

## Propiedades psicométricas de la versión peruana del Test de Desarrollo Motor Grueso – Tercera Edición

### Psychometric properties of the Peruvian version of the Gross Motor Development Test - Third Edition

\*Angel Anibal Mamani-Ramos, \*Edgar Froilan Damian-Nuñez, \*,\*\*\*Fred Torres-Cruz, \*Carlos Wyly Dextre-Mendoza, \*Mario Alcarraz-Curi, \*Jorge Alber Quisocala-Ramos, \*\*Yeferson Anibal Mamani-Cari, \*Fernando Renato Roncal-Serpa, \*\*\*Henry Quispe-Cruz, \*\*\*Alexander Paucar-Pancca, \*Paloma Magdalena Montoya-Castillo  
\*Universidad Nacional Mayor de San Marcos (Perú), \*\*Pontificia Universidad Católica del Perú (Perú), \*\*\*Universidad Nacional del Altiplano de Puno (Perú)

**Resumen.** El Test de Desarrollo Motor Grueso - Tercera Edición (TGMD-3) evalúa la competencia de las habilidades motrices fundamentales de niños de 3 a 10 años. El estudio tuvo como objetivos traducir el TGMD-3 al español peruano y analizar las propiedades psicométricas para niños peruanos de 6 a 10 años. Se realizó una traducción transcultural del TGMD-3 a fin de contar con una versión en español peruano. El estudio involucró 19 profesionales y 348 niños peruanos de 6 a 10 años reclutados de las tres regiones naturales del Perú (costa, sierra y selva). La concordancia de valoración de los ítems entre los expertos fue apropiada (Índice de Validez de Contenido > 94.4 %; Coeficiente Kappa > .802). El análisis factorial confirmatorio corroboró la validez del modelo de dos factores (chi-cuadrado sobre grados de libertad = 1.618; índice de ajuste comparativo = .905; error de aproximación cuadrático medio = .042; residuo cuadrático medio estandarizado = .056; índice de bondad de ajuste = .998), mostrando una validez de constructo aceptable. La validez concurrente ratificó las carencias o potencialidades motoras que presentaron los niños (-.68,  $p < .001$ ). Finalmente, los índices de fiabilidad obtenidos (consistencia interna = .851, test-retest = .951, inter-evaluador = .963, intra-evaluador = .983) cumplieron con los estándares más exigentes. Los resultados confirmaron que el TGMD-3 versión peruana es apropiado para evaluar la competencia de las habilidades motrices fundamentales de niños peruanos.

**Palabras clave:** Habilidades motrices fundamentales; desarrollo motor grueso; niños; psicometría.

**Abstract.** The Test of Gross Motor Development - Third Edition (TGMD-3) assesses the fundamental motor skills competence of children aged 3 to 10 years. The objectives of the study were to translate the TGMD-3 into Peruvian Spanish and to analyze its psychometric properties for Peruvian children aged 6 to 10 years. A cross-cultural translation of the TGMD-3 was carried out in order to have a Peruvian Spanish version. The study involved 19 professionals and 348 Peruvian children aged 6 to 10 years recruited from the three natural regions of Peru (coast, highlands and jungle). Item rating agreement among experts was appropriate (Content Validity Index > 94.4 %; Kappa Coefficient > .802). The confirmatory factor analysis corroborated the validity of the two-factor model (chi-square on degrees of freedom = 1.618; comparative fit index = .905; root mean square error of approximation = .042; standardized root mean residual = .056; goodness-of-fit index = .998), showing acceptable construct validity. Concurrent validity ratified the motor deficits or potentialities presented by the children (-.68,  $p < .001$ ). Finally, the reliability indices obtained (internal consistency = .851, test-retest = .951, inter-rater = .963, intra-rater = .983) met the most demanding standards. The results confirmed that the TGMD-3 Peruvian version is appropriate for assessing the fundamental motor skills competence of Peruvian children.

**Keywords:** Fundamental motor skills; gross motor development; children; psychometrics.

Fecha recepción: 27-05-23. Fecha de aceptación: 22-08-23

Angel Anibal Mamani-Ramos  
amamanir@unmsm.edu.pe

### Introducción

La evaluación oportuna del estado de las habilidades motrices fundamentales (HMF) permite tener una visión clara de la progresión secuencial (Goodway et al., 2019), retrasos o déficits motores (Hulsteen et al., 2020; Magistro et al., 2020) en niños de todas las culturas (Ulrich, 2017), tanto a profesores de Educación Física (EF), entrenadores, profesionales de la salud e investigadores. Las HMF son aquellas capacidades físicas que permiten a los niños moverse de diversas maneras (habilidades locomotoras) e interactuar con los objetos (habilidades manipulativas) (Bautista et al., 2020). Las HMF son los cimientos de los patrones de movimiento que conducen a la ejecución de secuencias de movimientos complejos y más avanzados del contexto (Hulsteen et al., 2020; Mamani-Ramos et al., 2021; Rey et al., 2020), esenciales para participar con éxito en actividades físico-deportivas, recreativas y dancísticas (Gallahue et al., 2012; Wagner et al., 2017) a lo largo de la vida (Estevan et al., 2017; Goodway et al., 2019; Mohammadi et al., 2018) con

el fin de contribuir al desarrollo sostenible de la salud pública (Eddy et al., 2020; Hulsteen et al., 2020).

La evaluación y el desarrollo de las HMF son factores clave de los programas de EF en educación preescolar y primaria en todo el mundo (Duncan et al., 2022; Magistro et al., 2020; Mohammadi et al., 2019). Por tal razón, los profesores de EF, principalmente, los profesionales de la salud y los investigadores requieren herramientas de evaluación válidas y fiables para diagnosticar, clasificar y monitorear los cambios del estado de las HMF de los niños (Asunta et al., 2017; Magistro et al., 2020). De acuerdo con los estudios de revisión sistemática que realizaron Eddy et al. (2020) y Hulsteen et al. (2020) respecto de la validez y la fiabilidad de herramientas de evaluación para medir las HMF en niños de edad escolar resalta el TGMD como una de las herramientas más investigadas que presenta resultados sólidos de validez y fiabilidad en sus tres ediciones. La primera edición del TGMD fue propuesto por Ulrich en la década de los 80 del siglo pasado (Ulrich, 1985). La segunda se publicó en el año 2000 (Ulrich, 2000). Y la tercera (y última edición) fue

