

Infant Motor Scale of Alberta: validation for a population of Southern Brazil

Escala Motora Infantil de Alberta: validação para uma população gaúcha

Escala motora infantil de Alberta: validación para una población de Rio Grande do Sul (Brasil)

Nadia Cristina Valentini¹, Raquel Saccani²

ABSTRACT

Objective: To translate, adapt and verify the validity of the motor criteria and the construct (internal consistence, discriminatory validity, correlation with other tests and predicted validity) of the Portuguese version of the Alberta Motor Infant Scale.

Methods: Cross-cultural validation study that enrolled 21 professionals and 561 children from the South of Brazil aged zero to 18 months (291 boys). The Alberta Infant Motor Scale, the Developmental Scale of Child Behavior and a questionnaire to control variables were used. Alberta Infant Motor Scale was translated, adapted and the content validity was analyzed. Professionals were then trained and the data were collected in daycares and participants' home. Scale's objectivity, reliability, criterion and construct were analyzed.

Results: The Portuguese version of Alberta Infant Motor Scale contained clear and pertinent motor criteria; it was reliable (total score, $p=0.88$; prone, $p=0.86$; supine, $p=0.89$; sitting, $p=0.80$ and standing, $p=0.85$), it had discriminative power (typical versus atypical development; score, $p<0.001$; percentile, $p=0.04$; classification criteria, chi-square: $p=0.047$) and, also showed temporal stability ($p=0.07$; $\rho=0.85$; $p<0.001$) and predictive power ($p<0.001$). The scale was moderately correlated with the Developmental Scale of Child Behavior ($\rho=0.342$; $p=0.03$), considering that the McNemar-Bowker proof showed differences between both methods ($p=0.047$).

Conclusions: The Portuguese version of Alberta Infant Motor Scale showed validity and reliability. It is a suitable tool for the assessment of motor acquisitions of Brazilian children, being valuable for researchers and practitioners.

Key-words: validation studies; motor skills; child development.

RESUMO

Objetivo: Traduzir, adaptar e verificar a validade de critérios motor e de constructo (consistência interna, validade discriminante, correlação com outros testes e validade preditiva) da versão em Português da *Alberta Motor Infant Scale*.

Métodos: Estudo de validação transcultural. Participaram 21 profissionais e 561 crianças do Rio Grande do Sul, com idades de zero a 18 meses (291 meninos). Foram utilizados os instrumentos: *Alberta Infant Motor Scale*, Escala do Desenvolvimento do Comportamento da Criança e um questionário para controle de variáveis. O estudo compreendeu as fases de tradução e adaptação da escala; análise da validade de conteúdo; treinamento de profissionais; coleta de dados no ambiente familiar e em creches; análise da objetividade e fidedignidade, e validade de critério e construto.

Resultados: A versão portuguesa da *Alberta Infant Motor Scale* continha critérios motores claros e pertinentes; apresentou ótima confiabilidade (escore total, $p=0,88$; prono, $p=0,86$; supino, $p=0,89$; sentado, $p=0,80$ e em pé, $p=0,85$) e poder discriminativo (desenvolvimento típico versus atípico; escore,

Instituição: Escola de Educação Física da Universidade de Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brasil

¹PhD em Health and Human Performance pela Auburn University, Alabama, Estados Unidos; Professora Adjunta da UFRGS; Porto Alegre, RS, Brasil

²Doutoranda em Ciências do Movimento Humano na Escola de Educação Física da UFRGS; Professora da Universidade de Caxias do Sul (UCS), Porto Alegre, RS, Brasil

Endereço para correspondência:

Nadia Cristina Valentini
Rua Felizardo Furtado, 750 – Jardim Botânico
CEP 90670-090 – Porto Alegre/RS
E-mail: nadiacv@esef.ufrgs.br

Fonte financiadora: CNPq, com bolsa produtividade em pesquisa

Conflito de interesse: nada a declarar

Recebido em: 24/6/2010
Aprovado em: 25/10/2010

$p < 0,001$; percentil, $p = 0,04$; critério de classificação, qui-quadrado: $p = 0,047$). Demonstrou correlação moderada com a Escala do Desenvolvimento do Comportamento da Criança ($\rho = 0,342$; $p = 0,03$), sendo que a prova de McNemar-Bowker demonstrou diferenças entre os dois métodos ($p = 0,047$). Foi ainda observada estabilidade temporal da escala ($p = 0,07$; $\rho = 0,85$, $p < 0,001$) e poder preditivo ($p < 0,001$).

