

ARTÍCULO ORIGINAL

Reproducibilidad y validez convergente de la Escala Abreviada del Desarrollo y una traducción al español del instrumento *Neurosensory Motor Development Assessment*

Claudia Milena Hormiga¹, Diana Marina Camargo², Luis Carlos Orozco³

¹ Maestría en Epidemiología, Universidad Industrial de Santander, Bucaramanga, Colombia

² Escuela de Fisioterapia, Universidad Industrial de Santander, Bucaramanga, Colombia

³ Escuela de Enfermería, Universidad Industrial de Santander, Bucaramanga, Colombia

Institución a la que debe atribuirse el trabajo: Universidad Industrial de Santander

Introducción. El desarrollo motor es la adquisición y evolución de habilidades motoras; su valoración permite detectar alteraciones y promueve una atención oportuna y adecuada. Se han diseñado varios instrumentos para valorar el desarrollo motor en la niñez temprana, entre los que se encuentra el test *Neurosensory Motor Development Assessment*. En Colombia no hay estudios publicados sobre las propiedades psicométricas de cualquiera de estas pruebas y la vigilancia del desarrollo motor se realiza con la *Escala Abreviada del Desarrollo*, de la cual no existen reportes de validez o reproducibilidad.

Objetivo. Evaluar la reproducibilidad del componente motor de la *Escala Abreviada del Desarrollo* y de una traducción al español del instrumento *Neurosensory Motor Development Assessment*, así como la validez convergente entre las dos pruebas.

Materiales y métodos. Se realizó un estudio de evaluación de tecnologías diagnósticas. La población estuvo conformada por 260 niños de 4 y 5 años de edad. Para el análisis se aplicaron coeficientes de correlación intraclase y los límites de acuerdo de Bland y Altman, así como el coeficiente de correlación de Spearman (r).

Resultados. El coeficiente de correlación intraclase y el promedio de las diferencias para el *Neurosensory Motor Development Assessment*, fueron 0,91 y 1,23, y para la *Escala Abreviada del Desarrollo*, 0,96 y 0,02, respectivamente. La validez convergente mostró un r de 0,51.

Conclusiones. Los dos instrumentos tienen buena reproducibilidad entre evaluadores. La convergencia entre las pruebas es moderada y posiblemente se explica por las diferencias en el enfoque de medición de cada instrumento.

Palabras clave: desarrollo infantil, destreza motora, preescolar, diagnóstico precoz, validez de las pruebas, reproducibilidad de resultados.

Reliability and convergent construct validity of the abridged development scale and a Spanish translation of the Neurosensory Motor Development Assessment

Introduction. Motor development is the acquisition and evolution of motor skills; its evaluation allows detection of abnormalities and permits a suitable opportunity for bringing them to medical attention. Several instruments have been designed to characterize motor development in early childhood, among them the Neurosensory Motor Development Assessment test. In Colombia, studies on the psychometric properties have not been published on any of these tests. Although surveillance of the motor development uses the Abridged Development Scale, characterization of its validity or reproducibility have not been undertaken previously.

Objective. A statistical comparison was undertaken to evaluate the reliability of the motor component of the Abridged Development Scale and a translation to Spanish of the Neurosensory Motor Development Assessment instrument was developed. The convergent construct validity between the two tests was assessed.

Materials and methods. An evaluation of diagnostic technologies was done. The sample consisted of 260 children, aged 4 to 5 yrs. The analysis included intraclass correlation

coefficients, the Bland and Altman's limits of agreement, and the Spearman's correlation coefficient (r).

Results. The intraclass correlation coefficients and the mean of differences for the Neurosensory Motor Development Assessment were of 0.91 and 1.23 and for the Development Scale 0.96 and 0.02, respectively. The convergent construct validity was 0.51.

Conclusions. Both instruments have good inter-rater reliability. The convergence between the tests was moderate, possibly explained by the differences in the approach to measurement of each instrument.

Key words: Child development, motor skills; child, preschool; early diagnosis, validity of tests, reproducibility of results.

El desarrollo motor es la adquisición y evolución de habilidades motoras durante la vida (1); sin embargo, los mayores cambios en este proceso se llevan a cabo durante los primeros años. Las alteraciones del desarrollo motor pueden darse por retardo o anormalidad e involucran un espectro amplio de diagnósticos y gravedad (2); incluso si son menores o leves, pueden tener un impacto en la habilidad del niño para aprender y participar con éxito en las actividades de la vida cotidiana en su casa y en el colegio (3).

Existe un creciente interés por el estudio del desarrollo motor en edad preescolar, puesto que niños aparentemente "normales" pueden presentar alteraciones leves o moderadas, inaparentes en edades anteriores (4). En estudios recientes se ha encontrado asociación entre una pobre coordinación motora en niños en edad preescolar y escolar con dificultades académicas (5), problemas psicológicos (6) y alteración deficitaria de atención e hiperactividad (7). Los niños con este tipo de alteraciones o anormalidades en el desarrollo motor pueden beneficiarse de una intervención oportuna, si se les identifica tempranamente (2).

En el estudio del desarrollo motor se pueden medir dos aspectos: calidad y cantidad. Desde el punto de vista neurológico, los aspectos cualitativos del movimiento reflejan la madurez e integridad del cerebro y pueden ejercer un rol importante en el

diagnóstico temprano de los trastornos del desarrollo; sin embargo, la medición de la "calidad" del movimiento todavía es un reto (8), debido a que ésta no constituye una característica única sino que involucra diferentes componentes, como el alineamiento corporal, el control de la postura, el equilibrio y la coordinación (9).

Actualmente, existen varios instrumentos diseñados para valorar el desarrollo motor en las edades que comprenden la niñez temprana, algunos muy reconocidos y utilizados en el panorama mundial. De estas pruebas, unas se enfocan específicamente en el desarrollo motor (especialmente a partir de los 4 años), como las escalas del desarrollo motor de Peabody (*Peabody Developmental Motor Scales*, PDMS) (10-12), el test de habilidades motoras de Bruininsk-Oseretsky (*Bruininsk-Oseretsky Test of Motor Proficiency*, BOTMP) (13-15) y la batería de valoración del movimiento de los niños (*Movement Assessment Battery for Children*, M-ABC) (13); y otras contemplan la valoración de otras áreas del desarrollo, como el test de tamización del desarrollo de Denver (*Denver Developmental Screening Test*, Denver II) (10,16), la segunda versión de las escalas del desarrollo de los infantes de Bayley (*Bayley Scales of Infant Development*, *BSID II*) (10) y el programa de desarrollo de Gessell (*Gessell Developmental Schedules*) (17). Sin embargo, ninguna de estas es reconocida como el método de referencia para valorar el desarrollo motor e identificar tempranamente alteraciones de este proceso (13,15,18); por el contrario, existen críticas relacionadas con su enfoque en la medición de aspectos cuantitativos del movimiento (4, 9,19,20) y se han reportado problemas en la metodología empleada para la evaluación de sus propiedades

Correspondencia:

Claudia Milena Hormiga, carrera 21 N° 152-30, Urbanización Parque San Agustín, Floridablanca, Santander.

Teléfonos: (7) 679 8632, (7) 638 4645 y (313) 830 6990; fax: (7) 679 8629.

claudihsanc@gmail.com

Recibido: 03/07/07; aceptado: 21/04/08

psicométricas, en especial, limitaciones relacionadas con el tamaño de la muestra y la aplicación de métodos analíticos deficientes (13).

Entre los instrumentos que miden el desarrollo motor, también se encuentra la prueba australiana *Neurosensory Motor Development Assessment* (NSMDA) (21), diseñado para identificar tempranamente alteraciones del desarrollo motor; este instrumento, a diferencia de la mayoría de los diseñados para medir este proceso, contempla aspectos cualitativos del movimiento, además de los aspectos cuantitativos, y se ha utilizado tanto en la práctica clínica como en la investigación pediátrica australiana (22-29). Si bien de este instrumento se ha reportado el acuerdo entre evaluadores (acuerdo observado > 0,80), la validez de criterio tipo predictivo y la validez convergente –evaluada mediante el análisis de factores confirmatorios, que mostró un único factor–, también presentan limitantes metodológicas en relación con la evaluación de dichas propiedades, como haber utilizado la misma prueba como referencia.

En Colombia no hay estudios publicados sobre la validez o la reproducibilidad de cualquiera de estos instrumentos. El seguimiento del desarrollo motor, enmarcado en el programa de crecimiento y desarrollo establecido en la Resolución 00412 del Ministerio de Salud (30), estipula la utilización de la Escala Abreviada del Desarrollo (31) en las instituciones prestadoras de servicios de salud del país, con el objetivo de detectar precozmente alteraciones en este proceso y promover una intervención oportuna y adecuada. Esta prueba valora todas las áreas del desarrollo y, en cuanto al desarrollo motor, se enfoca en la valoración de los aspectos cuantitativos del proceso. A pesar del uso generalizado de la Escala Abreviada del Desarrollo, no se cuenta con publicaciones sobre la evaluación de su validez o reproducibilidad.