publicada formalmente el año 2019 (Ulrich, 2019). La actualización del TGMD se realizó con un intervalo promedio de 15 años, conforme a las recomendaciones de la American Educational Research Association et al. (2014), quienes señalan que los tests educativos y psicológicos estandarizados referenciados a normas deben revisarse y actualizarse según este período de años para reflejar los cambios que experimentan las personas con el tiempo. Entre lo más destacado que se actualizó en el TGMD fue: (1) La denominación del subtest de control de objetos cambió por habilidades con la pelota, debido a que todas las habilidades se desarrollan con una pelota. (2) El aumento en el número de habilidades de los subtests, así entre el TGMD y el TGMD-3 existe una gran diferencia, ya que en el primero se tienen siete habilidades de locomoción y cinco de control de objetos; mientras que en este último se tienen seis habilidades de locomoción y siete habilidades con la pelota, a diferencia del TGMD-2 que contempla seis habilidades por cada subtest. (3) La muestra normativa, pues el TGMD-3 y el TGMD-2 cuentan con una muestra representativa de la población estadounidense según región geográfica, sexo, raza, nivel educativo de los padres, entre otros; lo que no se dio así en el TGMD, razón por la cual fue criticado. (4) Las puntuaciones normativas, dado que el TGMD-3 presenta puntuaciones actualizadas que garantizan una mayor exactitud en la interpretación de resultados que las puntuaciones del TGMD-2 y, más aún, del TGMD (Ulrich, 2019).

En definitiva, el TGMD-3 es una herramienta de evaluación de las HMF con mejores atributos que brinda confianza y que goza de reconocimiento por la comunidad científica (Garn y Webster, 2021; Mohammadi et al., 2018; Wagner et al., 2017) para su uso en entornos educativos, clínicos y de investigación (Brian et al., 2017; Rey et al., 2020). Sin embargo, el TGMD-3 como tal no puede ser utilizado libremente en otros países, debido a que la referencia normativa es aplicable a niños estadounidenses de edades comprendidas entre los 3 y 10 años (Brian et al., 2017). En esa línea, para considerar la misma referencia normativa y usar el TGMD-3 en otros países, antes se debe evaluar la validez y la fiabilidad del mismo (Estevan et al., 2017; Magistro et al., 2020; Mohammadi et al., 2018). Si los resultados son positivos, entonces, se puede considerar la misma referencia normativa y usar el test. Actualmente, el TGMD-3 evidencia validez y fiabilidad para evaluar la competencia de las HMF en niños de Alemania (Wagner et al., 2017), España (Estevan et al., 2017), Brasil (Valentini et al., 2017, 2021), Irán (Arman et al., 2021; Mohammadi et al., 2018), Italia (Magistro et al., 2020), Irlanda (Duncan et al., 2022), Canadá (Maiano et al., 2022), Eslovenia (Marinšek et al., 2023), entre otros países.

Desde esta perspectiva, y siguiendo las recomendaciones de Ulrich (2019) en cuanto a seguir investigando el TGMD-3 con diferentes muestras por medio de distintos procedimientos estadísticos se evaluó la validez y fiabilidad del TGMD-3 para niños peruanos, considerando que en el Perú no se cuenta con ninguna propuesta de evaluación que evidencie validez y fiabilidad para medir las

HMF en niños. Tampoco se cuenta con una herramienta de evaluación propuesta en otro contexto que evidencie validez y fiabilidad para el contexto peruano. En ese sentido, urge la necesidad de contar con una herramienta de evaluación válida y fiable que mida las HMF de niños peruanos, de manera que esta sea útil, por un lado, para diagnosticar y monitorear permanentemente el desarrollo de las HMF de los niños, sobre todo en las escuelas dentro de las clases de Educación Física (Brian et al., 2017; Stodden et al., 2008); y, por otro, para desarrollar programas de Educación Física más eficientes (Ulrich, 2019). Desde este panorama, el estudio tuvo como objetivos: (1) traducir el TGMD-3 al español peruano y (2) analizar las propiedades psicométricas mediante la validez de contenido, validez de constructo, validez concurrente y fiabilidad del TGMD-3 para niños peruanos de 6 a 10 años.

## Métodos

### Participantes

Participaron 19 profesionales: (1) cinco traductores bilingües; (2) 10 profesores universitarios expertos en conducta motriz (cinco doctores y cinco maestros) para evaluar la validez de contenido; (3) dos profesores de Educación Física (un doctor y un licenciado) para el proceso de validez aparente; y (4) dos profesores de Educación Física especialistas en conducta motriz (un doctor y un licenciado) para el proceso de fiabilidad inter-evaluador. Asimismo, participaron 348 niños con un desarrollo típico (51.44 % hombres y 48.56 % mujeres), de edades comprendidas entre los 6 y 10 años (edad hombres:  $M = 8.19$ ,  $DE = 1.32$ ; edad mujeres:  $M = 8.25$ ,  $DE = 1.32$ ), que fueron reclutados de las tres regiones naturales geográficas del Perú (costa, sierra y selva), tres departamentos (Lima, Puno y Loreto), tres ciudades y cuatro centros escolares (3 públicos y 1 privado). Para los distintos análisis de resultados, se utilizaron las siguientes submuestras elegidas al azar: 65 niños para la validez concurrente (36 hombres y 33 mujeres), 88 niños para la fiabilidad intra-evaluador (52 hombres y 36 mujeres) y 99 niños para la fiabilidad inter-evaluador y test-retest (53 hombres y 46 mujeres).

### Instrumentos

#### TGMD-3

Se utilizó el TGMD-3 desarrollado por Ulrich (2019) en niños estadounidenses pertenecientes a las regiones geográficas de oeste, medio oeste, noreste y sur. El test tiene por finalidad evaluar las HMF en niños de edades comprendidas entre los 3 años y cero meses (3-0), y los 10 años y 11 meses (10-11). El test tiene dos subtests (factores): el primero, subtest de locomoción, que mide seis habilidades motoras gruesas que implican movimientos fluidos y coordinados del cuerpo cuando el niño se mueve de un lugar a otro (correr, galope, salto en un pie, brinco, salto horizontal y salto lateral); el segundo, subtest de habilidades con la pelota, el cual mide siete habilidades motoras gruesas que demuestran dominio de los movimientos de golpeo, recepción y

