. L., da Silva Pacheco, S. C., Resende da Costa, L. M., & Kittel Ries, L. G. (2014). Influência da flexibilidade e sexo na postura de escolares. *Revista Paulista de Pediatria*, *32*(3), 223–228. https://doi.org/10.1590/0103-0582201432312

Cooper Institute for Aerobics Research. (1992). *The Prudential FITNESSGRAM test administration manual.* The Cooper Institute for Aerobics Research.

Delgado-Martín, J. L., Garoz-Puerta, L., Miguel-Tobal, F., & Martínez-de-Haro, V. (2020). Antropometría y fuerza, su influencia sobre el test sit and reach. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte, 20(78),369-380. https://doi.org/10.15366/rimcafd2020.78.012

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine*,*30*(1),1–15. https://doi.org/10.2165/00007256-200030010-00001

Hui, S. C., Yuen, P. Y., Morrow, J. R., & Jackson, A. W. (1999). Comparison of the Criterion-Related Validity of Sit-and-Reach Tests with and without Limb Length Adjustment in Asian Adults. *Research Quarterly for Exercise and Sport*, *70*(4), 401–406. https://doi.org/10.1080/02701367.1999.10608061

Kippers, V., & Parker, A. W. (1987). Toe-Touch Test. *Physical Therapy*, *67*(11), 1680–1684. https://doi.org/10.1093/ptj/67.11.1680

Kraus, H., & Eisenmenger-Weber, S. (1945). Evaluation of Posture Based on Structural and Functional Measurements. *Physical Therapy*, *25*(6), 267–271. https://doi.org/10.1093/ptj/25.6.267

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lexell, J., Flansbjer, U.-B. B., Holmbäck, A. M., Downham, D., Patten, C., & Lexell, J. (2005). RELIABILITY OF GAIT PERFORMANCE TESTS IN MEN AND WOMEN WITH HEMIPARESIS AFTER STROKE. *Journal of Rehabilitation Medicine*, *37*(2), 75–82. https://doi.org/10.1080/16501970410017215

Lopes, L., Póvoas, S., Mota, J., Okely, A. D., Coelho-e-Silva, M. J., Cliff, D. P., Lopes, V. P., & Santos, R. (2017). Flexibility is associated with motor competence in schoolchildren. *Scandinavian Journal of Medicine & Science in Sports*, *27*(12), 1806–1813. https://doi.org/10.1111/sms.12789

López-Miñarro, P., Sáinz de Baranda, P., Rodríguez-García, P., & Yuste, J. (2008). Comparison between sit-and-reach test and V sit-and-reach test in young adults. *Gazzetta Medica Italiana*, *167*(4), 135–142.

Matos-Duarte, M., & Berlanga, L. A. (2020). Efectos del ejercicio sobre la flexibilidad en personas mayores de 65 años. *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, *20*(80), 611-622. https://doi.org/10.15366/rimcafd2020.80.010

Mayorga-Vega, D., Merino-Marban, R., & Viciana, J. (2014). Criterion-Related Validity of Sit-and-Reach Tests for Estimating Hamstring and Lumbar Extensibility: a Meta-Analysis. *Journal of Sports Science & Medicine*, *13*(1), 1–14. https://doi.org/10.4100/jhse.2014.91.18

Mayorga-Vega, D., Viciana, J., Cocca, A., & Merino-Marban, R. (2014). Criterion-related validity of toe-touch test for estimating hamstring extensibility: A meta-analysis. *Journal of Human Sport and Exercise*, *9*(1), 188–200. https://doi.org/10.4100/jhse.2014.91.18

Muyor, J. M., Vaquero-Cristóbal, R., Alacid, F., & López-Miñarro, P. A. (2014). Criterion-Related Validity of Sit-and-Reach and Toe-Touch Tests as a Measure of Hamstring Extensibility in Athletes. *Journal of Strength and Conditioning Research*, *28*(2), 546–555. https://doi.org/10.1519/JSC.0b013e31829b54fb

Ortega, F. B., Cadenas-Sánchez, C., Sánchez-Delgado, G., Mora-González, J., Martínez-Téllez, B., Artero, E. G., Castro-Piñero, J., Labayen, I., Chillón, P., Löf, M., & Ruiz, J. R. (2015). Systematic Review and Proposal of a Field-Based Physical Fitness-Test Battery in Preschool Children: The PREFIT Battery. In *Sports Medicine* (Vol. 45, Issue 4, pp. 533–555). Springer International Publishing. https://doi.org/10.1007/s40279-014-0281-8

Pate, R., Oria, M., Pillsbury, L., Youth, C. on F. M. and H. O. in, Board, F. and N., & Medicine, I. of. (2012). Fitness Measures and Health Outcomes in Youth. In *Fitness Measures and Health Outcomes in Youth*. National Academies Press (US). https://doi.org/10.17226/13483

Sánchez-Lastra, M. A., Martínez-Lemos, R. I., Díaz, R., Villanueva, M., & Ayán, C. (2020). Efecto de un programa de natación en la condición física de preescolares (Effect of a swimming program on physical condition of preschoolers). *Retos*, (37), 48-53. https://doi.org/10.47197/retos.v37i37.69504

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

https://doi.org/10.1037/0033-2909.86.2.420

Vrbik, I., Sporiš, G., Štefan, L., Madić, D., Trajković, N., Valantine, I., & Milanović, Z. (2017). The Influence of Familiarization on Physical Fitness Test Results in Primary School-Aged Children. *Pediatric Exercise Science*, *29*(2), 278–284. https://doi.org/10.1123/pes.2016-0091

Número de citas totales / Total references: 27 (100%)
Número de citas propias de la revista / Journal's own references: 2 (7,4%)

Rev.int.med.cienc.act.fís.deporte - vol. 22 - número 88 - ISSN: 1577-0354