



Consejo General
de la **Psicología**
ESPAÑA



Miembro de la Federación Europea
de Asociaciones de Psicólogos

Conde Peñalver, 45. 5º Izq.
28006 Madrid – España
Tel.: +34 91 444 90 20
Fax: +34 91 309 56 15
secop@cop.es
www.cop.es

EVALUACIÓN DE LA BATERÍA DE EVALUACIÓN DEL MOVIMIENTO PARA NIÑOS-2 (MABC-2)

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Características	Descripción
Nombre del test	Batería de evaluación del movimiento para niños 2 (MABC-2)
Autor	Sheila E. Henderson, David A. Sugden y Anna L. Barnett
Autor de la adaptación española	Luis Miguel Ruiz Pérez, José Luis Graupera-Sanz, en colaboración con el Departamento de I+D de Pearson Clinical and Talent Assessment España: Elena de la Guía, Ana Hernández, Erica Paradell, Frédérique Vallar
Variable(s)	Competencia motriz
Área de aplicación	Psicología clínica, psicología educativa, neuropsicología e investigación
Soporte	Papel y lápiz y manipulativo

Valoración general

Características	Valoración	Puntuación
Materiales y documentación	Excelente	5
Fundamentación teórica	Excelente	5
Adaptación	Excelente	5
Análisis de los ítems	Buena	4
Validez: contenido	Buena	4
Validez: relación con otras variables	Buena	4
Validez: estructura interna	Buena-excelente	4,5
Validez: análisis del DIF	Adecuada	3
Fiabilidad: equivalencia	--	--
Fiabilidad: consistencia interna	Buena	4
Fiabilidad: estabilidad	Adecuada	3
Fiabilidad: TRI	--	--
Fiabilidad: inter-jueces	Excelente	5
Baremos e interpretación de puntuaciones	Buena-excelente	4,5

Comentarios generales

Las dificultades o limitaciones motrices de los niños en la etapa escolar pueden tener importantes consecuencias para su aprendizaje general y para su desarrollo socioemocional (Caçola, 2014). Por ello es necesario disponer de una prueba que permita evaluar sus potencialidades y limitaciones sin que el niño se sienta evaluado. En este sentido, la Batería de Evaluación del Movimiento para Niños-2 (MABC-2) se ha convertido en una de las baterías más utilizadas para identificar si “la adquisición y ejecución de habilidades motoras coordinadas está muy por debajo de lo esperado para la edad cronológica del individuo y la oportunidad de aprendizaje y el uso de las aptitudes. Las dificultades se manifiestan como torpeza así como lentitud e imprecisión en la realización de habilidades motoras” (Criterio A del DSM-5 para el Trastorno del Desarrollo de la Coordinación; 2014).

La MABC-2 ofrece un contexto de evaluación dinámico y atractivo para el niño. Además, la corrección es sencilla y el sistema de semáforo adoptado permite tener una imagen gráfica que resulta de gran ayuda en el proceso de interpretación de la puntuación Total del Test. Por otro lado, al permitir el registro de muchos y diferentes aspectos del comportamiento de un niño, favorece el desarrollo de programas de intervención ajustados a las limitaciones y potencialidades del niño. Asimismo, es una herramienta válida para evaluar la eficacia de programas de intervención en niños con dificultades motrices, observando los cambios en su rendimiento motor.

El test MABC-2 es un buen instrumento de medida que puede ser utilizado para evaluar la competencia motriz en niños en diferentes contextos (educativo, clínico e investigativo). De entre los puntos positivos del instrumento podría destacar los siguientes:

- Los materiales que proporciona. El test contiene un manual bien estructurado que describe con claridad los aspectos técnicos del mismo. Por su parte, los objetos con los que los niños tienen que trabajar (barritas perforadas, tornillos, clavijero, clavijas, soportes para equilibrio, etc.), así como el material para el evaluador (cronómetro, flexómetro, etc.), son de muy buena calidad.
- Hoja de respuestas para cada rango de edad. Las hojas de respuesta para cada rango de edad están impresas en diferentes colores y su diseño ha sido pensado para evitar o reducir errores en la recogida de datos.
- Manual de intervención ecológica. El test incluye un manual orientado a mejorar las habilidades motrices de los niños que puede ser de utilidad en contextos aplicados considerando el resultado de la evaluación generada por el test.
- Evaluación cualitativa y llamada de atención sobre aspectos que pueden afectar a la ejecución en el test. Cabe destacar la sensibilidad se aprecia por parte de los autores con relación al fenómeno de la medición psicológica. Por ello, en el manual del test se advierte a las personas que van a aplicarlo sobre la correcta interpretación de las puntuaciones que el test genera. Adicionalmente a los ítems propiamente dichos del test, el cuadernillo de registro contiene ítems destinados a que la persona que administra el test registre información cualitativa del comportamiento del niño. De este modo, la puntuación generada por el test es más rica a nivel informativo y se

espera, como indican los autores, que la interpretación producida por la herramienta sea interpretada más sensatamente.

