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

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

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

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

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

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

[3] Departamento de Didáctica de las Lenguas, Artes y Educación Física. Facultad de Educación, Universidad Complutense de Madrid (España) 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 (España) cayan@uvigo.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

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.

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

INTRODUCCIÓN

La flexibilidad se define como la propiedad intrínseca de los tejidos corporales, incluidos los músculos y los tejidos conectivos, que determina la amplitud de movimiento que se puede lograr sin lesiones en una articulación o grupo de articulaciones (Pate et al., 2012). La pérdida significativa de la flexibilidad se considera un agravante en la salud y la calidad de vida, aspectos que se ponen especialmente de manifiesto durante el envejecimiento (Matos-Duarte & Berlanga, 2020). Por ello, esta capacidad física es considerada como uno de los componentes de la condición física relacionada con la salud (Delgado-Marín et al., 2020), incluso desde la edad preescolar, relacionándose inversamente con el dolor (especialmente de la zona lumbar de la espalda) y con el riesgo de lesión (Pate et al., 2012). De hecho, en la edad infantil se ha observado que niveles reducidos de flexibilidad en los isquiotibiales afectan negativamente a las habilidades motoras (Lopes et al., 2017) y a la postura adoptada en las actividades de la vida diaria (Coelho et al., 2014). Por ello, se ha sugerido que su evaluación es importante en esta población (Ayán Pérez et al., 2018; Sánchez-Lastra et al., 2020). De esta manera, el poder detectar si existe una flexibilidad reducida de las extremidades inferiores podría considerarse una estrategia de salud preventiva en la etapa de Educación Infantil.

En la etapa de preescolar, se recomiendan la realización de pruebas de campo para evaluar los componentes de la condición física relacionados con la salud, como la flexibilidad. Estas son fáciles de utilizar, el equipamiento y el coste es mínimo y un gran número de participantes puede ser evaluado en un periodo de tiempo relativamente corto. Sin embargo, antes de que una prueba de campo pueda ser considerada como una herramienta de medición útil, sus propiedades

psicométricas, especialmente los niveles de fiabilidad y validez, deben ser previamente demostrados.

Existen dos pruebas que son habitualmente empleadas para valorar la flexibilidad de las extremidades inferiores y que son fáciles de utilizar en el entorno escolar; el "V sit-and-reach" (VS) y el "Toe-Touch" (TT). El VS fue propuesto por la asociación norteamericana "Presidente' Challenge", como una prueba para mediar la flexibilidad de los músculos de la parte baja de la espalda y de los isquiotibiales (Pate et al., 2012). Es una versión modificada del "Sit-and-Reach," cuya ventaja radica en que no se necesita un cajón para su desarrollo, requerido en la prueba original (Cooper Institute for Aerobics Research, 1992). El TT, originalmente conocido como "Fingertip-to-floor" , fue diseñado como parte de una batería de pruebas dirigidas a la valoración muscular (Kraus & Eisenmenger-Weber, 1945) y ha sido utilizado para evaluar la extensibilidad de los músculos isquiotibiales (Kippers & Parker, 1987).

Las propiedades psicométricas del VS y del TT han sido analizadas previamente, si bien en adultos jóvenes y pacientes con dolor lumbar (Ayala et al., 2012; Mayorga-Vega, Merino-Marban, et al., 2014; Mayorga-Vega, Viciana, et al., 2014). Sin embargo, la investigación sobre las propiedades psicométricas de las pruebas de flexibilidad cuando son realizadas por preescolares son escasas. Así, tras revisar la literatura, parece que únicamente se ha realizado un estudio a este respecto, centrado sobre la prueba "Sit-and-reach" (Ayán Pérez et al., 2018). Por lo tanto, sus propiedades psicométricas en el contexto preescolar se desconocen, mostrando la necesidad de ampliar el conocimiento científico a este respecto.

OBJETIVO

El objetivo de esta investigación fue evaluar la fiabilidad y la validez de las pruebas VS y TT cuando son realizadas por niños en edad en preescolar.