Conclusões: A versão em português da *Alberta Infant Motor Scale* demonstrou ser válida e fidedigna na avaliação das aquisições motoras de crianças brasileiras, o que a torna um instrumento útil a várias áreas da pesquisa científica e da clínica.

Palavras-chave: estudos de validação; destreza motora; desenvolvimento infantil.

RESUMEN

Objetivo: Traducir, adaptar y verificar la validez y clareza de criterios motor y constructo (consistencia interna, validez discriminante, correlación con otras pruebas y validez predictiva) de la versión en Portugués de la *Alberta Motor Infant Scale*.

Métodos: Estudio de validación transcultural, descriptivo y transversal. Participaron 21 profesionales y 561 niños de Rio Grande do Sul, con edades entre 0 y 18 meses (291 niños y 270 niñas). Se utilizaron los instrumentos: Alberta Infant Motor Scale (AIMS), Escala del Desarrollo del Comportamiento del Niño (EDCC) y un cuestionario para control de variables. El estudio comprendió las etapas de traducción y adaptación de la escala; análisis de la validez de contenido; entrenamiento de profesionales; recolección de datos en el ambiente familiar y en guarderías; análisis de la objetividad y fidedignidad, validez de criterio y constructo.

Resultados: Los resultados indican que la versión portuguesa de la AIMS contiene criterios motores claros y pertinentes; presenta óptima confiabilidad (escore total $\alpha = 0,88$) (prono, $\alpha = 0,86$; supino, $\alpha = 0,89$; sentado, $\alpha = 0,80$ y en pie, $\alpha = 0,85$) y poder discriminativo (típicas x atípicas - escore, $p < 0,001$; percentil, $p < 0,04$; criterio de clasificación, $\chi^2 = 6,03$; $p = 0,047$). Demostró correlación moderada con la EDCC ($\rho = 0,342$; $p = 0,03$), siendo que la prueba de McNemar Bowker demostró diferencias entre los dos métodos ($p = 0,047$). Se observó además estabilidad temporal de la escala ($p = 0,07$) ($\rho = 0,85$; $p < 0,001$) y poder predictivo ($p < 0,001$).

Conclusiones: La versión en portugués de la AIMS demostró ser válida y fidedigna en la evaluación de las ad-

quisiciones motoras de niños brasileños, lo que la convierte en un instrumento significativamente útil a varias áreas de la investigación científica y clínica.

Palabras clave: estudios de validación; destreza motora; desarrollo infantil.

Introduction

During the first years of life, motor behaviors provide information on the integrity and functionality of other systems, and changes in these systems become apparent over time⁽¹⁾. Motor delays are the first signs of possible developmental disorders. Vulnerable children exposed to risk factors may resist the negative effects of this exposure if diagnosed early, since learning is enhanced as a result of brain plasticity⁽²⁾. Early identification of development and motor function levels and introduction of intervention strategies may optimize the outcome for patients, allowing for proper decision making about strategies for optimal performance⁽³⁾.

The evaluation of development is ineffective when clinical investigation is the only method used^(4,5). Reliable scales, with good sensitivity and specificity, should be used to detect neuropsychomotor abnormalities^(5,6). In Brazil, the challenge of diagnosing developmental changes is aggravated by the lack of normative data and instruments standardized and validated for early childhood. The Alberta Infant Motor Scale (AIMS) is an example of an instrument used for Brazilian children without proper validation^(5,7-9). The use of this instrument without considering the required cross-cultural adaptations may lead to erroneous categorizations of motor developmental delays⁽¹⁰⁾.

The AIMS is an observational assessment tool used to measure gross motor maturation that evaluates the sequence of motor development and postural control of antigravity muscles in term and preterm infants in four positions: prone, supine, sitting, and standing⁽¹¹⁾. The clinical feasibility and psychometric properties of the AIMS have made it a valuable tool for identifying motor delays and abnormalities, providing information to health professionals and family about acquisition of motor skills, monitoring motor performance over time, detecting subtle changes, and evaluating the effectiveness of interventions in children with neuropsychomotor delays or disorders^(11,12). The theoretical construct of the AIMS establishes this scale as a supporting tool for research, clinical practice, and intervention actions^(10,13-22).

