

## Estudio sobre la idoneidad de dos test de valoración de la amplitud de movimiento en personas sedentarias con discapacidad intelectual

### Study on the suitability of two test for the assessment of the range of motion in sedentary people with intellectual disabilities

Ruth Cabeza-Ruiz, Pedro Tomás Gómez-Píriz

Universidad de Sevilla (España)

**Resumen.** Introducción: Las personas con discapacidad intelectual (DI) muestran, en general, menores niveles de condición física (CF) que las personas sin discapacidad. Sin embargo, la mayoría de test utilizados para su valoración han sido diseñados para población normalizada y no muestran su conveniencia para ser utilizados con personas con DI. Entre las capacidades condicionales una de las que habitualmente es valorada es la amplitud de movimiento (ADM). El objetivo de esta investigación fue conocer la idoneidad de dos test de evaluación de la ADM a través de la valoración de sus propiedades psicométricas en personas con DI leve-moderada. Los test seleccionados fueron el sit and reach (SAR) y el deep trunk flexion (DTF). Método: Se llevó a cabo el estudio con un diseño test-retest, en el que participaron 60 personas con DI: se distribuyeron aleatoriamente 15 mujeres ( $37.07 \pm 11.60$  años) y 15 hombres ( $37.47 \pm 11.62$ ) para valorar el DTF y 15 mujeres ( $36.87 \pm 11.72$ ) y 15 hombres ( $36.40 \pm 8.57$ ) para el SAR. Para conocer la idoneidad de ambos test, se calculó la fiabilidad, la viabilidad, el error estándar de la medida (EEM) y el mínimo cambio detectable (MCD). Resultados: La fiabilidad del DTF y del SAR fue buena-alta. Sin embargo, los gráficos de Bland-Altman mostraron sesgos para el SAR. La viabilidad del SAR fue baja, con menos del 50% de los participantes realizando la prueba correctamente. El EEM y el MCD mostraron mayor variabilidad de puntuaciones del SAR respecto al DTF. Conclusiones: el DTF se ha mostrado como el más recomendable de los dos test para la valoración de la ADM en personas con DI. Sin embargo, futuras investigaciones deben orientarse a la creación de test específicos diseñados para personas con DI.

**Palabras claves:** psicometría, discapacidad intelectual, flexibilidad, adultos, test.

**Abstract.** Introduction: People with intellectual disabilities (ID) show, in general, lower levels of physical fitness (PF) than people without disabilities. However, most of the test used for their assessment have been designed for the general population and do not show their suitability to be used with people with ID. Among the conditional capabilities, one of the most commonly assessed is the range of motion (ROM). The aim of this research was to assess the psychometric properties of the deep trunk flexion (DTF) and the sit and reach test (SAR) in people with mild-moderate ID in order to find out whether they are suitable for assessing this population. Method: The study was carried out with a test-retest design, in which 60 people with ID participated: 15 women ( $37.07 \pm 11.60$  years) and 15 men ( $37.47 \pm 11.62$ ) were randomly distributed to assess the DTF, and 15 women ( $36.87 \pm 11.72$ ) and 15 men ( $36.40 \pm 8.57$ ) were randomly distributed to assess the SAR. Reliability, feasibility, standard error of measurement (EEM) and minimum detectable change (MDC) were calculated to determine the suitability of both tests. Results: The reliability of the DTF and SAR was good to high. However, Bland-Altman plots showed biases for the SAR. The feasibility of the SAR was low, with less than 50% of participants performing the test correctly. The EEM and the MDC also showed higher variability for SAR with respect to the DTF. Conclusions: The DTF has been shown to be the more suitable of the two tests for the assessment of ROM in people with ID. However, future research should be oriented towards the creation of specific test for people with ID.

**Key words:** psychometrics, intellectual disability, flexibility, adults, test.

## Introducción

Un buen desarrollo de todas las capacidades condicionales es importante para la salud. Entre ellas la popular y generalmente denominada flexibilidad o amplitud de movimiento (ADM) es una de las capacidades que tiene un rol relevante en la vida cotidiana ya que facilita el desarrollo de las actividades diarias de manera autónoma y funcional. La ADM representa directamente la salud de los sistemas osteoarticular y muscular (Geraldés, et al., 2008), y está relacionada con el riesgo de sufrir caídas (Guimarães & Farinatti, 2005).

Un buen desarrollo de esta cualidad permite realizar tareas como agacharse, calzarse, vestirse o ducharse con solvencia. La ADM se ve mermada a causa del sedentarismo, entre otros factores, debido al escaso uso de las articulaciones y al menor desarrollo de la capacidad de elongación muscular. Una disminución de la ADM en la cadera se encuentra relacionada también con una menor salud de la espalda (Kennedy & Noh, 2011). En diversos estudios los hallazgos confirman que las personas mayores activas presentan mejores resultados en pruebas de ADM que personas mayores sedentarias (Matos-Duarte et al., 2017; 2022).