En reconocimiento de la gran influencia que el área motora ejerce sobre el desarrollo integral de los niños y que únicamente mediante un proceso adecuado de valoración del desarrollo motor es posible diseñar estrategias de intervención pertinentes y oportunas, se considera necesario disponer de instrumentos válidos y reproducibles que permitan la identificación temprana de

cualquier característica en el desarrollo motor que se considere fuera de la variabilidad normal de este proceso.

De los antecedentes presentados se plantea la comparación del componente motor de la Escala Abreviada del Desarrollo con una traducción al español del instrumento *Neurosensory Motor Development Assessment* (NSMDA) en población preescolar de 4 y 5 años, para establecer la validez convergente entre las dos pruebas (32) y determinar la reproducibilidad intra y entre evaluadores del NSMDA y la reproducibilidad entre evaluadores de la Escala Abreviada del Desarrollo (32).

El NSMDA fue escogido porque contempla aspectos cualitativos del movimiento, además de los aspectos cuantitativos, lo cual constituye una característica distintiva del mismo, y por su amplia utilización tanto en la práctica clínica como en investigación pediátrica australiana; además, se contó con la autorización y disponibilidad del mismo, concedidas por su autora, quien facilitó las instrucciones y el material bibliográfico, y apoyó permanentemente al equipo investigador.

Es importante mencionar que en la fase de planeación del estudio se consultaron autores de otros instrumentos ampliamente utilizados, como las *Peabody Developmental Motor Scales*, y se encontró que los derechos sobre el instrumento pertenecían a sus quienes lo comercializan, por lo que su adquisición implicaba un alto costo para el equipo investigador.

Materiales y métodos

Se realizó un estudio de evaluación de tecnologías diagnósticas (33). La población estuvo conformada por una muestra seleccionada por conveniencia, de 260 niños y niñas de 4 y 5 años de edad, escolarizados y procedentes de los hogares infantiles “Fe y alegría” y “Bambi” del Instituto Colombiano de Bienestar Familiar y del Colegio Maiporé, instituciones educativas ubicadas en la zona norte de Bucaramanga. Tanto los representantes legales de las instituciones como los padres o acudientes de los niños, firmaron el consentimiento informado previo al ingreso de éstos al estudio. Las valoraciones se llevaron a cabo entre agosto y noviembre de 2005.

El tamaño de la muestra se calculó siguiendo la metodología propuesta por Kraemer (34). En relación con la reproducibilidad entre evaluadores, se tuvo en cuenta que permitiera obtener un coeficiente de correlación intraclase (CCI) de 0,80, asumiendo una hipótesis nula de 0,70, con un poder de 80% y un nivel de significancia de 5%. Para la validez convergente, se estableció un poder de 80% y un nivel de significancia de 5%, asumiendo que la reproducibilidad de un instrumento es de 0,70 y la del otro es de 0,80. Con base en estos cálculos, el número de niños requeridos fue de 160; sin embargo, dado que la Escala Abreviada del Desarrollo cuenta con unos ítems específicos para los niños de 4 años y otros para los niños de 5 años, se decidió involucrar en el estudio a 320 niños (160 de 4 años y 160 de 5 años), para el análisis de reproducibilidad y de validez convergente para cada edad.

Variables e instrumentos de medición

La población del estudio se caracterizó teniendo en cuenta las variables de sexo, edad y el desarrollo motor medido por los dos instrumentos: la Escala Abreviada del Desarrollo (31) y una traducción al español del instrumento *Neurosensory Motor Development Assessment* (21).

La Escala Abreviada del Desarrollo valora el desarrollo de los niños entre 0 y 84 meses y consta de cuatro áreas: motora gruesa, motriz fino-adaptativa, audición y lenguaje, y personal-social; cada una cuenta con ítems específicos por rangos de edad, con escala de medición nominal dicotómica (ausente-presente), de los cuales se evaluaron los correspondientes a la edad del niño. El total de cada área se obtiene de la suma acumulativa de sus ítems, puesto que se toman como realizados los ítems anteriores al primero ejecutado (31).

El NSMDA valora el desarrollo sensorio-motor y consta de seis áreas: motora gruesa, motora fina, patrones primitivos, neurológica, postura y equilibrio, y sensorio-motora. Esta prueba cuenta con ítems específicos por rangos de edad (los 4 y 5 años se contemplan en un solo rango etario) cuya escala de medición es ordinal: 1: definitivamente anormal (respuesta hiperactiva o hipoactiva); 2: respuesta retardada, inmadura o

moderadamente anormal (lenta o exagerada); 3: dentro del rango esperado para la edad; 4: por encima del desempeño promedio. El total de cada área se obtiene de la suma de sus ítems y con la suma de todas las áreas se obtiene la calificación total (21). De este instrumento se evaluaron 79 ítems correspondientes al rango de edad; se excluyeron los ítems como el nistagmo, debido a falta de claridad en el criterio de la respuesta normal para la edad de los niños valorados y al procedimiento (estímulo vestibular rotatorio) que constituye una prueba traumática porque puede provocar malestar tipo náuseas en el participante. El rango de los puntajes para los niños de 4 a 5 años es de 79 a 261 puntos, con un promedio de 182.

Cada instrumento cuenta con un punto de corte para establecer los niños con retardo motor, en el caso del NSMDA, y los niños con señal de alerta en su desarrollo, en la Escala Abreviada del Desarrollo. Para la prueba australiana, este punto corresponde a un puntaje total de 217 (calificación total por debajo del 25% de la calificación esperada para la edad) (21) y para la Escala Abreviada del Desarrollo, es de 26 en la motricidad gruesa y de 23 en la motricidad fino-adaptativa (31). En el anexo 1 se presenta un cuadro comparativo de los dos instrumentos escogidos en el presente estudio, en relación con las áreas que evalúa, los ítems y la escala de medición.

Procedimientos

El NSMDA fue traducido al español por un docente del Instituto de Lenguas de la Universidad Industrial de Santander; esta versión fue complementada con una propuesta de traducción independiente realizada por un profesional de la fisioterapia con experiencia en la valoración del desarrollo motor, debido al manejo de términos técnicos incluidos en el instrumento. La traducción definitiva al español se obtuvo en consenso, mediante la discusión de las dos propuestas mencionadas. No se realizó el proceso de equivalencia translingüística, puesto que la terminología técnica del instrumento hace necesario contar con una persona que domine los idiomas español e inglés, de preferencia que este último sea su lengua materna, y así mismo, que

esté familiarizada con el desarrollo motor y su valoración, lo cual representa una gran dificultad en el medio en el que se realizó el estudio.

Una vez dispuestos los dos instrumentos, dos fisioterapeutas con 3 y 10 años experiencia en el área pediátrica, respectivamente, fueron entrenadas como evaluadoras con el objetivo de estandarizar la aplicación de los mismos. Posteriormente, se llevó a cabo una prueba piloto en la que participaron 30 niños (15 por cada rango de edad), cuyos resultados permitieron estimar el tiempo promedio de aplicación de los instrumentos y ajustar el tamaño de muestra requerido para evaluar la reproducibilidad intraevaluadores del NSMDA.

En la fase de recolección del estudio, las valoraciones se realizaron en las instituciones educativas durante el horario escolar habitual en condiciones ambientales similares y se procedió con la administración de los instrumentos según las instrucciones y recomendaciones provistas para cada uno (21,31), las cuales contemplan los procedimientos de evaluación y los criterios de respuesta de cada ítem, y aspectos generales referentes al enfoque de las pruebas, el tiempo de aplicación, la calificación y el registro de los datos, y las condiciones ambientales y de los niños en que deben llevarse a cabo las valoraciones.

El orden de los instrumentos fue constante; se inició con la valoración de las áreas motoras de la Escala Abreviada del Desarrollo y posteriormente se aplicó el NSMDA, con el objetivo de no inducir la calificación con la primera, dado el mayor número de variables que contempla el NSMDA. La sección correspondiente al tono muscular en el NSMDA fue valorada directamente por cada fisioterapeuta.

En la fase de recolección de datos, las dos evaluadoras llevaron a cabo y de manera simultánea la valoración del desarrollo motor con cada instrumento, para lo cual se asignó en forma aleatoria a la fisioterapeuta encargada de su conducción; ambas evaluadoras diligenciaron la hoja de calificación de cada instrumento.

Para evaluar la reproducibilidad intraevaluadores del NSMDA, 30 niños fueron valorados por segunda vez por una de las evaluadoras; estos

niños fueron escogidos en forma aleatoria a lo largo del período de recolección de datos, para minimizar el efecto del aprendizaje de la evaluadora. La segunda valoración se llevó a cabo con un intervalo mínimo de 5 días y máximo de 10 días después de la valoración inicial y en condiciones similares. Finalizando la fase de recolección de datos, una evaluadora no continuó su participación en el estudio, por lo que sólo 260 niños fueron valorados por las dos evaluadoras (109 de 4 años y 151 de 5 años) con los dos instrumentos.