MATERIALES Y MÉTODOS

DISEÑO

Esta investigación tuvo un diseño psicométrico en el que se realizaron dos pruebas de valoración de la flexibilidad en dos ocasiones, tras un periodo previo de familiarización, con un intervalo de una semana, al objeto de identificar su fiabilidad test-retest. Para determinar la validez de las mismas, sus resultados se contrastaron con los obtenidos en una prueba considerada como Gold Standard. El análisis de las propiedades psicométricas de las pruebas realizados se llevó a cabo con datos registrados durante el tercer trimestre escolar del curso 18/19 y el segundo trimestre del curso académico 19/20.

PARTICIPANTES

Los participantes de este estudio pertenecían a tres centros de Educación Infantil situados en zonas urbanas del norte de España. Los criterios de inclusión para

participar fueron: a) tener entre 3 y 5 años de edad y b) no tener problemas médicos que pudieran dificultar la realización de las pruebas propuestas en la investigación. Aquellos niños que presentasen algún tipo de discapacidad intelectual o física que les impidiese comprender el protocolo de las pruebas o realizar las mismas correctamente, fueron excluidos de la investigación. Previa a la realización del estudio, se obtuvo el consentimiento informado por escrito de los padres o tutores de todos los niños que participaron en la investigación. Un total de 158 participantes (49,3% niñas) se presentaron voluntarios para el estudio. El diseño del mismo fue aprobado por el Comité de Ética de la Facultad de Ciencias de la Educación y el Deporte de la Universidad de Vigo.

VALORACIONES

Antropometría. El peso corporal (kg) y la altura (cm) fueron evaluados empleando una báscula digital y un tallímetro portátil. El índice de masa corporal (IMC) de cada niño se calculó utilizando la siguiente formula: peso/altura^2 (kg/m^2).

V-Sit and Reach. Para la realización del VS se pidió a los niños que se sentasen en el suelo, con las rodillas extendidas y los pies separados unos 30 cm, para formar una posición de letra uve (V). Se colocó una regla entre las piernas con la marca de los 9 cm ubicada en la línea de los talones. Se pidió a los niños que, poniendo las palmas hacia abajo y una mano encima de la otra, avanzasen lentamente lo más lejos posible, deslizando sus manos a lo largo de la regla y manteniendo la posición durante 2 segundos (Hui et al., 1999).

Toe-Touch. El TT se realiza utilizando el cajón de la prueba “Sit-and-reach” con una marca situada a 15 cm, representando el punto en el que las yemas de los dedos de los niños estaban en línea con los dedos de los pies. Los niños se situaban de pie encima del cajón, con las rodillas extendidas y los pies separados a la anchura de las caderas. Se les indicó que se inclinaran lo máximo posible hacia abajo, manteniendo las rodillas, los brazos y los dedos completamente extendidos y manteniendo la posición durante aproximadamente 5 segundos (Muyor et al., 2014).

Para las pruebas VS y TT, la posición final que alcanzaron los niños se consideró como la puntuación de la prueba y se registró el punto más alejado alcanzado con las puntas de los dedos. El mejor de los dos intentos en cada prueba se usó para el análisis. En ambas pruebas se eligió la marca de 15 cm como punto de partida de la medición en vez de 0, para evitar en la medida de lo posible puntuaciones negativas, observadas previamente al realizar este tipo de pruebas (Kippers & Parker, 1987). Además, de acuerdo con este procedimiento, se intenta que las puntuaciones de ambas pruebas sean siempre positivas, lo que facilita la presentación y tratamiento de los datos (Ayala et al., 2012). En ambas pruebas un evaluador supervisó la ejecución correcta de las mismas, mientras que otro evaluador registraba las marcas obtenidas (Ayán Pérez et al., 2018).

Prueba de elevación pasiva de la pierna recta (PSLR). Esta prueba se realizó estando los niños en posición supina sobre una colchoneta (Hui et al., 1999). Para su desarrollo, el eje de un goniómetro digital se alineó con el eje de la

articulación de la cadera, mientras que el brazo fijo del goniómetro se colocó en línea con el tronco, y el brazo móvil se situó en línea con el fémur. Se les pidió a los niños que elevasen su pierna preferida, manteniendo la rodilla extendida hasta que ellos mismos, o el examinador, percibieran tensión. Para la correcta ejecución de la prueba, un evaluador asistía en la elevación de la pierna, observando que la rodilla permaneciese en extensión, mientras que otro evaluador registraba el ángulo máximo alcanzado con el goniómetro (grados) desde el punto de máxima flexión de la cadera.

