REVISTA CHILENA DE PEDIATRÍA
www.revistachilenadepediatria.cl
www.scielo.cl

Rev Chil Pediatr. 2019;90(5):522-532
DOI: 10.32641/rchped.v90i5.881

ORIGINAL ARTICLE

Age band 1 of the movement assessment battery for children –2. Reliability of the spanish version

Batería para la evaluación del movimiento en niños –2– banda 1. Confiabilidad de la versión en español

Gloria Isabel Niño-Cruz[a], Diana Marina Camargo-Lemos[a], Laura Isabel Velásquez-Escobar[a], Jeimy Katherine Rodríguez-Ortiz[a], María Solange Patiño-Segura[a]

[a]Universidad Industrial de Santander. School of Physical Therapy. Bucaramanga, Colombia

Received: 17-8-2018; Approved: 13-5-2019

Abstract

Introduction: The development of motor skills influences the capacity of the child to interact with the environment. Thus, several instruments have been created for their assessment. **Objective:** To evaluate the internal consistency, reproducibility, and agreement level of age band 1 of the Movement Assessment Battery for Children – 2 in a preschool group. **Patients and Method**: Assessment study of diagnostic tests with 29 preschoolers, selected by convenience, enrollments in an educational institution in Bucaramanga city, Colombia. For the inter-evaluators reproducibility assessment, three evaluators watched each video independently. In the intra-evaluator reproducibility assessment, each evaluator watched the same video on two different occasions. The internal consistency, the intra- and inter-evaluator reproducibility, and the agreement level were determined using Cronbach's alpha coefficient, the Intraclass Correlation Coefficient (ICC), and the Bland and Altman limits of agreement method, respectively. **Results:** Internal consistency of the total test for each of the three evaluators was higher than 0.60. Very good reliability was found for all items, domains, and total score of age band 1 of MABC-2 (ICC ≥ 0.85), as well as good limits of agreement. **Conclusions:** age band 1 of MABC-2 Spanish version is an instrument with adequate reliability psychometric properties that can be used for the motor skills development evaluation in preschoolers.

Keywords:
Motor Skills;
Child Development;
Reproducibility of Results;
Children;
Preschoolers

Correspondence:
Gloria Isabel Niño-Cruz
ginc_s@hotmail.com

How to cite this article: Rev Chil Pediatr. 2019;90(5):522-532. DOI: 10.32641/rchped.v90i5.881

Introduction

Preschool phase is a critical period in the comprehensive development of the child since its important changes inherent to growth, the formation of personality bases and habits essential to life, as well as the fundamental motor skills development (FMS), among other aspects[1,2].

Problems related to the performance of FMS in childhood negatively influence the child's ability to explore his/her own body and the participation in family and educational contexts, affecting his/her ability to interact with other children and caregivers in the short and long term[3]. Therefore, the early detection of possible motor development disorders in early childhood is key, highlighting the need for easily applied tests and with adequate psychometric properties for the screening of FMS[4].

There is a wide variety of instruments for FMS measuring, some of them evaluate the movement pattern (muscle strength, range of motion, and execution speed, among others). Other ones are focused on quantitative indicators of motor skills (reaction time, correct or incorrect attempts), and also on the result of movement evaluation, focused on postural alignment and substitution mechanisms[5-8].

The Movement Assessment Battery for Children, second edition (MABC-2), by Henderson, Sudgen and Bernett[9], 2007, is an instrument containing an FMS performance test organized in three bands corresponding to different age ranges (3 to 6, 7 to 10, and 11 to 16 years). Its objective is to screen for the identification of children with suspected developmental coordination disorder. Due to its easy application, the European normative data availability, the fine motor skills movement quality analysis in children under five years of age, among other aspects, the MABC-2 has been an instrument widely used to assess the motor skill performance, both in the clinical and in the investigative context, including its application in healthy children with movement impairments[10-13].

Thus, several studies have been published in the last few years aimed at the psychometric properties assessment of the MABC-2. Among the results, it is worth to mention that the reproducibility intra- and inter-evaluators (assessed with the Intraclass Correlation Coefficient - ICC) of the MABC-2 present coefficients higher than 0.75[10-13]. However, the potential use of the MABC-2 as a screening instrument in children with possible movement impairments in South America has been limited since the reproducibility findings of the MABC-2 have been reported in different sociocultural contexts such as Greece[14], China[15], Brazil[16], and the United States[10], among others.

Therefore, it is necessary to provide scientific evidence that verifies the reliability of MABC-2 in Spanish. Considering this issue, the objective of this study was to evaluate the internal consistency, the reproducibility (intra- and inter-evaluators), and the level of agreement of the MABC-2 first age band in a group of preschoolers enrolled in an educational institution from Bucaramanga (Colombia).

Patients and Method

An evaluation study of MABC-2 band 1 diagnostic tests was performed. The study population was comprised of 29 preschoolers selected by convenience, clinically healthy, of both sexes, and aged between three to six years, enrolled in an educational institution of Bucaramanga, Colombia in 2014.

Those children whose caregivers accepted their participation in the study were included. Children with skeletal muscle, cognitive, neurological, genetic, visual, auditory and/or metabolic dysfunctions affecting human body movement and requiring specialized medical care were excluded.