Las personas con discapacidad intelectual (DI) muestran, en general, menores niveles de ADM que las personas sin discapacidad de su mismo sexo y edad (Hilgenkamp, et al., 2012). Aunque aún no se ha escl-

Fecha recepción: 08-07-21. Fecha de aceptación: 11-02-22

Pedro Tomás Gómez Piriz  
ptgomez@us.es

recido hasta qué punto la propia discapacidad es la causa de esta menor ADM, es importante señalar que existe una alta prevalencia de sedentarismo en personas con DI jóvenes y adultas (Sundahl, et al., 2016), pudiendo constituir una de las principales causas de esta menor ADM y de su condición física (CF) general.

Valorar la CF relacionada con la salud de las personas con DI es necesario pues permite conocer el grado de autonomía y de salud física en el que se encuentran y así proporcionarles los apoyos necesarios para recuperarla, mejorarla o mantenerla. Sin embargo, los estudios que muestran los niveles de CF en población con DI que incluyen test de valoración de la ADM se han llevado a cabo, en general, utilizando instrumentos o test de valoración diseñados para personas sin discapacidad (Cuesta-Vargas, et al., 2011). Estos test, al no estar adaptados, pueden resultar muy complejos de realizar desde un punto de vista técnico para las personas con DI (Cabeza-Ruiz, 2020), que en ocasiones no comprenden el protocolo o no pueden realizar las pruebas con solvencia, por lo que se podrían estar obteniendo datos falsos con respecto al estado de forma real en el que se encuentran (Hilgenkamp, et al., 2012; Wouters, et al., 2017; Oppewal & Hilgenkamp, 2020). Por ello, se hace necesario realizar pruebas psicométricas que evalúen la idoneidad de los test normalizados al emplearlos con personas con DI.

La psicometría permite la construcción y validación de los instrumentos de medición. Evaluando las características psicométricas de esos instrumentos se puede conocer si son herramientas fiables y viables para una determinada población (Ginty, 2013; Oppewal & Hilgenkamp, 2020). La fiabilidad inter-sesión de un instrumento se refiere a su capacidad para obtener los mismos resultados al medir el mismo fenómeno en diferentes momentos. Para su cálculo se utiliza el coeficiente de correlación intraclase (CCI), como medida de asociación y los gráficos de Bland-Altman, como medida de concordancia (Aldridge, et al., 2017; Atkinson & Nevill, 1998). El cálculo de la fiabilidad conlleva el cálculo del error estándar de la medida (EEM), que depende de la fiabilidad y de la desviación estándar, y se expresa en las mismas unidades de escala que los test. El mínimo cambio detectable (MCD) es la mínima diferencia que debe tener una persona en las puntuaciones de un test para poder concluir que ese cambio no es fruto de errores de medida y, por tanto, que es un cambio debido a un factor externo. La viabilidad, finalmente, es el índice de casos exitosos, es decir, el porcentaje de personas que realizan la prueba adecuadamente, lo

que incluye no solo la comprensión de la misma, sino la correcta ejecución del test (Gidron, 2013; Oppewal & Hilgenkamp, 2020) y que influirá, por tanto, en la fiabilidad del instrumento con esa población determinada.

Dos de los tests más utilizados en la valoración de la ADM en población normalizada son el sit and reach (SAR) y el deep trunk flexion (DTF) (Bajo, 2003; Zurita-Ortega, et al., 2008; Delgado-Valdivia, et al., 2009). El uso de estas herramientas viene determinado porque los procedimientos son sencillos, fáciles de administrar, requieren una formación mínima de los evaluadores y permiten realizar muchas mediciones en poco tiempo lo cual puede ser muy útil para los trabajadores de centros de atención a personas con DI. Sin embargo, es necesario que en la investigación se justifique la conveniencia de estos test para la valoración de la ADM a través del análisis de diferentes propiedades psicométricas complementarias que permitan realizar una comparación entre ambas herramientas. Sin embargo, hasta la fecha, no se han realizado estudios que valoren las características psicométricas de estos test, tan populares, en personas con DI.

El objetivo de esta investigación fue conocer las propiedades psicométricas (fiabilidad inter-sesión y concordancia-, viabilidad, EEM y MCD) del DTF y el SAR en personas sedentarias con DI leve-moderada con el fin de conocer si son idóneos para evaluar a este colectivo.

## Metodología

Se llevó a cabo un estudio transversal, con un diseño de test-retest en un intervalo de 2 semanas en el que participaron 5 centros de atención a personas con discapacidad intelectual.