PROCEDIMIENTO

Para identificar la fiabilidad del test-retest del VS y del TT, ambas pruebas se realizaron en tres ocasiones, siguiendo procedimientos previamente indicados al respecto (Ayán Pérez et al., 2018). Las pruebas se realizaron en grupos de 15 participantes, durante las dos sesiones de psicomotricidad semanales que tenían lugar en los gimnasios de los centros educativos, en periodos de cuatro semanas. Durante la primera semana se realizaron las valoraciones antropométricas y se explicaron los protocolos de las tres pruebas. Se animó a los niños/as a que realizasen las mismas tras observar su ejecución por parte de su profesora y se les permitió la ejecución de varios intentos al objeto de familiarizarse con las mismas y evitar un efecto aprendizaje. A lo largo de las dos sesiones de la siguiente semana se realizaron las pruebas VS (primera sesión) y TT (segunda sesión), considerándose esta fase como “test”. Durante la tercera semana se llevó a cabo la prueba PSLR, cuyos resultados sirvieron para determinar la validez de las pruebas VS y TT, dado que se considera la medida “Gold Standard” en este tipo de investigaciones,(Mayorga-Vega, Viciana, et al., 2014). La cuarta semana sirvió de fase “retest”, repitiéndose la organización de la segunda semana. Todas las valoraciones fueron precedidas de una actividad de calentamiento dinámico que incluía carrera, saltos y cuadrupedia y realizadas por un profesor de Educación Infantil Graduado en Ciencias de la Actividad Física y el Deporte en presencia de la profesora de psicomotricidad de cada uno de los centros implicados en el estudio.

El cálculo del tamaño de la muestra indicó un mínimo requerido de 108 participantes para conseguir un nivel de potencia estadística de 0,90, con un tamaño de efecto anticipado (d de Cohen) de 0,4 y un nivel alfa (valor p) de 0,05.

ANÁLISIS ESTADÍSTICO

La media y la desviación estándar (SD) se calcularon para describir las características de la muestra. La prueba *t de Student* para muestras independientes se usó para determinar las diferencias entre los grupos de sexo, junto con la evaluación de normalidad utilizando la prueba Kolmogorov-Smirnov, Shapiro-Wilk y la homocedasticidad a través de la prueba de Levenne (Atkinson & Nevill, 1998). Los tamaños del efecto de las diferencias de sexo fueron estimados por la *d de Cohen,* siguiendo los valores clásicos (Lakens, 2013): 0,2 efecto pequeño , 0,5 efecto mediano y 0,8 efecto grande. La comparación entre los grupos de edad se realizó mediante ANOVA de un factor y Games-Howell post hoc.

El coeficiente de correlación intraclase (ICC) entre el test y el re-test y la *prueba t* de muestras relacionadas, incluido el tamaño del efecto Cohen, se usó para determinar la fiabilidad relativa. Los resultados de ICC fueron interpretados usando estos rangos: 0,90-1 excelente, 0,80-0,89 buena, 0,70-0,79 aceptable y <0,69 baja fiabilidad. El error estándar de la medida (*Estándar Error Measurament*, *SEM*) y el cambio mínimo detectable (*Mínimal Detectable Change, MDC*) se utilizaron para evaluar la fiabilidad absoluta (Shrout & Fleiss, 1979). Fueron calculados de la siguiente forma: SEM = [SD·√(1-ICC)] y MDC = [1,96·SEM·√2(1-ICC)] donde SD representa la desviación estándar de las diferencias entre el test y re-test y 1,96 es la puntuación correspondiente para intervalo de confianza del 95%. Además, SEM y MDC se expresaron como un porcentaje [SEM% = (SEM/$media_{test\text{-}retest}$)·100] y [MDC% = (MDC/$media_{test\text{-}retest}$)·100]. Estos valores se interpretaron de la siguiente manera: MDC% > 30% bajo, 10-30% aceptable, y <10% excelente (Ayán Pérez et al., 2018). Los valores de SEM% por debajo del 10% se interpretaron como una excelente fiabilidad absoluta (Lexell et al., 2005). Por último, para evaluar la variabilidad del test-retest y el posible sesgo sistemático, se utilizaron gráficos Bland-Altman, que muestran diferencias del test-retest respecto a la media test-retest y representan el límite de concordancia del 95% (*limits of agreemente* ,LOA).