Estimates of reliability and validity for the AIMS have yet to be made for Brazilian children. Instruments used

to assess motor development, standardized in the source country, may have their results affected due to adaptation to other environments and diverse socioeconomic, ethnic, and cultural backgrounds^(18,21). In general, researchers tend to employ international norms and standards when assessing motor development in Brazilian children, thus hindering the generalization of results. There is a need for studies verifying the psychometric properties of the AIMS, since satisfactory levels of validity and reliability may not be achieved when the scale is used in culturally distinct populations. The objective of this study was to translate, adapt, and verify the validity of motor and construct criteria (internal consistency, discriminant validity, correlation with other tests, and predictive validity) of the Brazilian version of the AIMS.

Method

This study was approved by the Research Ethics Committee of Universidade Federal do Rio Grande do Sul, Brazil. A total of 21 professionals (4 translators, 9 physiotherapists, 6 physical educators, 1 nurse, and 1 pediatrician) participated in the process of translation, adaptation, and content validity of the Brazilian version of the AIMS. The sample was composed of 561 preterm and term children (291 boys and 270 girls) from nursery schools and primary health care units who were divided in a representative manner according to age group (Table 1).

Children were selected consecutively after permission was received from the institutions and informed consent was obtained from the parents/legal guardians. Inclusion criteria were age between 0 and 18 months and no previous participation in intervention programs. Exclusion criteria were musculoskeletal disorders, such as fractures, peripheral nerve injury, and musculoskeletal infection, among others. Sample size calculation was performed using Programs for Epidemiologists, version 4.0. For a 95% confidence level, a response rate of 50% and a 4% margin of error, approximately 600 children should be evaluated.

The AIMS is an observational assessment tool used to measure the development of term and preterm infants, from 38 weeks gestational age to 18 months corrected age⁽¹¹⁾, comprising 58 items divided into subscales (prone, supine, sitting, and standing) that describe spontaneous movement and motor skills. The examiner observes the child, taking into account aspects related to weight bearing, posture, and antigravity movements^(11,12). The scale presents raw scores, percentiles, and categorization of motor performance

Table 1 – General distribution of the study sample according to age group (trimester) and sex

Age at 1st assessment (months)	Female n (%)	Male n (%)
0-3	34 (12.6)	44 (15.1)
4-6	53 (19.6)	54 (18.6)
7-9	46 (17.0)	51 (17.5)
10-12	58 (21.5)	37 (12.7)
13-15	49 (18.1)	61 (21.0)
16-18	30 (11.1)	44 (15.1)

classified as: normal (above 25%); suspect (between 25 and 5%); and abnormal (below 5%)^(11,12). We also applied the Child Behavior Development Scale (CBDS)⁽²³⁾, standardized for Brazilian children aged 1 to 12 months, in which behaviors are classified as: axial and appendicular; spontaneous and stimulated; and communicative and non-communicative. This scale allows an estimate of the degree to which the child's behavior is consistent with expectations for his or her age level.

Translators and specialists in motor behavior translated and evaluated AIMS motor criteria. The translators worked first individually and then as a committee. The raters assigned scores for clarity and relevance individually.

After permission was received from the institutions, the parents/legal guardians were sent a consent form. The application of the scale took approximately 20 minutes per child in their homes or institutions of origin, with minimal manipulation. The sessions were videotaped for later analysis by independent evaluators. To analyze the temporal stability of the scale, part of the sample was retested (n=259) within 15 days.

To ensure an accurate and valid version of the instrument, a preliminary version (translation) was prepared, evaluated, and adapted to the Brazilian culture, including the analysis of content validity, objectivity, reliability, construct validity, and criterion validity⁽²⁴⁻²⁸⁾. Translations were performed (two independent bilingual translators) that resulted in two Portuguese-language versions. Based on the two translations, the instrument was back-translated into English by two other bilingual professionals, resulting in two new English-language versions of the test. The final version was established in consultation with the translators and the researchers.

Content validity was established by three raters with expertise in the area. Content validity index (CVI) and kappa coefficient of agreement were calculated and face validity

was determined by health professionals (4 PhDs, 4 Masters, 1 specialist, 8 graduates). Each rater received the adapted version and used a 5-point Likert scale to assign scores regarding clarity and relevance of motor criteria.

