

Escala Motora Infantil de Alberta: validação para uma população gaúcha

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

Nadia Cristina Valentini¹, Raquel Saccani²

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, $p<0,001$; 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.

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:

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

$p=0.047$) 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.

Introdução

Nos primeiros anos de vida, os comportamentos motores representam a integridade e a funcionalidade dos demais sistemas, cujas alterações tornam-se aparentes com o passar do tempo⁽¹⁾. Atrasos motores são as primeiras manifestações de possíveis desordens do desenvolvimento. A criança vulnerável exposta a fatores de risco pode resistir aos efeitos negativos dessa exposição se diagnosticada precocemente, uma vez que as aprendizagens são potencializadas em decorrência da plasticidade cerebral⁽²⁾. Com a identificação precoce dos níveis de desenvolvimento e função motora, intervenções planejadas otimizam o prognóstico, possibilitando a adequada tomada de decisão quanto a estratégias para o melhor desempenho⁽³⁾.

A avaliação do desenvolvimento é ineficaz quando utilizada somente a investigação clínica^(4,5). Escalas confiáveis, com comprovada sensibilidade e especificidade, devem ser empregadas para identificar alterações neuropsicomotoras^(5,6). No Brasil, o desafio do diagnóstico de alterações no desenvolvimento é agravado pela escassez de dados normativos e de instrumentos padronizados e validados na primeira infância. A *Alberta Infant Motor Scale* (AIMS) é um exemplo de uso de instrumentos em crianças brasileiras sem a necessária validação^(5,7-9). O uso desse instrumento sem considerar as necessárias adaptações em diferentes culturas pode conduzir a categorizações errôneas de atrasos no desenvolvimento motor⁽¹⁰⁾.

A AIMS é um instrumento observacional da motricidade ampla, que avalia a sequência do desenvolvimento motor e o controle da musculatura antigravitacional nas posturas prono, supino, sentado e de pé, de crianças a termo e pré-termo⁽¹¹⁾. A praticidade e as características psicométricas da AIMS a tornam uma ferramenta valiosa para identificar atrasos ou anormalidades, fornecer informações a profissionais da saúde

e familiares sobre aquisições de habilidades, acompanhar o desempenho ao longo do tempo, detectar mudanças sutis, avaliar a eficácia de intervenções em crianças com disfunções e atraso neuropsicomotor^(11,12). O constructo teórico da escala a configura como um instrumento de apoio à pesquisa, à prática clínica e à ação interventiva^(10,13-22).

As estimativas de confiabilidade e validade da AIMS ainda não foram feitas para crianças brasileiras. Instrumentos que avaliam o desenvolvimento motor, padronizados no país de origem, podem sofrer interferência em seus resultados frente à adaptação a outro meio e a fatores socioeconômicos, étnicos e culturais diversos^(18,21). Em geral, pesquisadores, ao avaliarem o desenvolvimento de crianças brasileiras, relacionam-no às normas e categorias internacionais, dificultando a generalização de resultados. Observa-se a necessidade de estudos para verificar as propriedades psicométricas da AIMS; a validade e confiabilidade da escala podem não se manter frente a uma população culturalmente distinta. Este estudo teve como objetivo traduzir, adaptar e verificar a validade dos critérios motores e de constructo (consistência interna, validade discriminante, correlação com outros testes e validade preditiva) da versão brasileira da AIMS.

Método

O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal do Rio Grande do Sul (UFRGS). Participaram 21 profissionais (quatro tradutores, nove fisioterapeutas, seis educadores físicos, um enfermeiro e um pediatra) nos processos de tradução, adaptação e validação de conteúdo da versão em língua portuguesa da AIMS, além de 561 crianças pré-termo e a termo (291 meninos) provenientes de escolas de educação infantil e Unidades Básicas de Saúde (UBS), distribuídas de forma representativa em cada faixa etária (Tabela 1).

Tabela 1 - Distribuição geral da amostra segundo faixa etária (trimestre) e gênero

Idade na 1ª avaliação (meses)	Sexo feminino n (%)	Sexo masculino 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)

As crianças foram incluídas de forma consecutiva, mediante autorização das instituições e da assinatura do termo de consentimento pelos responsáveis, obedecendo a critérios de inclusão (idade entre zero e 18 meses e não participação em programas de intervenção) e exclusão (afecções osteomioarticulares como fraturas, lesão nervosa periférica, infecção osteomuscular, entre outras). O cálculo amostral foi realizado no *Programs for Epidemiologists* versão 4.0. Para um nível de confiança de 95%, uma proporção de respostas em 50% e uma margem de erro de 4%, seria necessário avaliar, aproximadamente, 600 crianças.