Los criterios de validez se evaluaron por medio del coeficiente de correlación de Pearson (PCC), que se interpretó de la siguiente manera: >0,8 correlación muy alta, 0,79-0,60 correlación moderada alta, 0,59-0,30 correlación aceptable, y <0,29 correlación baja (Chan, 2003). Todos los análisis estadísticos se realizaron con el Paquete Estadístico para las Ciencias Sociales (SPSS Versión 25,0, Chicago, IL) y Microsoft Excel 2011.

RESULTADOS

Del total de participantes inicialmente reclutados, 26 de los mismos no completaron todas las pruebas de valoración por no haber asistido a clase en alguno de los momentos de evaluación. Por lo tanto, la muestra final del estudio estuvo conformada por 132 prescolares (68 niños y 64 niñas, con una edad media de 4,76±0,85 años), quienes completaran todas las mediciones y proporcionasen datos válidos para el análisis (Tabla 1). No se observaron dificultades a la hora de comprender o ejecutar las pruebas por parte de los niños y tampoco se observaron efectos adversos derivados de su realización.

Tabla 1. Valores descriptivos (media ± SD) y comparación entre grupos por sexo (*t* de Student y tamaño del efecto *d* de Cohen).

	Total (n =132)				Niños (n =68)		Niñas (n =64)		*t* de Student, p	*d* de Cohen
Variables	Media	DT	Máx.	Mín.	Media	DT	Media	DT		
Edad (años)	4,76	0,85	2,89	5,87	4,76	0,81	4,75	0,89	0,952	0,011
Peso corporal (kg)	18,66	3,19	12,80	36,30	18,73	2,75	18,50	3,57	0,681	0,071
Altura(cm)	106,36	6,68	92	121	107,32	6,86	105,20	6,31	0,068	0,316
IMC (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
Goniómetro (PSLRT) (º)	91,85	13,26	61,50	122,00	87,88	13,44	95,97	11,88†	0,000	-0,611

*IMC = Índice de masa corporal; VS = V sit and reach; TT = Toe Touch. †Valores significativamente mayores que el otro sexo ($p < 0,05$).

Los resultados obtenidos indicaron la existencia de diferencias significativas en los niveles de flexibilidad a favor de las niñas. Las puntuaciones obtenidas en el VS y el PSLR mostraron una relación inversa con la edad, tendencia que no se observó cuando se utilizó el TT como referencia (Tabla 2).

Tabla 2. Valores descriptivos (media ± SD) y comparación entre grupos por edad (ANOVA y Games-Howell post hoc).

	3 años (*n* =26)	**4 años (*n* =38)**	**5 años (*n* =68)**	**ANOVA**		**Tamaño del efecto (d de Cohen)**		
Variables	M ± SD	M ± SD	M ± SD	F	*p*	3-4 años	4-5 años	3-5 años
Edad (años)	3,43 ± 0,29†‡	4,37 ± 0,24‡	5,49 ± 0,24	841,92	0,000	3,57	4,67	7,84
Peso corporal (kg)	17,00 ± 2,27‡	18,07 ± 2,39‡	19,61 ± 3,55	8,00	0,001	0,46	0,52	0,90
Altura (cm)	99,21 ± 4,94†‡	105,80 ± 5,07‡	109,37 ± 5,90	32,50	0,000	1,32	0,65	1,87
IMC (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
Goniómetro (PSLRT) (º)	96,63 ± 12,77‡	96,14 ± 15,16‡	87,68 ± 10,93	7,82	0,001	-0,04	-0,65	-0,76

*IMC = Índice de masa corporal; VS = V sit and reach; TT = Toe Touch. †Diferencias significativas entre los grupos de edad de 3 y 4 años (p < 0,05).

El VS mostró una excelente fiabilidad (ICC = 0,91-0,93), excepto para el grupo de cinco años (ICC = 0,88), mientras que la fiabilidad del TT fue excelente en todos los casos (ICC = 0,93-0,97). La fiabilidad absoluta fue excelente, con SEM y MDC con un rango de valores de 2,7 a 5,1% y de1,9 a 4,9% para el VS y de 2,4 a 5,1% y de 1,1 a 3,5% para el TT (Tabla 3).