Por su parte, también se pueden destacar algunos aspectos que podrían mejorarse en futuras versiones del test.

- En el Manual se establece que cuando una prueba se completa con el lado derecho y con el lado izquierdo se deben sumar las puntuaciones escalares de cada segmento corporal y dividirlos entre dos. No se explica por qué se ha adoptado esa decisión, faltaría un argumento teórico para tal propuesta metodológica.

- Evidencias sobre la validez circunscritas a España. Como se ha indicado más arriba, creo que se podrían haber aportado más evidencias sobre la validez y fiabilidad de las puntuaciones que genera el test en el contexto en el que ha sido aplicado para realizar la adaptación.

- Se podrían llevar a cabo estudios desde óptica la TRI.

De manera general el test es un instrumento de calidad que tiene una potencialidad muy importante para ahondar en la comprensión de los trastornos psicológicos relacionados con la psicomotricidad. Por ello, su uso tanto clínica como educativamente así como en el ámbito de la investigación psicológica podría ser de utilidad para abordar algunos de los problemas más acuciantes a los que se enfrenta nuestra sociedad y que tienen que ver con la motricidad de los niños.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

- Batería de evaluación del movimiento para niños-2 (MABC-2).

1.2. Nombre del test en su versión original:

- Movement Assessment Battery for Children-2 (Movement ABC-2).

1.3. Autor del test original:

- Sheila E. Henderson, David A. Sugden y Anna L. Barnett.

1.4. Autor de la adaptación española:

- Luis Miguel Ruiz Pérez, José Luis Graupera-Sanz, en colaboración con el Departamento de I+D de Pearson Clinical and Talent Assessment España: Elena de la Guía, Ana Hernández, Erica Paradell y Frédérique Vallar.

1.5. Editor del test en su versión original:

- Pearson Education, Inc.

1.6. Editor de la adaptación española:

- Pearson Educación.

1.7. Fecha de publicación del test original:

- 2007.

1.8. Fecha de la publicación del test en su adaptación española:

- 2012.

1.9. Fecha de la última revisión del test en su adaptación española:

- 2012.

1.10. Área general de la(s) variable(s) que pretende medir el test:

- Competencia motriz.

1.11. Breve descripción de la(s) variable(s) que pretende medir el test:

- La Batería de Evaluación del Movimiento en Niños-2 (MABC-2) es un test estandarizado que requiere realizar una serie de pruebas motrices siguiendo unas pautas estrictamente establecidas, que permiten identificar y describir a los niños con dificultades motrices. Se ha convertido en una de las baterías más utilizadas para identificar si “la adquisición y ejecución de habilidades motoras coordinadas está muy por debajo de lo esperado para la edad cronológica del individuo y la oportunidad de aprendizaje y el uso de las aptitudes. Las dificultades se manifiestan como torpeza así como lentitud e imprecisión en la realización de habilidades motoras” (Criterio A del DSM-5 para el Trastorno del Desarrollo de la Coordinación; 2014). Por otro lado, al permitir el registro de muchos y diferentes aspectos del comportamiento de un niño, favorece el desarrollo de programas de intervención ajustados a las limitaciones y potencialidades del niño. En este sentido, también es una herramienta válida para evaluar la eficacia de programas de intervención en niños con dificultades motrices, observando los cambios en su rendimiento motor.
- El test está dividido en tres rangos de edad (4-6 años; 7-10 años; 11-16:11 años). Para cada rango de edad se aplican 8 ítems agrupados en tres dimensiones: Destreza manual, Puntería y atrape y Equilibrio (estático y dinámico). En el caso de la dimensión de Destreza manual, las tres pruebas que la integran han sido diseñadas para mostrar cómo afronta el niño las demandas especiales y temporales que plantean las tareas manuales. En concreto, permiten analizar la precisión con la que el niño ejecuta acciones bastante exigentes (p. e., el trazo entre dos líneas) y además observar cómo actúa el niño bajo presión (en dos tareas se añaden instrucciones relacionadas con la velocidad). Por otro lado, las dos pruebas de Puntería y atrape suponen una combinación compleja de movimientos finos y gruesos. En concreto integra dos componentes motores básicos de esta habilidad motriz: la precisión al recibir un objeto móvil y la precisión para dar en una diana. Tanto en las pruebas de Destreza manual como en las de Puntería y atrape se pueden observar el esfuerzo empleado por el niño en cada tarea y la capacidad para controlar la fuerza ejercida. Por último, las

pruebas de Equilibrio permiten observar la capacidad de estabilización del cuerpo que tiene el niño, que es fundamental para el control adaptativo del movimiento. Las tres pruebas que componen esta dimensión permiten observar la capacidad del niño para mantenerse en equilibrio estático y en equilibrio dinámico cuando debe ejecutar movimientos lentos y precisos.