lanzamiento (bateo a dos manos de una pelota inmóvil, golpe con raqueta a una mano luego del rebote de la pelota, rebote de la pelota con una mano en el mismo lugar, atrapar con dos manos, patear una pelota inmóvil, lanzamiento por encima de la cabeza y lanzamiento por debajo de la cadera). Cada una de las 13 habilidades tienen entre tres a cinco criterios de desempeño. Estos criterios son componentes conductuales (ítems) que representan el patrón de movimiento apropiado de la habilidad. La evaluación comprende: (1) una demostración y descripción verbal precisa de la habilidad por realizar por el examinador, (2) un ensayo de práctica del niño para corroborar que entendió lo que debe hacer (si es necesario realizar un ensayo adicional) y (3) dos ejecuciones formales de la habilidad. Cada habilidad incluye varios componentes conductuales (ítems). Si el niño realiza un componente conductual (ítem) correctamente, el examinador debe puntuarlo con uno y si no lo realiza correctamente debe puntuarlo con cero. La suma de los criterios de desempeño de cada habilidad comprende las puntuaciones brutas de cada subtest. La puntuación total del subtest de locomoción oscila entre 0 y 46, y la puntuación total del subtest habilidades con la pelota oscila entre 0 y 54. La puntuación del compuesto (combinación de las puntuaciones de los subtest) oscila entre 0 y 100. El TGMD-3 cumplió con los estándares de fiabilidad (coeficiente alfa .93; test-retest .90; inter-evaluador .96; intra-evaluador .98) y validez (análisis factorial confirmatorio:  $\chi^2/gl = 3.730$ ; TLI = .962; NFI = .958; RMSEA = .056) de un instrumento. Por lo tanto, quienes utilicen el test pueden tener confianza en los resultados.

#### *Cuestionario de Observación Motora para Profesores (MOQ-T)*

El MOQ-T evalúa la conducta motora gruesa (funcionamiento motor general) y fina (escritura a mano) de niños con riesgo de trastorno del desarrollo de la coordinación, de edades entre 5 y 11 años. El cuestionario está compuesto por 18 ítems. Cada ítem es valorado en una escala de cuatro puntos (1 = nunca es cierto, 2 = rara vez es cierto, 3 = casi siempre es cierto, 4 = siempre es cierto). Las puntuaciones más altas reflejan un mayor riesgo de problemas motores. El cuestionario presentó validez ( $\chi^2/gl = 2.899$ ; TLI = .964; RMSEA = .047) y fiabilidad (general  $\alpha = .96$ , funcionamiento motor general  $\alpha = .89$ , escritura a mano  $\alpha = .63$ ) (Asunta et al., 2017).

El objetivo principal del TGMD-3 es identificar a los niños que presentan un retraso significativo con respecto de sus compañeros en el desarrollo de la motricidad gruesa (Ulrich, 2019). En esa perspectiva, el MOQ-T también identifica el retraso o no del desarrollo motor de los niños, y fue considerado para la validez concurrente, pero solo el factor de funcionamiento motor general (14 ítems).

#### *Adaptación transcultural*

El proceso de adaptación transcultural del TGMD-3 (formato de registro del examinador) del idioma inglés (Ulrich, 2019) al español peruano se desarrolló siguiendo la línea de Beaton et al. (2000).

Fase 1, traducción inicial: 3 traductores bilingües independientes (2 dos traductores expertos en Educación Física y conducta motriz, y 1 traductor que desconocía la temática y el objetivo del test), cuyo idioma nativo fue el español, tradujeron el TGMD-3 del inglés al español.

Fase 2, síntesis de las traducciones: los traductores y el investigador principal se reunieron para sintetizar las versiones traducidas. Cada punto de traducción diferente fue discutido y resuelto por consenso, completando así la versión común del TGMD-3.

Fase 3, traducción inversa: 2 traductores bilingües independientes que desconocían el TGMD-3 original, cuyo idioma nativo fue el español, tradujeron la versión común en español al inglés.

Fase 4, comité de expertos: el investigador principal, 2 investigadores colaboradores, 1 corrector de texto y 5 traductores (de ida y vuelta) revisaron las traducciones y resolvieron por consenso todas las discrepancias para preservar la equivalencia (semántica, idiomática, experiencial y conceptual) entre la versión original y la versión meta del TGMD-3. Luego de este proceso, esta versión consensuada fue enviada a 10 profesores expertos en conducta motriz para evaluar la validez de contenido (simplicidad, claridad del lenguaje y relevancia). Finalmente, el investigador principal y los expertos revisaron y resolvieron por consenso pequeñas discrepancias. Al concluir esta fase se logró obtener la versión prefinal del TGMD-3.

Fase 5, prueba de la versión prefinal: el investigador principal y dos profesores de Educación Física aplicaron el TGMD-3 a 35 niños entre 6 y 10 años (siete por cada año). Antes, durante y después de la aplicación del TGMD-3 no se encontraron inconvenientes.

Fase 6, finalización de la versión peruana del TGMD-3: todos los profesionales participantes en la adaptación transcultural (Inglés-Español) del TGMD-3, dirigido por el investigador principal, generaron el TGMD-3 versión peruana (TGMD-3-PE).

#### *Procedimientos*

Este estudio fue ganador del concurso del Programa de Proyectos de Investigación para Grupos de Investigación 2020 (Resolución Rectoral N.º 01686-R-20), promovido por la Universidad Nacional Mayor de San Marcos (UNMSM), Lima, Perú. Asimismo, fue aprobado por el Comité de Ética en Investigación de la Facultad de Medicina de la UNMSM con código N.º 0068-2022, y respetó los principios éticos fundamentales para la investigación con seres humanos establecido en la Declaración de Helsinki (World Medical Association, 2013). También, cuenta con la autorización del director de las escuelas primarias participantes. Los padres/tutores fueron informados del estudio por carta. Todos los padres/tutores que estuvieron de acuerdo firmaron un formulario de consentimiento informado, mientras que los niños dieron su asentimiento verbal para participar en el estudio. Los datos de los niños fueron asociados a un número de serie con el fin de garantizar su anonimato.

Los datos fueron recogidos entre el 20 de junio y el 25 de noviembre de 2022, en horario escolar, durante las clases de EF (duración de la clase entre 80 a 90 minutos), con la presencia de los profesores. Se solicitó a los padres/tutores completar un cuestionario sociodemográfico (condición del centro escolar, grado de estudio del menor, sexo, fecha de nacimiento). El equipo responsable de administrar el TGMD-3 estuvo integrado por dos aplicadores principales (responsables de la demostración y la descripción verbal de las habilidades por realizar), un camarógrafo (responsable de la grabación y codificación de videos vista sagital, con una cámara Sony FDR-AX40/BC Serie 3437809, un trípode Vivitar VIV-VPT-6072 y una tarjeta Sandisk SDS-DUNC-128G), un asistente de materiales (responsable de acomodar y facilitar materiales oportunamente) y un organizador de niños (responsable de mantener el orden y animar a los niños para que observen y descansen). Todo el equipo responsable fue capacitado en dos talleres de dos horas por el investigador principal. En el primer taller, se les brindó instrucciones sobre el protocolo del test y responsabilidad en su función. Y en el segundo, se hizo una simulación de aplicación del test con cinco niños de siete años, donde se reforzó algunos detalles que no estaban muy claros.