Tabla 3. Fiabilidad de las pruebas VS y TT según el sexo y los grupos de edad.

		IC al 95% para CCI		Diferencia Test-retest	*t* Student			SEM		MDC	
	ICC	Límite inferior	Límite superior	M ± SD	*t*	*p*	Muestra	Puntuación	%	Puntuación	%
Todos (*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
Niños (*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
Niñas (*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 años (*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 años (*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 años (*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

El análisis gráfico de la fiabilidad (test-retest) tanto del VS como del TT se muestran en la Figura 1.

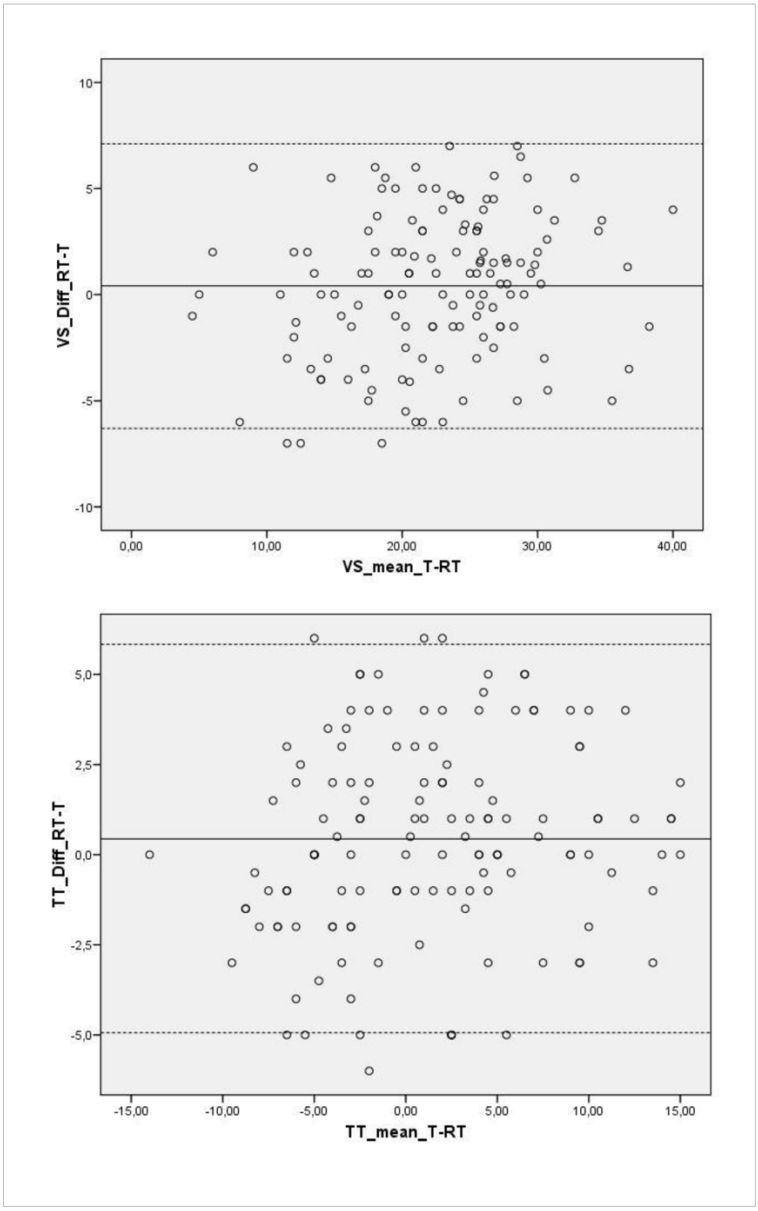

Figura 1. Gráficos de Bland-Altman para VS y TT.
Notas: La línea continua representa el valor medio y las líneas discontinuas superior e inferior que representan los límites (1,96 SD).

En la Tabla 4 se muestran los datos relativos a la validez de las pruebas VS y TT.