A AIMS é um instrumento de observação que avalia o desenvolvimento dos recém-nascidos a termo e pré-termo, a partir de 38 semanas de idade gestacional até 18 meses de idade corrigida⁽¹¹⁾. É composta por 58 itens agrupados em subescalas (prono, supino, sentado e em pé) que descrevem a movimentação espontânea e as habilidades motoras. O examinador observa a criança, levando em consideração aspectos da superfície do corpo que sustenta o peso, a postura e os movimentos antigravitacionais^(11,12). A escala apresenta escores brutos, percentis e categorização do desempenho motor em: normal (>25%); suspeito (entre 25 e 5%); anormal (<5%)^(11,12). Também foi aplicada a Escala de Desenvolvimento do Comportamento da Criança (EDCC)⁽²³⁾, padronizada para crianças brasileiras, entre um e 12 meses, que categoriza os comportamentos em: axial e apendicular, espontâneo e estimulado, e em comunicativo e não comunicativo. Essa escala possibilita estimar se os comportamentos estão de acordo com o esperado para a idade.

Tradutores e doutores em comportamento motor realizaram a tradução e a avaliação dos critérios motores da AIMS. O trabalho dos tradutores foi primeiramente individualizado e, posteriormente, em comitê. As pontuações para clareza e pertinência pelos juízes foram individuais.

Após o consentimento das instituições, foi enviado aos responsáveis, o termo de consentimento livre e esclarecido. A aplicação da escala levou aproximadamente 20 minutos por criança em suas residências ou instituições de origem, com o mínimo de manipulação; o procedimento foi filmado para posterior análise por avaliadores independentes. Para analisar a estabilidade temporal da escala, parte da amostra foi retestada (n=259) no intervalo de 15 dias.

A fim de assegurar uma versão válida e fiel do instrumento, foi realizada a preparação da versão preliminar (tradução), a avaliação e a adaptação para realidade cultural brasileira, além das análises de validade de conteúdo, objetividade, fidedignidade, de constructo e de critério⁽²⁴⁻²⁸⁾. Foram feitas traduções (dois tradutores bilíngues independentes) que determinaram

duas versões em português. A partir dessas duas traduções, o instrumento foi reconvertido ao idioma de origem por dois outros profissionais bilíngues, obtendo-se duas novas versões em inglês do teste. A elaboração da versão final foi em comitê de tradutores acompanhados pelas pesquisadoras.

A validade de conteúdo foi determinado por três juízes, com conhecimento específico na área. Nesses dados foram calculados o Índice de Validade de Conteúdo (IVC), o coeficiente de concordância Kappa e a análise de validade aparente ou de face, determinada por profissionais da área da saúde (quatro doutores, quatro mestres, um especialista e oito graduados). Cada avaliador recebeu a versão adaptada e utilizou uma escala Likert de 5 pontos para atribuir escores no que se refere à clareza e à pertinência de critérios motores.

Para verificar a reprodutibilidade, avaliadores treinados e com experiência no uso do instrumento por mais de dois anos analisaram a concordância interavaliadores e intra-avaliador por meio do coeficiente de correlação intraclassas (ICC), teste de Friedman e de Wilcoxon.

Para a fidedignidade teste-reteste, 259 crianças foram retestadas, num intervalo máximo de 15 dias, com avaliação da estabilidade das pontuações após um lapso de tempo, utilizado-se do teste *t* de Wilcoxon, o coeficiente de correlação de Spearman, o coeficiente de concordância kappa e o teste de McNemar-Bowker.

Para a análise da validade de critério, representada pela validade preditiva, 28 crianças foram acompanhadas mensalmente, de forma longitudinal, com cinco avaliações (G1: cinco meninos e três meninas com idades na primeira avaliação entre um e seis meses); 20 crianças foram avaliadas com a AIMS e retestadas após seis meses (G2: 11 meninos e nove meninas com idades entre um e 11 meses). Os dados foram analisados com o teste Friedman, qui-quadrado e o coeficiente de correlação de Pearson.