Regarding reproducibility, well trained evaluators, experienced in the use of the instrument for more than two years, analyzed inter and intrarater agreement using the intraclass correlation coefficient (ICC), Friedman test, and Wilcoxon test.

As for test-retest reliability, 259 children were retested, within a 15-day interval, and scoring stability was evaluated after some period of time using the Wilcoxon *t* test, Spearman's correlation coefficient, kappa coefficient of agreement, and McNemar-Bowker test.

For the analysis of criterion validity, expressed as predictive validity, 28 children were longitudinally followed up monthly, including 5 assessments (G1: 5 boys and 3 girls aged 1 to 6 months at 1st assessment); 20 children were evaluated with the AIMS and retested after 6 months (G2: 11 boys and 9 girls aged 1 to 11 months). Data were analyzed using the Friedman test, chi-square test, and Pearson's correlation coefficient.

For construct validity, the following procedures were performed: (1) Analysis of internal consistency using the Cronbach's alpha index as an indicator for the test and items ($n=561$); (2) Analysis of discriminant validity using the Student *t* test, Pearson's chi-square test, and Pearson's correlation coefficient to compare preterm infants ($n=124$, aged 0 to 15 months, including 62 atypical – extreme premature – infants) with typical infants (born at term), matched for chronological age; (3) Correlation of AIMS with CBDS rating criteria of 40 children, 36 preterm and 4 term infants aged 0 to 12 months, using Spearman's correlation coefficient, Kendall's coefficient, kappa coefficient of agreement, and McNemar-Bowker test.

The tests were selected according to data distribution (parametric or nonparametric data) and type of variable (qualitative or quantitative variables). For correlation coefficients, values were considered as follows: above 0.60 (strong); 0.30-0.60 (moderate); and below 0.30 (poor)⁽²⁹⁾. For Cronbach's alpha, values above 0.80 were considered excellent, above 0.70 good, and 0.60-0.70 acceptable⁽²⁷⁾. Significance level was set at $p \leq 0.05$.

Results

In the unified version, the semantics of the items was maintained and only a few changes were necessary, such as the

replacement of some words with more commonly used synonyms. Correction of technical terms used in the translations and adaptation of motor descriptors for each item were required to allow a better understanding by the target public.

As for content validity, with respect to face validity, 100% of health professionals assigned a Likert scale score of 5 in the analysis of AIMS items. Regarding content validity (Table 2), CVI agreement values ranged from $\alpha=66.7$ to $\alpha=92.8$ for clarity, with values above 0.98 for relevance. For both clarity and relevance, a significant and strong agreement was observed among raters (kappa coefficient, $p < 0.05$). Regarding objectivity, ICC ranged from $\alpha=0.86$ to $\alpha=0.99$, indicating strong agreement among raters (Table 3). No significant differences ($p < 0.05$) were found among raters. The analysis of intrarater reliability showed strong agreement, with ICC ranging from $\alpha=0.915$ to $\alpha=0.993$.

Regarding test-retest reliability, percentiles ranged from 0 to 100 in the test and retest, with no significant changes ($p=0.07$); a significant and strong positive correlation was detected ($\rho=0.85$; $p < 0.001$). The rating of motor performance ranged from 1 to 3 in the test and retest, and a significant and strong agreement was detected (kappa=0.680; $p < 0.001$). The McNemar-Bowker test revealed no differences in test-retest ratings ($p=0.40$). Regarding internal consistency, Cronbach's alpha coefficient for the four postures showed excellent reliability ($\alpha=0.88$). Values remained high in each posture, as follows: prone, $\alpha=0.86$; supine, $\alpha=0.89$; sitting, $\alpha=0.80$; and standing, $\alpha=0.85$. According to raw scores and percentiles, reliability was confirmed in the four postures ($\alpha=0.72$) (Table 4).

In the analysis of discriminant validity, term infants had significantly higher scores than preterm infants, both in total score ($p < 0.001$) and percentiles ($p=0.04$). A significant and strong correlation was found between scores and percentiles in both groups ($p < 0.001$). When comparing rating criteria, the term group was significantly associated with the normal range, whereas the preterm group was associated with suspected delay or delay ($p=0.047$).

A significant and moderate positive correlation was also detected between the AIMS and CBDS, confirmed by Kendall's coefficient ($W=0.319$; $p=0.02$) and kappa coefficient (0.309; $p=0.003$). The McNemar-Bowker test ($p=0.047$) revealed significant differences between instruments: the AIMS rated a greater number of children as "normal" and the CBDS rated significantly more cases as "delay".