Análisis de la información

Se aplicaron medidas de tendencia central y dispersión para caracterizar la población, previa evaluación de su normalidad (35). La reproducibilidad intra y entre evaluadores se estableció con la metodología de Bland y Altman (36), mediante la cual se obtiene una estimación, con el 95% de confianza, de la diferencia entre los puntajes obtenidos en dos mediciones realizadas, y se obtuvieron los coeficientes de correlación de Lin (37) e intraclase (32). La concordancia entre las mediciones para determinar retardo motor, previa categorización del puntaje obtenido en cada prueba, se realizó con el coeficiente kappa de Cohen (38) con base en los puntos de corte señalados (217 para el NSMDA) y (23 en la motricidad fina y 26 para la motricidad gruesa para la escala abreviada), atendiendo a la interpretación de Landis y Koch referenciados por Kramer (33) de considerar una concordancia buena a los coeficientes entre 0,60 y 0,80 y una concordancia casi perfecta a los coeficientes superiores a 0,80. Además, se determinó la concordancia de cada ítem de los dos instrumentos mediante el índice kappa ponderado (39) y se calculó el error estándar de las medias (*Standard Error of the Means, SEM*) (40), el cual permite determinar la cercanía de las mediciones con su valor real (con una probabilidad de 68% y 96%, el verdadero valor del puntaje de cada prueba se encuentra dentro de ± 1 SEM y ± 2 SEM de la medición original, respectivamente).

La validez convergente se evaluó aplicando el coeficiente de correlación de Spearman entre los puntajes obtenidos en los dos instrumentos. La

base de datos fue digitada por duplicado en el programa Epi-Info, versión 6.04, (41) y posteriormente fueron validadas; una vez depurada, se exportó al programa STATA versión 9 (42) para su análisis definitivo.

Consideraciones éticas

La investigación siguió las recomendaciones para la investigación en seres humanos establecidas en la Declaración de Helsinki y en la Resolución 008430 del Ministerio de Salud de Colombia; según esta última, esta investigación se consideró de riesgo mínimo. Se respetaron los principios básicos de la ética biomédica (43).

Resultados

Se evaluaron 260 niños, 109 (41,9%) de 4 años y 151 (58,1%) de 5 años; el 41,3% (45/109) correspondió al sexo femenino y el 49,1% (74/151) al masculino, respectivamente. De los niños valorados, 115 (44,23%) estudiaban en el colegio

Maiporé, 99 (38,08%) en el hogar infantil Fe y Alegría y los 46 restantes (17,69%) en el hogar Bambi.

La descripción general de los puntajes obtenidos para cada instrumento y por evaluador se muestra en el cuadro 1. Al separar por edad, el promedio de los puntajes de los niños de 4 años fue menor, incluso en el NSMDA que contempla sólo un conjunto de ítems para los 4 y 5 años.

Reproducibilidad entre evaluadores

En los cuadros 2 y 3 se muestran los hallazgos de la evaluación de la reproducibilidad. En general, se obtuvieron CCI altos (>0,90). De acuerdo con la metodología de Bland y Altman, para el NSMDA el acuerdo entre los evaluadores fue de $\pm 8,7$ puntos al incluir la totalidad de los niños (44); considerando el rango de valores posibles que puede tomar la prueba (182), el promedio de las diferencias representa el 0,67% de este rango y

Cuadro 1. Puntajes obtenidos por los dos evaluadores al aplicar los instrumentos evaluados (n=260).

Edad	NSMDA		EAD	
	Evaluador 1	Evaluador 2	Evaluador 1	Evaluador 2
4 años	215,9 \pm 8,7* [186 – 237]**	218,1 \pm 8,9 [184 – 241]	54,3 \pm 3,2 [47 – 60]	54,3 \pm 3,2 [48 – 60]
5 años	223,0 \pm 10,4 [188 – 249]	223,6 \pm 10,5 [182 – 243]	57,8 \pm 2,5 [48 – 60]	57,8 \pm 2,5 [48 – 60]
Total	220,1 \pm 10,3 [186 – 249]	221,3 \pm 10,2 [182 – 243]	56,3 \pm 3,3 [47 – 60]	56,3 \pm 3,3 [48 – 60]

*(Promedio \pm DE)

** Rango

Cuadro 2. Reproducibilidad entre evaluadores del NSMDA.

Edad	CCI IC (95%)*	Coficiente de Lin	Promedio de las diferencias	Límites 95%	Kappa 0,5
4 años (n=109)	0,90 (0,90 – 0,95)	0,90 (0,86 – 0,94)	2,16 \pm 3,27**	-4,26; 8,57	0,65
5 años (n=151)	0,88 (0,84 – 0,91)	0,88 (0,84 – 0,91)	0,56 \pm 5,05	-9,35; 10,46	0,73
Total (n=260)	0,91 (0,88 – 0,93)	0,89 (0,87 – 0,92)	1,23 \pm 4,46	-7,51; 9,97	0,71

CCI: coeficiente de correlación intraclase

*: IC: intervalo de confianza

**: promedio \pm DE

el acuerdo, el 5%. Se encontraron 14 datos (5,4%), fuera de los límites de acuerdo del 95% (figura 1).

El acuerdo de la EAD fue de $\pm 1,8$. Teniendo en cuenta el rango de valores posibles de la escala, la diferencia y el nivel de acuerdo encontrados representan el 0,02% y el 3% de este rango, respectivamente. Sólo se encontraron 15 observaciones (5,76%) fuera de los límites de acuerdo del 95%, al analizar la totalidad de niños (figura 2).

El estadístico kappa de Cohen mostró una concordancia buena para el instrumento australiano y el área motora gruesa de la Escala abreviada

del desarrollo; la concordancia obtenida para el área motriz fino-adaptativa de la escala abreviada fue casi perfecta. En relación con la concordancia de los ítems, 34% de los ítems del NSMDA tuvo un índice kappa ponderado superior a 0,80 y 33% tuvo un índice entre 0,6 y 0,8. En especial, tuvieron mejor concordancia los ítems de las áreas motora gruesa y fina, el tono muscular, las reacciones protectoras y de equilibrio. De los ítems de la escala abreviada, todos tuvieron un índice superior a 0,80. Los resultados de la evaluación de la concordancia entre evaluadores de cada uno de los ítems de los dos instrumentos,

Cuadro 3. Reproducibilidad entre evaluadores de la EAD.

Edad	CCI* IC (95%)**	Coeficiente de Lin	Promedio de las diferencias	Límites 95%	Kappa 0,5	
					MG	MFA
4 años (n=109)	0,97 (0,95 – 0,98)	0,96 (0,95 – 0,97)	0,01 \pm 0,83 [^]	-1,62; 1,64	0,81	0,97
5 años (n=151)	0,92 (0,89 – 0,94)	0,91 (0,89 – 0,94)	0,02 \pm 0,99	-1,93; 1,97	0,75	1,0
Total (n=260)	0,96 (0,95 – 0,97)	0,96 (0,95 – 0,97)	0,02 \pm 0,93	-1,81; 1,84	0,79	0,97

MG: motricidad gruesa

MFA: motricidad fino-adaptativa

*CCI: coeficiente de correlación intraclase

**IC: intervalo de confianza

[^]: promedio \pm DE

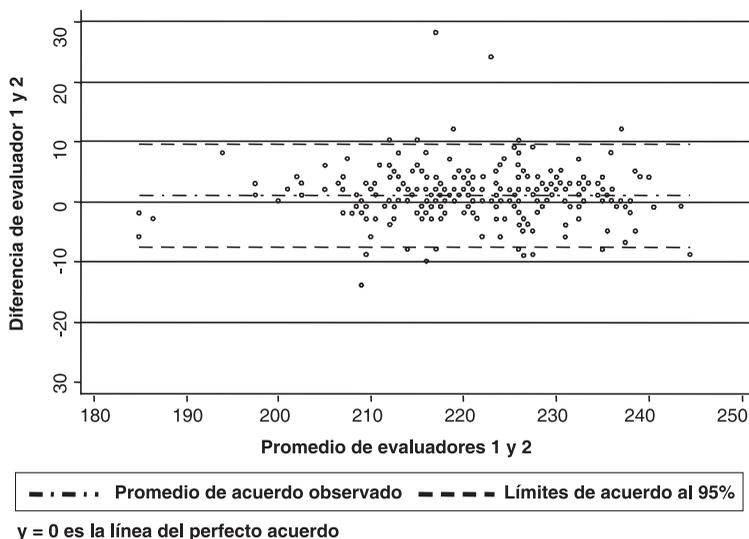


Figura 1. Límites de acuerdo de Bland y Altman, puntajes de los evaluadores en el NSMDA, niños de 4 y 5 años.

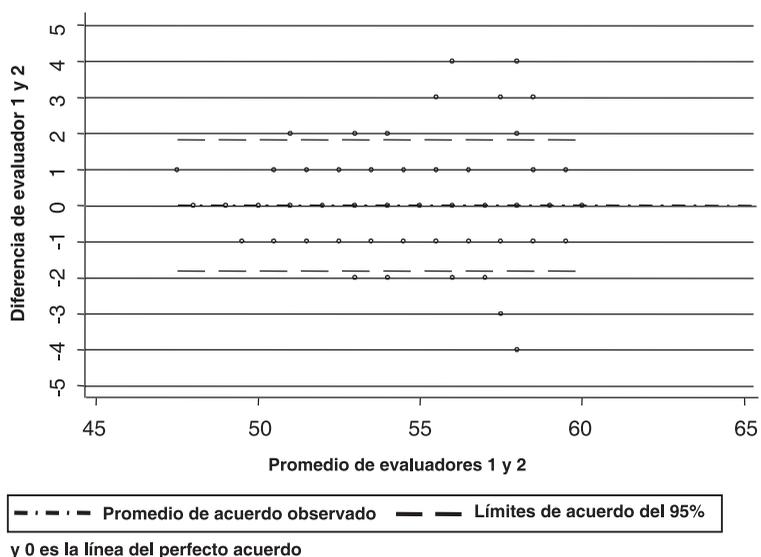


Figura 2. Límites de acuerdo de Bland y Altman, puntajes de los evaluadores en la Escala Abreviada del Desarrollo, niños de 4 y 5 años.

se presentan en el anexo 2. Finalmente, el error estándar de la medición fue de 3,1 para el NSMDA y de 0,7 para la escala abreviada.