- La Batería MABC-2 se administra de forma individual. El tiempo de aplicación de estas ocho pruebas es de 20 a 40 minutos dependiendo de la edad del niño y del grado de dificultad experimentado, así como de la experiencia del examinador.

1.12. Áreas de aplicación:

- Psicología clínica, psicología educativa, neuropsicología, psicología del deporte, e investigación.

1.13. Formato de los ítems:

- Respuesta dicotómica y de elección múltiple.

1.14. Número de ítems:

Observaciones cuantitativas:

- Dimensión de Destreza manual: 3 pruebas de ejecución.
- Dimensión de Puntería y atrape: 2 pruebas de ejecución.
- Dimensión de Equilibrio: 3 pruebas de ejecución.

Observaciones cualitativas:

- Dimensión de Destreza manual: insertar o voltear clavijas (15 ítems); enhebrar cuentas (15 ítems), entrelazar el cordel (15 ítems), montar un triángulo (13 ítems); dibujar el trazado (14 ítems).
- Puntería y atrape: atrapar el saquito (14 ítems), atrapar con dos manos (15 ítems), atrapar la pelota con una mano (15 ítems); lanzar el saquito a una diana (11 ítems).
- Equilibrio: equilibrio sobre una pierna (7 ítems), equilibrio sobre un soporte (7 ítems), equilibrio sobre dos soportes (7 ítems); andar de puntillas (11 ítems), andar hacia delante talón-punta (10 ítems), andar hacia atrás talón-punta (10 ítems); saltar sobre alfombrillas (14 ítems), saltar a la pata coja en línea (14 ítems), saltar a la pata coja en zigzag (13 ítems).
- Factores no motores que podrían afectar al movimiento: 13 ítems.
- Factores físicos que podrían afectar al movimiento: 6 ítems.

1.15. Soporte:

- Papel y lápiz y manipulativo.

1.16. Cualificación requerida para el uso del test de acuerdo con la documentación aportada:

- Nivel A.

1.17. Descripción de las poblaciones a las que el test es aplicable:

- Edad: de 4 a 16 años.
- Nivel educativo correspondiente a la edad.

- Aplicable a grupos clínicos: el ámbito de aplicación de la Batería MABC-2 en grupos clínicos es más amplio: niños con Trastorno por Déficit de Atención con Hiperactividad, niños con Trastorno Específico del Lenguaje, niños con Trastorno del Espectro Autista, niños con Síndrome de Tourette, niños con Síndrome de Williams, niños con Síndrome de Turner, niños con Síndrome de X Frágil, niños con hipotiroidismo congénito, niños con hiperlaxitud articular, niños con trastornos hereditarios del tejido conectivo (Síndrome de Marfan, Síndrome de Ehlers-Danlos), niños cuya movilidad se ha visto limitada durante periodos de tiempo largos (recuperación de cirugía cardíaca, quimioterapia para leucemia linfoblástica aguda).

1.18. Existencia o no de diferentes formas del test y sus características:

- Hay tres versiones del Test en función del rango de edad. Para cada rango de edad se ha mantenido la misma estructura del Test, es decir, ocho pruebas organizadas en tres dimensiones: Destreza manual, Puntería y atrape y Equilibrio. Sin embargo, el contenido de las pruebas varía:

1) Rango de edad 1 (4-6 años). En este caso se han introducido ciertas modificaciones en algunas de las pruebas en función de si los niños tenían 4 años o 5-6 años. Los ítems que integran el Test para este rango son:

Para la dimensión de Destreza manual:

- Introducir monedas. 6 monedas para los niños de 4 años, 12 para los de 5-6 años.
- Enhebrar cuentas. 6 cuentas para los 4 años, 12 cuentas para los 5-6 años.
- Dibujar el trazado. Se utiliza un rotulador.

Para la dimensión de Puntería y atrape:

- Atrapar el saquito. Los niños de 4 años pueden realizar el atrape contra su cuerpo, los de 5-6 años deben atraparlo en el aire sin emplear su cuerpo.
- Lanzar el saquito a una diana. En este caso se considera válido que el saquito caiga dentro de la alfombrilla aunque quede fuera de la diana.