The MABC-2 was bought, translated, re-translated into English and adapted by a physiotherapist (MSPS) with 32 years of experience and training in human kinetics. Three evaluators participated in the study, two of them were students of 7th level of physiotherapy trained during two academic semesters by a physiotherapist master student, trained during six months by the previously mentioned expert physiotherapist. This training included the theoretic and conceptual bases related to motor development, skills, and control during the first years of life. In addition, the evaluators received practical training which included the observation of six videos of children aged three to six years in which MABC-2 was applied and then directly assessed five more preschoolers.

Study variables were the age in years and months and sex (male/female). Anthropometric variables were also registered such as weight, height, and Body Mass Index (IBM), classified according to the Center for Disease Control and Prevention[17] recommendations.

Eight motor skills (items) constitute the age band 1 of the MABC-2 that were evaluated grouped into three attributes: manual dexterity - MD (inserting coins, threading beads, drawing a trail), aiming and catching - AC (throwing a beanbag to a target, catching a beanbag), and balance - B (walk heels raised along a line, one-leg balance, jumping five mats).

Each of the skills was scored according to the criteria described in the instrument's manual, obtaining a raw score of each item (e.g. Time inserting coins). Then, this score was standardized and, finally, the sum

of those scores was turned into percentiles for each attribute (MD, AC, and B) and for the total test score[9].

Additionally, the participants that present movement impairments were identified, considering the cut-off points for the total test score. This classification system called 'traffic light' establishes three zones: red (significant movement difficulty), yellow (mild movement difficulty), and green (no detection of movement difficulty)[9].

Procedure

Before applying the test, a room was adapted inside the educational institution to evaluate the participants where the location and installation height of two video cameras were standardized, as well as the organization of elements and materials for the evaluation. The evaluation was carried out individually by the same evaluator in all cases (a student of 7th level of physiotherapy) who initially determine the age in years and months according to the record of the educational institution and also established the dominant hand of the participants.

Then, the evaluator demonstrated each of the activities and allowed the participant to practice them, and then record through two video cameras, previously mentioned, the execution of the FMS of the study participants. The order of the test follows the protocol described by the authors of MABC-2 and the average administration time was 35 minutes per participant.

To evaluate the inter-evaluators reproducibility, the three evaluators watch each video independently and fulfill the test evaluation format. In the intra-evaluator reproducibility evaluation, each evaluator watches the same video in two different moments with an interval of at least 20 days between each evaluation to reduce recall bias.

This work followed the Declaration of Helsinki international recommendations which establishes the ethical principles for medical research involving human beings[18] and the national current regulations established by the Resolution 008430, 1993 of the Colombian Ministry of Health[19]. In addition, it is approved by the Scientific Research Ethics Committee of the Industrial University of Santander, and all parents of the evaluated children signed informed consent.

Analysis

First, the sociodemographic and anthropometric characteristics of the sample were described, and subsequently, the distribution of the data was analyzed. With this objective, each of the motor skills scores obtained in the first evaluation made by each evaluator was analyzed graphically and through statistical characteristics such as bias and kurtosis. Thus, it was considered as a criterion to identify if the variable had normal distribution when the bias was close to zero and when the kurtosis was close to three[20,21].

Additionally, to describe the FMS distribution, the average and standard deviation were determined in each measure made by the evaluators. Once the sample had been characterized, the Cronbach's α was calculated, a test that estimates the weighted average of the correlations between the items of an instrument in order to establish internal consistency. The amplitude of its value ranges from zero to one, indicating a poor relationship between the items that make up the scale when it is lower than 0.5, acceptable when it is between 0.5 and 0.7, and a good one when it is higher than 0.75[22,23].

In order to determine the reproducibility of the movement impairments classification system, 'traffic lights', the weighted Kappa test was used which compares the proportion of the observed concordance with the expected one between two different measures or observations. Kappa values range from zero to one, indicating poor reproducibility when they are lower than 0.4, good when they are between 0.4 and 0.75, and excellent when the estimation is higher than 0.75[22,23].

The intra- and inter-evaluators reproducibility evaluation was carried out through the $ICC_{2,1}$ which quantifies how similar two variables measurements are in ordinal or ratio scale applied to the same people[21,23]. When the value is close to 1.0, the reproducibility is almost perfect, therefore, coefficients higher than 0.75 indicate a good reproducibility between measurements[24].

Likewise, the intra- and inter-evaluators level of agreement was established by applying the Bland-Altman method of limits of agreement. Graphic analysis that shows the difference between two measurements and their average. Its interpretation is based on the differences average and their limits of agreement, thus, as the average closes to zero and the amplitude of its limits is low, a good agreement is established between the two measurements[22,25].

Each of the raw, standardized, and percentile scores were analyzed for both items and attributes, as well as the total test score. The database was typed in Excel twice and then its definitive validation and analysis were carried out in the software STATA 12.1[26].

Results

29 preschoolers participated aged between three and six years, 12 (41.4%) were male, 6 (20.7%) present overweight, and 5 (17.2%) were obese. 55.2% of the caregivers reported incomes between one and two current legal minimum wages (Table 1).