Para a validade do constructo, foram utilizados os seguintes procedimentos: (1) análise da consistência interna, utilizando como indicador do teste e dos itens o índice alfa de Cronbach (n=561); (2) validade discriminante, utilizando teste *t* de Student, qui-quadrado de Pearson e coeficiente de correlação de Pearson para comparar crianças pré-termo (n=124, na faixa etária de zero a 15 meses, sendo 62 atípicas – prematuras extremas) com crianças típicas (a termo), pareadas segundo a idade cronológica; (3) correlação dos critérios de classificação da AIMS com EDCC com 40 crianças, incluindo pré-termos (36) e a termo (4), com idade entre zero e 12 meses, por meio do coeficiente de correlação de Spearman, coeficiente de Kendall, coeficiente de concordância de Kappa e prova de McNemar-Bowker.

A escolha dos testes foi feita de acordo com a distribuição dos dados (paramétricos ou não paramétricos) e tipo de variável (qualitativas ou quantitativas). Para os coeficientes de correlação, foram considerados os valores acima de 0,60 (forte); os valores entre 0,30 e 0,60 (moderada) e os inferiores a 0,30 (pobre)⁽²⁹⁾. Para o alfa de Cronbach, foram considerados ótimos os valores acima de 0,80; bons, acima de 0,70; e aceitáveis entre 0,60 e 0,70⁽²⁷⁾. O nível de significância adotado foi de $p \leq 0,05$.

Resultados

Na versão unificada, manteve-se a semântica dos itens e poucas mudanças foram necessárias – apenas substituições de palavras menos usadas por sinônimos mais frequentes no

cotidiano. Correções de termos técnicos utilizados na tradução e adequação dos descritores motores de cada item também foram realizadas para a melhor compreensão do público-alvo.

Quanto à validade de conteúdo no que se refere à validade aparente, 100% dos profissionais da área da saúde atribuíram escore 5 na escala Likert para a análise dos itens da AIMS. Quanto à validade de conteúdo (Tabela 2), para clareza, os valores de concordância do IVC variaram de $\alpha=66,7$ até $\alpha=92,8$ e, para pertinência, os valores foram superiores a 0,98. Para a clareza e pertinência (IVC), observou-se concordância forte e significativa entre os juízes (coeficiente Kappa $p < 0,05$). Quanto à objetividade, os coeficientes de correlação intraclasse oscilaram entre $\alpha=0,86$ e $\alpha=0,99$, indicando forte concordância entre avaliadores (Tabela 3).

Tabela 2 - Índice de Validade de Conteúdo e coeficiente de concordância Kappa quanto à clareza e à pertinência da versão brasileira da *Alberta Infant Motor Scale*

	IVC (%)	Kappa (IC95%)	Valor p
<i>Clareza</i>			
Juiz 1 vs juiz 2 vs juiz 3	66,7	-	-
Juiz 1 vs juiz 2	92,8	0,870 (0,746-0,994)	<0,001
Juiz 1 vs juiz 3	73,2	0,752 (0,579-0,925)	0,002
Juiz 2 vs juiz 3	67,8	0,514 (0,278-0,745)	0,028
<i>Pertinência</i>			
Juiz 1 vs juiz 2 vs juiz 3	98,2	-	-
Juiz 1 vs juiz 2	100,0	0,935 (0,845-1,023)	<0,001
Juiz 1 vs juiz 3	98,2	0,824 (0,685-0,973)	<0,001
Juiz 2 vs juiz 3	98,2	0,824 (0,685-0,973)	<0,001

IVC: Índice de Validade de Conteúdo.

Tabela 3 - Análise da objetividade da versão brasileira da *Alberta Infant Motor Scale* inter e intra-avaliadores

Variáveis	Avaliador (A)				
	A1 vs A2	A1 vs A3	A2 vs A3	A1 vs A1	A1 vs A2 vs A3
	ICC (α) Wilcoxon (p)	ICC (α) Wilcoxon (p)	ICC (α) Wilcoxon (p)	ICC (α) Wilcoxon (p)	ICC (α) Friedman (p)
Prono	0,99 -1,342 (0,180)	0,99 -1,186 (0,236)	0,99 -1,508 (0,132)	0,97 -0,351 (0,726)	0,997 2,24 (0,289)
Supino	0,92 -0,756 (0,450)	0,86 -1,000 (0,317)	0,87 -1,265 (0,206)	0,915 -1,342 (0,180)	0,917 5,25 (0,07)
Sentado	0,99 0,0 (1,00)	0,93 -1,552 (0,121)	0,88 -1,035 (0,301)	0,993 -1,134 (0,257)	0,954 2,33 (0,311)
Em pé	0,99 -0,466 (0,641)	0,98 -0,36 (0,972)	0,97 -0,492 (0,623)	0,983 -0,320 (0,749)	0,988 0,157 (0,925)
Escore total	0,99 -1,469 (0,142)	0,96 -0,494 (0,622)	0,97 -0,571 (0,568)	0,987 -0,755 (0,450)	0,983 0,70 (0,705)

ICC: coeficiente de correlação intraclases.