The analysis of predictive validity in the group of children followed up longitudinally revealed a significant and strong

Table 2 – Content validity index (CVI) and kappa coefficient of agreement regarding clarity and relevance of the Brazilian version of the Alberta Infant Motor Scale

Comparisons	CVI	Kappa (95%CI)	p
Clarity			
Rater 1 x Rater 2 x Rater 3	66.7%	-	-
Rater 1 x Rater 2	92.8%	0.870 (0.746; 0.994)	<0.001
Rater 1 x Rater 3	73.2%	0.752 (0.579; 0.925)	0.002
Rater 2 x Rater 3	67.8%	0.514 (0.278; 0.745)	0.028
Relevance			
Rater 1 x Rater 2 x Rater 3	98.2%	-	-
Rater 1 x Rater 2	100.0%	0.935 (0.845; 1.023)	<0.001
Rater 1 x Rater 3	98.2%	0.824 (0.685; 0.973)	<0.001
Rater 2 x Rater 3	98.2%	0.824 (0.685; 0.973)	<0.001

IVC: Índice de Validade de Conteúdo.

Table 3 – Intra and interrater analysis of the objectivity of the Brazilian version of the Alberta Infant Motor Scale

Variables	Rater (R)				
	R1 x R2	R1 x R3	R2 x R3	R1 x R1	R1 x R2 x R3
	ICC (α) Wilcoxon (p)	ICC (α) Wilcoxon (p)	ICC (α) Wilcoxon (p)	ICC (α) Wilcoxon (p)	ICC (α) Friedman(p)
Prone	0.99	0.99	0.99	0.97	0.997
	-1.342 (0.180)	-1.186 (0.236)	-1.508 (0.132)	-0.351 (0.726)	2.24 (0.289)
Supine	0.92	0.86	0.87	0.915	0.917
	-0.756 (0.450)	-1.000 (0.317)	-1.265 (0.206)	-1.342 (0.180)	5.25 (0.07)
Sitting	0.99	0.93	0.88	0.993	0.954
	0.0 (1.00)	-1.552 (0.121)	-1.035(0.301)	-1.134 (0.257)	2.33 (0.311)
Standing	0.99	0.98	0.97	0.983	0.988
	-0.466 (0.641)	-0.36 (0.972)	-0.492 (0.623)	-0.320 (0.749)	0.157 (0.925)
Total Score	0.99	0.96	0.97	0.987	0.983
	-1.469 (0.142)	-0.494 (0.622)	-0.571 (0.568)	-0.755 (0.450)	0.70 (0.705)

ICC = intraclass correlation coefficient.

Table 4 – Median (25th; 75th percentile) of total score, percentile, and categorization for reliability and predictive, discriminant, and concurrent validity of the Brazilian version of the Alberta Infant Motor Scale

		Total score	Percentile	Categorization
Reliability (n=259)	Test	-	43 (14; 71)	3 (2; 3)
	Retest	-	61 (23; 71)	3 (2; 3)
Predictive validity G1 (n=8)	A1	9 (7.25; 17.75)	51.5 (23.5; 64.5)	3 (2.25; 3)
	A2	15 (10; 27.75)	42 (22.25; 81)	3 (2; 3)
	A3	29 (21.25; 37.5)	60.5 (17.75; 95)	3 (2.25; 3)
	A4	37.5 (31; 49.25)	86.5 (26.5; 92.75)	3 (2.25; 3)
	A5	49.5 (39; 52)	73.5 (62.25; 87.75)	3 (3; 3)
G2 (n=20)	A1	22 (10.25; 33.75)	34.50 (22; 60)	3 (2; 3)
	A2	51.5 (48.25; 58)	67.50 (48; 71)	3 (3; 3)
Discriminant validity (n=124)	AT	32.50 (22; 48.25)	52.50 (23.50; 71.50)	3 (2.75; 3)
	PT	18 (11; 27.75)	36.40 (11.50; 63.25)	3 (2; 3)
Concurrent validity (n=80)	AIMS	-	-	2 (2; 3)
	CBDS	-	-	2 (1; 2)

AIMS = Alberta Infant Motor Scale; CBDS = Child Behavior Development Scale.