## Participantes

60 adultos con DI de ambos sexos (30 mujeres y 30 hombres) participaron en el estudio (Tabla 1). Los participantes se distribuyeron en dos grupos homogéneos, cada grupo de 30 personas hizo solo una de las pruebas. La realización de ambos test por un mismo grupo no fue contemplada para evitar confusión e interferencias en la realización de ambas técnicas por parte de los participantes. La distribución fue aleatoria y estratificada por sexo. Un grupo compuesto por 15 mujeres ( $37.07 \pm 11.60$  años) y 15 hombres ( $37.47 \pm 11.62$  años) realizaron el DTF y otro compuesto por 15 mujeres ( $36.87 \pm 11.72$  años) y 15 hombres ( $36.40 \pm 8.57$  años) el SAR. Los criterios de inclusión en el estudio fueron: (i) tener diagnosticada DI por el órgano oficial regula-

dor del gobierno competente, (ii) estar institucionalizados en centros de atención diurna, (iii) no participar con regularidad en actividades deportivas, (iv) no tener DI asociada con síndromes de naturaleza cromosómica (como el síndrome de Down), y (v) poder desplazarse de manera autónoma.

Tabla 1.

*Características antropométricas de la muestra (Media y DE)*

	Peso (Kg)	Talla (m)	IMC	Cintura (cm)
H (30)	79.96 (15.25)	1.66 (.09)	28.89 (4.70)	77.44 (44.22)
M (30)	69.51 (15.67)	1.56 (.05)	28.63 (6.46)	94.00 (12.30)

Con el fin de conocer el nivel de DI de los participantes se solicitó información al servicio psicológico de los centros participantes. Todos los participantes fueron descritos como personas con DI leve-moderada según estos servicios.

Para participar en el estudio, los familiares responsables de las personas voluntarias y ellas mismas, tras ser informadas por el equipo investigador, debían firmar un consentimiento de participación. De la misma manera, se requirió a los participantes una autorización médica donde se reflejó su aptitud para realizar actividad física sin riesgos para su salud. El estudio fue aprobado por el Comité Ético de Investigación Biomédica de Andalucía y siguió las directrices de la Declaración de Helsinki de la Asociación Médica Mundial (2013).

### **Material e instrumentos**

Se utilizó el cajón o banco de medición (10 x 10 x 55 cm, altura, anchura y largo, respectivamente) Baseline modelo 12-1085 (Estados Unidos), se añadió un tablero extra por delante de la misma de 50 cm para poder medir a aquellos participantes que no fueran capaces de llegar al principio de la plataforma milimetrada en la ejecución de los test. Esta plataforma adicional estaba igualmente milimetrada y disponía de un cursor que valoraba en cm la distancia obtenida, y mantenida durante 3", desde la posición de partida.

### **Procedimiento**

Ambas pruebas fueron seleccionadas por su simplicidad técnica. Aun siendo test muy popularizados, requieren para realizarlos de ajustes posturales complejos, sobre todo para personas con DI. En ambos casos los test implican la capacidad de movimiento del raquis y la cadera durante la flexión profunda del tronco, involucrando también la capacidad de elongación de la musculatura isquiosural del sujeto y el grado de movilidad de la zona pélvica.

Previo al inicio de las valoraciones, se realizaron ejer-

cicios generales de activación (10'). En ambas mediciones se permitió realizar 2-3 ensayos previos y 2 intentos posteriores, registrando el mejor de ellos para el análisis estadístico. El protocolo de cada test fue explicado y demostrado a cada una de las personas participantes con el fin de maximizar el nivel de comprensión del mismo. Debido a las características de las personas con DI, se utilizaron estrategias de motivación básicas para la realización de las tareas, teniendo en cuenta la información y asesoramiento aportados por los profesionales del centro. Los test fueron realizados en salas acondicionadas (temperatura de 20°-25°), suelo aislante y en la misma franja horaria, coincidiendo con las últimas horas de la mañana (12-14h)

### **Deep trunk flexion**

Para llevar a cabo el DTF, el participante se colocaba en bipedestación con las piernas separadas a la anchura de las caderas. Para que supiera exactamente cómo adquirir la postura inicial, se colocaron unas huellas de goma sobre las que colocar los pies. Desde la bipedestación, el sujeto flexionaba el tronco progresivamente, introduciendo los brazos entre las piernas hacia abajo y atrás, empujando con los dedos centrales de ambas manos el cursor de la barra milimetrada, con el fin de conseguir la mayor distancia posible (hasta un máximo de 50 cm). Se permitió flexionar levemente las rodillas, tal como indican Delgado-Valdivia et al. (2009). Todos los participantes llegaron correctamente al cursor (punto 0).

### **Sit and reach**

El test clásico SAR se realizó desde la posición de sentado en una superficie horizontal con las piernas juntas, las rodillas en extensión y los pies apoyados sobre la parte frontal del banco de medición. Desde esa posición, el sujeto iba flexionando el tronco progresivamente y adelantando los brazos en dirección a las puntas de los pies, empujando con los dedos centrales de ambas manos el cursor de la barra milimetrada del cajón de medición, hasta conseguir la mayor distancia posible. No se permitió flexionar las rodillas (Ayala, et al., 2012). Los datos obtenidos en el SAR fueron almacenados en forma de números positivos, en el caso de que el participante pudiera empujar el cursor del banco, y en números negativos en el caso de que solo alcanzara la plataforma adicional (Tabla 2).