Table 1. Socio-economic and demographic characteristics of study population

Characteristics		Boys (n: 12)	Girls (n: 17)	Total (n: 29)
Age (years)	mean ± SD	4.5 ± 0.8	4.6 ± 0.9	4.5 ± 0.9
	(Range)	(3 - 5.7)	(3 - 6.0)	(3 - 6.0)
Height (cm)	mean ± SD	105.3 ± 7.6	106.3 ± 8.7	106.0 ± 8.2
	(Range)	(93.1 - 116.7)	(93.0 - 126)	(93.0 - 126)
Weight (kg)	mean ± SD	18.5 ± 3.9	18.5 ± 4.1	18.5 ± 4.0
	(Range)	(13.1 - 25)	(13.4 - 26.3)	(13.1 - 26.3)
BMI	N (%)			
Underweight		1 (8.3)	0	1 (3.5)
Normal weight		6 (50.0)	11 (64.8)	17 (58.6)
Overweight		3 (25.0)	3 (17.6)	6 (20.7)
Obese		2 (16.7)	3 (17.6)	5 (17.2)
Family income (minimum wage) N (%)				
	< 1	4 (33.3)	0	4 (1.4)
	1-2	6 (50.0)	10 (58.8)	16 (55.2)
	2-3	2 (16.7)	7 (41.2)	9 (43.4)

SD: Standard deviation

Regarding the FMS distribution assessed with the MABC-2, considering the graphic analysis as well as the bias and kurtosis characteristics, it was possible to observe that only the raw score of each attribute (the raw score sum of the items comprising the attribute) is normally distributed (missing data).

The internal consistency evaluation showed a Cronbach's α for the test total higher than 0.60 for each of the evaluators (evaluator 1: 0.65; evaluator 2: 0.63; evaluator 3: 0.66) showing an acceptable internal consistency of the instrument. Regarding the reproducibility of the movement impairments classification system 'traffic lights', excellent results were obtained with Kappas equal to 1.0 for all evaluators.

Tables 2 to 4 show the intra- and inter-evaluator reproducibility findings where the good ICC results between 0.85 and 0.99 for all items and attributes stand out. Total scores also present excellent ICC, showing a high reproducibility level.

Regarding the intra-evaluator level of agreement, there were difference averages close to zero with narrow limits (Table 5), and an even distribution of points around the mean without more than 5% of the data outside the 95% limits, both for items and for total attributes and scores for the whole instrument. Figure 1 shows a graphic analysis example using the Bland-Altman method.

Discussion

The objective of this study was to assess the internal consistency, the level of agreement, and the intra- and inter-evaluators reproducibility of the band 1 of MABC-2 in a sample of preschoolers enrolled in an educational institution from Bucaramanga, Colombia.

The findings of this study showed a high frequency of obesity (17%) in the assessed sample. Although previous scientific literature confirms the negative impact of overnutrition on the motor skill performance in preschoolers[27], the influence of this characteristic in the band 1 reliability of MABC-2 was controlled since 58.6% of the preschoolers had a healthy weight, showing that in relation to BMI, the sample was not homogeneous, thus validating the found results[28].

Regarding the distribution results of the FMS assessed with the MABC-2, two characteristics should be considered. The first one, related to the sample size (29 participants), and the second one, related to the sampling method used in this study (convenience sampling). Due to these two aspects, the variability of the sample individuals was lower, therefore, the data tend not to be distributed normally, explaining the findings[20,21].

Besides that, the FMS scores obtained in this study clearly show a better performance in the manual dexterity (65.8 ± 29) and balance (52.7 ± 25.9) attributes

Table 2. Inter-evaluator reproducibility. Intraclass Correlation Coefficient (ICC) and 95% confidence interval (95% CI)

Attributes / Item	ICC (95% IC)		
	Evaluator 1	Evaluator 2	Evaluator 3
Manual Dexterity			
Posting coins preferred hand[a]	0.98 (0.97 - 0.99)	0.99 (0.99 - 1.0)	0.98 (0.97 - 0.99)
Posting coins non-preferred hand[a]	0.99 (0.99 - 1.0)	0.99 (0.99 - 0.99)	0.99 (0.99 - 1.0)
Threading beads[a]	0.99 (0.99 - 0.99)	0.99 (0.99 - 0.99)	0.99 (0.99 - 0.99)
Drawing trail[b]	0.99 (0.98 - 0.99)	0.99 (0.98 - 0.99)	0.99 (0.98 - 0.99)
Component score	0.97 (0.95 - 0.99)	0.99 (0.99 - 1.0)	0.98 (0.97 - 0.99)
Standard score	0.98 (0.96 - 0.99)	0.98 (0.97 - 0.99)	0.98 (0.96 - 0.99)
Percentile	0.90 (0.80 - 0.95)	0.92 (0.88 - 0.94)	0.90 (0.83 - 0.95)
Aiming and catching			
Catching beanbag[b]	0.99 (0.98 - 0.99)	0.99 (0.98 - 0.99)	0.99 (0.98 - 0.99)
Throwing beanbag onto mat[b]	0.99 (0.97 - 0.99)	0.91 (0.74 - 0.97)	0.91 (0.74 - 0.96)
Component score	0.98 (0.95 - 0.99)	0.98 (0.97 - 0.99)	0.98 (0.97 - 0.99)
Standard score	0.96 (0.91 - 0.98)	0.97 (0.97 - 0.98)	0.96 (0.92 - 0.97)
Percentile	0.96 (0.91 - 0.98)	0.95 (0.93 - 0.99)	0.97 (0.92 - 0.97)
Balance			
One-leg balance best leg[a]			
One-leg balance other leg[a]	0.99 (0.99 - 1.0)	0.99 (0.99 - 1.0)	0.99 (0.99 - 1.0)
Walking heels raised[b]	0.99 (0.99 - 1.0)	0.99 (0.99 - 1.0)	0.99 (0.99 - 1.0)
Jumping on mats[b]	0.99 (0.99 - 0.99)	0.99 (0.98 - 0.99)	0.99 (0.99 - 0.99)
Component score	0.85 (0.67 - 0.92)	0.87 (0.71 - 0.94)	0.85 (0.70 - 0.90)
Standard score	0.97 (0.93 - 0.98)	0.97 (0.93 - 0.98)	0.97 (0.93 - 0.98)
Percentile	0.99 (0.97 - 0.99)	0.98 (0.96 - 0.99)	0.99 (0.98 - 0.99)
Total MABC-2			
Total score	0.98 (0.95 - 0.99)	0.98 (0.96 - 0.99)	0.98 (0.96 - 0.99)
Standard score			
Percentile	0.98 (0.96 - 0.99)	0.99 (0.98 - 0.99)	0.98 (0.97 - 0.99)