El test fue aplicado entre 10 a 15 niños en dos clases de EF (clase 1, locomoción, tiempo aproximado 15 a 20 minutos; clase 2, habilidades con la pelota, tiempo aproximado 35 a 40 minutos) en los polideportivos de las escuelas participantes (cubierto y abierto). Previo a la aplicación del test se realizaron varios ensayos grabados con 2, 5, 10, 15, 20 y 25 niños; de lo cual resultó más apropiado, cómodo y ecológico (la vivencia de los niños es similar a un ambiente natural de una clase de Educación Física) aplicar el test entre 10 a 15 niños, aunque excepcionalmente se puede aplicar el test completo en una sola clase a niños de 8, 9 y 10 años (tiempo aproximado entre 50 a 60 minutos), siempre y cuando estos colaboren.

Durante la administración del test, cada una de las habilidades fue demostrada y descrita verbalmente por uno de los aplicadores principales. Luego, los niños fueron invitados por el aplicador a practicar una vez, seguido de las dos ejecuciones formales. Tras la aplicación del test y con base en el contenido de los videos, el investigador principal (con sólida experiencia en la evaluación de las habilidades del TGMD-2 a través del video y en vivo) evaluó cada una de las 13 habilidades (intento 1 y 2) de la muestra total para la validez de constructo (348) del TGMD-3-PE durante tres meses (septiembre, octubre y noviembre de 2022). Una submuestra de 99 niños repitió el test después de dos semanas para el análisis de fiabilidad test-retest. Otra submuestra también de 99 niños fue elegida para el análisis de fiabilidad inter-evaluador. La evaluación de esta submuestra estuvo a cargo de otro investigador, quien fue capacitado y contaba también con una sólida experiencia en la evaluación de las habilidades del TGMD-2 mediante video. Para la fiabilidad intra-evaluador fueron elegidos 88 niños para ser evaluados por segunda vez por el investigador principal después de dos meses. Por último, para la validez concurrente fueron

elegidos 65 niños para ser evaluados adicionalmente con el MOQ-T, evaluación que estuvo a cargo por los profesores de EF y de aula (las discrepancias fueron resueltas por consenso), ambos tenían dominio del idioma inglés.

### *Análisis de datos*

Para analizar la validez de contenido (simplicidad, claridad del lenguaje y relevancia) de las habilidades y criterios de desempeño del TGMD-3-PE, se utilizó el Índice de Validez de Contenido (IVC) (Tristán-López, 2008) y el Coeficiente Kappa (Neuendorf, 2002). El porcentaje del IVC mínimo aceptable de acuerdo entre los expertos fue de 80 % (Polit et al., 2007), y el valor del Coeficiente Kappa mínimo aceptable fue de .75 para considerar una concordancia excelente entre los expertos (Landis y Koch, 1977). Los datos atípicos de la muestra total fueron analizados mediante el método de distancia de Mahalanobis (Hair et al., 2019). Los datos que presentaron una  $p < .001$  fueron considerados como atípicos. La asimetría y curtosis se utilizaron para identificar la distribución normal de los datos, teniendo como índice mínimo aceptable el  $\pm 2$  (Gravetter y Wallnau, 2014). La estadística descriptiva fue analizada a través de la media (además de la desviación estándar) y correlación (coeficiente Spearman). La media permitió presentar las puntuaciones brutas de locomoción, habilidades con la pelota y compuesto en cinco intervalos de edad (seis, siete, ocho, nueve y 10 años) y la correlación permitió identificar la magnitud del efecto (.0 a .1 = muy pequeña o insustancial, .1 a .3 = pequeña, .3 a .5 = moderada, .5 a .7 = grande, .7 a .9 = muy grande, .9 a 1 = perfecta) (Hopkins, 2016) entre la edad y el desempeño de todas las puntuaciones brutas.

Para verificar la validez factorial del TGMD-3-PE, se realizó un análisis factorial confirmatorio (AFC) a través del método de Estimación de Mínimos Cuadrados Ponderados Diagonalmente (DWLS) con la matriz de correlación polidimensional, tomando como referencia cinco índices de ajuste. Además de los cortes que evidenciaron un buen ajuste del modelo:  $\chi^2/gl \leq 3$  (Hair et al., 2019),  $CFI \geq .90$  (Hair et al., 2010),  $RMSEA \leq .050$  (Browne y Cudeck, 1993),  $SRMR \leq .060$  (Brown, 2015), e  $GFI > .90$  (Hair et al., 2019). El tamaño de las cargas factoriales y la correlación entre factores fueron valorados con base en las magnitudes de efecto descritas líneas arriba, tal cual lo hizo el autor original del TGMD-3 (Ulrich, 2019).

Para verificar la validez concurrente se utilizó el coeficiente de correlación de Spearman, prueba que permitió medir la relación entre las puntuaciones brutas del desarrollo motor grueso total del TGMD-3-PE y el MOQ-T (factor de funcionamiento motor general). La magnitud del efecto del resultado fue valorada al igual que las demás correlaciones.

Para verificar la fiabilidad del TGMD-3-PE se utilizó tres tipos de varianzas de error: contenido, tiempo y calificador (Anastasi y Urbina, 1998). El coeficiente alfa se utilizó para medir la consistencia interna del test mediante método de Cronbach (Breakwell et al., 2006). El test-retest se utilizó para medir la estabilidad temporal que posee el test

mediante el coeficiente de correlación de Spearman (Anastasi y Urbina, 1998). La calificación inter-evaluador se utilizó para medir la coherencia de las puntuaciones por dos evaluadores que calificaron de forma independiente a los mismos participantes y la calificación intra-evaluador se utilizó para medir el grado de coherencia de las puntuaciones obtenidas en la evaluación del test en dos ocasiones (dos meses después) por el mismo evaluador, ambas pruebas se analizaron mediante el coeficiente intraclass (Rousson et al., 2002). Se estableció una fiabilidad de  $\geq .80$  como adecuada y  $\geq .90$  como estándar de fiabilidad (Nunnally y Bernstein, 1994).

El tratamiento estadístico se desarrolló mediante el software Microsoft Excel (IVC), IBM SPSS 26 (Distancia de Mahalanobis, media, desviación estándar, coeficiente Spearman para la correlación entre la edad y el desempeño de las puntuaciones brutas, alfa de Cronbach) y R 4.2.2. (Coeficiente Kappa, asimetría y curtosis, AFC, coeficiente Spearman para validez concurrente y test-retest, y coeficiente de correlación intraclass).