Reproducibilidad intraevaluadores del NSMDA

El promedio del puntaje obtenido por los 30 niños en la primera valoración fue de $220,9 \pm 8,0$ y en la segunda fue de $223,6 \pm 8,3$. De manera consistente, en la segunda valoración se obtuvieron puntajes superiores. El CCI fue de 0,81 y el estadístico kappa de Cohen mostró una concordancia de 0,62 al incluir la totalidad de los niños. Al evaluar los límites de acuerdo de Bland y Altman, se presenta un promedio de la diferencia de 2,67, IC95% (-6,09, 11,42), al incluir la totalidad de los niños valorados. El promedio de la diferencias representó el 1,46% de la suma que puede tomar el instrumento. Ningún dato se encuentra por fuera de los límites de acuerdo del 95% (cuadro 4, figura 3).

Validez convergente entre el NSMDA y la EAD

El diagrama de dispersión que relaciona la media de las valoraciones de los dos evaluadores en cada instrumento muestra una relación lineal positiva entre los mismos (figura 4), hallazgo uniforme al evaluar la relación de los instrumentos por separado en cada evaluador y al restringir los

puntajes a cada rango de edad. El coeficiente de correlación de Spearman fue de 0,51 al incluir los 260 niños, de 0,39 en los niños de cuatro años y de 0,42 en los de cinco años. Al correlacionar los puntajes de la escala abreviada con los de las áreas del instrumento australiano correspondientes a la motricidad gruesa y fina, también se observa una relación lineal positiva entre los puntajes; el coeficiente de correlación de Spearman fue de 0,61.

Discusión

Las decisiones en el manejo clínico de un individuo, así como la calidad de una investigación, están condicionadas por la calidad de las mediciones; por esta razón, la evaluación de las pruebas diagnósticas es primordial tanto en las investigaciones epidemiológicas como en la práctica clínica (45,46).

La importancia de diseñar y utilizar instrumentos estandarizados, reproducibles y válidos que permitan valorar el desarrollo motor, ha sido ampliamente reconocida desde hace varias décadas (47), pues de ello depende el diagnóstico precoz de alteraciones sensorio-motoras susceptibles de intervención por parte del fisioterapeuta o el terapeuta ocupacional, con lo cual se promueve la aplicación de estrategias de intervención más específicas.

Cuadro 4. Reproducibilidad intraevaluadores del NSMDA y concordancia entre las mediciones (n total=30).

Edad	CCI IC (95%)	Coefficiente de Lin IC (95%)	Promedio de las diferencias IC (95%)	Límites 95%	Kappa 0,5
4 años (n=14)	0,85 (0,59 - 0,94)	0,84 (0,66 - 1,01)	1,29 ± 3,45**	-5,48; 8,05	0,65
5 años (n=15)	0,77 (0,30 - 0,92)	0,76 (0,54 - 0,97)	3,88 ± 4,99	-5,01; 13,66	0,60
Total (n=30)	0,81 (0,55 - 0,91)	0,81 (0,67 - 0,93)	2,67 ± 4,47	-6,09; 11,42	0,62

CCI: coeficiente de correlación intraclase

*IC: intervalo de confianza

** : promedio ± DE

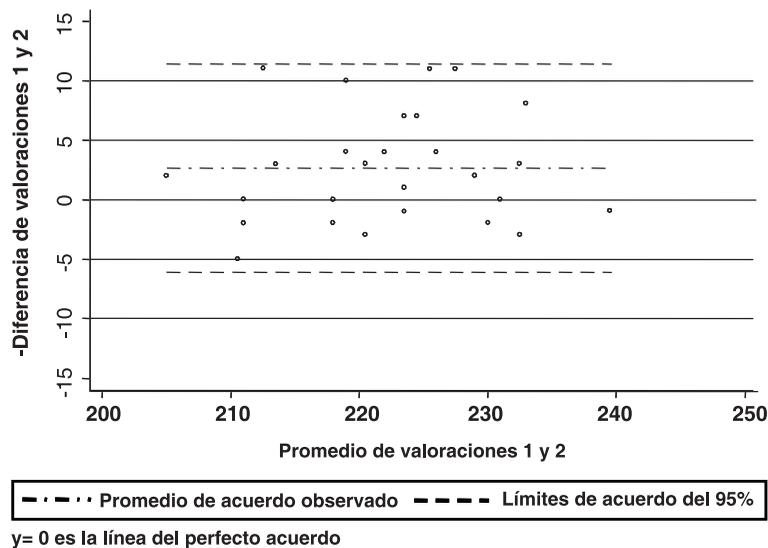


Figura 3. Límites de acuerdo de Bland y Altman, puntajes de las dos valoraciones con el NSMDA en niños de 4 y 5 años.

La mayoría de instrumentos de valoración del desarrollo motor en la niñez temprana disponibles actualmente, han sido creados para evaluar la aparición de comportamientos motores (*milestones*) y utilizan las normas establecidas por Gessell y Amatruda (48); es decir, se centran en la adquisición de habilidades motoras (cantidad) y prestan poco interés a sus componentes (calidad) (9). La falta de instrumentos que midan los aspectos cualitativos del movimiento da como resultado interpretaciones subjetivas e individualizadas que dificultan la comunicación con otros profesionales y promueven el uso de registros “anecdóticos” para

documentar el estado y progreso en el desarrollo motor del niño (47).

En general, las pruebas disponibles presentan deficiencias en la evaluación de su reproducibilidad y validez (cuadro 5). Del *Test Gessell Developmental Schedules* actual, por ejemplo, se encuentran principalmente reportes sobre la reproducibilidad y la validez de predicción de las versiones anteriores del instrumento. De la actual versión, Campbell (17) refiere un estudio en el que se evaluó la reproducibilidad inter-evaluador, pero solamente reporta un porcentaje de acuerdo del 93,7% para la escala total.

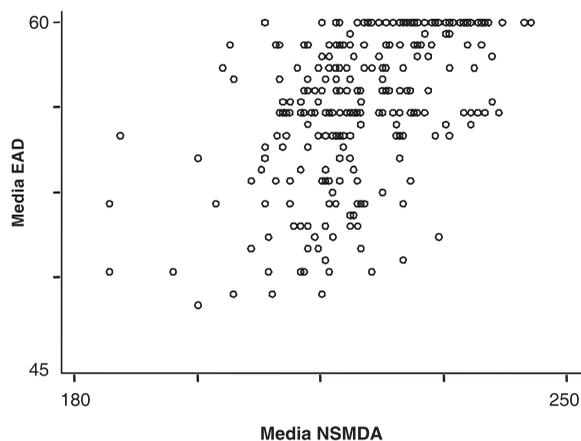


Figura 4. Diagrama de dispersión de la media de los puntajes de cada instrumento, niños de 4 y 5 años (n=260).

El autor del *Denver Developmental Screening Test-Denver II*, prueba utilizada ampliamente en Estados Unidos, argumenta que la validez del instrumento se basa en su estandarización y normatización (16); por esta razón, no existen estudios publicados en los que este instrumento se compare con otro y de esta manera se explore la convergencia o acuerdo entre los mismos. En una revisión, Breneman (10) reporta los datos sobre reproducibilidad publicados en el manual del instrumento, pero sólo menciona unos datos de reproducibilidad entre observadores y no especifica el tipo de coeficiente de correlación ni la metodología empleada (cuadro 5).

La segunda versión de *Bayley Scales of Infant Development-BSID II* es otro instrumento muy empleado en Norteamérica para evaluar el desarrollo de los niños. Sin embargo, existen críticas sobre su enfoque en aspectos cuantitativos; incluso algunos investigadores reportan que los niños con patrones anormales de movimiento pueden aprobar los ítems (9). Existen reportes sobre reproducibilidad test-retest de las subescalas en niños entre los 0 y 12 meses y entre los 24 y 26 meses (cuadro 5).

En relación con la evaluación de las propiedades de las pruebas que valoran solamente el desarrollo motor en la niñez temprana, Hinderer, referenciado por Schdmith (12), refiere la reproducibilidad intra y entre evaluadores reportada en el manual del *Test Peabody Developmentl Motor Scales -PDMS*

(entre 0,80 y 0,99) y discute que las autoras no especificaron los coeficientes de correlación empleados. Asimismo, llama la atención sobre las edades de los 36 niños que participaron en dicho estudio, las cuales estaban distribuidas en todo el rango que contempla la prueba (entre los 0 y 7 años), de tal manera que el rango de las calificaciones obtenidas y la variabilidad entre las mismas fueron amplios, condiciones que afectan los coeficientes de correlación haciendo que tiendan a tomar valores cercanos a uno (50).