### **Análisis de datos**

Todas las pruebas estadísticas se realizaron separan-

do a mujeres y hombres en dos grupos siguiendo las indicaciones de la Comisión Europea para la integración de la dimensión de género en la investigación (European Commission, 2012), teniendo en cuenta que los resultados entre hombres y mujeres podrían diferir. Para la estimación de la fiabilidad o grado de asociación de las variables, se calculó el CCI (modelo mixto de dos factores, acuerdo absoluto) y el 95% de los intervalos de confianza (IC). Los resultados se interpretaron siguiendo las indicaciones de Shrout & Fleiss (1979): valores entre .90-.99 muestran alta fiabilidad, de .80-.89 buena fiabilidad, de .70-.79 fiabilidad media y, puntuaciones iguales o inferiores a 0.69, fiabilidad baja. Además del CCI, se realizaron pruebas de comparación de medias para conocer posibles diferencias entre el test y el retest. Con este fin los datos de los participantes fueron previamente analizados para conocer la normalidad de la distribución. Se realizaron contrastes no paramétricos para conocer posibles diferencias entre condiciones en medidas relacionadas (test-retest).

Además del CCI, para el análisis de la fiabilidad se construyeron los gráficos de Bland-Altman, que muestran el nivel de concordancia de las medidas entre el test y el retest para cada sujeto, así como los límites de acuerdo (LdA) que representan los límites superior e inferior del rango medio del 95% de los datos observados. Los gráficos se utilizan para examinar la distribución de los datos alrededor del 0 y para detectar outliers y sesgos.

También se calculó el EEM para conocer la fiabilidad absoluta (Gempp, 2006). Para una precisión aceptable  $EEM < \text{desviación estándar (DE)}/2$  (Boer & Moss, 2016). Se mostró la relación, en %, de EEM con respecto al promedio para tener una visión clara de esta variable según los datos obtenidos. Se valoró el MCD para conocer qué mínimo cambio sería necesario en la medida para no considerar la modificación como efecto de errores de medida (Atkinson & Nevill, 1998). El nivel de significación establecido fue de  $p < .05$ .

Para la valoración de la viabilidad o tasa de término de ambos test, se calculó el porcentaje de éxito de la muestra, es decir, el número de personas que realizaban correctamente la prueba. Para ello se siguieron las indicaciones de Wouters et al. (2017), i.e.: no viable  $< 50\%$ , medianamente viable  $50 - 75\%$  y viable  $> 75\%$ .

Todas las pruebas estadísticas se realizaron con SPSS en su versión 25, excepto el 95% de LdA, EEM y MCD que se valoraron en la hoja de cálculo de Microsoft Excel v. 16.46.

## Resultados

Los resultados obtenidos en las pruebas se muestran en la Tabla 2 y en las figuras (1A-1B-2A-2B).

### Fiabilidad

Los resultados de CCI muestran una fiabilidad buena-alta tanto en el DTF como en el SAR ( $CCI > .80$ ), y tanto en el grupo de hombres como en el de mujeres. Las pruebas no paramétricas no detectaron diferencias significativas entre test y retest en el DTF. Sin embargo, en el SAR, los datos mostraron diferencias significativas entre test y retest en el caso de las mujeres (Tabla 2). En el DTF, los datos de variabilidad fueron medios pero la DE y el 95% IC mostraron una variabilidad alta en los resultados del SAR.

Tabla 2.

Resultados de la valoración de la fiabilidad de los test DTF y SAR en los grupos de hombres y mujeres						
Test	n	TEST Media (DE) cm	RE-TEST Media (DE) cm	CCI (95% IC)	Wilcoxon (sig)	
Sit and reach	M	15 1.40 (10.03)	-7.33 (9.91)	.85 (.623-.949)	.001*	
	H	15 -14.53 (12.76)	-13.77 (13.67)	.81 (.531-.933)	.977	
Deep trunk flexibility	M	15 34.47 (7.17)	34.77 (6.99)	.91 (.767-.971)	.669	
	H	15 32.63 (9.80)	34.93 (9.18)	.88 (.699-.908)	.221	

M, mujeres; H, hombres; DE, desviación estándar; CCI, coeficiente de correlación intraclase; IC, intervalo de confianza; sig, nivel de significación (p).

Los gráficos de Bland-Altman reflejan el nivel de concordancia entre las puntuaciones del test y el retest tanto para el DTF (figuras 1.A y 1.B) como para el SAR (figuras 2.A y 2.B). Los LdA, que permiten examinar la distribución de los datos alrededor del 0, muestran diferencias entre SAR y DTF, siendo mejores los datos para el segundo. También se observan diferencias entre hombres y mujeres.