## Resultados

### Validez de contenido

La concordancia entre expertos se verificó a través del IVC y Coeficiente Kappa. Para la primera prueba, el resultado del TGMD-3-PE fue muy relevante, con valores superiores al 80 % de acuerdo entre los expertos (simplicidad 94.5 %, claridad del lenguaje 95.6 %, relevancia 96.3 %). El resultado del Coeficiente Kappa para el TGMD-3-PE presentó una concordancia excelente entre los expertos (simplicidad .803, claridad del lenguaje .862, relevancia .898) ( $p < .001$ ) (Tabla 1). La validez aparente desarrollada por dos profesores de EF presentó un resultado de 98 % de concordancia entre ellos.

Tabla 2.

Medias de las puntuaciones brutas (y desviaciones estándar) por edad para los sub-test y compuesto del TGMD-3-PE

Edad (en años)	n	Locomoción		Habilidades con la pelota		Compuesto	
		Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres
		M (DE)	M (DE)	M (DE)	M (DE)	M (DE)	M (DE)
6	45	33.82 (5.22)	32.96 (7.01)	39.18 (5.93)	32.61 (6.16)	73.00 (8.52)	65.57 (10.33)
7	66	33.08 (6.60)	34.65 (5.89)	40.95 (5.45)	35.00 (6.92)	74.03 (10.13)	69.65 (10.25)
8	76	34.83 (6.78)	35.05 (6.55)	41.74 (5.82)	36.78 (5.78)	76.57 (11.13)	71.83 (9.38)
9	89	34.63 (6.17)	36.33 (5.02)	44.35 (5.03)	39.60 (4.68)	78.98 (8.54)	75.93 (8.12)
10	72	37.27 (4.59)	37.56 (5.40)	45.89 (3.82)	41.75 (5.40)	83.17 (6.41)	79.31 (8.81)
Correlación con la edad		.217**	.232**	.402**	.464**	.377**	.444**

\*\* La correlación es significativa en el nivel .01 (bilateral).

Abreviaturas: n = tamaño de la muestra; M = media; DE = desviación estándar.

### Validez de constructo

El modelo del TGMD-3-PE de dos factores, siguiendo la línea del TGMD-3 original (Ulrich, 2019), presentó un buen ajuste ( $\chi^2/\text{gl} = 1.618$ ; CFI = .905; RMSEA = .042; SRMR = .056; GFI = .998), lo que confirma la validez del modelo

Tabla 3.

Índices de ajuste del AFC del TGMD-3-PE

Modelo	Índices de ajuste						
	$\chi^2$	gl	$\chi^2/\text{gl}$	CFI	RMSEA	SRMR	GFI
TGMD-3-PE (dos factores)	103.538	64	1.618	.905	.042	.056	.998

Abreviaturas:  $\chi^2/\text{gl}$  = chi-cuadrado sobre grados de libertad; CFI = índice de ajuste comparativo; RMSEA = error de aproximación cuadrático medio; SRMR = residuo cuadrático medio estandarizado; GFI = índice de bondad de ajuste.

Tabla 1.

IVC y Coeficiente de concordancia de Kappa

Criterios	IVC%	Kappa	p (Kappa)
Simplicidad	94.5	.803	.000
Claridad del lenguaje	95.6	.862	.000
Relevancia	96.3	.898	.000

Abreviaturas: % = porcentaje; p = significancia.

### Análisis exploratorio de datos

De acuerdo con el análisis de datos realizado con el método de distancia de Mahalanobis, se encontraron cinco datos atípicos ( $p < .001$ ) de 353, por lo que fueron eliminados, quedando así 348 participantes para la validez de constructo. Los valores de asimetría (-.30 a -1.87) se encontraron en el rango aceptable de una distribución normal ( $\pm 2$ ), excepto en una habilidad (atrapar con dos manos = -2.10). En cambio, los valores de curtosis no se encontraron en el rango aceptable (2.41 a 8.25). En general, los datos de las habilidades del TGMD-3-PE no presentaron una distribución normal.

### Estadística descriptiva

Las medias de las puntuaciones brutas de locomoción, habilidades con la pelota y compuesto se incrementaron a medida que los niños tenían más edad, excepto en las puntuaciones de locomoción entre hombres de 6 y 7 años (los niños de seis años presentaron una ligera ventaja de .47 años frente a los niños de siete años), y 8 y 9 años (los niños de ocho años presentaron una ligera ventaja de .02 años frente a los niños de nueve años). Asimismo, se encontraron correlaciones moderadas en habilidades con la pelota (hombres .402, mujeres .464) y compuesto (hombres .377, mujeres .444), y correlaciones pequeñas en locomoción (hombres .217, mujeres .232) (Tabla 2).

(Tabla 3). El tamaño de las cargas factoriales de las habilidades oscilaron entre grandes (.55), moderadas (.35 a .49) y pequeñas (.24), todas ellas estadísticamente significativas ( $p < .001$ ). La correlación entre factores (sub-tests de locomoción y control de objetos) fue grande (.56) (Figura 1).

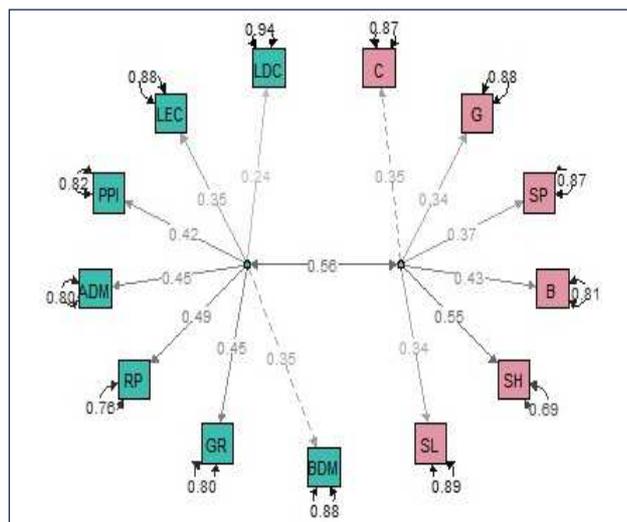


Figura 1. Cargas factoriales y correlación entre factores

Abreviaturas: C = Carrera; G = Galope; SP = Salto en un pie; B = Brinco; SH = Salto horizontal; SL = Salto lateral; BDM = Bateo a dos manos de una pelota inmóvil; GR = Golpe con raqueta a una mano luego del rebote de la pelota; RP = Rebote de la pelota con una mano en el mismo lugar; ADM = Atrapar con dos manos; PPI = Patear una pelota inmóvil; LEC = Lanzamiento por encima de la cabeza; LDC = Lanzamiento por debajo de la cadera.

**Validez concurrente**

Los resultados de la validez concurrente del TGMD-3-PE se expresaron entre la relación de la puntuación bruta del compuesto de este y la puntuación bruta del factor de funcionamiento motor general del MOQ-T, presentando una correlación negativa grande (-.608,  $p < .001$ ).