En los estudios en que se evalúa la reproducibilidad del instrumento se aprecian factores limitantes en relación con el tamaño de la muestra, como en el de Boulton y colaboradores, en el que tan sólo se evaluaron 12 niños con parálisis cerebral de leve a moderada (11); asimismo, se aprecia limitación en el análisis, como en el estudio de Schmith y colaboradores (12), en el que se obtuvieron coeficientes de correlación intraclass altos, pero no se emplearon otras técnicas de análisis, como la metodología de Bland y Altman (36).

De otro lado, la reproducibilidad del BOTMP, uno de los instrumentos más populares y empleado por los fisioterapeutas de Estados Unidos y Canadá, también genera inquietudes; en el estudio de Moore y colaboradores, referenciado en la revisión de Wiert (13), los autores reportan coeficientes de 0,76 para todo el instrumento; sin embargo, los coeficientes de los ítems variaron entre 0,0 y 0,79; el de 0,0 correspondió a un ítem de coordinación bilateral.

Otra prueba muy utilizada, especialmente en Europa, es el M-ABC, revisión del TOMI (*Test of Motor Impairment*), cuya información respecto a sus propiedades psicométricas corresponden al TOMI (cuadro 5); es decir, no se cuenta con datos sobre las propiedades del M-ABC, a pesar de los cambios significativos en el sistema de calificación en la revisión del TOMI al M-ABC (13).

Un instrumento recientemente diseñado con el propósito de valorar aspectos cualitativos y cuantitativos del movimiento en los niños de 5 a 6 años de edad, es el test motor de Maastricht (*Maastricht Motor Test*, MMT), del cual existen reportes de la reproducibilidad, empleando

Cuadro 5. Reproducibilidad y validez de instrumentos seleccionados que evalúan el desarrollo motor en la niñez temprana.

Instrumento, autor	Reproducibilidad	Validez
Programa de desarrollo de Gessell, Arnold Gesell (10,17)	REE (n=48 niños entre 16 semanas y 21 meses): porcentaje de acuerdo = 93,7% para el total de la escala	
DENVER II, Frankenburg, Dodds y Archer (10)	REE (n=38 niños entre 0 y 6 años) (10): entre 0,95 y 1,00; promedio de reproducibilidad = 0,99 RTT (lapso de 7 a 10 días): entre 0,5 y 1,00; promedio de reproducibilidad = 0,90 No se especifica el tipo de coeficiente empleado.	De acuerdo con el autor, la validez del test recae sobre su estandarización, no en su correlación con otros tests (16).
<i>Bayley Scales of Infant Development II</i> -BSID II, Nancy Bayley (10)	Escala motora (49) RIE: r=0,77 (niños de 1-12 meses) RIE: r=0,79 (niños de 24-36 meses) REE: r=0,75	Validez concurrente con (20): o Escala Bayley original: r=0,63 o <i>McCarthy Scales of Children's Abilities Motor Subscale</i> : r=0,59 o <i>Wechsler Preschool and Primary Scale of Intelligence - Revised</i> : r=0,41 o <i>Differential Abilities Scales</i> : r=0,24 a 0,35 o <i>Pre-School Language Scale-3</i> : r=0,37 o PDMS II (n=38 niños de 2 años de edad): r=0,87; CC kappa entre 0,22 y 0,35 o PDMS II (54) (n=110 niños entre 3 y 41 meses): r=entre 0,85 para el agarre y r=0,97 para la locomoción; CC kappa: 0,13 entre BSID II y la escala motora gruesa del PDMS II, 0,02 entre BSID II y la escala motora fina del PDMS II, 0,09 entre BSID II y escala total del PDMSII
<i>Peabody Developmental Motor Scales</i> -PDMS, Rebecca Fewel y Rhona Folio (10)	Reporte en el manual (12): RIE (EMG) = 0,95 (EMF) = 0,80 REE (EMG) = 0,97 (EMF) = 0,94 No se especifica el tipo de coeficiente empleado. EMG (n=35 niños de 4 y 5 años) : CCI= 0,84 a 0,94 (12)	o Validez de criterio con el Bayley (20: CC kappa entre 0,22 y 0,35; r entre 0,49 y 0,87 o Validez de criterio con <i>The Vulpe Assessment Battery</i> (n=13 niños): r=0,97 (55)
<i>Bruininsk-Oseretsky Test Motor Proficiency</i> – BOTMP, Robert H. Bruininks (10)	RTT (n= 32 niños de 5 años): 0,76 para la escala total; para los subtests desde 0,0 hasta 0,76 (13) REE (n=30): r=0,77 a 0,97	Validez convergente: un solo factor of (15) Consistencia entre test para evaluar el DCD (15): o BOTMP vs M-ABC: CC kappa entre 0,07 y 0,43 o BOTMP vs DCDQ: CC kappa entre 0,28 y 0,44
<i>Movement Assessment Battery for Children</i> - M-ABC (13)	No conocidas; existe reporte del test antecesor (TOMI-H).	No conocidas; existen reportes del test antecesor (TOMI).

Cuadro 5 (continuación).

Instrumento, autor	Reproducibilidad	Validez
	RTT: CCI > 0,70. REE: CCI ≥ 0,70	
<i>Maastricht Motor Test</i> - MMT, Mariëlle Kroes (19)	RIE (n=42 niños de 5 años): CCI = 0,96 (escala total), CCI = 0,92 (aspectos cualitativos) CCI = 0,97 (aspectos cuantitativos)	Validez de criterio tomando como método de referencia una valoración médica: especificidad=70%, sensibilidad=86% (punto de corte 108 puntos)
<i>Neurosensory Motor Development Assessment</i> - NMDA, Yvonne Burns	Acuerdo observado: 0,8 (22) **REE: CCI=0,91 Promedio de diferencias = 1,23; LA: -7,51;9,97 CC kappa= 0,71 **RIE: CC= 0,81	Validez convergente: un solo factor (23) **Validez convergente en comparación con la EAD: $\rho=0,51$
Escala Abreviada del Desarrollo - EAD (31)	**REE: CCI= 0,96 Promedio de diferencias = 0,02; LA: -1,81;1,84 CC kappa= EMG: 0,74; EMF: 0,97	

RIE: reproducibilidad intraevaluador; REE: reproducibilidad entre evaluadores; RTT: reproducibilidad test-retest; EGM: escala motora gruesa; EMF: escala motora fina; LA: límites de acuerdo de Bland y Altman; r: coeficiente de correlación de Pearson; CCI: coeficiente de correlación intraclase; CC Kappa: coeficiente de correlación kappa; $\hat{\rho}$: coeficiente de correlación de Spearman; DCD: trastorno de coordinación en el desarrollo; DCDQ: *Developmental Coordination Disorder Questionnaire*; **: Resultados obtenidos en el presente estudio.

coeficientes de correlación intraclase y validez de criterio (cuadro 5) (19).

Se observa, entonces, que en la evaluación de la reproducibilidad de las pruebas, con frecuencia se ha aplicado el coeficiente de correlación de Pearson, que solamente indica la relación lineal entre las variables, pero que, en realidad, no mide su verdadero acuerdo (51). En otros estudios revisados, únicamente se emplea el coeficiente de correlación intraclase, que también tiene dificultades en su método de cálculo (11,12). No se aprecia la utilización de la metodología propuesta por Bland y Altman, que expresa de manera más adecuada el acuerdo de las mediciones, porque examina la magnitud de discrepancias entre cada par de observaciones, en la misma escala de medición de la variable original, lo cual permite un mejor entendimiento del análisis desde la perspectiva clínica (52).

En cuanto a la evaluación de la validez, se ha tratado de establecer la validez concurrente (un tipo de validez de criterio) en estudios en los que se comparan instrumentos con objetivos similares,

pero no se reconoce alguno como el referente y simplemente se evalúa la correlación entre los dos, lo que en realidad corresponde a una evaluación de validez convergente entre dos instrumentos (53), como la evaluada en el presente estudio.

En el estudio realizado por Provost y colaboradores (20), en el que se evalúa la validez de criterio del *Bayley Scales of Infant Development* con el *Peabody Developmental Motor Scales*, por ejemplo, se referencian diversos estudios en los que a estas pruebas se les ha evaluado su validez concurrente con otras menos conocidas y se reportan los respectivos coeficientes de correlación de Pearson (cuadro 5); la misma metodología es empleada en el estudio mencionado y otro de la misma autora (54), en el que nuevamente compara las dos escalas mencionadas. En otros estudios, sin embargo, se ha sugerido al Peabody (55) y al BOTM (15) como la prueba de referencia, pero de manera similar se ha evaluado la correlación entre las pruebas, sin realizar cálculos propios de este tipo de validez como la sensibilidad, especificidad y los valores de predicción.