Para la dimensión de Equilibrio:

- Equilibrio sobre una pierna.
- Andar de puntillas.
- Saltar sobre alfombrillas. Solo los niños de 5-6 años deben hacerlo con continuidad en los saltos.

2) Rango de edad 2 (7-10 años). Se introducen ciertas modificaciones en algunas de las pruebas en función de si los niños tenían 7-8 años o 9-10 años. Los ítems que integran el Test para este rango son:

Para la dimensión de Destreza manual:

- Insertar clavijas.
- Entrelazar el cordel.
- Dibujar el trazado. Se utiliza un bolígrafo.

Para la dimensión de Puntería y atrape:

- Atrapar con dos manos. Para los niños de 7-8 años con bote en el suelo, para los de 9-10 años sin bote en el suelo.
- Lanzar el saquito a una diana. Debe caer dentro de la diana, no es válido si cae en otra parte de la alfombrilla.

Para la dimensión de Equilibrio:

- Equilibrio sobre un soporte.
- Andar hacia delante talón-punta.
- Saltar a la pata coja en línea.

3) *Rango de edad 3 (11-16:11 años)*. Los ítems que integran el Test para este rango son:

Para la dimensión de Destreza manual:

- Voltar clavijas.
- Montar un triángulo.
- Dibujar el trazado. Se utiliza un bolígrafo.

Para la dimensión de Puntería y atrape:

- Atrapar la pelota con una mano.
- Lanzar a una diana.

Para la dimensión de Equilibrio:

- Equilibrio sobre dos soportes.
- Andar hacia atrás talón-punta.
- Saltar a la pata coja en zigzag.

1.19. Procedimiento de corrección:

- Manual.

1.20. Puntuaciones:

- En función de la dimensión y del ítem a ejecutar se debe anotar la mano preferida y utilizada (introducir monedas, insertar clavijas, voltear clavijas, dibujar el trazado, lanzar el saquito a una diana, lanzar a una diana), el tiempo empleado (introducir monedas, enhebrar cuentas, equilibrio sobre una pierna, insertar clavijas, entrelazar el cordel, equilibrio sobre un soporte, voltear clavijas, montar un triángulo, equilibrio sobre dos soportes), el número de errores (dibujar el trazado), el número de atrapes correctos (atrapar el saquito, atrapar con dos manos, atrapar la pelota con una mano), el número de lanzamientos correctos (lanzar el saquito a una diana, lanzar a una diana), el número de pasos consecutivos correctos (andar de puntillas, andar hacia delante talón-punta, andar hacia atrás talón-punta), el número de saltos consecutivos correctos (saltar sobre alfombrillas, saltar a la pata coja en línea, saltar a la pata coja en zigzag). Además, en las pruebas de Destreza manual y de Puntería y atrape cuando el niño falla o no realiza la tarea adecuadamente habrá que indicar una “F” y en todas ellas, incluidas las de equilibrio se marcará “R” cuando el niño rechazó llevar a cabo la prueba e “I” si la prueba es inapropiada para el niño.

- El procedimiento de corrección a seguir es el siguiente:
- En primer lugar se extrae la puntuación directa obtenida por el niño en cada prueba, que corresponde al mejor ensayo realizado. El segundo paso a seguir es convertir las puntuaciones directas en puntuaciones escalares que se han obtenido en base a la edad cronológica del niño para cada uno de los ítems que integran el Test. A continuación, para obtener la puntuación escalar y percentil de cada dimensión se deben sumar las puntuaciones escalares de los ítems de cada dimensión. Es decir, para la dimensión de Destreza manual hay que sumar las puntuaciones escalares de las tres pruebas de destreza manual. Para la dimensión de Puntería y atrape se deben sumar las puntuaciones de las dos pruebas que integran esta dimensión. Para la dimensión de Equilibrio se deben sumar las puntuaciones escalares de las tres pruebas que conforman ésta. Una vez obtenido el resultado de esa suma se consultan la tabla de baremos donde aparecen la puntuación escalar y percentil que corresponde a cada valor.
- Por otro lado, además el Test ofrece una puntuación Total que resulta de la suma de las puntuaciones escalares de las ocho pruebas y cuyos baremos para la obtención de la puntuación escalar y percentil total se recogen de nuevo en una única tabla de puntuaciones para todos los grupos de edad. Para esta puntuación Total la Batería MABC-2 utiliza puntos de corte mediante un sistema de semáforo, de forma que cualquier niño cuya puntuación Total del Test es igual o inferior a un percentil 5 probablemente presente una dificultad motriz importante. Mientras que un niño con puntuación situada entre el percentil 6 y el percentil 15 se situaría en la categoría de riesgo y cualquier niño con una puntuación entre el percentil 16 o por encima de él es improbable que tenga dificultades de movimiento.
- Ajustes específicos antes de proceder a convertir las puntuaciones directas en puntuaciones escalares:
 - En aquellos casos en los que un niño falla en los dos ensayos de los que consta cada prueba, se le asigna una puntuación escalar de 1.
 - Cuando una prueba se completa con el lado derecho y con el lado izquierdo se deben sumar las puntuaciones escalares de cada segmento corporal y dividir las entre dos.