**Fiabilidad**

El TGMD-3-PE presentó una fiabilidad adecuada en el compuesto, locomoción y habilidades con la pelota ( $\geq .80$ ) en consistencia interna (.851, .826, .801, con  $p < .001$ ), test-retest (.951, .937, .951, con  $p < .001$ ), inter-evaluador (.963, .950, .966, con  $p < .001$ ) e intra-evaluador (.983, .988, .982). Con el valor de las tres últimas pruebas también se cumplió con los estándares de fiabilidad ( $\geq .90$ ) (Tabla 4).

Tabla 4. Fiabilidad del TGMD-3 según tres tipos de fiabilidad

TGMD-3-PE	Coeficiente Alfa	Tipos de coeficiente de fiabilidad		
		Test-retest	Calificación	
			Inter-evaluador	Intra-evaluador
Compuesto	.851	.951	.963	.983
Locomoción	.826	.937	.950	.988
Habilidades con la pelota	.801	.951	.966	.982

**Discusión**

Los objetivos de este estudio fueron: (1) traducir el TGMD-3 al español peruano y (2) analizar las propiedades psicométricas mediante la validez de contenido, validez de constructo, validez concurrente y fiabilidad del TGMD-3 para niños peruanos de 6 a 10 años. Respecto del primer objetivo, la traducción del TGMD-3 implicó la traducción del formato completo del examinador. Dicho formato incluye seis secciones: (1) datos de identificación, (2) rendimiento en los subtests, (3) rendimiento integral, (4) términos descriptivos, (5) lineamientos de administración y puntuación, y (6) registro del rendimiento de los ítems. Dentro

de esta última sección, se encuentran las 13 habilidades y sus respectivos criterios de desempeño, además de los materiales e instrucciones. La traducción se desarrolló siguiendo las seis fases (1 traducción inicial, 2 síntesis de las traducciones, 3 traducción inversa, 4 comité de expertos, 5 prueba de la versión prefinal, y 6 finalización de la versión peruana) para el proceso de adaptación transcultural propuesto Beaton et al. (2000), e involucró la participación de 19 profesionales (entre traductores bilingües, profesores universitarios y de EF especialistas en conducta motriz), además del investigador principal, dos investigadores colaboradores y un corrector de texto. Todo el trabajo e interacción realizada propiciaron la versión final del TGMD-3-PE. Es cierto que la validez y fiabilidad del TGMD-3-PE apunta directamente a las habilidades y criterios de desempeño (además de los materiales e instrucciones); sin embargo, un test como este no podría ser utilizado solo con esa información por los interesados, motivo por el que la traducción abarcó el formato del examinador en su integridad.

Con relación al segundo objetivo, la validez de contenido (simplicidad, claridad del lenguaje y relevancia) de las habilidades y criterios de desempeño del TGMD-3-PE evidenció resultados sólidos en la concordancia entre los expertos, tanto en el IVC (Polit et al., 2007), Coeficiente Kappa (Landis y Koch, 1977) y validez aparente; lo que confirma que el TGMD-3-PE tiene un contenido apropiado para evaluar las HMF de niños; por lo tanto, se cumplió satisfactoriamente con el primer paso crítico previo a la validez de constructo (Haynes et al., 1995).

La validez de constructo del TGMD-3-PE se realizó siguiendo el modelo de dos factores del autor original del TGMD-3 (Ulrich, 2019), a través del AFC. Los valores obtenidos en los índices de ajuste confirmaron la validez del modelo de dos factores del TGMD-3-PE (Brown, 2015; Browne & Cudeck, 1993; Hair et al., 2010, 2019). Todas las cargas factoriales son positivas y significativas. La correlación entre factores es adecuada para un modelo de dos factores. Estos dos últimos resultados proporcionan un apoyo adicional al modelo (Cronbach, 1989). El resultado del AFC del modelo de dos factores del TGMD-3-PE está alineado con el estudio original del TGMD-3 (Ulrich, 2019) y con estudios previos de validez del TGMD-3 realizado en otros países, como en Alemania (Wagner et al., 2017), España (Estevan et al., 2017), Brasil (Valentini et al., 2017), Irán (Mohammadi et al., 2018), Italia (Magistro et al., 2020), Canadá (Maïano et al., 2022), entre otros. Otro aspecto que corrobora la validez del modelo de dos factores del TGMD-3-PE son las puntuaciones brutas y las correlaciones entre las puntuaciones brutas y la edad. Las puntuaciones brutas del compuesto, habilidades con la pelota y locomoción de mujeres se incrementa con la edad, lo que indica que las HMF siguen una trayectoria de desarrollo normal, que mejora a medida que se tiene más años (Allen et al., 2017). Lo que llama la atención en estos resultados son las puntuaciones de locomoción de hombres, entre los niños de 6 y 7 años, y 8 y 9 años. Los niños de 7 y 9 años presentan una ligera ventaja frente a sus menores; esto puede deberse

a la cantidad y calidad de experiencias motrices adquiridas por algunos niños (Temple y Foley, 2017). Esta situación también se percibe en otros estudios similares (Valentini et al., 2017; Wagner et al., 2017). En general, las puntuaciones brutas de locomoción de hombres se incrementan, los niños de 8 y 9 años presentan mayor puntuación que los de 6 y 7 años, y los de 10 años presentan mayor puntuación que los de 8 y 9 años. Las correlaciones entre las puntuaciones brutas y la edad van de moderado a pequeño.

La validez concurrente confirma que tanto el TGMD-3-PE como el MOQ-T (factor de funcionamiento motor general) miden de forma similar el desarrollo motor grueso total; es decir, las carencias o potencialidades de las HMF que presentaron los niños pudieron observarse en ambos instrumentos. Por lo tanto, este resultado ratifica la pertinencia del TGMD-3-PE para evaluar las HMF de niños.

Respecto de la fiabilidad, los resultados del coeficiente alfa, test-retest, inter-evaluador e intra-evaluador confirman que el TGMD-3-PE cumple con los estándares de fiabilidad más exigentes (Nunnally y Bernstein, 1994), al igual que el TGMD-3 original (Ulrich, 2019). La consistencia interna refleja una alta homogeneidad de las habilidades y criterios de desempeño (Breakwell et al., 2006). El test-retest confirma la estabilidad temporal de los resultados (Anastasi y Urbina, 1998). La calificación inter-evaluador e intra-evaluador confirman un acuerdo sustancial entre examinadores y entre las evaluaciones del mismo examinador (Rousson et al., 2002). Los resultados de validez y fiabilidad del TGMD-3-PE garantizan para que este se utilice con confianza. La fortaleza principal del estudio es que incluyó a niños de las tres regiones naturales del Perú; es decir, de la costa, la sierra y la selva. Además, incluyó a niños que cursan sus estudios del nivel primaria en instituciones públicas y privadas, con lo que se contó con la participación de niños que pertenecen a distintos niveles socioeconómicos. En ese sentido, conforme a los resultados, el TGMD-3-PE puede ser utilizado en niños de la costa, la sierra o la selva del contexto peruano y de distintos niveles socioeconómicos.