Tabla 4. Grado de asociación entre la goniometría (prueba de elevación pasiva de la pierna recta) y VS y TT por sexo y edad. †

PCC goniómetro	VS				TT			
	Test		Retest		Test		Retest	
Grupos	PCC	*p*	PCC	*p*	PCC	*p*	PCC	*p*
Todos(*n* = 132)	*r* = 0,620†	0,000	*r* = 0,647†	0,000	*r* = 0,618†	0,000	*r* = 0,609†	0,000
Niños (*n* = 68)	*r* = 0,581†	0,000	*r* = 0,628†	0,000	*r* = 0,617†	0,000	*r* = 0,619†	0,000
Niñas (*n* = 64)	*r* = 0,568†	0,000	*r* = 0,571†	0,000	*r* = 0,529†	0,000	*r* = 0,526†	0,000
3 años (*n* = 26)	*r* = 0,538†	0,005	*r* = 0,398†	0,044	*r* = 0,372	0,061	*r* = 0,320	0,111
4 años (*n* = 38)	*r* = 0,572†	0,000	*r* = 0,618†	0,000	*r* = 0,727†	0,000	*r* = 0,727†	0,000
5 años(*n* = 68)	*r* = 0,624†	0,000	*r* = 0,705†	0,000	*r* = 0,655†	0,000	*r* = 0,661†	0,000

PCC= Coeficiente de correlación de Pearson' †=Valor significativo de PCC (p < 0,05)

Se observaron correlaciones estadísticamente significativas entre ambas pruebas y el PSLR para todos los grupos analizados, excepto para la asociación ente el TT y el PSLR en el grupo de 3 años. El grado de correlación fue moderadamente alto en general. Las asociaciones entre el VS y la prueba PSLR fueron consideradas aceptables para el grupo de 3 años (r = 0,398-0,538). El VS y el TT mostraron una validez moderadamente alta para los grupos de 4 años (r = 0,572-0,727) y 5 años de edad (r = 0,624-0,705).

DISCUSIÓN

La evidencia científica actual sobre las propiedades psicométricas de las pruebas de flexibilidad cuando son realizadas por preescolares es escasa. Este estudio tuvo como objetivo proporcionar información sobre la fiabilidad relativa y absoluta, así como sobre la validez de dos de las pruebas de valoración de flexibilidad más utilizadas.

La literatura indica que los niveles de flexibilidad están altamente influenciados por el sexo y la edad, incluso desde edades tempranas (Ayán Pérez et al., 2018). En el presente estudio, los resultados obtenidos en ambas pruebas confirman que las niñas son más flexibles que los niños. Además, las puntuaciones del VS indicaron una relación inversa entre la edad y la flexibilidad. Sin embargo, esta tendencia no se observó cuando se analizaron los resultados del TT. Estos hallazgos podrían indicar que la prueba VS es más sensible para medir la flexibilidad en edades tempranas.

Tanto el VS como el TT mostraron una excelente fiabilidad relativa. Si bien no existe evidencia científica previa sobre los valores de fiabilidad para ambas pruebas en preescolares, estos datos pueden ser comparados con los reportados para población adulta. En el caso del VS, los valores de fiabilidad han sido por lo general buenos o excelentes en adultos (ICC = 0,86-0,99) (Ayala & de Baranda, 2012). Sobre el TT, Vrbik informó de una consistencia interna que oscila entre 0,96 y 0,98, al ser realizada por niños de Educación Primaria (Vrbik et al., 2017). La fiabilidad relativa encontrada en este estudio para ambas pruebas fue ligeramente más elevada que la demostrada por el “Sit-and-Reach” original en preescolares (ICC = 0,74-0,94) (Ayán Pérez et al., 2018; Ortega et al., 2015).

Cabe destacar que, desde un punto de vista estadístico, el uso exclusivo del ICC para evaluar la fiabilidad puede considerarse insuficientemente preciso, ya que se ve afectado por la heterogeneidad de la muestra y no indica la presencia de un posible sesgo sistemático (Hopkins, 2000). Por lo tanto, la cantidad de error aleatorio que ocurre cuando un solo individuo se prueba varias veces (SEM), así como la cantidad mínima de cambio necesaria para superar el error de medición (MDC), fueron incluidos en el análisis estadístico como una estrategia adicional para identificar la fiabilidad absoluta de ambas pruebas. A este respecto, se observó que cambios de entorno a un 2,68% (unos 0,60 cm) para el VS y de 2,29% (unos 0,38 cm) para el TT, indicarían la existencia de cambios de magnitud considerable en la extensibilidad isquiosural entre ambas mediciones (test-retest). Los resultados obtenidos apuntan a que el VS y el TT presentan una excelente fiabilidad absoluta, resultado que ha sido previamente observado

para el “Sit-and-Reach” cuando es realizado por niños en edad preescolar con un cajón adaptado (Ayán Pérez et al., 2018).