En el presente estudio se evaluaron la reproducibilidad y la validez convergente de dos instrumentos que intentan medir el desarrollo motor y tienen por objetivo identificar tempranamente alteraciones en este proceso: la Escala Abreviada del Desarrollo, empleada en Colombia, y la valoración australiana *Neurosensory Motor Development Assessment*, utilizada en la práctica clínica y en investigación. De la escala abreviada no se cuenta con reportes de evaluación de sus propiedades psicométricas; del NSMDA, aunque existen reportes de buena reproducibilidad, validez convergente y criterio, la evaluación de sus propiedades psicométricas no ha contado con la metodología más apropiada; por ejemplo, se encuentran reportes de “alta” reproducibilidad entre evaluadores basados solamente en el porcentaje de acuerdo observado. Asimismo, se han realizado varios estudios con el fin de evaluar su validez de criterio para pronosticar alteración motora en los meses 1, 4, 8 y 12 de edad, para lo cual emplean como referencia la misma prueba en una valoración posterior a los 2 o 4 años de edad y no una prueba diferente (23-25).

Los resultados de este estudio mostraron dos instrumentos con muy buena reproducibilidad entre evaluadores; ligeramente mayor para la Escala Abreviada de Desarrollo (CCI de 0,96) comparada con el NSMDA, de 0,91; ambos estimados contaron con intervalos de confianza estrechos. El promedio de las diferencias entre evaluadores y el acuerdo en cada instrumento fue muy pequeño (representaron menos del 5% del rango de valores posibles dentro de la escala de cada prueba). Así mismo, el error estándar de las mediciones fue pequeño (3,08 en el NSMDA y 0,65 en la Escala Abreviada de Desarrollo) lo cual es otro hallazgo que ratifica la precisión de las pruebas y fortalece su utilización en la práctica clínica.

La reproducibilidad intraevaluadores del instrumento australiano fue buena (0,81) aunque menor que entre evaluadores; además, a diferencia de esta última, el intervalo de confianza del estimado fue amplio debido al pequeño tamaño de la muestra empleada para evaluar esta propiedad. De manera comparativa con el acuerdo entre evaluadores, el promedio de las diferencias entre la primera y la segunda valoración fue mayor

($2,67 \pm 4,47$) que en las diferencias entre los dos evaluadores ($1,23 \pm 4,46$), ya que la evaluadora tendió a asignar puntajes mayores en la segunda valoración, hecho que puede atribuirse al conocimiento de las pruebas tanto por parte del niño como del evaluador.

Cabe resaltar que en la valoración del comportamiento motor de los niños se reconoce que existen diversas fuentes de error que pueden conducir a la baja reproducibilidad de las mediciones, inclusive la misma situación del “examen”, puede ocasionar que el niño modifique sus respuestas frente a las preguntas o requisitos de la prueba (47); por ello, la utilización de técnicas como la filmación para evaluar esta propiedad puede ser adecuada y, de esta manera, se garantiza que se valora el mismo desempeño.

En el estudio de Kroes y colaboradores (19), en el que se evaluó la reproducibilidad del *Maastricht Motor Test*, se apreció que los CCI de dos mediciones, realizadas con un intervalo de un mes entre la primera y la segunda, tendieron a tomar valores inferiores que los CCI de dos mediciones, de las cuales, la segunda se realizó a un video de la valoración inicial. Los autores discuten que este fenómeno pudo ocurrir debido a cambios reales en el comportamiento de los niños durante el mes transcurrido entre las valoraciones o deberse a aprendizaje del examen por cuenta de los niños, por lo que en la segunda valoración mostraron un mejor desempeño. Sin embargo, en el presente estudio el tiempo transcurrido entre las valoraciones fue de 10 días como máximo, por lo que no se considera que hayan ocurrido cambios reales en el comportamiento de los niños valorados.

Es importante señalar que en este estudio no se evaluó la reproducibilidad intraevaluadores de la escala abreviada; se priorizó la evaluación de la reproducibilidad entre evaluadores debido a que esta prueba es utilizada para la identificación precoz de alteraciones por parte del personal encargado del programa de crecimiento y desarrollo, el cual contempla dos valoraciones por año para los niños de 4 y 5 años (30). El tiempo transcurrido entre cada prueba puede ser superior a 6 meses, período en el que pueden ocurrir cambios verdaderos en el comportamiento de los

niños. Sin embargo, la falta de evaluación de la reproducibilidad intraevaluadores puede considerarse como una limitación del estudio.

En el análisis de la concordancia de cada instrumento, se apreció que el componente fino de la escala abreviada tuvo los mayores índices ($\kappa=0,97$ en la totalidad de los niños y $\kappa=1$ en los niños de 5 años), mientras que el NSMDA tuvo índices buenos pero inferiores. Debe tenerse presente que los puntos de corte establecidos en la escala abreviada para determinar el nivel de alerta (26 para la motricidad gruesa y 23 para la motricidad fina) corresponden al rango de los 49 a los 60 meses, y puesto que este instrumento tiene unos ítems específicos para los cuatro años y otros para los cinco, es posible que los puntos de corte sean inadecuados para los niños de mayor edad, ya que se espera que obtengan puntajes más altos.

En conclusión, los dos instrumentos comparados en este estudio demostraron una buena reproducibilidad al ser aplicados por fisioterapeutas con experiencia en el área pediátrica; sin embargo, es conveniente reconocer que no cuentan con el mismo enfoque de valoración. Estas diferencias pueden ocasionar discrepancias en la interpretación del desempeño de un niño y, por consiguiente, orientar conductas de intervención diferentes. La Escala Abreviada del Desarrollo se dirige a evaluar el comportamiento voluntario y rítmico, enfocándose en determinar la presencia o ausencia de habilidades motoras (aspectos cuantitativos), mientras que el NSMDA, además de determinar la presencia de habilidades motoras, intenta valorar la calidad en su ejecución. Por ejemplo, la valoración del área motora fina en la Escala Abreviada del Desarrollo no hace referencia a los patrones de movimiento de las manos, sino al producto de tareas realizadas, como dibujar una escalera, las cuales pudieran realizarse empleando diversos movimientos; mientras en el NSMDA, este tipo de motricidad se dirige al movimiento de las manos y los dedos en la ejecución de dos tareas: agarre del lápiz y enhebrado.

De otro lado, en relación con la validez convergente entre los dos instrumentos, se encontró una relación moderada al incluir la totalidad de los

ítems ($r=0,51$); sin embargo, al correlacionar los puntajes de la escala abreviada con los puntajes de las áreas del instrumento australiano correspondientes a la motricidad gruesa y fina, el coeficiente es mayor ($r= 0,61$), fenómeno que puede explicarse por una mayor relación entre ítems que miden movimientos similares. El resultado de la convergencia entre las pruebas es reflejo de las discrepancias en el enfoque de la valoración de cada uno discutido previamente.

Es importante enfatizar que la identificación temprana de las alteraciones en el desarrollo motor demanda la existencia de instrumentos de medición apropiados, que no constituyan simplemente listas de chequeo, sino que valoren los componentes del movimiento necesarios para alcanzar habilidades motoras específicas y funcionales, ya que los hallazgos de estas mediciones orientan y definen las conductas de intervención (9).

La Escala Abreviada del Desarrollo es un instrumento de rápida aplicación que proporciona una visión general de la motricidad de los niños y puede ser de utilidad en pesquisas rápidas del desarrollo motor; sin embargo, en la aplicación e interpretación de sus resultados en la práctica clínica, se debe tener presente que su enfoque de valoración se limita a aspectos cuantitativos del desarrollo motor, por lo que este instrumento no se debe utilizar como una herramienta diagnóstica por el fisioterapeuta, ya que pueden pasar desapercibidas alteraciones leves de este proceso.

El NSMDA es una prueba más compleja de la cual se destaca su meticulosidad en la valoración de estímulos sensoriales necesarios para un adecuado control del movimiento y una aproximación a la medición de la calidad del mismo; aunque, precisamente debido a la cantidad de ítems que evalúa, su aplicación consume gran cantidad de tiempo y depende de mucha colaboración del niño, en especial, para la valoración de los enderezamientos, lo cual puede restringir su uso en la práctica clínica.

Cabe llamar la atención sobre las teorías del control motor y desarrollo motor actuales (48,56,57), que establecen que para la realización de movimiento

coordinado y funcional se requiere de una armoniosa integración de estímulos nerviosos sensoriales y motores, la cual se logra y refina con la maduración de estructuras nerviosas y osteomusculares, y a la influencia del ambiente y la experiencia del niño con este último.

Este conocimiento ha propiciado que instrumentos recientes de valoración del desarrollo motor, como la *Alberta Infant Motor Scale* (9) dirigida a niños entre 0 y 18 meses, se centren en la observación de los movimientos voluntarios y rítmicos, ya que su adecuada realización supone un funcionamiento apropiado de las estructuras involucradas en el control del movimiento, lo cual no desconoce principios como la ganancia gradual de control de postura y la adquisición de patrones de movimiento cada vez más complejos y funcionales.

Al tener en cuenta estas consideraciones teóricas y prácticas para la aplicación de un instrumento que mida el desarrollo motor y que sea útil en el contexto clínico, resulta coherente orientar la valoración del desarrollo motor en la niñez temprana hacia la medición de la motricidad gruesa y fina, como se contempla en la escala abreviada, pero haciendo énfasis tanto en la cantidad como en la calidad de los movimientos, como lo proponen otros instrumentos como el NSMDA.