Não foram encontradas diferenças significativas ($p < 0,05$) entre avaliadores. A análise de confiabilidade intra-avaliador demonstrou concordância forte, com ICC variando de $\alpha = 0,92$ a $\alpha = 0,99$.

Para a fidedignidade teste-reteste, os percentis variaram de 0 a 100 no teste e reteste, sem alteração significativa ($p = 0,07$); foi detectada correlação positiva, de grau forte e significativa ($\rho = 0,85$; $p < 0,001$). A classificação de desempenho motor variou entre 1 e 3 tanto no teste quanto no reteste, com concordância significativa de grau forte ($\kappa = 0,680$; $p < 0,001$). O resultado do teste de McNemar-Bowker não evidenciou diferenças nas classificações do teste-reteste ($p = 0,40$). Quanto à consistência interna, o coeficiente de confiabilidade alfa de Cronbach para as quatro posturas demonstrou ótima confiabilidade ($\alpha = 0,88$). Em cada postura, os valores se mantiveram altos (prono, $\alpha = 0,86$; supino, $\alpha = 0,89$; sentado, $\alpha = 0,80$; e em pé, $\alpha = 0,85$). Para o escore bruto e o percentil, a confiabilidade foi confirmada nas quatro posturas ($\alpha = 0,72$) (Tabela 4).

Na análise discriminante, os escores das crianças a termo se mostraram significativamente mais elevados que os das crianças pré-termo, tanto no escore total ($p < 0,001$) como no percentil ($p = 0,04$). Foi identificada correlação forte e significativa entre os escores e percentis nos grupos ($p < 0,001$). Na comparação dos critérios de classificação, o grupo a termo se mostrou significativamente associado à classificação normal, enquanto o grupo pré-termo se associou às classificações de suspeita de atraso e atraso ($p = 0,047$).

Foi detectada ainda correlação significativa e positiva de grau moderado ($\rho = 0,342$; $p = 0,03$) entre a AIMS e a EDCC, confirmada pelo coeficiente de Kendall ($W = 0,319$; $p = 0,02$) e pelo Kappa ($0,309$; $p = 0,003$). A prova de McNemar-Bowker ($p = 0,047$) demonstrou diferenças relevantes entre os dois instrumentos: a AIMS classificou um número maior de crianças como “normal” e a EDCC classificou significativamente mais casos como “atraso”.

Na validade preditiva, com o grupo de crianças acompanhadas longitudinalmente, a correlação conjunta foi significativa e positiva de grau forte para as variáveis consideradas: escore final (ICC = 0,96 e variação: 0,89-0,99; $p < 0,0001$); percentil (ICC = 0,88 e variação: 0,67-0,97; $p < 0,0001$) e classificação de desempenho (ICC = 0,83 e variação: 0,49-0,96; $p = 0,001$), indicando que os valores elevados na avaliação 1 se mostraram correlacionados a valores elevados nas demais avaliações e assim sucessivamente. Nas análises comparativas, o teste de Friedman detectou diferença significativa ($p < 0,001$) nos escores nas avaliações de 1 a 5. Os escores da avaliação 5 foram mais elevados que os da avaliação 4, que, por sua vez, foram maiores que os da avaliação 3, que, por sua vez, foram maiores que os escores das avaliações 1 e 2. Os valores de percentis em todas as avaliações apresentaram aumento progressivo significativo ($p = 0,049$). Nas comparações entre as classificações nas avaliações de 1 a 5, não foi detectada diferença estatística significativa ($p = 0,37$).