Es fundamental exponer las limitaciones de este estudio. La principal limitación es no haber incorporado en el estudio a niños de 3 a 5 años que estudian en preescolar; por tanto, no se ha confirmado la validez y fiabilidad del TGMD-3-PE para todos los grupos de edad que considera el TGMD-3, que es de 3 a 10 años (Ulrich, 2019). En ese sentido, se presenta una validez y fiabilidad del TGMD-3-PE solo para niños de 6 a 10 años que cursan sus estudios en primaria, por lo que en posteriores estudios se debería cubrir este vacío. Asimismo, otra limitación del estudio es no incluir a niños con discapacidades, tal como se consideró en el estudio original (Ulrich, 2019). Solo se incluyeron en el estudio a niños con desarrollo típico; por lo tanto, se recomienda evaluar las propiedades psicométricas del TGMD-3 en niños peruanos que presentan discapacidades.

## Conclusión

El estudio incorpora información nueva e importante

sobre las propiedades psicométricas del TGMD-3-PE. Los resultados de validez y fiabilidad confirman que el TGMD-3-PE es apropiado para evaluar la competencia de las HMF de niños peruanos. Asimismo, el TGMD-3-PE puede ser un instrumento de mucha utilidad para profesores de EF, entrenadores, profesionales de la salud e investigadores para identificar retrasos o déficits motores y la competencia de las HMF de los niños.

## Reconocimientos

El estudio está financiado por la Universidad Nacional Mayor de San Marcos, aprobado mediante Resolución Recctoral N.º 01686-R-20 con código de proyecto E20061491.

Este estudio se llevó a cabo en varias instituciones educativas del Perú. Los autores agradecen a las siguientes personas por su contribución al desarrollo y logro de este estudio:

Lidia Bellido: Directora de la institución educativa primaria Villa del Lago del distrito Puno, Puno.

Rober Cruz: Director de la institución educativa N.º 60756 Claverito del distrito Iquitos, Loreto.

Luis Flores: Director de la institución educativa N.º 2031 Virgen de Fátima del distrito San Martín de Porres, Lima.

Hermana Zenobia Gamarra: Promotora del colegio de los Sagrados Corazones Belén del distrito San Isidro, Lima.

Yvonne Vasquez: Profesora de Educación Física de la institución educativa N.º 60756 Claverito del distrito Iquitos, Loreto.

## Referencias

- Allen, K. A., Bredero, B., Van, T., Ulrich, D. A., & Simons, J. (2017). Test of Gross Motor Development-3 (TGMD-3) with the Use of Visual Supports for Children with Autism Spectrum Disorder: Validity and Reliability. *Journal of Autism and Developmental Disorders*, 47(3), 813–833. <https://doi.org/10.1007/s10803-016-3005-0>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.
- Anastasi, A., & Urbina, S. (1998). *Test psicológicos*. Prentice Hall.
- Arman, M., Barnett, L. M., Bowe, S. J., Bahram, A., & Kazemnejad, A. (2021). The Validity and Reliability of Scales to Measure Perceived Movement Skill Competence in Iranian Young Children. *Journal of Motor Learning and Development*, 9(1), 58–79. <https://doi.org/10.1123/jmld.2019-0023>
- Asunta, P., Viholainen, H., Ahonen, T., Cantell, M., Westerholm, J., Schoemaker, M. M., & Rintala, P. (2017). Reliability and validity of the Finnish version of the motor observation questionnaire for teachers.

- Human Movement Science*, 53, 63–71. <https://doi.org/10.1016/j.humov.2016.12.006>
- Bautista, A., Moreno-Núñez, A., Vijayakumar, P., Quek, E., & Bull, R. (2020). Gross motor teaching in preschool education: where, what and how do Singapore educators teach? (Enseñanza de la motricidad gruesa en educación infantil: ¿dónde, qué y cómo enseñan las maestras en Singapur?). *Journal for the Study of Education and Development*, 43(2), 443–482. <https://doi.org/10.1080/02103702.2019.1653057>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Breakwell, G. M., Hammond, S., Fife-Schaw, C., & Smith, J. A. (2006). *Research methods in psychology*. Sage Publications, Inc.
- Brian, A., Taunton, S., Lieberman, L. J., Haibach-Beach, P., Foley, J., & Santarossa, S. (2017). Psychometric Properties of the Test of Gross Motor Development-3 for Children With Visual Impairments. *Adapted Physical Activity Quarterly*, 35(2), 145–158. <https://doi.org/10.1123/apaq.2017-0061>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage Publications, Inc.
- Cronbach, L. J. (1989). Construct validation after thirty years. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy. Proceedings of a symposium in honor of Lloyd G. Humphreys* (pp. 147–171). University of Illinois Press.
- Duncan, M. J., Martins, C., Ribeiro Bandeira, P. F., Issartel, J., Peers, C., Belton, S., O'Connor, N. E., & Behan, S. (2022). TGMD-3 short version: Evidence of validity and associations with sex in Irish children. *Journal of Sports Sciences*, 40(2), 138–145. <https://doi.org/10.1080/02640414.2021.1978161>
- Eddy, L. H., Bingham, D. D., Crossley, K. L., Shahid, N. F., Ellingham-Khan, M., Otteslev, A., Figueredo, N. S., Mon-Williams, M., & Hill, L. J. B. (2020). The validity and reliability of observational assessment tools available to measure fundamental movement skills in school-age children: A systematic review. *PLOS ONE*, 15(8), e0237919. <https://doi.org/10.1371/journal.pone.0237919>
- Estevan, I., Molina-García, J., Queralt, A., Álvarez, O., Castillo, I., & Barnett, L. (2017). Validity and Reliability of the Spanish Version of the Test of Gross Motor Development–3. *Journal of Motor Learning and Development*, 5(1), 69–81. <https://doi.org/10.1123/jmld.2016-0045>
- Gallahue, D., Ozumn, J., & Goodway, J. (2012). *Understanding motor development: Infants, children, adolescents and adults*. McGraw-Hill.
- Garn, A. C., & Webster, E. K. (2021). Bifactor structure and model reliability of the Test of Gross Motor Development — 3rd edition. *Journal of Science and Medicine in Sport*, 24(1), 67–73. <https://doi.org/10.1016/j.jsams.2020.08.009>
- Goodway, J. D., Ozmun, J. C., & Gallahue, D. L. (2019). *Understanding Motor Development: Infants, Children, Adolescents, Adults*. Jones & Bartlett Learning. [https://books.google.com.pe/books?id=h5KwDwAAQBAJ&printsec=frontcover&hl=es&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](https://books.google.com.pe/books?id=h5KwDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false)
- Gravetter, F. J., & Wallnau, L. B. (2014). Essentials of the statistics for the behavioral science. In *Wadsworth Cengage Learning*. Cengage Learning.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis*. Pearson Prentice Hall.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2019). *Multivariate data analysis*. Cengage Learning EMEA.
- Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238–247. <https://doi.org/10.1037/1040-3590.7.3.238>
- Hopkins, W. G. (2016). *A New View of Statistics*. Retrieved March 12. <https://www.sportsci.org/resource/stats/index.html>
- Hulsteen, R. M., Barnett, L. M., True, L., Lander, N. J., del Pozo Cruz, B., & Lonsdale, C. (2020). Validity and reliability evidence for motor competence assessments in children and adolescents: A systematic review. *Journal of Sports Sciences*, 38(15), 1717–1798. <https://doi.org/10.1080/02640414.2020.1756674>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Magistro, D., Piumatti, G., Carlevaro, F., Sherar, L. B., Esliger, D. W., Bardaglio, G., Magno, F., Zecca, M., & Musella, G. (2020). Psychometric proprieties of the Test of Gross Motor Development–Third Edition in a large sample of Italian children. *Journal of Science and Medicine in Sport*, 23(9), 860–865. <https://doi.org/10.1016/j.jsams.2020.02.014>
- Mañano, C., Morin, A. J. S., April, J., Webster, E. K., Hue, O., Dugas, C., & Ulrich, D. (2022). Psychometric Properties of a French-Canadian Version of the Test of Gross Motor Development – Third Edition (TGMD-3): A Bifactor Structural Equation Modeling Approach. *Measurement in Physical Education and Exercise Science*, 26(1), 51–62. <https://doi.org/10.1080/1091367X.2021.1946541>
- Mamani-Ramos, A. A., Dextre-Mendoza, C. W., Lava Gálvez, J. J., Flores, G. T., Quispe Cruz, L. M., Torres-Cruz, F., Quisocala Mamani, J. A., & Fuentes-López, J. D. (2021). Gross motor development in preschoolers on the islands of Lake Titicaca (3810