Varios estudios han analizado la validez del VS (Mayorga-Vega, Merino-Marban, et al., 2014) y el TT (Mayorga-Vega, Viciana, et al., 2014) usando la prueba PSLR como medida de fiabilidad, reportando grados de correlaciones entre moderados y aceptables (VS, r = 0,44-0,65; TT, r = 0,54-0,79). En este estudio, ambas pruebas mostraron correlaciones moderadas, que fueron incluso superiores a las encontradas en el “Sit-and-Reach” en población preescolar (r=0,352-0,51)[2]. Estos hallazgos concuerdan con los de Hui et al. (1999), quienes hallaron que el VS muestra mayores niveles de validez que el “Sit-and-Reach” en adultos jóvenes. Por el contrario, otros autores han encontrado que el “Sit-and-Reach” mostró una mayor validez que el VS en una población similar (López-Miñarro et al., 2008). Adicionalmente, y en contraste con estas ideas, Ayala y de Baranda (2012) informaron que el “Sit-and-Reach” tiene un coeficiente de validez más alto que el TT en adultos jóvenes activos.

En la presente investigación, a pesar de encontrarse una validez moderada para el TT, debe tenerse en cuenta que esta prueba no mostró una asociación significativa con la prueba PSLR en el grupo de 3 años, lo que indica una falta de validez de criterio en este grupo. A este respecto, se ha observado que la distancia de los dedos al suelo puede estar influenciada por características antropométricas tales como el tamaño del brazo o el rango de movilidad de la columna vertebral (Carregaro et al., 2007). En base a esto, se podría plantear la hipótesis de que, modificando las dimensiones de cajón, se lograría que los niños de menor estatura alcanzasen mejores resultados, tal y como se ha observado para el “Sit-and-Reach” (Ayán Pérez et al., 2018). Futuras investigaciones podrían profundizar en esta idea.

Los resultados de este estudio dejan dos datos de especial interés. En primer lugar, indican que la fiabilidad absoluta y relativa de las pruebas VV y TT es elevada para niños de 3 a 5 años. Este es un dato relevante ya no solo desde el punto de vista psicométrico, sino porque aporta información sobre el nivel de condición física en niños de 3 años, aspecto relativamente poco estudiado en la literatura (Arufe-Giráldez, 2020). En segundo lugar, después de analizar el grado de asociación entre el VS y el TT con la prueba PSLR, parece ser que ambas pruebas demuestran una validez de criterio moderada, excepto en el caso del TT cuando es realizado por niños de 3 años. Considerando todo esto y teniendo en cuenta que la ejecución del VS es más sencilla, pudiera ser aconsejable el realizar esta prueba para estimar la extensibilidad de los isquiotibiales en preescolares.

A pesar de la originalidad de este estudio, hay una serie de limitaciones que merecen ser destacadas. En primer lugar, el TT y el VS son medidas indirectas de flexibilidad, ya que implican el movimiento del todo el cuerpo. Este aspecto favorece que su validez puede estar influenciada por factores antropométricos que no fueron evaluados. Debido a ello, el uso de otro enfoque metodológico para estimar los criterios relativos de validez, como medir el rango de movimiento de la cadera por medio de una radiografía o usando un inclinómetro, habría permitido una mejor identificación de esta propiedad psicométrica. En segundo

lugar, debe tenerse en cuenta que el análisis de los subgrupos estuvo limitado por el pequeño tamaño de la muestra.

CONCLUSIONES

Los resultados de esta investigación indican que el VS y el TT son pruebas de valoración de la flexibilidad de las extremidades inferiores fiables para ser empleadas en el marco de la educación preescolar, mostrando además una validez moderada. El VS puede ser una prueba preferible sobre el TT, si se considera la relación inversa entre sus resultados y la edad de los participantes, así como la validez moderada presentada para todos los grupos de edad.

REFERENCIAS BIBLIOGRÁFICAS

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%)