1.21. Escalas utilizadas:

- Centiles, CI de desviación ($M=10$; $dt=3$) y puntuaciones estandarizadas normalizadas.

1.22. Posibilidad de obtener informes automatizados:

- No.

1.23. Tiempo estimado para la aplicación del test:

- Individual: de 20 a 40 minutos.

1.24. Documentación aportada por el editor:

- Manual y libros o artículos complementarios.

1.25. Precio de un juego completo de la prueba:

- 699,95 Euros. (consultado el 24 de mayo de 2016)

1.26. Precio y número de ejemplares del paquete de cuadernillos:

- Cuadernillo (25 unidades) de anotación rango 1 (4 a 6 años): 42,45 Euros.
- Cuadernillo (25 unidades) de anotación rango 2 (7 a 10 años): 42,45 Euros.
- Cuadernillo (25 unidades) de anotación rango 3 (11 a 16 años): 42,45 Euros.
(Consultado el 24 de mayo de 2016).

1.27. Precio y número de ejemplares del paquete de hojas de respuesta:

- No aplicable.

1.28. Precio de la corrección y/o elaboración de informes por parte del editor:

- No aplicable.

2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

2.1. Aspectos generales

Contenido	Valoración	Puntuación
2.1. Calidad de los materiales del test	Excelente	5
2.2. Calidad de la documentación aportada	Excelente	5
2.3. Fundamentación teórica	Excelente	5
2.4. Adaptación del test	Buena	4
2.5. Desarrollo de los ítems del test	Buena	4
2.6. Calidad de las instrucciones para el participante	Excelente	5
2.7. Calidad de las instrucciones (administración, puntuación, interpretación)	Excelente	5
2.8. Facilidad para registrar las respuestas	Excelente	5
2.9. Bibliografía del manual	Excelente	5
2.10. Datos sobre el análisis de los ítems	Adecuada	3

2.11. Validez

2.11.1. Evidencias de validez de contenido

Contenido	Valoración	Puntuación
2.11.1.1. Calidad de la representación del contenido o dominio	Buena	4
2.11.1.2. Consultas a expertos	Amplia muestra sistematizado	5

2.11.2. Evidencias de validez basadas en la relación entre las puntuaciones del test y otras variables

2.11.2.1. Evidencias de validez basadas en la relación con otras variables

Contenido	Valoración	Puntuación
2.11.2.1. Diseños empleados	Correlaciones con otros tests, diferencias de grupos	
2.11.2.1.1. Tamaño de las muestras	Varios estudios con muestras de tamaño moderado o con alguna muestra grande y otras pequeñas	4
2.11.2.1.2. Procedimiento de selección de las muestras	Aleatorio, con muestras representativas de la población objetivo	
2.11.2.1.3. Calidad de los tests empleados como criterio o marcador	Buena	4
2.11.2.1.4. Promedio de las correlaciones con otros tests que miden constructos similares	Adecuada	3
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos no relacionados	-	-
2.11.2.1.6. Resultados de la matriz multirrasgo-multimétodo	-	-
2.11.2.1.7. Resultados de las diferencias intergrupo	Buena	4

2.11.2.2. Evidencias de validez basadas en la relación entre las puntuaciones del test y un criterio

Contenido	Valoración	Puntuación
2.11.2.2.1. Criterios empleados	Amplia variedad de estudios de investigación en los que relacionan las medidas del Test con las de otros test motores y escalas de observación	
2.11.2.2.2. Calidad de los criterios empleados	Buena	4
2.11.2.2.3. Relación temporal entre test y criterio	Concurrente	
2.11.2.2.4. Tamaño de las muestras	Un estudio con una muestra moderada o varios con muestras pequeñas	2
2.11.2.2.5. Procedimiento de selección de las muestras	Aleatorio, con muestras representativas de la población objetivo	
2.11.2.2.6. Promedio de las correlaciones del test con los criterios	Buena	3

2.11.3. Evidencias de validez basadas en la estructura interna

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Buena	4
2.11.3.2. Funcionamiento diferencial de los ítems	-	-