[a]Seconds; [b]Number.

compare with the studies of Logan[10] (2011) and Ellinoudes[14] (2011) carried out in the United States and Greece respectively, who found percentiles between 18.4 ± 16.7 and 25.6 ± 61.5 in manual dexterity and 38.1 ± 32.7 and 32.7 ± 6.5 in balance. These differences may be explained by the variability in the socio-cultural and environmental contexts of the preschoolers, associated with games that allow better self-organization of the movement, enhancing by the experience and the preference in the execution of these skills that contribute to a better performance quality[29].

The internal consistency results for the test total showed acceptable coefficients for each of the evaluators (α between 0.63 and 0.66). These findings may be explained in part by the small number of items that has the performance test of the MABC-2 and by the internal correlation between them. The study carried out by Civetta[30] in 2008 showed an improvement in the internal consistency of the MABC-1 by removing the item with the lowest correlation (drawing a trail) thus increasing the α up to 0.70 [30].

Other publications that exclusively evaluate this feature in the band 1 of MABC-2, such as Ellinoudes[14] in 2011 and Hua[15] in 2013, reported coefficients

Table 3. Scores and intra-evaluator reproducibility in the first measure. Intraclass Correlation Coefficient (ICC) and 95% confidence interval (95% CI)

Attributes / Item	Measure 1			ICC (95% CI)
	Evaluator 1 Mean ± SD	Evaluator 2 Mean ± SD	Evaluator 3 Mean ± SD	
Manual Dexterity				
Posting coins preferred hand[a]	12.9 ± 4.0	13.3 ± 4.0	12.8 ± 4.1	0.98 (0.95 - 0.99)
Posting coins non-preferred hand[a]	15.9 ± 5.3	15.9 ± 5.4	15.6 ± 5.4	0.99 (0.98 - 0.99)
Threading beads[a]	44 ± 34.8	42.8 ± 33.3	41.4 ± 31.7	0.99 (0.89 - 0.97)
Drawing trail[b]	2.9 ± 3.5	2.4 ± 3.0	3.0 ± 2.8	0.94 (0.88 - 0.97)
Component score	32.1 ± 6.8	31.4 ± 6.4	31.6 ± 6.0	0.93 (0.89 - 0.96)
Standard score	11.7 ± 3.2	11.3 ± 3.1	11.6 ± 2.9	0.88 (0.96 - 0.93)
Percentile	65.8 ± 29.6	64.8 ± 30.0	65.1 ± 28.9	0.99 (0.98 - 0.99)
Aiming and catching				
Catching beanbag[b]	5.7 ± 2.8	5.6 ± 2.7	5.9 ± 2.9	0.97 (0.95 - 0.99)
Throwing beanbag onto mat[b]	2.7 ± 1.8	2.4 ± 1.7	2.2 ± 1.6	0.91 (0.83 - 0.95)
Component score	16.4 ± 4.0	16.1 ± 3.7	15.6 ± 3.9	0.98 (0.95 - 0.99)
Standard score	8.5 ± 2.9	8.0 ± 2.8	7.9 ± 3.0	0.93 (0.88 - 0.96)
Percentile	34.5 ± 26.7	33.7 ± 25.8	34.0 ± 27.0	0.99 (0.97 - 0.99)
Balance				
One-leg balance best leg[a]	12.0 ± 10.0	12.3 ± 10.1	11.8 ± 10.2	0.99 (0.99 - 1.0)
One-leg balance other leg[a]	8.5 ± 7.7	8.9 ± 7.8	8.4 ± 8.0	0.99 (0.99 - 1.0)
Walking heels raised[b]	12.5 ± 3.5	12.8 ± 3.2	12.7 ± 3.3	0.90 (0.82 - 0.94)
Jumping on mats[b]	4.7 ± 0.6	4.8 ± 0.5	4.9 ± 0.4	0.90 (0.84 - 0.96)
Component score	31.1 ± 5.1	29.8 ± 4.8	31.0 ± 4.4	0.90 (0.83 - 0.95)
Standard score	10.4 ± 2.7	10.1 ± 2.2	10.2 ± 2.4	0.96 (0.92 - 0.98)
Percentile	52.7± 25.9	52.3 ± 25.9	52.3 ± 25.8	0.98 (0.98 - 0.99)
Total MABC-2				
Total score	79.6 ± 11.2	79.0 ± 11	80.0 ± 11.0	0.98 (0.97 - 0.99)
Standard score	10.3 ± 2.5	9.8 ± 2.3	10.2 ± 2.3	0.96 (0.93 - 0.98)
Percentile	53.9 ± 24.5	53.6 ± 24.1	53.6 ± 24.7	0.99 (0.98 - 0.99)

[a]Seconds; [b]Number. SD: Standard deviation.

between 0.44 to 0.70, attributing these results to the same reason (small number of items) and empirically demonstrating the internal correlation by removing the FMS with the lowest correlation, which in their cases was drawing a trail, therefore increasing the Cronbach's α [14,15].