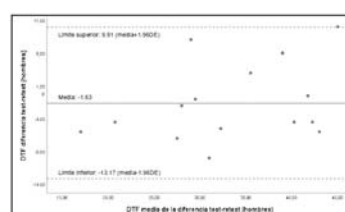


Figura 1.A. Gráfico Bland-Altman para DTF (Hombres)

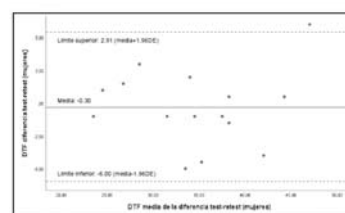
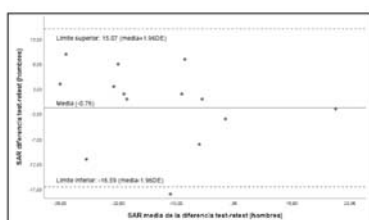


Figura 1.B. Gráfico Bland-Altman para DTF (Mujeres)

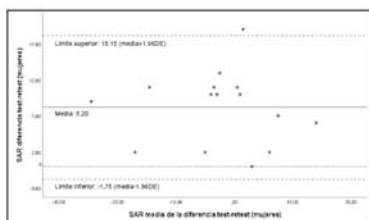
### EEM, MCD y viabilidad

Los datos de EEM reflejan que en todos los casos se



Límites de acuerdo inferior y superior. SAR, sit and reach, DE, desviación estándar, Media, media de la diferencia entre puntuaciones.

Figura 2.A. Gráfico Bland-Altman para SAR (Hombres)



Límites de acuerdo inferior y superior. SAR, sit and reach, DE, desviación estándar, Media, media de la diferencia entre puntuaciones.

Figura 2.B. Gráfico Bland-Altman para SAR (Mujeres)

ajustan a las consideraciones de Boer y Moss (2016) ( $EEM < DE/2$ ). El MCD fue  $> 10$  cm en el SAR y  $< 9$  cm en el DTF. Los resultados mostraron bajos índices de viabilidad en el SAR ( $< 50\%$ ) y mayores del  $75\%$  en el DTF (Tabla 3). Ningún hombre pudo alcanzar el punto 0 en la ejecución del SAR.

Tabla 3.

*EEM, MCD y viabilidad de los test DTF y SAR en los grupos de hombres y mujeres*

Test	n	EEM (cm, % del promedio)	MCD (cm, % del promedio)	Viabilidad %	
Sit and reach	M	15	3.80 (51.84%)	10.52 (143.52%)	33.3
	H	15	5.71 (41.47%)	15.84 (115.03%)	46.7
Deep trunk flexibility	M	15	2.06 (5.92%)	5.72 (16.45%)	86.7
	H	15	3.23 (9.25%)	8.96 (25.65%)	80.0

M, mujeres; H, hombres; EEM, error estándar de la medida; MCD, mínimo cambio detectable.

## Discusión

El principal hallazgo de este estudio es que el test DTF muestra unos resultados que lo hacen ser más apropiado que el SAR para valorar la ADM (cadera, espalda y hombros) en las personas con DI leve-moderada.

Aunque el CCI ha mostrado niveles buenos-altos en ambos test, el CCI por sí solo no puede establecer la fiabilidad de un instrumento, sino que es necesario complementar su uso con otros análisis estadísticos (Aldridge, et al., 2017). En este sentido, el DTF ha obtenido mejores resultados en las pruebas complementarias. En primer lugar, no se encontraron diferencias significativas en la comparación entre el test y del retest, a diferencia del SAR, lo que evidencia inconsistencia del instrumento en la medición inter-sesiones. La discordancia entre los valores de CCI y el resultado de los contrastes no paramétricos puede deberse a la variabilidad de los datos de la muestra (Weir, 2005). En el presente estudio se homogeneizó la muestra en cuanto a edad y sexo.

Sin embargo, la diversidad de la expresión de la discapacidad en el colectivo de personas con DI puede dar como resultado puntuaciones caracterizadas por la variabilidad de los datos que ofrecen valores de CCI elevados que pueden llevar a error sobre su idoneidad. En los estudios previos sobre CF de las personas con DI se observa que las respuestas de los participantes presentan mucha variabilidad (Ayán-Pérez, et al., 2016; 2017) incluyendo aquellos que incluyen pruebas de ADM (Cuesta-Vargas, et al., 2011; Hilgenkamp, et al., 2012b).

La mayor variabilidad de los resultados en el SAR respecto del DTF hace pensar que las personas con DI que formaron parte del estudio podían estar enfrentando dificultades a la hora de realizar el test. Estos hallazgos pueden observarse en la DE. En el caso del SAR, esta medida de dispersión es cercana o superior a la media del grupo, tanto en hombres y mujeres, lo que indica valores muy dispersos. Sin embargo, el DTF muestra una DE más ajustada a la media.