positive joint correlation for the following variables: final score (ICC=0.96, range: 0.89-0.99; $p < 0.0001$); percentile (ICC=0.88, range: 0.67-0.97; $p < 0.0001$); and performance rating (ICC=0.83, range: 0.49-0.96; $p = 0.001$), indicating that increased values in the 1st assessment were correlated with increased values in the subsequent assessment, and so on. In the comparative analysis, the Friedman test revealed a significant difference in scores from assessments 1 to 5 ($p < 0.001$). Scores from assessment 5 were higher than those from assessment 4, which in turn were higher than those from assessment 3, which in turn were higher than those from assessments 1 and 2. The percentile values in all assessments showed a significant and progressive increase ($p = 0.049$). No statistically significant differences were observed when ratings from assessments 1 to 5 were compared ($p = 0.37$).

The analysis of children ($n = 20$) evaluated twice within a 6-month interval revealed a significant and strong positive correlation between scores 1 and 2 ($r = 0.730$; $p < 0.0001$). The analysis of percentiles ($r = 0.22$; $p = 0.347$) and ratings ($\rho = 0.26$; $p = 0.269$) revealed no significant results in both assessments. A significant difference was found between mean values of scores 1 and 2 ($p < 0.001$): the mean value observed in total score 2 was higher than that of total score 1. All children had higher values at assessment 2 than at assessment 1. When comparing percentiles 1 vs. 2, percentiles at assessment 2 were significantly higher ($p < 0.01$) than those at assessment 1. A direct comparison of percentiles at both assessments revealed that, out of 20 children, 18 (90.0%) had higher percentiles at assessment 2 and 2 (10.0%) had lower scores at assessment 2. The agreement analysis of performance rating (1x2) showed good (kappa=0.458) and significant ($p = 0.03$) results.

Discussion

The analysis of the psychometric properties of the scale indicated good reliability, similar to that obtained in the Canadian population⁽¹¹⁾. The independent back-translation resulted in a unified version in Portuguese of the AIMS, the *Escala Motora Infantil de Alberta* (EMIA). A comparison among the four translated versions, discussion, and analysis in consultation with all translators allowed a reduction of biases inherent in the process carried out with only one translator.

Regarding face validity, there was unanimous opinion among the professionals involved that the content was

appropriate to assess motor skills of children in different postures. In the analysis of content validity, in relation to agreement among raters regarding CVI for clarity and relevance, there was strong agreement among raters, confirmed by kappa coefficient of agreement with similar responses⁽²⁹⁾. The results indicate that the Brazilian version of the AIMS (EMIA) showed excellent content validity, with clear and relevant criteria, adequate representation of items in relation to concepts, and theoretical relevance.

The results also demonstrated strong agreement among raters (ICC: 0.86-0.99), high reliability, and agreement with values proposed by the author of the scale, who suggests values above 0.80⁽¹²⁾, which appears to be sufficient to classify our results as correct⁽²⁷⁾.

Regarding temporal stability of the scale, percentiles and performance rating criteria remained constant after some period of time. High percentiles and ratings in the test were significantly associated with high values in the retest. The Brazilian version of the AIMS (EMIA) has reliability and temporal stability, with data showing high correlation and no significant differences between test and retest⁽²⁹⁾. The phenomenon of acquiescence (whether positive or negative) was not observed, indicating reliable data⁽²⁵⁾. The correlation values for percentiles ($\rho = 0.85$) and rating criteria (kappa=0.68) could have been even higher if a shorter time interval between test and retest had been employed, considering a 7-day period as a reference interval between assessments⁽¹¹⁾.

The Brazilian version of the AIMS (EMIA) has high internal consistency; the values obtained using Cronbach's alpha (0.72-0.89) reflect a profile of high homogeneity among variables. The Cronbach's alpha coefficient should be at least 0.60, and the larger the sample, the more difficult it becomes to obtain consistent results⁽²⁷⁾. Given the sample size, our results show homogenous items representing a same trait, measuring the same construct.

The Brazilian version of the AIMS (EMIA) was shown to be valid to distinguish atypical behaviors, because percentiles and total score were significantly different in the groups of preterm and term infants investigated. The results indicate that a high total score was associated with high percentile values, with a positive correlation between two measurements of a same construct, thus confirming convergent validity⁽²⁷⁾. Considering rating criteria, discriminant validity was confirmed, establishing a difference between term and preterm groups. Results from previous studies confirm the

ability of the scale to diagnose motor disorders during the first years of life^(15,16,18,22,30).