Additionally, the results of the study carried out by McCrae[31] et al. in 2010 showed that factors specific to the individual (age), as well as the type of task performed influence the Cronbach's α. According to McCrae[31], internal consistency may be different in each population group and is vulnerable to random errors derived from the characteristics of each task. This may be explained by the Cronbach's α magnitude in studies such as those of Wuang[32] et al. (2012), and Valentini[16] et al. (2014) who reported coefficients higher than 0.70 when evaluating the internal consistency of bands 2 and 3 of the MABC-2, which consider other ages and other types of skills.

Regarding the intra- and inter-evaluators reproducibility, when comparing our results with those reported in the literature, the ICC for the final score were similar (ICC 0.85-0.97) showing a good reproducibility when applying two measures with this instrument[12,15].

Table 4. Scores and intra-evaluator reproducibility in the second measure. Intraclass Correlation Coefficient (ICC) and 95% confidence interval (95% CI)

Attributes / Item	Measure 2			ICC (95% CI)
	Evaluator 1 Mean ± SD	Evaluator 2 Mean ± SD	Evaluator 3 Mean ± SD	
Manual Dexterity				
Posting coins preferred hand[a]	12.8 ± 4.0	13.4 ± 4.0	12.8 ± 4.1	0.98 (0.97 - 0.99)
Posting coins non-preferred hand[a]	15.8 ± 5.3	16.0 ± 5.2	15.6 ± 5.3	0.99 (0.99 - 0.99)
Threading beads[a]	44 ± 34.7	42.9 ± 33.3	41.3 ± 31.6	0.98 (0.96 - 0.99)
Drawing trail[b]	3.1 ± 3.5	2.3 ± 2.9	2.9 ± 2.6	0.96 (0.93 - 0.98)
Component score	32.1 ± 6.1	31.5 ± 6.3	31.9 ± 6.0	0.96 (0.93 - 0.98)
Standard score	11.7 ± 3.0	11.4 ± 3.1	11.6 ± 3.0	0.93 (0.87 - 0.96)
Percentile	62.3 ± 29.1	64.7 ± 30.0	65.2 ± 28.9	0.96 (0.93 - 0.97)
Aiming and catching				
Catching beanbag[b]	5.7 ± 2.8	5.6 ± 2.5	5.8 ± 2.9	0.98 (0.97 - 0.99)
Throwing beanbag onto mat[b]	2.8 ± 1.8	2.4 ± 1.8	2.2 ± 1.8	0.91 (0.77 - 0.96)
Component score	16.6 ± 4.1	16.0 ± 3.5	15.3 ± 4.1	0.95 (0.90 - 0.97)
Standard score	8.4 ± 2.7	8.0 ± 2.6	7.8 ± 2.9	0.93 (0.87 - 0.96)
Percentile	34.3 ± 25.1	34.1 ± 25.8	33.8 ± 27.4	0.98 (0.97 - 0.99)
Balance				
One-leg balance best leg[a]	12.1 ± 9.9	12.1 ± 10.1	11.7 ± 10.4	0.99 (0.99 - 1.0)
One-leg balance other leg[a]	8.6 ± 7.7	8.9 ± 7.9	8.6 ± 8.2	0.99 (0.99 - 1.0)
Walking heels raised[b]	12.5 ± 3.4	13.0 ± 3.1	12.7 ± 3.1	0.95 (0.91 - 0.97)
Jumping on mats[b]	4.8 ± 0.4	4.8 ± 0.4	4.8 ± 0.4	0.88 (0.84 - 0.98)
Component score	31.2 ± 5.0	30.0 ± 4.8	31.0 ± 4.1	0.90 (0.81 - 0.95)
Standard score	10.3 ± 2.8	10.2 ± 2.2	10.4 ± 2.3	0.94(0.90 - 0.97)
Percentile	51.5 ± 26.1	52.1 ± 25.9	52.2 ± 26.0	0.99 (0.98 - 0.99)
Total MABC-2				
Total score	80.0 ± 11.2	79.0 ± 11	80.0 ± 10.1	0.99 (0.98 - 0.99)
Standard score	10.3 ± 2.6	9.8 ± 2.3	10.2 ± 2.4	0.96 (0.93 - 0.98)
Percentile	53.5 ± 26.2	54.0 ± 24.0	53.5 ± 25.0	0.99 (0.98 - 0.99)

[a]Seconds; [b]Number.

It is important to note that in most of the reviewed studies, the attribute with the highest reproducibility was the manual dexterity one (ICC 0.66-0.90), contrary to our results, which is attributable to the applied evaluation method since video evaluation allows the evaluator to observe the FMS performed by the child under exactly the same conditions, increasing the likelihood of designating a closer score. In contrast, the evaluation through direct observation implies an increased modification probability in all sources of variation related to the evaluator, the test, the evaluated subject, and the context[33].

We could not compare the reproducibility of the movement impairments classification system, 'traffic lights' since there were no similar results in the reviewed literature. However, it is possible that the excellent reproducibility in the Kappas is derived from the intra-evaluator reproducibility obtained with the raw scores, which in turn reflected the sum of the items that determined the total test score, thus generating the movement impairments classification established by the evaluators.