- m.a.s.l.), Puno, Peru. *Retos*, 39, 592–597. <https://doi.org/10.47197/retos.v0i39.79522>
- Marinšek, M., Bedenik, K., & Kovač, M. (2023). Psychometric Properties of the Slovenian Version of the Test of Gross Motor Development–3. *Journal of Motor Learning and Development*, 1–15. <https://doi.org/10.1123/jmld.2022-0055>
- Mohammadi, F., Bahram, A., Khalaji, H., Ulrich, D. A., & Ghadiri, F. (2018). Evaluation of the Psychometric Properties of the Persian Version of the Test of Gross Motor Development–3rd Edition. *Journal of Motor Learning and Development*, 7(1), 106–121. <https://doi.org/10.1123/jmld.2017-0045>
- Mohammadi, F., Bahram, A., Khalaji, H., Ulrich, D. A., & Ghadiri, F. (2019). Evaluation of the psychometric properties of the Persian version of the test of Gross Motor Development-3rd edition. *Journal of Motor Learning and Development*, 7(1), 106–121. <https://doi.org/10.1123/JMLD.2017-0045>
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Sage Publications.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hili, Inc.
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467. <https://doi.org/10.1002/nur.20199>
- Rey, E., Carballo-Fazanes, A., Varela-Casal, C., & Abelairas-Gómez, C. (2020). Reliability of the test of gross motor development: A systematic review. *PLOS ONE*, 15(7), e0236070. <https://doi.org/10.1371/journal.pone.0236070>
- Rousson, V., Gasser, T., & Seifert, B. (2002). Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Statistics in Medicine*, 21(22), 3431–3446. <https://doi.org/10.1002/sim.1253>
- Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A Developmental Perspective on the Role of Motor Skill Competence in Physical Activity: An Emergent Relationship. *Quest*, 60(2), 290–306. <https://doi.org/10.1080/00336297.2008.10483582>
- Temple, V. A., & Foley, J. T. (2017). A Peek at the Developmental Validity of the Test of Gross Motor Development–3. *Journal of Motor Learning and Development*, 5(1), 5–14. <https://doi.org/10.1123/jmld.2016-0005>
- Tristán-López, a. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Avances En Medición*, 6, 37–48. [https://www.humanas.unal.edu.co/lab\\_psicometria/application/files/9716/0463/3548/VOL\\_6.\\_Articulo\\_4\\_Indice\\_de\\_validez\\_de\\_contenido\\_37-48.pdf](https://www.humanas.unal.edu.co/lab_psicometria/application/files/9716/0463/3548/VOL_6._Articulo_4_Indice_de_validez_de_contenido_37-48.pdf)
- Ulrich, D. A. (1985). *Test of Gross Motor Development TGMD*. Pro-Ed. file:///C:/Users/Anibal/Downloads/TGMDManual.pdf
- Ulrich, D. A. (2000). *Test of Gross Motor Development (Second Edition)*. Pro-Ed.
- Ulrich, D. A. (2017). Introduction to the Special Section: Evaluation of the Psychometric Properties of the TGMD-3. *Journal of Motor Learning and Development*, 5(1), 1–4. <https://doi.org/10.1123/jmld.2017-0020>
- Ulrich, D. A. (2019). *Test of Gross Motor Development (Third Edition)*. Pro-Ed.
- Valentini, N. C., Nobre, G. C., Zanella, L. W., Pereira, K. G., Albuquerque, M. R., & Rudisill, M. E. (2021). Test of Gross Motor Development–3 Validity and Reliability: A Screening Form. *Journal of Motor Learning and Development*, 9(3), 438–455. <https://doi.org/10.1123/jmld.2020-0061>
- Valentini, N. C., Zanella, L. W., & Webster, E. K. (2017). Test of Gross Motor Development—Third Edition: Establishing Content and Construct Validity for Brazilian Children. *Journal of Motor Learning and Development*, 5(1), 15–28. <https://doi.org/10.1123/jmld.2016-0002>
- Wagner, M. O., Webster, E. K., & Ulrich, D. A. (2017). Psychometric Properties of the Test of Gross Motor Development, Third Edition (German Translation): Results of a Pilot Study. *Journal of Motor Learning and Development*, 5(1), 29–44. <https://doi.org/10.1123/jmld.2016-0006>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>