A comparison between the Brazilian version of the AIMS (EMIA) and CBDS categorizations revealed a moderate correlation ($\rho=0.34$; $\kappa=0.30$; $W=0.31$), indicating poor concurrent validity, which is consistent with findings from studies comparing the AIMS with the Test of Infant Performance (correlations between 0.20 and 0.67)⁽¹⁷⁾ and the AIMS with the Daily Activities of Infant Scale⁽⁵⁰⁾. Campos *et al* observed moderate agreement of the AIMS with the Bayley scale at 5 months ($k=0.503$) and poor agreement at 10 months ($k=0.209$)⁽⁵⁾. In contrast, international studies^(11,13,15,21) and a Brazilian study⁽¹⁴⁾ showed a high correlation of the AIMS with other motor scales. In the present study, the results obtained may be due to the fact that the CBDS has fewer items related to motor function (15 items) compared to the Brazilian version of the AIMS (EMIA) (58 items), thus compromising the reliability of results. Although the CBDS is applied with similar goals, this scale also evaluates social and cognitive aspects, thus providing a reduced number of items in the motor subscale. Both instruments have advantages and disadvantages. Therefore, professionals should be able to determine which one better meets their needs.

Our results suggest that the Brazilian version of the AIMS (EMIA) has predictive validity, since ratings were similar longitudinally (G1) and pre- and post-tests (G2). These findings are consistent with a study for validation of the instrument in the Canadian population⁽¹²⁾, but different from a study conducted in Taiwan, which found limited predictive validity for this scale to evaluate preterm infants⁽²¹⁾.

The increase in percentile values over time, within the age groups studied, highlights a greater sensitivity and accuracy of the Brazilian version of the AIMS (EMIA) to evaluate children aged 3 to 9 months, as demonstrated in a previous study⁽¹⁷⁾.

It can be concluded that the procedures performed to analyze the psychometric properties of the Brazilian version of the AIMS (EMIA) showed that the back-translation was efficient to avoid biases resulting from misunderstandings in English. In addition, the scale was recognized by experts as an efficient tool to evaluate motor function in children and proved to be a reliable and consistent instrument, with significant predictive and discriminant power. It is noteworthy, however, that the results for concurrent validity were limited, thus warranting a comparison of the AIMS with motor tests other than the CBDS. Regarding predictive validity, a longer follow-up period could have yielded different results, favoring the predictive power of the instrument analyzed in this study.

The results obtained can be reproduced in daily practice, confirming the reliability and content, criterion, and construct validity of the Brazilian version of the AIMS (EMIA), thus encouraging professionals to use this scale in the evaluation and planning of intervention programs. We highlight the need for normative studies involving Brazilian children, since the use of categorizations from other populations may not be suitable to interpret data obtained in Brazilian children, assuming that these categorizations may vary given the diversity of socioeconomic and cultural factors in the country.

References

1. Wijnhoven TM, de Onis M, Onyango AW, Wang T, Bjoerneboe GE, Bhandari N *et al*. Assessment of gross motor development in the WHO Multicentre Growth Reference Study. *Food Nutr Bull* 2004;25:S37-45.
2. Mahoney G, Robinson C, Perales F. Early motor intervention: the need for new treatment paradigms. *Infants and Young Children* 2004;17:291-300.
3. Valentini NC. A influência de uma intervenção motora no desempenho motor e na percepção de competência de crianças com atrasos motores. *Rev Paul Educ Fis* 2004;16:61-75.
4. Santos RS, Araújo AP, Porto MA. Diagnóstico precoce de anormalidades no desenvolvimento em prematuros: instrumentos de avaliação. *J Pediat (Rio J)* 2008;84:289-99.
5. Campos D, Santos DC, Gonçalves VM, Goto MM, Arias AV, Brianeze AC *et al*. Agreement between scales for screening and diagnosis of motor development at 6 months. *J Pediat (Rio J)* 2006;82:470-4.
6. Spittle AJ, Doyle LW, Boyd RN. A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. *Dev Med Child Neurol* 2008;50:254-66.
7. Campos D, Santos DC, Gonçalves VM, Montebelo MI, Goto MM, Gabbard C. Postural control of small for gestational age infants born at term. *Rev Bras Fisioter* 2007;11:7-12.
8. Chagas PS, Mancini MC, Fonseca ST, Soares TB, Gomes VP, Sampaio RF. Neuromuscular mechanisms and anthropometric modifications in the initial stages of independent gait. *Gait Posture* 2006;24:375-81.
9. Zajonz R, Müller AB, Valentini NC. The influence of environmental factors in motor and social performance of children from the suburb of Porto Alegre. *Maringá* 2008;19:159-71.
10. Lopes VB, Lima CD, Tudella E. Motor acquisition rate in Brazilian infants. *Inf Child Dev* 2009;18:122-32.
11. Piper MC, Darrah J. *Motor assessment of the developing infant*. Philadelphia: Saunders; 1994.
12. Piper MC, Pinnell LE, Darrah J, Maguire T, Byrne PJ. Construction and validation of the Alberta Infant Motor Scale (AIMS). *Can J Public Health* 1992;83:S46-50.
13. Uesugui M, Tokuhisa K, Shimada T. The reliability and validity of the Alberta Infant Motor Scale in Japan. *J Phys Ther Sci* 2008;20:169-75.