A contribution of this work is the analysis according to intra- and inter-evaluator, which was not pre-

Table 5. Agreement limits of Bland and Altman intra-evaluator. Averages of the differences between the first and Second measurement and 95% agreement limits

Attributes / Item	Bland and Altman limits of agreement		
	Evaluator 1	Evaluator 2	Evaluator 3
Manual Dexterity			
Posting coins preferred hand[a]	0.03 (-0.8 ; 0.9)	-0.07 (-1 ; 0.8)	0.03 (-0.4 ; 0.4)
Posting coins non-preferred hand[a]	0.1 (-0.5 ; 0.7)	-0.10 (-0.9 ; 0.7)	0.03 (-0.6 ; 0.7)
Threading beads[a]	0.1 (-0.9 ; 1.1)	-0.07 (-1 ; 0.8)	0.07 (-0.8 ; 1)
Drawing trail[b]	-0.1 (-0.8 ; 0.6)	0.07 (-0.8 ; 1)	0.07 (-0.1 ; 1.2)
Component score	-0.07 (-2.8 ; 2.7)	-0.03 (-1 ; 1)	-0.2 (-2.1 ; 1.7)
Standard score	0.03 (-1.2 ; 1.3)	-0.07 (-1.5 ; 1.3)	-0.07 (-1.5 ; 1.4)
Percentile	3.4 (-1.0 ; 2.4)	3.0 (-2.8 ; 1.0)	3.2 (-1.2 ; 2.0)
Aiming and catching			
Catching beanbag[b]	-0.07 (-0.8 ; 0.7)	0.03 (-0.8 ; 0.9)	0.07 (-0.8 ; 1.0)
Throwing beanbag onto mat[b]	-0.1 (-0.7 ; 0.5)	0.07 (-0.7 ; 0.4)	0.5 (-1.1 ; 2.1)
Component score	-0.2 (-2 ; 1.5)	0.1 (-1. ; 1.2)	0.2 (-1.7 ; 2.2)
Standard score	0.07 (-1.6; 1.7)	0.06 (-1.6 ; 1.3)	0.1 (-1.1; 1.4)
Percentile	-0.5 (-5.3; 3.0)	-0.4 (-4.9 ; 4.0)	-0.8 (-7.0 ; 8.1)
Balance			
One-leg balance best leg[a]	-0.07 (-0.1 ; 0.8)	0.1 (-0.7 ; 0.9)	0.03 (-1.2 ; 1.3)
One-leg balance other leg[a]	-0.07 (-0.6 ; 0.4)	-0.03 (-0.9 ; 0.8)	-0.1 (-1.2 ; 1.0)
Walking heels raised[b]	-0.03 (-0.7 ; 0.6)	-0.2 (-1.1 ; 0.7)	-0.03 (-0.3 ; 0.4)
Jumping on mats[b]	-0.07 (-0.8 ; 0.7)	-0.07 (-0.9 ; 0.8)	-0.03 (-0.8 ; 0.7)
Component score	-0.03 (-3.0 ; 2.6)	-0.03 (-4.0 ; 3.6)	-0.2 (-1.6 ; 1.2)
Standard score	0.07 (-0.8 ; 1.0)	-0.1 (-1 ; 0.9)	-0.2 (-1.3 ; 0.9)
Percentile	1. 2 (-0.4 ; 1.6)	1.0 (-1.0 ; 1.3)	1.1 (-0.6 ; 1.8)
Total MABC-2			
Total score	-0.4 (-4.4 ; 3.7)	0.03 (-1.1 ; 1.1)	0.1 (-1.1 ; 1.3)
Standard score	0 (-1.2 ; 1.2)	0.03 (-0.7 ; 0.8)	0 (-1.0 ; 1.0)
Percentile	-0.4 (-5.1 ; 4.0)	-0.4 (-4.1 ; 3.5)	0.3 (-4.3 ; 6.0)

[a]Seconds; [b]Number. SD: Standard deviation.

viously reported in the literature. The results showed narrow limits and an average of differences close to zero for the items and attributes of the MABC-2 included in the raw scores of the instrument. These findings are useful since, with this methodology, it is possible to determine if the changes generated from an intervention go beyond the measurement variability[22].

This aspect is very relevant since between the age of three and six, the changes detected in the motor skill performance in band 1 of the MABC-2 are minor compared with those changes during the first year of life and also, as raw scores and their respective percentiles are standardized, there is a tendency to homogenize the results in the motor skill performance of the child, therefore, the raw score would be more useful for detecting the change generated from an intervention[5-9].