Para finalizar la discusión, se recomienda continuar con el estudio de instrumentos estandarizados que valoren el desarrollo motor de los niños, debido a la importancia de este proceso en el desarrollo en general. Asimismo, se considera que la valoración del desarrollo motor debe ser idealmente conducida por un profesional de la salud cuyo objeto de estudio sea el movimiento corporal humano y preferiblemente con experiencia en el área pediátrica, ya que son las personas con la formación académica requerida para evaluar aspectos cualitativos del movimiento como la alineación, el control postural, el equilibrio, la coordinación y las habilidades funcionales.

Agradecimientos

Los autores agradecen la amplia disposición y colaboración de Yvonne Burns, quien facilitó el instrumento *Neurosensory Motor Development Assessment* con fines de investigación en nuestro

medio. Así mismo, expresan su agradecimiento a las fisioterapeutas Fanny Ortiz y Sulys Margarita Trujillo, quienes estuvieron a cargo de la valoración de los niños participantes.

Conflicto de intereses

Los autores manifiestan que en la realización del presente estudio no se presentaron conflictos de interés que pudieran afectar o sesgar los resultados o su interpretación.

Financiación

La financiación de esta investigación se realizó con recursos propios.

Referencias

1. **American Physical Therapy Association.** What types of tests and measures do physical therapist use? (A guide to physical therapist practice). *Phys Ther.* 2001; 81:51-103.
2. **American Academy of Pediatrics.** Developmental surveillance and screening of infants and young children. Committee on children with disabilities. *Pediatrics.* 2001;108:192-5.
3. **Goyen TA, Liu K.** Longitudinal motor development of «apparently normal» high-risk infants at 18 months, 3 and 5 years. *Early Hum Dev.* 2002;70:103-15.
4. **Kroes M.** Early motor, psychosocial and behavioural characteristics of attention-deficit/hyperactivity disorder (thesis). Maastricht: Maastricht University; 2000.
5. **Sullivan MC, McGraft MM.** Perinatal morbidity, mild motor delay and later school outcomes. *Dev Med Child Neurol.* 2003;45:104-12.
6. **Merrell KW, Holland ML.** Social-emotional behavior of preschool-age children with and without developmental delays. *Res Dev Disabil.* 1997;18:393-405.
7. **Kroes M, Kessels AG, Kalff AC, Feron FJ, Vissers YL, Jolles J, et al.** Quality of movement as predictor of ADHD: results from a prospective population study in 5-and 6-year-old children. *Dev Med Child Neurol.* 2002;44:753-60.
8. **Largo RH, Fischer JE, Rousson V.** Neuromotor development from kindergarten age to adolescence: developmental course and variability. *Swiss Med Wkly.* 2003;133:193-9.
9. **Piper MC, Darrah J.** Motor assessment of the developing infant. Philadelphia: W.B. Saunders Company; 1994.
10. **Brenneman SK.** Assessment and testing of infant and child development. En: Tecklin JS, editor. *Pediatric physical therapy.* 3ª edición. Philadelphia: Lippincott Williams & Wilkins; 1999. p. 28-70.

11. **Boulton JE, Kirsch SE, Chipman M, Etele E, White AN, Pape KE.** Reliability of the Peabody Developmental Gross Motor Scale in children with cerebral palsy. *Phys Occup Ther Pediatr.* 1995;15:35-51.
12. **Schmidh LS, Westcott SI, Crowe TK.** Interrater reliability of the gross motor scale of the Peabody Developmental Motor Scales with 4- and 5- year-old children. *Pediatr Phys Ther.* 1993;5:169-75.
13. **Wiat L, Darrah J.** Review of four tests of gross motor development. *Dev Med Child Neurol.* 2001;43:279-85.
14. **Flegel J, Kolobe T.** Predictive validity of the test of infant motor performance as measured by the Bruininks-Oseretsky test of motor proficiency at school age. *Phys Ther.* 2002;82:762-71.
15. **Crawford SG, Wilson BN, Dewey D.** Identifying developmental coordination disorder: Consistency between tests. *Phys Occup Ther Pediatr.* 2001;20:29-50.
16. **Frankenburg WK.** Developmental surveillance and screening of infants and young children. *Pediatrics.* 2002;109:144-5.
17. **Stengel TJ.** Assessing motor development in children. En: Campbell SK, editor. *Pediatric neurologic physical therapy.* Second edition. Illinois: Churchill Livingstone; 1993. p. 33-65.
18. **Rodger S, Ziviani J, Watter P, Ozanne A, Woodyatt G, Springfield E.** Motor and functional skills of children with developmental coordination disorder: A pilot investigation of measurement issues. *Hum Mov Sci.* 2003;22:461-78.
19. **Kroes M, Vissers YL, Sleijpen FA, Feron FJ, Kessels AG, Bakker E, et al.** Reliability and validity of a qualitative and quantitative motor test for 5 to 6 year old children. *Eur J Paediatr Neurol.* 2004;8:135-43.
20. **Provost B, Crowe T, McClain C.** Concurrent validity of the Bayley Scales of Infant Development II Motor Scale and the Peabody Developmental Motor Scales in two-year-old children. *Phys Occup Ther Pediatr.* 2000; 20:5-18.
21. **Burns YR.** *Physiotherapy assessment for infants and young children.* Brisbane: Publishing Company Pty Ltd; 1991.
22. **Burns YR, Ensbey RM, Norrie MA.** The Neurosensory Motor Development Assessment. Part 1: Development and administration of the test. *Aust J Physiother.* 1989; 35:141-9.
23. **Burns YR, Ensbey RM, Norrie MA.** The Neurosensory Motor Development Assessment. Part 2: Predictive and concurrent validity. *Aust J Physiother.* 1989;35:151-7.
24. **Burns YR, O'Callaghan M, Tudehope DI.** Early identification of cerebral palsy in high risk infants. *Aust Paediatr J.* 1989;25:215-9.
25. **Burns YR, Mohay H, Croker A.** The predictive value of developmental testing of children under the age of 2 years. *Physiother Theory Pract.* 1987;3:2-10.
26. **MacDonald JA, Burns YR, Mohay HA.** Characteristics of neuro-sensory-motor performance of very low birth weight and high-risk infants at six years of age. *NZ Journal of Physiotherapy.* 1991; 19:17-20
27. **Burns YR, Ensbey R, O'Callaghan.** Motor abilities at eight to ten years of children born weighing less than 1,000 gr. *Physiotherapy.* 1999;85:360-9.
28. **Gray PH, O'Callaghan J, Mohay HA, Burns YR, King JF.** Maternal hypertension and neurodevelopmental outcome in very preterm infants. *Arch Dis Child Fetal Neonatal Ed.* 1998;79:F88-93.
29. **Burns Y, O'Callaghan M, McDowell B, Rogers Y.** Movement and motor development in ELBW infants at 1 year is related to cognitive and motor abilities at 4 years. *Early Hum Dev.* 2004;80:19-29.
30. **Ministerio de Salud.** Resolución 412 de 2000. Santa Fe de Bogotá: Ministerio de Salud; 2000.
31. **Grupo Interinstitucional de Programas de Salud del Niño en Antioquia.** *Salud Integral para la Infancia (SIPI). Crecimiento y desarrollo por grupos de edad. Tomo II.* Medellín: Servicio Seccional de Salud de Antioquia; 1996.
32. **Staquet M, Hays R, Fayer P.** Assessing reliability and validity of measurement in clinical trials. En: *Quality of life assessment in clinical trials. Methods and practice.* New York: Oxford University Press; 1998. p. 171-5.
33. **Kraemer HC.** *Evaluating medical test.* California: Sage Publications; 1992.
34. **Kraemer HC, Thieman S, Denenberg VH.** Correlation coefficients. En: *How many subjects? Statistical power analysis in research.* London: Sage Publications; 1987.
35. **Norman GR, Streiner DL.** *Bioestadística.* Madrid: Mosby/Doyma Libros; 1996. p. 14-22.
36. **Bland JM, Altman DG.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-10.
37. **Lin L I-K.** A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989;45:255-68.
38. **Kraemer HC, Bloch D.** Kappa coefficients in epidemiology an appraisal of a reappraisal. *J Clin Epidemiol.* 1998;41:959-68.
39. **Kraemer HC, Periyakoil V, Noda A.** Kappa coefficients in medical research. *Stat Med.* 2002;21: 2109-29.
40. **Domholdt E.** *Rehabilitation research. Principles and applications.* 3rd edition. St. Louis, Mo: Elsevier Saunders; 2005. p. 245-64.
41. **Dean AG, Dean JA, Coulumbier D, Brendel KA, Smith DC, Burton AH, et al.** *Epi Info, Versión 6.01: a*