Quanto à análise referente às crianças ($n = 20$) avaliadas duas vezes num intervalo de seis meses, observou-se

Tabela 4 - Medianas (percentil 25; 75) nas variáveis de escore total, percentil e categorização para fidedignidade, validade preditiva, discriminante e concorrente da versão brasileira da *Alberta Infant Motor Scale*

		Escore total	Percentil	Categorização
Fidedignidade (n=259)	Teste	-	43 (14; 71)	3 (2; 3)
	Reteste	-	61 (23; 71)	3 (2; 3)
V. Preditivo G1 (n=8)	A1	9 (7; 18)	52 (24; 65)	3 (2; 3)
	A2	15 (10; 28)	42 (22; 81)	3 (2; 3)
	A3	29 (21; 38)	61 (18; 95)	3 (2; 3)
	A4	38 (31; 49)	87 (27; 93)	3 (2; 3)
	A5	50 (39; 52)	74 (62; 88)	3 (3; 3)
G2 (n=20)	A1	22 (10; 34)	35 (22; 60)	3 (2; 3)
	A2	52 (48; 58)	68 (48; 71)	3 (3; 3)
V. Discriminante (n=124)	AT	33 (22; 48)	53 (24; 72)	3 (3; 3)
	PT	18 (11; 28)	36 (12; 63)	3 (2; 3)
V. Concorrente (n=80)	AIMS	-	-	2 (2; 3)
	EDCC	-	-	2 (1; 2)

V: valor; A: avaliador; AT: a termo; PT: pré-termo; AIMS: *Alberta Infant Motor Scale*; EDCC: Escala de Desenvolvimento do Comportamento da Criança.

correlação significativa, positiva e de grau forte entre os escores 1 e 2 ($r=0,730$; $p<0,0001$). Os resultados referentes ao percentil ($r=0,22$; $p=0,347$) e à classificação ($\rho=0,26$; $p=0,269$) nas duas avaliações não foram significativos. Foi detectada diferença significativa entre as médias dos escores 1 e 2 ($p<0,001$): a média observada no escore total 2 se mostrou mais elevada que a média no escore total 1. Todas as crianças apresentaram valores mais elevados na avaliação 2 comparada à 1. Nas comparações entre os percentis 1 *versus* 2, os percentis da avaliação 2 se mostraram significativamente mais elevados ($p<0,01$) que os da avaliação 1. No confronto direto dos percentis das duas avaliações, notou-se que, das 20 crianças, 18 (90,0%) apresentaram percentis mais elevados na avaliação 2 e duas (10,0%) mostraram escores menores na avaliação 2. A análise de concordância da classificação de desempenho (1 *versus* 2) foi boa ($\kappa=0,458$) e significativa ($p=0,03$).

Discussão

As propriedades psicométricas da escala apresentaram adequada confiabilidade, similares às analisada em população canadense⁽¹¹⁾. A tradução reversa e independente resultou em uma versão unificada em língua portuguesa, a Escala Motora Infantil de Alberta (EMIA). A comparação realizada entre as quatro versões traduzidas, a discussão e a análise em comitê possibilitaram a redução de vieses inerentes ao processo realizado com um só tradutor.

Quanto à validade de conteúdo aparente, os profissionais foram unânimes em considerar o conteúdo adequado para avaliar as habilidades motoras das crianças nas diferentes posturas. Na análise da validade de conteúdo, no que se refere à concordância das avaliações dos juízes quanto ao IVC para clareza e pertinência, observa-se concordância forte entre os juízes, confirmada pelo coeficiente de concordância Kappa com semelhança nas respostas⁽²⁹⁾. Os resultados indicam que a versão da EMIA apresentou índices ótimos de validade de conteúdo, com critérios claros e pertinentes, adequada representação dos itens em relação aos conceitos e relevância teórica.

Os resultados demonstram também concordância forte entre os avaliadores (ICC entre 0,86 e 0,99), altos índices de confiabilidade e concordância com o que foi proposto pela autora da escala, a qual sugere valores acima de 0,80⁽¹²⁾, sendo os mesmos suficientes para considerar os resultados obtidos como corretos⁽²⁷⁾.

Quanto à estabilidade temporal da escala, os percentis e o critério de classificação de desempenho mantiveram-se constantes após intervalo de tempo. Percentil e classificação elevados no teste se mostraram significativamente associados a valores elevados do reteste. A EMIA possui confiabilidade e estabilidade temporal; os dados demonstraram alta correlação e diferenças não significativas entre teste-reteste⁽²⁹⁾. Em nenhum caso ocorreu o fenômeno da aquiescência (positiva ou negativa), o que indica serem os dados confiáveis⁽²⁵⁾. Os valores de correlação para o percentil ($\rho=0,85$) e para o critério de classificação ($\kappa=0,68$) poderiam ter sido ainda mais elevados se o intervalo de tempo entre teste e reteste fosse menor, considerando como referência um intervalo de sete dias entre as avaliações⁽¹¹⁾.