14. Almeida KM, Dutra MV, Mello RR, Reis AB, Martins PS. Concurrent validity and reliability of the Alberta Infant Motor Scale in premature infants. *J Pediatr (Rio J)* 2008;8:442-8.
15. Heineman KR, Bos AF, Hadders-Algra M. The Infant Motor Profile: a standardized and qualitative method to assess motor behaviour in infancy. *Dev Med Child Neurol* 2008;50:275-82.
16. Tse L, Mayson TA, Leo S, Lee LL, Harris SR, Hayes VE *et al*. Concurrent validity of the Harris Infant Neuromotor Test and the Alberta Infant Motor Scale. *J Pediatr Nurs* 2008;23:28-36.
17. Campbell SK, Kolobe TH, Wright BD, Linacre JM. Validity of the Test Infant Motor Performance for prediction of 6-, 9- and 12-month scores on the Alberta Infant Motor Scale. *Dev Med Child Neurol* 2002;44:263-72.
18. Fleuren KM, Smit LS, Stijnen T, Hartman A. New reference values for the Alberta Infant Motor Scale need to be established. *Acta Pædiatr* 2007;96:424-7.
19. Charitou S, Simitsopoulou A, Kontogianni V, Skordilis EK, Koutsouki D. Application of the Alberta Infant Motor Scale in Greece: a preliminary study. *Inq Sport Phys Ed* 2007;5:273-82.
20. Bartlett DJ, Fanning JE. Use of the Alberta Infant Motor Scale to characterize the motor development of infants born preterm at eight months corrected age. *Phys Occup Ther Pediatr* 2003;23:31-45.
21. Jeng SF, Yau KI, Chen LC, Hsiao SF. Alberta infant motor scale: reliability and validity when used on preterm infants in Taiwan. *Phys Ther* 2000;80:168-78.
22. Darrah J, Piper M, Watt MJ. Assessment of gross motor skills of at-risk infants: predictive validity of the Alberta Infant Motor Scale. *Dev Med Child Neurol* 1998;40:485-91.
23. Pinto EB, Vilanova LC, Vieira RM. O desenvolvimento do comportamento da criança no primeiro ano de vida: padronização de uma escala para a avaliação e o acompanhamento. São Paulo: Casa do Psicólogo; 1997.
24. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000;25:3186-91.
25. Cronbach LJ. Fundamentos da testagem psicológica. 5ª ed. Porto Alegre: Artes Médicas; 1996.
26. Fachel JM, Camey S. Avaliação psicométrica: a qualidade das medidas e o entendimento dos dados. In: Cunha JA (Org). *Psicodiagnóstico-V*. 5ª ed. Porto Alegre: Artmed; 2003. p. 159-70.
27. Pasquali L. Parâmetros psicométricos dos testes psicológicos. São Paulo: Casa do Psicólogo; 2001.
28. Vallerand RJ. Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: implications pour la recherche en langue française. *Can Psychol* 1998;30:662-80.
29. Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: Artmed; 2003.
30. Bartlett DJ, Fanning JK, Miller L, Conti-Becker A, Doralp S. Development of the daily activities of infants scale: a measure supporting early motor development. *Dev Med Child Neurol* 2008;50:613-7.