En la información complementaria que aportan los gráficos de Bland-Altman (figuras 1.A.a 2.B.) pueden observarse aspectos que esclarecen la que aporta el indicador CCI. Los plots de Bland-Altman el 0 se establece como valor de referencia. Esto es así porque si la diferencia entre una puntuación en el pretest y en el posttest es 0, es que la repetibilidad de la medida es perfecta (Cabeza-Ruiz, et al., 2020). En el presente estudio, el primero de los aspectos a señalar es que en la figura 2.A. (SAR en hombres) se observa que los LdA son muy amplios (15.07—16.59), es decir, se alejan mucho del 0, lo que indica heterocedasticidad de los datos. En términos prácticos esto informa de que el SAR no se mostró como un instrumento fiable pues la concordancia entre las mediciones (test-retest) es muy inexacta. En el caso de las mujeres, sobre la base de los test de t de student intermedios necesarios para el cálculo de los gráficos de Bland-Altman, la prueba de SAR presentó diferencias estadísticamente significativas ( $p=.001$ ) entre la media de la diferencia ( $=8.20$ , figura 2.B) y el 0, que se establece como valor de referencia. Estos datos indicaron que no existió consistencia entre las puntuaciones obtenidas entre el test y el retest. Cuanto más próxima fue la media de la diferencia al valor 0, mejor se mostró la repetibilidad de la medida en general (Aldridge, et al., 2017). En esta figura también se pudo observar un sesgo importante, pues todos los valores se encontraron por encima del 0, ninguno negativo, lo que indica que todas las participantes disminuyeron su puntuación en el retest, excepto una persona que obtuvo la misma marca. Estos resultados pudieron ser



debidos a diferentes motivos o errores de medida derivados del uso de un test que resultó poco cómodo de realizar para estas personas como se analizará en la discusión de los resultados de viabilidad.

En el caso del DTF, al observar los gráficos, se aprecian datos por encima y por debajo de la media de la diferencia, ambas próximas a 0, así como puntuaciones repartidas uniformemente a ambos lados del cero, lo que muestra una variabilidad de error aleatorio en la que las mediciones no difieren en promedio. Solo se observó un outlier en el grupo de mujeres (Figura 1.B.). Tanto en SAR como en DTF los LdA fueron más estrechos en el caso de las mujeres, lo que indican mejores valores para ellas respecto de sus compañeros.

En general, el SAR se utiliza con frecuencia en estudios con personas sin DI (Ayala, et al., 2012). Sin embargo, más de la mitad de los participantes de este estudio tuvieron problemas para la ejecución correcta del SAR a pesar de las indicaciones de los evaluadores (SAR <50% vs DTF >75%). Las personas evaluadas no podían mantener la posición de partida exigida en el test debido a alteraciones de la estabilidad del equilibrio en sedestación por la acción limitadora de la musculatura isquiosural, las propias limitaciones de movilidad en la cadera, o al volumen del abdomen, entre otros. Es el caso del grupo de hombres, que no podían adoptar correctamente la posición de partida y alcanzar el cajón. Esta afirmación es producto de la constatación de la necesidad de muchos participantes de apoyar al menos una mano en el suelo en los diversos intentos para no caer hacia atrás, o hacia al lado, durante la ejecución del test. Por el contrario, la viabilidad fue mayor para el DTF. Se mostró como un test con menos exigencias en su técnica de ejecución, además permitió tener una gran estabilidad del equilibrio debido a la flexión de rodillas o la separación en los apoyos de ambos pies. También puede facilitar dicha ejecución su similitud con actividades de la vida cotidiana (Delgado-Valdivia, et al., 2009).

Por último, los resultados obtenidos tanto en EEM como en MCD y su comparación con las medias y DE de los resultados de la investigación (Tabla 2) deben ser estudiados en su conjunto para conocer los errores en la medida. Como se ha indicado anteriormente, el MCD es un indicador relevante a tener en cuenta a la hora de valorar la fiabilidad de un instrumento. Conocer este índice evitará, en estudios experimentales, caer en el error de afirmar que existe un incremento en la capacidad cuando en realidad se trata de un error de las medidas no asociado al efecto de un factor. Similares conclusiones se obtuvieron en una revisión realizada para per-

sonas mayores en donde se encontraron diversos estudios con fiabilidad baja o muy baja del SAR como predictor de la capacidad de elongación muscular (Mayorga-Vega, et al., 2015). Con relación a esto, hay estrategias que podrían aumentar la viabilidad y la fiabilidad del test. El control de las medidas antropométricas es, como bien se sabe en la literatura específica (Mayorga-Vega, et al, 2015), un factor determinante para la fiabilidad y la validez de los datos en los trabajos de investigación. En el DTF y, sobre todo, en el SAR la diferencia entre la longitud de los brazos (distancia acromio-punta de los dedos) y la de las piernas (distancia cadera-talón) podrían conocerse previamente con el fin de ponderar los resultados, solucionando en parte el problema (Gómez-Píriz, 2011). Se propone para futuros análisis controlar estos factores antropométricos y relacionarlos con los resultados obtenidos, así como evaluar otros test de ADM que tengan en cuenta las características particulares de las personas con DI y que estén adaptados o diseñados para ellas.