A EMIA possui elevada consistência interna; os valores obtidos por meio do alfa de Cronbach (0,72 a 0,89) refletem um perfil de alta homogeneidade entre as variáveis. O coeficiente alfa de Cronbach deve ser, no mínimo, de 0,60 e, quanto maior a amostra, mais difícil se torna obter resultados concordantes⁽²⁷⁾. Tendo em vista o tamanho amostral, os resultados mostram que os itens são homogêneos e que constituem a representação de um mesmo traço e medem o mesmo constructo.

Verificou-se que a EMIA é válida para discriminar comportamentos atípicos, pois os percentis e o escore total foram significativamente diferentes nos grupos de neonatos pré-termo e a termo investigados. Os resultados indicam que as pontuações elevadas no escore total se mostraram associadas às pontuações elevadas do percentil, com correlação positiva entre duas medidas de um mesmo constructo, confirmando a validade convergente⁽²⁷⁾. Considerando o critério de classificação, a validade discriminante se confirma, diferenciando os grupos a termo e pré-termo. Resultados de outros estudos comprovam a capacidade da escala de diagnosticar problemas motores nos primeiros anos de vida^(15,16,18,22,30).

As comparações realizadas entre as categorizações da EMIA e da EDCC revelaram correlação moderada ($\rho=0,34$; $\kappa=0,30$; $W=0,31$), indicando validade concorrente insatisfatória, similar aos estudos com a AIMS e o *Test of Infant Performance* (correlações entre 0,20 a 0,67)⁽¹⁷⁾ e da AIMS com a *Daily Activities of Infants Scale*⁽³⁰⁾. Campos *et al* observaram concordância moderada da AIMS com a escala Bayley aos cinco meses ($k=0,503$) e fraca aos dez meses ($k=0,209$)⁽⁵⁾. Em contraponto, estudos internacionais^(11,13,15,21) e nacional⁽¹⁴⁾ demonstram alta correlação da AIMS com outras escalas motoras. No presente

estudo, o resultado pode ser decorrente do fato de a EDCC possuir um número menor de itens referentes à motricidade (15) quando comparada a EMIA (58), comprometendo, assim, a confiabilidade dos resultados. Embora seja aplicada com objetivos similares, a EDCC avalia também aspectos sociais e cognitivos, o que torna a subescala motora simplificada em número de itens. Ambos os instrumentos têm vantagens e desvantagens; cabe aos profissionais determinar o melhor para cada necessidade.

Os resultados sugerem que a EMIA possui valor preditivo, uma vez que as classificações foram semelhantes longitudinalmente (G1) e entre pré e pós-testes (G2), similar ao observado no estudo de validação canadense⁽¹²⁾; porém, divergente da avaliação em Taiwan, que detectou validade preditiva limitada dessa escala em crianças pré-termo⁽²¹⁾. O aumento dos percentis ao longo do tempo, nas faixas etária da amostra, ressalta a maior sensibilidade e o rigor da EMIA para avaliar crianças entre três e nove meses, similar a estudo prévio⁽¹⁷⁾.

Pode-se concluir que os procedimentos realizados para a análise das propriedades psicométricas da EMIA

evidenciaram que a tradução invertida mostrou-se eficiente para evitar vieses da compreensão da língua inglesa. A escala foi reconhecida pelos especialistas como um instrumento eficiente para avaliar motricidade de crianças e mostrou-se fidedigna, consistente e com importante poder preditivo e discriminante. Ressalta-se, porém, que os resultados para a validade concorrente foram limitados, indicando ser relevante a comparação da AIMS com outros testes motores. Quanto à validade preditiva, uma frequência maior de acompanhamento poderia alterar os resultados, favorecendo o poder preditivo do instrumento em questão.

Os resultados obtidos repercutem na prática cotidiana, confirmando a confiabilidade, a validade de conteúdo, de critério e de constructo da EMIA, encorajando profissionais a usá-la para avaliar e planejar programas de intervenção. Destaca-se a necessidade de estudos normativos para as crianças brasileiras, já que o uso de categorizações de outras populações para interpretar dados obtidos em crianças do Brasil pode não ser adequado, supondo-se que estes variam frente à diversidade de fatores socioeconômicos e culturais.

Referências bibliográficas

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 Pediatr (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 Pediatr (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. Zajonc 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.