- word processing, database, and statistics program for epidemiology on microcomputers. Atlanta: Centers for Disease Control and Prevention; 1995.
42. **StataCorp.** Stata Statistical Software. Release 9.0. College Station, TX: StataCorp LP; 2005.
 43. **Coughlin SS, Beauchamp TL.** Ethics and epidemiology. New York: Oxford University Press; 1996.
 44. **Bland M, Altman D.** Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8: 135-60.
 45. **Latour J, Abaira V, Cabello JB, López J.** Métodos de investigación en cardiología clínica. *Rev Esp Cardiol.* 1997;50:117-28.
 46. **Bruns DE.** The STARD initiative and the reporting of studies of diagnostic accuracy. *Clin Chem.* 2003;49:18-20.
 47. **Campbell SK.** Measurement in developmental therapy: past, present and future. En: Miller LJ, editor. *Developing norm-referenced standardized tests.* New York: Haworth Press; 1989.
 48. **Shumway-Cook A, Woollacott MH.** Motor control. Theory and practical applications. 2nd. edition. Baltimore: Lippincott Williams & Wilkins; 2001. p. 1-25.
 49. **Koseck K.** Review and evaluation of psychometric properties of Revised Bayley Scales of Infant Development. *Pediatr Phys Ther.* 1999;11:198-204.
 50. **Keating J, Matyas T.** Unreliable inferences from reliable measurements. *Aust J Physiother.* 1998;44:5-10.
 51. **Guyatt G, Walter S, Norman G.** Measurement change over time: Assessing the usefulness of evaluates instruments. *J Chronic Dis.* 1987;40:171-8.
 52. **Patton N, Aslam T, Murray G.** Strategies to assess reliability in ophthalmology. *Eye.* 2006;20:749-54.
 53. **Hays RD, Anderson RT, Revicki D.** Assessing reliability and validity of measurement in clinical trials. En: Staquet M, Hays R, Fayer P, editores. *Quality of life assessment in clinical trials. Methods and practice.* New York: Oxford University Press; 1998. p. 171-5.
 54. **Provost B, Heimerl S, McClain C, Kim N-H, López BR, Kodituwakku P.** Concurrent validity of the Bayley Scales of Infant Development II Motor Scale and the Peabody Developmental Motor Scales-2 in Children with Developmental Delays. *Pediatr Phys Ther.* 2004;16:149-56.
 55. **Jain M, Turner D, Worrel T.** The Vulpe Assessment Battery and the Peabody Developmental Motor Scales: a preliminary study of concurrent validity between gross motor sections. *Phys Occup Ther Pediatr.* 1994;14: 23-32
 56. **Hadders-Algra M.** The Neuronal Group Selection Theory: a framework to explain variation in normal development. *Dev Med Child Neurol.* 2000;42:566-72.
 57. **Hadders-Algra M.** Variability in the infant motor behavior: A hallmark of the healthy nervous system. *Infant Behav Dev.* 2002;25:433-51.

Anexo 1. Comparativo de la Escala Abreviada del Desarrollo (EAD) y el *Neurosensory Motor Development Assessment* (NSMDA)

INSTRUMENTO		NSMDA							
EAD - componente motor		Matricidad gruesa	Matricidad fino-adaptativa	Matricidad gruesa	Matricidad fina	Patrones primitivos	Neurológica	Reacciones posturales	Sensorio motor
Áreas correspondientes a los 4 y 5 años de edad	<ul style="list-style-type: none"> • Camina en línea recta. • Salta tres o más pasos en un pie. • Hace rebotar y agarra la pelota. • Salta a pies juntillas cuerda a 25 cm. • Hace caballito alternando los pies. • Salta desde 60 cm de altura. 	<ul style="list-style-type: none"> • Dibuja figura humana II. • Agrupa por color y forma. • Dibuja escalera (imita). • Agrupa por color, forma y tamaño. • Reconstruye escalera de 10 cubos. • Dibuja casa. 	<ul style="list-style-type: none"> • Calidad del sentado • Calidad de la marcha • Calidad de la carrera • Calidad del equilibrio unipodal • Equilibrio y coordinación (marcha en línea recta) • Salto en el sitio • Salto en los dos pies 	<ul style="list-style-type: none"> • Agarre de lápiz • Enhebrado 	<ul style="list-style-type: none"> • Patrones que interfieren con los patrones flexores • Patrones que interfieren con los patrones extensores 	<ul style="list-style-type: none"> • Tono en MMSS y MMII • Reflejos osteotendinosos en MMSS y MMII • Clonus en MMSS y MMII • Temblor en MMSS y MMII 	<ul style="list-style-type: none"> • Enderzamientos ópticos horizontales y verticales • Enderzamientos vestibulares y horizontales y verticales • Landau • Colocación en MMSS y MMII • Soporte en MMSS y MMII • Traslado de peso • Reacciones protectoras en MMSS y MMII • Reacciones de equilibrio en posición sedente y de pie 	<ul style="list-style-type: none"> • Localización táctil en MMSS, MMII y cara • Seguimiento ocular • Propiocepción en brazos y manos 	
Escala de medición	Nominal Dicotómica Sí-NO	Ordinal de 1 a 3 ó de 1 a 4							

MMSS: Miembros superiores; MMII: Miembros inferiores.

Anexo 2. Concordancia entre las mediciones inter-evaluador de cada ítem, *Neurosensory Motor Development Assessment*.

Área	Ítem	Kappa ponderado
Motricidad gruesa	Sentado	0,7457
	Marcha	0,7981
	Carrera	0,7530
	Equilibrio unipodal	0,8527
	Marcha en línea recta	0,7639
	Salto en sitio (en un pie)	0,8490
	Atrapado de balón	0,9229
Motricidad fina	Agarre de lápiz	0,8271
	Enhebrado de cuentas	0,8957
Patrones primitivos	Patrón 1	0,5853
	Patrón 2	0,3367
Neurológico	Tono MSI	0,8770
	Tono MSD	0,9021
	Tono MII	0,7917
	Tono MID	0,8062
	Reflejos tendinosos MSI	0,5638
	Reflejos tendinosos MSD	0,5794
	Reflejos tendinosos MII	0,7969
	Reflejos tendinosos MID	0,7696
	Clono MSI	-
	Clono MSD	-
	Clono MII	1,0000
	Clono MID	1,0000
	Temblor MSI	0,4979
	Temblor MSD	0,5037
	Temblor MII	0,4616
	Temblor MID	0,4607
Postural y equilibrio	Enderezamiento óptico vertical adelante	0,2804
	Enderezamiento óptico vertical atrás	0,6582
	Enderezamiento óptico vertical izquierdo	0,5387
	Enderezamiento óptico vertical derecha	0,6823
	Enderezamiento óptico horizontal prono	0,8383
	Enderezamiento óptico horizontal supino	0,6432
	Enderezamiento óptico horizontal izquierdo	0,8184
	Enderezamiento óptico horizontal derecho	0,7631
	Enderezamiento vestibular vertical adelante	0,5377
	Enderezamiento vestibular vertical atrás	0,6320
	Enderezamiento vestibular vertical izquierdo	0,6792
	Enderezamiento vestibular vertical derecha	0,5880
	Enderezamiento vestibular horizontal prono	0,6911
	Enderezamiento vestibular horizontal supino	0,7958
	Enderezamiento vestibular horizontal izquierdo	0,8203
	Enderezamiento vestibular horizontal derecho	0,8415
	Landau	0,8414
	Colocación MSI	1,0000
	Colocación MSD	1,0000
	Colocación MII	0,7363
	Colocación MID	0,5089
	Soporte MSI	0,8166
	Soporte MSD	0,8576
Soporte MII	0,7853	
Soporte MID	0,7910	

Área	Ítem	Kappa ponderado
	Traslado de peso superior	0,8292
	Traslado de peso inferior	0,3710
	Protectoras MMSS adelante	0,7753
	Protectoras MMSS atrás	0,8800
	Protectoras MMSS izquierda	0,8497
	Protectoras MMSS derecha	0,9194
	Protectoras MMII adelante	0,7590
	Protectoras MMII atrás	0,8529
	Protectoras MMII izquierda	0,8277
	Protectoras MMII derecha	0,8822
	Equilibrio sedente hacia adelante	0,5812
	Equilibrio sedente hacia atrás	0,7528
	Equilibrio pie hacia adelante	0,7640
	Equilibrio pie hacia atrás	0,7615
Sensorio-motor	Localización táctil MSI	0,5596
	Localización táctil MSD	0,4525
	Localización táctil MII	0,5156
	Localización táctil MID	0,4865
	Localización táctil facial izquierda	0,5282
	Localización táctil facial derecha	0,5922
	Seguimiento ocular izquierda	0,9125
	Seguimiento ocular derecha	0,9120
	Propiocepción MSI	0,6328
	Propiocepción MSD	0,6119
	Propiocepción mano izquierda	0,5959
	Propiocepción mano derecha	0,6526

Instrumento: Escala Abreviada del Desarrollo

Área	Ítem	Kappa ponderado
Motricidad gruesa	Camina en línea recta.	- - - -
	Salta en un pie.	0,8361
	Rebota y agarra la pelota.	0,8567
	Salta a pies juntilla.	0,8111
	Hace caballito.	0,8699
	Salta desde altura.	0,8612
Motricidad fina	Dibuja figura humana.	0,9710
	Agrupar por color y forma.	0,9121
	Dibuja escalera.	0,9409
	Agrupar por color, forma y tamaño.	0,9461
	Reconstruye escalera.	0,8719
	Dibuja casa.	0,9197

MMSS: miembros superiores; MMII: miembros inferiores; MSD: miembro superior derecho; MSI: miembro superior izquierdo; MII: miembro inferior izquierdo; MID: miembro inferior derecho.