Ha sido una limitación general no contar con estudios previos similares con los que comparar los resultados con test específicos para personas con DI y no ponderarlos con respecto a las características particulares de los sujetos de investigación. Por otro lado, en muestras pequeñas, los datos de los gráficos de Bland-Altman son muy sensibles a los cambios. Además, haber realizado ambos test a todos los grupos hubiera arrojado resultados más sólidos.

## Conclusiones

Teniendo en cuenta los resultados del estudio, se puede concluir que el DTF se ha mostrado como un instrumento más adecuado para la valoración de la ADM en personas con DI. Se hace evidente la necesidad de incluir la valoración de las propiedades psicométricas de los instrumentos de valoración utilizados en los estudios realizados en personas con DI, tanto experimentales como descriptivos.

## Agradecimiento

Los autores quieren agradecer a las personas participantes, a sus familias, a todas las voluntarias de la investigación, así como al personal de los centros diurnos, su participación en el estudio.

## Conflicto de intereses

No existen

## Referencias

- Aldridge VK, Dovey, T.M., & Wade, A. (2017). Assessing Test-Retest Reliability of Psychological Measures Persistent Methodological Problems. *European Psychologist*, 22, 207-218. <https://doi.org/10.1027/1016-9040/a000298>.
- Atkinson, G., & Nevill, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sport Medicine*, 26(4), 217-38. <https://doi.org/10.2165/00007256-199826040-00002>
- Ayala, F., Sainz de Baranda, P., de Ste Croix, M., & Santonja, F. (2012). Fiabilidad y validez de las pruebas sit-and-reach: revisión sistemática. *Revista Andaluza de Medicina del Deporte*, 5(2), 57-66. [https://doi.org/10.1016/S1888-7546\(12\)70010-2](https://doi.org/10.1016/S1888-7546(12)70010-2)
- Ayán-Pérez, C., Martínez-Lemos, I., Lago-Ballesteros, J., Cancela Carral, J.M., & Loira-Camiña, N. (2016). Reliability and Validity of Physical Fitness Field-Based Test in Down Syndrome: A Systematic Review. *Journal of Policy and Practice in Intellectual Disabilities*, 13(2):142-156. <https://doi.org/10.1111/jppi.12155>
- Ayán-Pérez, C., Martínez-Lemos, R.I., & Cancela-Carral, J.M. (2017). Reliability and convergent validity of the 6-min run test in young adults with Down syndrome. *Disability and Health Journal*, 10(1):105e113. <https://doi.org/10.1016/j.dhjo.2016.07.004>.
- Bajo, S. (2003). La flexibilidad y la educación física escolar: evolución y aplicación en la escuela. En: M Guillén del Castillo, (Ed). *Medicina deportiva y educación física en edad escolar*. (pp. 421-40). Universidad de Córdoba, Servicio de Publicaciones.
- Boer, P.H., & Moss, S.J. (2016). Test-retest reliability and minimal detectable change scores of twelve functional fitness test in adults with Down syndrome. *Research in Developmental Disabilities*, 48, 176-185. <https://doi.org/10.1016/j.ridd.2015.10.022>
- Cabeza-Ruiz, R. (2020). Considerations for the Design of a Physical Fitness Battery to Assess Adults with Intellectual Disabilities: Preliminary Reference Values for the SAMU DIS-FIT Study. *International Journal of Environmental Research and Public Health*, 17, 9280. <https://doi.org/10.3390/ijerph17249280>
- Cabeza-Ruiz, R., Sánchez-López, A.M., Trigo, M.E., & Gómez-Píriz, P.T. (2020). Feasibility and reliability of the Assessing Levels of Physical Activity health-related fitness test battery in adults with intellectual disabilities. *Journal of Intellectual Disability Research*, 64, 612–628. <https://doi.org/10.1111/jir.12756>
- Cuesta-Vargas, A.I., Paz-Lourido, B., & Rodríguez, A. (2011). Physical fitness profile in adults with intellectual disabilities: Differences between levels of sport practice. *Research in Developmental Disabilities*, 32, 788-794. <https://doi.org/10.1016/j.ridd.2010.10.023>.
- Delgado-Valdivia, O., Martín-Cañada, M.A., Zurita-Ortega, F., Antequera-Rodríguez, J.J. & Fernández-Sánchez, M. (2009) Changes in flexibility according to gender and educational stage. *Apunts Med Esport*, 44, 161, 10-7. [https://doi.org/10.1016/s1886-6581\(09\)70103-3](https://doi.org/10.1016/s1886-6581(09)70103-3)
- European Commission. Directorate-General for Research and Innovation. (2012). *El género en la investigación: manual*. Publications Office. <https://data.europa.eu/doi/10.2777/23655>
- Gempp, R. (2006). El error estándar de medida y la puntuación verdadera de los test psicológicos: Algunas recomendaciones prácticas. *Terapia Psicológica*, 24(2), 117–130. Recuperado Abril 13, 2021, de <https://www.redalyc.org/articulo.oa?id=78524201>
- Geraldes, A.A.R., Albuquerque, R.B., Soares, R.M., Carvalho, J. & Farinatti, P.T.V. (2008). Correlação entre flexibilidade das articulações glenoumerais e coxofemorais e o desempenho funcional de idosas fisicamente ativas. *Brasilian Journal of Physical Therapy*, 12(4). <https://doi.org/10.1590/S1413-35552008000400005>
- Gidron, Y. (2013). Reliability and Validity. In: Gellman MD, Turner JR (eds) *Encyclopedia of Behavioral Medicine* (pp. 1643-1644). Springer.
- Ginty, A.T. (2013) Psychometric Properties. In: Gellman M.D., Turner J.R. (eds) *Encyclopedia of Behavioral Medicine* (pp. 1563–1564). Springer.
- Gómez-Píriz, P.T. (2011). *El Entrenamiento Deportivo en el Siglo XXI*. Fundación Alcala.
- Guimarães, J.M.N. & Farinatti, P.T.V. (2005). Análise descritiva de variáveis teóricamente associadas ao risco de quedas em mulheres idosas. *Revista Brasileira de Medicina do Esporte*, 11(5), 299-305. <https://doi.org/10.1590/S1517-86922005000500011>
- Hilgenkamp, T.I.M., van Wijck, R., & Evenhuis, H.M. (2012). Low physical fitness levels in older adults with ID: Results of the HA-ID study. *Research in Developmental Disabilities*, 33, 1048–1058. <https://doi.org/10.1016/j.ridd.2012.01.013>
- Hilgenkamp, T.I.M., van Wijck, R., & Evenhuis, H.M.