2.11.4. Acomodaciones en la administración del test

Contenido	Valoración	Puntuación
2.11.4. El manual del test informa sobre las acomodaciones en la administración del test	No	

Comentarios generales sobre evidencias de validez

El Manual de la MABC-2 recoge información sobre distintos estudios de validación llevados a cabo con las versiones del Test. Se presentan diferentes estudios donde se recogen evidencias de validez en su versión original, y española. El tamaño de las muestras utilizadas en los diferentes estudios oscilan desde pequeños ($N = 20$) hasta muestras grandes ($N = 786$). Para la tipificación general de la versión española del MABC-2 se utilizó una muestra de 520 participantes de edades comprendidas entre los 4:00 y los 16:11 años. Para la configuración de la muestra se tomó como referencia los datos del censo de población tomando como referencia las variables sociodemográficas de edad, género, nivel de estudios de los padres y zona geográfica.

Con respecto a la primera versión del Test se aporta diversa evidencia basada en estudios de correlación. Las correlaciones con otros Test motores de ámbito similar oscilan entre valores moderados y elevados, tanto para la versión española de la Batería como para la versión británica. En concreto, en población española se comparó con las medidas de la ECOMI, que en su concepción es más similar a la Lista de Observación Conductual del MABC que al Test de ejecución. Sin embargo, pese a que los coeficientes de correlación obtenidos fueron moderados, en todos los casos fueron estadísticamente significativos. En otro de los estudios realizado con la versión española del Test se utilizaron las medidas obtenidas en el Test Stay-in-Step para compararlas con la puntuación Total del MABC y se correlacionaron los ítems de ambas pruebas. Los resultados mostraron coeficientes de correlación elevados en todos los casos. Por otro lado, los estudios recogidos en los que se utilizó la versión británica ponen de manifiesto resultados similares a los obtenidos con la

versión española. En este caso se comparó la ejecución del MABC con el BOT. En uno de estos estudios se analizó también la validez concurrente usando los rangos percentiles de ambos tests (Croce et al., 2001). Para el rango de edad 1 del MABC se llevó a cabo un estudio en el que la puntuación total del MABC se comparó con la puntuación total de las escalas PDMS-2, obteniendo un coeficiente de correlación bastante elevado. También se recogen comparaciones con tests más específicos como el KTK, un test alemán diseñado para identificar dificultades en los niños relacionadas con su agilidad motriz y su equilibrio dinámico, o con el Test VMI que permite evaluar la función manual y aspectos perceptivos. Las correlaciones con las medidas de KTK fueron elevadas, sin embargo con el Test VMI fueron moderadas, probablemente debido a su alta especificidad.

Por otro lado, el Manual también recoge estudios que se han llevado a cabo con la primera versión para evidenciar la capacidad del Test para discriminar entre grupos clínicos y no clínicos. Los resultados de estos estudios alcanzan valores muy aceptables, pero sólo para aquellos niños que están situados en el extremo más grave de la escala. Con población española se utilizó una muestra de 786 participantes de 4 a 12 años de edad, que en base a la puntuación total obtenida en el MABC se clasificaron como grupo sin riesgo de problemas de coordinación motriz (puntuación percentil superior a 15) y grupo con riesgo de problemas de coordinación motriz (puntuación percentil igual o inferior a 15). Se utilizó como criterio las medidas de la ECOMI y se encontraron diferencias significativas en todos los rangos de edad para ambos grupos.

Con respecto a la validez de la Batería MABC-2 se recoge evidencia de la relevancia del contenido a través de las conclusiones del panel de expertos sobre el contenido de la Batería, considerando que “el Test revisado representaba el ámbito motor especificado por los autores y lo cubría adecuadamente” (p. 153). La calidad de la representación de contenido es adecuada. Este aspecto queda bien argumentado en el Manual cuando se explican las diferentes dimensiones que integran el Test. Estas conclusiones se complementan con la opinión de especialistas que han utilizado el Test en sus ámbitos. En cuanto a la validez de criterio se reporta un estudio con una muestra de participantes pequeña. Se correlacionaron las puntuaciones de los ítems de la dimensión de Destreza manual del MABC-2 con las medidas del Test de la Figura Humana, obteniendo un coeficiente de correlación alto. Además, con respecto a la capacidad discriminante del Test, el Manual recoge dos estudios llevados a cabo con población clínica: uno con participantes que cumplen los criterios de TDC y otro con participantes diagnosticados con Síndrome de Asperger. En ambos casos, el Test ha sido capaz de detectar adecuadamente las dificultades motrices de los participantes.