It is worth to mention that the Test of Gross Motor Development 2 (TGMD-2) is the most widely reported instrument in the literature for documenting intervention experiences aimed at promoting motor development in the preschool stage, which evaluates 12 gross motor skills (six of locomotor and six of object control skills), therefore the evaluation of experiences aimed at improving the motor performance of fine motor skills

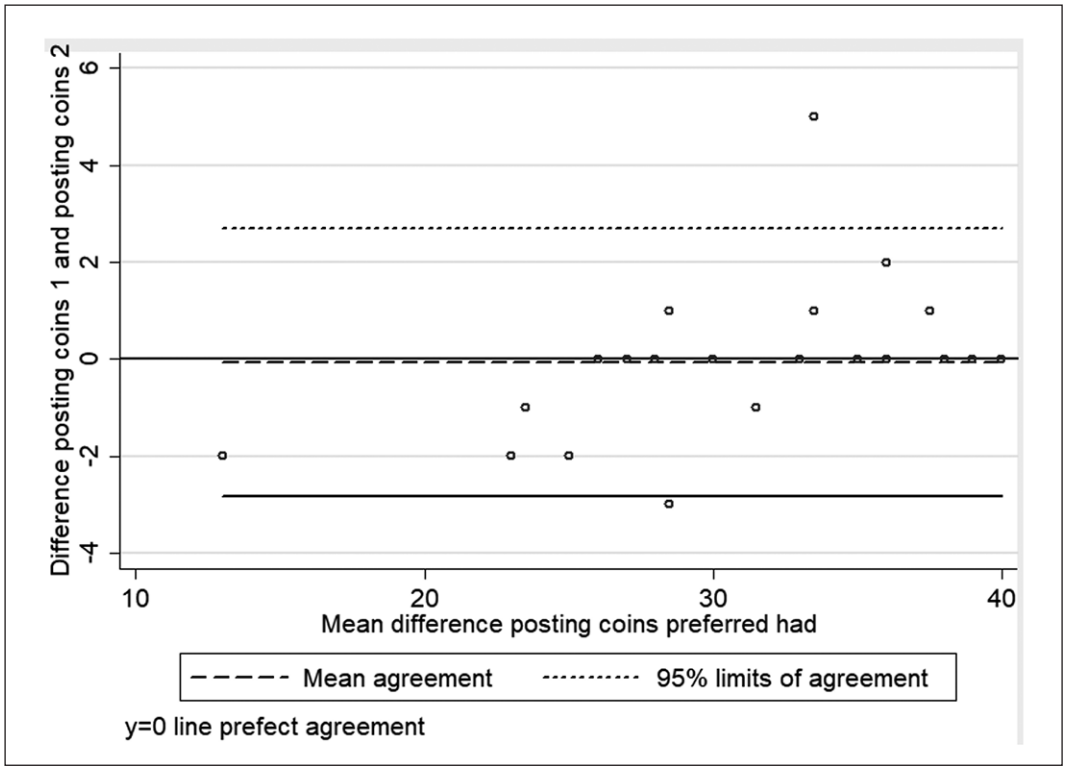

Figure 1. Example Bland-Altman plot posting coins preferred hand.

is limited[34]. This limitation would be overcome with the application of MABC-2, and thus could be a valuable tool to evaluate FMS from the attribute of manual dexterity included among its attributes[9].

The results found in this study should be cautiously interpreted due to the following limitations. The sampling type (convenience sampling) and the small sample size tend to decrease the variability of the sample participants as well as the data distribution restricting the application of highly sensitive tests to central tendency and dispersion measurements[20,21] that would provide a better accuracy understanding of the individual MABC-2 scores, such as measurement standard error and coefficient of variance. However, it is important to emphasize that we implemented analyses with tests that allow us to evaluate reliability, using the relevant statistics in the evaluated sample[22-25].

In addition, the participants were contacted once due to the study type (cross-sectional), limiting the determination of other important properties in an instrument such as the minimum detectable difference, highly useful in the clinical context. For this reason, future studies need to include the application of MABC-2 in the same population over a period of time (longitudinal analysis)[21]. However, this study provides a starting point for research experiences interested in studying motor development since it evaluates the three reliability components of an instrument, such as the internal consistency, the reproducibility, and the level of agreement, which in the published literature have not been reported together in a study so far.

One of the strengths of this work is the participation of a researcher in human kinetics highly experienced, as well as undergraduate and graduate physical therapy students trained in the public health and child development areas, contributing to the basis of knowledge and work of professions such as kinesiology and physiotherapy, in priority areas for the new millennium in the preschool population[35,36].

Based on this evaluation, we can conclude that the band 1 of the MACB-2 in Spanish is an instrument with adequate psychometric properties of reliability. It is important to conduct research that also demonstrates the validity and change sensitivity of the instrument in a larger sample, which would provide an instrument for diagnosis and evaluation of motor development in preschoolers.

Ethical Responsibilities

Human Beings and animals protection: Disclosure the authors state that the procedures were followed according to the Declaration of Helsinki and the World Medical Association regarding human experimentation developed for the medical community.

Data confidentiality: The authors state that they have followed the protocols of their Center and Local regulations on the publication of patient data.

Rights to privacy and informed consent: The authors have obtained the informed consent of the patients and/or subjects referred to in the article. This document is in the possession of the correspondence author.

Conflicts of Interest

Authors declare no conflict of interest regarding the present study.