- (2012b). Feasibility and reliability of physical fitness test in older adults with intellectual disability: A pilot study. *Journal of Intellectual & Developmental Disability*, 37(2), 158–162. <https://doi.org/10.3109/13668250.2012.681773>
- Kennedy, D.J., & Noh, M.Y. (2011). The role of core stabilization in lumbosacral radiculopathy. *Physical Medicine and Rehabilitation Clinics of North America*, 22(1), 91-103. <https://doi.org/10.1016/j.pmr.2010.12.002>
- Matos-Duarte, M., Martínez de Haro, V., Sanz-Arribas, I., & Berlanga, L.A. (2022). El estilo de vida como condicionante de la flexibilidad del adulto mayor. *Retos*, 43, 283-289. <https://doi.org/10.47197/retos.v43i0.88752>
- Matos-Duarte, M., Martínez de Haro, V., Sanz-Arribas, I., Andrade, A.G.P., & Chagas, M.H. (2017). Estudio longitudinal de la flexibilidad funcional en mayores físicamente activos. *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, 17(65), 121-137. <http://dx.doi.org/10.15366/rimcafd2017.65.008>
- Mayorga-Vega, D., Viciano-Ramirez, J., Cocca, A., Becerra-Fernandez, C., & Merino-Marban, R. (2015). Criterion-related validity of sit-and-reach test for estimating hamstring and low back flexibility among elderly: A systematic review. *Journal of Sport and Health Research*, 7(1), 1-10. Recuperado Jun 20, 2021, de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3918544/>
- Oppewal, A., & Hilgenkamp, T.I.M. (2020). Adding meaning to physical fitness test results in individuals with intellectual disabilities. *Disability and Rehabilitation*, 42(10), 1-8. <https://doi.org/10.1080/09638288.2018.1527399>
- 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>
- Sundahl, L., Zetterberg, M., Wester, A., Rehn, B., & Blomqvist, S. (2016). Physical activity levels among adolescent and young adult women and men with and without intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 29, 93-98. <https://doi.org/10.1111/jar.12170>
- Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231–240.
- Wouters, M., Van Der Zanden, A.M., Evenhuis, H.M., & Hilgenkamp, T.I.M. (2017). Feasibility and reliability of test measuring health-related physical fitness in children with moderate to severe levels of intellectual disability. *American Journal of Intellectual and Developmental Disabilities*, 122(5), 422-438. <https://doi.org/10.1352/1944-7558-122.5.422>
- Zurita-Ortega, F.; Romero-Cerezo, C.; Ruiz-Rodríguez, L.; Martínez-Martínez, A.; Fernández-García, R., & Fernández-Sánchez, M. (2008). Influencia de las alteraciones raquídeas en la flexibilidad de los escolares. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 8(32), 282-298. Recuperado Jun 20, 2021, de <https://cdeporte.rediris.es/revista/revista32/artalteraciones87.htm>