En la versión española también se analizó la validez estructural del Test a través de un análisis factorial confirmatorio, utilizando la muestra de tipificación (N = 520). Se puso a prueba el modelo tridimensional del Test (destreza manual, puntería y atrape y equilibrio),

mostrando un ajuste general excelente. Asimismo, se comprobó la invarianza de medición entre los distintos grupos y se encontró que con respecto al género los índices de ajuste fueron muy similares a los del modelo general. Los resultados fueron similares cuando las comparaciones se hicieron por zonas geográficas. Sin embargo, no se recoge información sobre la variable edad.

Como comentario general sobre las evidencias de validez del test son adecuadas, si bien hubiese sido deseable recoger datos del test en la muestra española que hubieran servido para aportar evidencias sobre la validez de las inferencias sobre las puntuaciones que genera el test en el contexto español.

2.12. Fiabilidad

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Varios coeficientes de fiabilidad para cada escala así como para diferentes grupos. Error de medida para diferentes grupos.	

2.12.2. Equivalencia formas paralelas

Contenido	Valoración	Puntuación
2.12.2.1. Tamaño de las muestras	-	-
2.12.2.2. Puesta a prueba de los supuestos de paralelismo	-	-
2.12.2.3. Promedio de coeficientes de equivalencia	-	-

2.12.3. Consistencia interna

Contenido	Valoración	Puntuación
2.12.3.1. Tamaño de las muestras	Un estudio con una muestra grande	3
2.11.3.2. Coeficientes de consistencia interna presentados	Coeficiente alfa o KR-20	
2.11.3.3. Promedio de los coeficientes de consistencia	Adecuada	3

2.12.4. Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.12.4.1. Tamaño de las muestras	Un estudio con una muestra moderada o varios con muestras pequeñas	2
2.12.4.2. Coeficientes de estabilidad	Buena	4

2.12.5. Cuantificación de la precisión mediante TRI

Contenido	Valoración	Puntuación
2.12.5.1. Tamaño de las muestras	-	-
2.12.5.2. Coeficientes proporcionados	-	-
2.12.5.3. Tamaño de los coeficientes	-	-

2.12.6. Fiabilidad inter-jueces

Contenido	Valoración	Puntuación
2.12.6.1. Tipos de coeficientes presentados	Kappa y correlación intraclass	
2.12.6.2. Promedio de los coeficientes	Excelente	5

Comentarios generales sobre fiabilidad

La información presentada con relación a la fiabilidad del test (original y adaptada) es buena. Los valores obtenidos y presentados sobre el nivel de fiabilidad pueden considerarse apropiados.

En la adaptación española de la Batería MABC-2 la fiabilidad de las puntuaciones se calculó a través del coeficiente alfa con la muestra de niños de 4:00 a 16:11 años de edad (N = 520), mostrando una consistencia interna adecuada (>.70). El coeficiente ligeramente más elevado se obtuvo para el rango de edad 1 (4:00 a 6:11 años), con un valor de alfa de .81. En los tres rangos de edad y en la puntuación Total del Test los límites inferiores del intervalo de confianza del coeficiente alfa fueron superiores a .70. Además de los coeficientes de alfa también se calculó el error típico de medida, siendo en todos los casos inferior a 1.44. Estos resultados fueron similares a los obtenidos con un grupo reducido de la muestra de tipificación británica.

Con respecto a los coeficientes de estabilidad, se recogen en el Manual diversos estudios realizados tanto con la primera versión como con la actual. Respecto a la fiabilidad test-retest:

- Croce et al. (2001), en una muestra de 106 niños de 5 a 12 años de edad, se examinaron dos veces con un intervalo de una semana entre la primera y la segunda evaluación. El promedio de CCI entre los 4 rangos de edad fue de .95 (.92 - .98).
- Chow y Henderson (2003), con una muestra de 138 niños de 4 a 6 años de edad, a los que se les evaluó en dos ocasiones con una diferencia de 2-3 semanas. El promedio de CCI de los 8 ítems del Test fue de .77 (.64 - .86).
- Van Waelvelde et al. (2007), con una muestra de 37 niños de 4-5 años de edad, seleccionados por sus maestros por mostrar las “peores” habilidades motrices de la clase. Los niños fueron evaluados en tres ocasiones diferentes dejando 21 días entre una evaluación y la siguiente. En este caso se obtuvo un CCI de .88 para la puntuación Total del Test.
- Visser et al. (2004), con una muestra de 55 niños de 3 años, que fueron evaluados en dos ocasiones diferentes con un intervalo de 11 días entre la primera y la segunda evaluación. Los coeficientes de correlación de Pearson para las pruebas de enhebrar cuentas, dibujar el trazado y equilibrio sobre una pierna alcanzaron el criterio de .70; sin embargo, para las pruebas de introducir monedas (.65) y atrapar el saquito (.68) los coeficientes fueron marginales y para el ítem de lanzar el saquito a una diana el coeficiente fue de .49.
- Chow, Chan, Chain y Lau (2002), con una muestra de 31 adolescentes, obtuvieron un promedio de CCI .77 (.62 - .86) para seis de las pruebas que se han mantenido en la versión actual del MABC-2.
- Faber et al. (2004), seleccionaron a 30 jóvenes de edades comprendidas entre los 18 y los 28 años de una muestra general de 64. 14 de ellos fueron examinados en dos momentos diferentes por el mismo evaluador y 16 por un evaluador diferente. Los resultados mostraron un CCI para la puntuación Total de .79 (intraexaminador) y .79 (interexaminador).
- Estudio de tipificación de Gran Bretaña: se llevó a cabo un estudio en paralelo con 60 niños de 3 a 16 años (20 niños en cada rango de edad), que fueron evaluados en dos ocasiones con un intervalo de 1-2 semanas entre la primera y la segunda evaluación. Los coeficientes de correlación de Pearson para cada una de las dimensiones y para la puntuación Total fueron superiores a .70.

Por último, en relación a la fiabilidad interobservadores, el Manual recoge los siguientes estudios:

- Chow, Chan, Chain y Lau (2002), con una muestra de 31 adolescentes, obtuvieron un promedio de CCI excelente .97 (.92 - 1.00) para siete de las pruebas que se han mantenido en la versión actual del MABC-2.
- Chow y Henderson (2003), en este estudio los CCI para la fiabilidad interobservador fue superior a .95 en siete de los 8 ítems.

- Bom Fiers et al. (2007), en este estudio una muestra de 9 niños de edades comprendidas entre los 4 y los 12 años fueron evaluados por 131 fisioterapeutas (cada fisioterapeuta evaluó a 2 niños). Los coeficientes Kappa oscilaron entre .94 y 1.00 para los nueve niños.

En suma, la información presentada con relación a la fiabilidad del test es adecuada. Futuros estudios podrían analizar la fiabilidad de las puntuaciones en nuevas muestras en población española e incorporar análisis estadísticos basados en la Teoría de Respuesta al Ítem y la función de información.

2.13. Baremación e interpretación de las puntuaciones

2.13.1. Interpretación normativa de las puntuaciones

Contenido	Valoración	Puntuación
2.13.1.1. Calidad de las normas	Varios baremos dirigidos a diversos estratos poblacionales	4
2.13.1.2. Tamaño de las muestras	Moderado	3
2.13.1.3. Aplicación de estrategia de tipificación continua	No	
2.12.1.4. Procedimiento de selección de las muestras	Aleatorio, con muestras representativas de la población objetivo	
2.12.1.5. Actualización de baremos	Excelente	5

2.13.2. Interpretación referida a criterio

Contenido	Valoración	Puntuación
2.13.2.1. Adecuación del establecimiento de los puntos de corte	-	-
2.13.2.2. Procedimiento empleado para fijar los puntos de corte		-
2.13.2.3. Procedimiento de obtención del acuerdo inter-jueces		-
2.13.2.4. Valor del coeficiente de acuerdo inter-jueces	-	-

Comentarios generales sobre baremación e interpretación de las puntuaciones

El Test de la Batería MABC-2 utiliza diversos baremos en función de la edad de los participantes. En concreto, se proporcionan tablas con las puntuaciones escalares equivalentes a las puntuaciones directas de cada ítem del test para cada edad, excepto para el rango de 4:00 a 4:11 años que se divide en dos períodos (4:00 a 4:5 años; 4:6 a 4:11 años). Además, el manual también recoge dos tablas más: una donde se recogen las puntuaciones escalares y percentiles equivalentes a las puntuaciones de las tres dimensiones, y otra con las puntuaciones escalares y percentiles equivalentes a la puntuación Total del Test, independientemente del grupo de edad.

El procedimiento seguido para la selección de las muestras de tipificación fue aleatorio, con muestras representativas. Se tomaron los datos del Censo de población de 2001 (INE, 2004) para calcular la proporción de los grupos, en función de los rangos de edad utilizados en la Batería MABC-2. Se formó un panel de expertos de 21 especialistas distribuidos en las cuatro zonas geográficas consideradas (Norte, Este, Centro, Sur). Los especialistas seleccionados disponían de una plantilla de los niños y adolescentes que debían seleccionar en base a los criterios de estratificación establecidos: edad, género, nivel educativo de los padres y zona geográfica.