References

1. Largo R, Fischer J, Rousson V. Neuromotor development from kindergarten age to adolescence developmental course and variability. Swiss Med Wkly. 2003;133:193-99.
2. Hardy LL, Farrell L, Macniven R, Howlett S. Fundamental movement skills among Australian preschool children.J Sci Med Sport. 2010;13: 503-8.
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. Dua T, Tomlinson M, Tablante E, et al. Global research priorities to accelerate early child development in the sustainable development era. Lancet Glob Health. 2016;4(12):e887-9.
5. Gallahue DL, Ozmun JC. Motor development: a theoretical model. En: Understanding motor development: infants, children, adolescents, adults. Boston: McGraw-Hill Interamericana; 2006. p. 46-63.
6. Henderson A, Pehoski C. Development of hand skills. En: Hand function in the child. Waltham: Mosby; 2006. p. 143-60.
7. Hadders-Algra M. Variation and variability: keywords in human motor development. Phys Ther. 2010;90:1823-37.
8. VanSant AF. Life-Span development in functional task.Phys Ther. 1990; 70:788-98.
9. Henderson S, Sugden DA, Barnett A.Movement Assessment Battery for Children-Second Edition [Movement ABC-2]. London, UK: The Psychological Corporation;2007.
10. Logan SW, Robinson LE, Getchell N. The comparison of performances of preschool children on two motor assessments. Percept Mot Skills. 2011;6:715-23.
11. Brown T, Lalor A. The Movement Assessment Battery for Children-Second edition (MABC-2): A review and critique. Phys Occup Ther Pediatr. 2009;29(1):86-103.
12. Smits-Engelsman BC, Niemeijer AS, van Waelvelde H. Is the Movement Assessment Battery for Children 2 edition a reliable instrument to measure motor performance in 3 year old children?.Res Dev Disabil. 2011;32:1370-7.
13. Bouwien CM, Smits-Engelsman, M, Fiers S. Interrater Reliability of the movement assessment battery for children. Phys Ther. 2008; 88:286-94.
14. Ellinoudis T, Evaggelinou C, Kourtessis T, Konstantinidou Z, Venetsanou F, Kambas A. Reliability and validity of age band 1 of the Movement Assessment Battery for Children-second edition. Res Dev Disabil. 2011;32:1046-51.
15. Hua J, Gu G, Meng W, Wu Z. Age band 1 of the Movement Assessment Battery for Children-Second Edition: exploring its usefulness in mainland China. Res Dev Disabil. 2013;34(2):801-8.
16. Valentini NC, Ramalho MH, Oliveira MA. Movement Assessment Battery for Children-2: Translation, reliability, and validity for Brazilian children. Res Dev Disabil. 2014; 35:733-40.
17. Centro de Control y Prevención de Enfermedades. División de Nutrición, Actividad Física y Obesidad. Cálculo del Índice de masa corporal en niños. [acceso el 28 de marzo de 2014]. Disponible en: http://www.cd.cdc.gov/dnpabmi/Calculator.aspx
18. Asociación Médica Mundial. Declaración de Helsinki. [acceso el 28 de marzo de 2014]. Disponible en: http://www.wma.net/es/30publications/10policies/b3/17c_es.pdf
19. República de Colombia. Ministerio de Salud. Resolución No. 8430 de 1993. [acceso 28 de marzo de 2014]. Disponible en: https://www.invima.gov.co/images/pdf/medicamentos/resoluciones/etica_res_8430_1993.pdf
20. Altman, Douglas G. Practical statistics for medical research. London (UK): Chapman & Hall, 1991.
21. Kirkwood B R, Sterne JAC. Medical statistics. Oxford (UK): Blackwell Science, 2003.
22. Orozco LC. Medición en Salud. En: Diagnóstico y evaluación de resultados. Un manual crítico más allá de lo básico. Bucaramanga: Publicaciones UIS; 2010. p. 84-93.
23. Kraemer H. Kappa coefficients in epidemiology. An appraisal of a reappraisal. J Clin Epidemiol. 1988;41:959-68.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.
25. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;8:307-10.
26. StataCorp. 2011. Stata: Release 12. Statistical Software. College Station, TX: Stata Corp LP.
27. Musalek M, Kokstejn J, Papez P, et al. Impact of normal weight obesity on fundamental motor skills in pre-school children aged 3 to 6 years. Anthropol Anz. 2017;74(3):203-12.
28. Ercan I, Yazici B, Ocakoglou G, Sigirli D, Kan I. Review of reliability and factors affecting the reliability.2007. Retrieved from the InterStat website: http://interstat.statjournals.net/YEAR/2007/articles/0704008.pdf.
29. Goodway J, Robinson L, Crowe H. Gender differences in fundamental motor skills in disadvantaged preschoolers from two geographical regions. RQES 2010;8:17-24.
30. Civetta, LR, Hillier, S. The Developmental coordination disorder questionnaire and Movement Assessment Battery for Children as a diagnostic method in Australian children. Ped Phys Ther. 2008;20:39-46.
31. McCrae RR, Kurtz JE, Yamagata S, Terracciano A. Internal consistency, retest reliability, and their implications for personality scale validity. Pers Soc Psychol Rev. 2011;15:28-50.
32. Wuang Y, Su J, Su C. Reliability and responsiveness of the Movement Assessment Battery for Children-Second Edition Test in children with developmental coordination disorder. Dev Med Child Neurol. 2012;54:160-5.
33. Domholdt E. Measurement Theory. En: Rehabilitation Research. Principles

and applications. 2. Ed. Elsevier Saunders: St Louis, 2005.
34. Ulrich DA. The test of Gross motor development. 2. ed., Danver: Austins, 2000.
35. Dean E. Physical therapy in the 21st century (Part I): Toward practice informed by epidemiology and the crisis of the lifestyle condition. Physioter Theory Pract. 2009;25:330-53.
36. Presidencia de la República de Colombia. Fundamentos Técnicos de la Estrategia de Atención Integral a la Primera Infancia. [Internet] 2012 [acceso el 26 de noviembre de 2012]. Disponible:http://www.deceroasiempre.gov.co/QuienesSomos/Documents/1.Fundamentos-Tecnicos.pdf