

Análise da validade dos itens do *Movement Assessment of Infants - MAI* - para crianças pré-termo

Item validity analysis of Movement Assessment of Infants - MAI - to preterm infants

Tatiana Teixeira Barral de Lacerda ¹

Lívia de Castro Magalhães ²

^{1,2} Escola de Educação Física, Fisioterapia e Terapia Ocupacional. Universidade Federal de Minas Gerais. Av. Antônio Carlos 6627. Pampulha. Belo Horizonte, MG, Brasil. CEP: 31.270-90.1
E-mail: tatianabarral@terra.com.br

Abstract

Objectives: to assess the validity of the Movement Assessment of Infants (MAI) test, comprised of 65 items, grouped into four sections, used to detect cerebral palsy.

Methods: the sample included 73 preterm babies, assessed between four and eight months of corrected age. Individual assessments were made for the two ages, using Rasch's statistical model.

Results: MAI is an easy scale for the level of ability of the babies assessed, including both age groups, four and eight months old. Some of the items do not fit the one-dimensional expectations of the statistical model, compromising test construction validity.

Conclusions: differences in the difficulty level of the items at four and eight months old reaffirm the need of different performance profiles for each age. Notwithstanding the restrictions, the test has positive aspects and review suggestions are offered aiming at validity improvement.

Key words Validity of tests, Musculoskeletal development, Cerebral palsy

Resumo

Objetivos: examinar a validade de constructo do Movement Assessment of Infants (MAI), teste composto por 65 itens, agrupados em quatro seções, que é usado para detectar paralisia cerebral.

Métodos: a amostra incluiu 73 bebês pré-termo brasileiros, avaliados aos quatro e aos oito meses de idade corrigida. Foram feitas análises individuais para as duas idades, com uso do modelo estatístico Rasch.

Resultados: o MAI é uma escala fácil para o nível de habilidade dos bebês avaliados, tanto aos quatro quanto aos oito meses de idade. Alguns itens não se enquadram nas expectativas de unidimensionalidade do modelo estatístico, o que compromete a validade de constructo do teste.

Conclusões: diferenças no nível de dificuldade dos itens aos quatro e oito meses reaffirmam a necessidade de perfis distintos de desempenho para cada idade. Apesar das limitações, o teste tem aspectos positivos, sendo apresentadas sugestões para revisão, visando melhorar sua validade.

Palavras-chave Validade dos testes, Desenvolvimento musculoesquelético, Paralisia cerebral

Introdução

O desenvolvimento tecnológico nas unidades de Terapia Intensiva Neonatal (UTIN) tem levado à sobrevivência de recém-nascidos com alto risco para transtornos neurológicos. O desenvolvimento motor é um parâmetro importante para a avaliação desses bebês, já que atraso ou desvio desse desenvolvimento pode ser o primeiro indicador de desordens neuromotoras, como a paralisia cerebral (PC).¹ A identificação precoce da PC e do atraso no desenvolvimento é meta importante para os profissionais de reabilitação, uma vez que a intervenção precoce pode minimizar deficiências motoras comumente encontradas naquela população.

Considerando a importância da intervenção precoce, nos últimos anos foram criados vários testes para detectar, o mais cedo possível, sinais de anormalidade no desenvolvimento infantil.²⁻⁶ Dentre esses testes, destacamos o Movement Assessment of Infants (MAI), que é uma avaliação da função neuromotora da criança durante o primeiro ano de vida, desenvolvido especialmente para detecção de disfunção motora em bebês de alto risco.⁴ O MAI foi publicado em 1980 e, desde então, vem sendo amplamente usado por médicos e terapeutas.⁷ A principal meta do MAI é identificar transtornos motores, antes dos 12 meses de idade, e monitorar os efeitos da reabilitação.^{2,8,9} O teste não tem escore normativo, mas foram criados perfis de desempenho para quatro, seis e oito meses,^{1,7} sendo atribuídos pontos de risco quando a resposta da criança é diferente da do padrão esperado para a idade. Segundo uma série de estudos realizada pelos autores do teste,⁴ pontuação acima de sete sugere risco aumentado de distúrbios neuromotores. No entanto, em estudos subsequentes, Schneider *et al.*¹⁰ concluíram que sete pontos é um número muito baixo, e recomendam que o critério de risco deveria ser acima de 10 pontos.

As propriedades psicométricas do MAI, na avaliação de bebês de risco, foram documentadas por vários autores.^{4,8,9} Harris *et al.*⁸ examinaram a correlação inter e intra-examinadores do MAI e encontraram valores modestos ($r = 0,72$ e $r = 0,76$, respectivamente). Haley *et al.*,⁹ por sua vez, avaliaram a confiabilidade dos itens individuais do teste e apontaram que muitos deles apresentam valores questionáveis, demonstrando que os itens que requerem manipulação mínima do bebê apresentam melhor concordância.

Swanson *et al.*⁷ administraram o MAI em 160 bebês prematuros, aos quatro e aos oito meses de idade corrigida, e examinaram sua validade para prever o desenvolvimento neuromotor aos 18

meses de idade, utilizando a Escala Motora do Teste de Bayley. Usando ponto de corte igual a 10 pontos de risco, aos quatro meses de idade corrigida, a sensibilidade foi 83%, especificidade 78%, valor preditivo positivo 59% e valor preditivo negativo 85%. Aos oito meses, a sensibilidade e o valor preditivo negativo aumentaram para 96% e 91%, respectivamente, enquanto a especificidade e o valor preditivo positivo diminuíram para 64% e 52%, respectivamente.

Existe um número considerável de trabalhos sobre o MAI, publicados principalmente na década de 1980, que assinalam seus pontos fortes e fracos, havendo indicações de que o instrumento é útil para a detecção de PC.^{7,10,11} Apesar de ter sido bastante estudado, observa-se que ainda não foi feita uma análise da validade dos itens e de constructo do teste, sendo esse o objetivo deste estudo.

Este estudo dá continuidade a trabalhos anteriores de análise da validade do MAI para crianças pré-termo brasileiras. Em um primeiro estudo, Magalhães *et al.*¹² examinaram o impacto de fatores de risco biológico nos escores do MAI, que foi aplicado aos quatro meses em uma amostragem de 170 bebês pré-termo, sendo examinada a relação entre o número de pontos de risco obtidos no MAI e sete variáveis biológicas (i.e., sexo, idade gestacional, intercorrências neonatais, escores de Apgar, peso ao nascimento, idade e doenças maternas). Os resultados apontaram influência significativa das três primeiras variáveis sobre os escores do MAI, indicando que intercorrências perinatais se correlacionam de maneira importante com o desempenho neuromotor de crianças prematuras aos quatro meses de idade corrigida.

Cardoso *et al.*¹³ examinaram a validade preditiva do MAI para crianças brasileiras. A amostra incluiu 89 crianças nascidas prematuramente (idade gestacional menor ou igual a 32 semanas e peso inferior a 1500 kg), avaliadas pelo MAI aos quatro e aos oito meses de idade corrigida, e diagnóstico neurológico confirmado entre os dois e sete anos. A validade preditiva foi estimada pela utilização dos valores de especificidade, sensibilidade, valor preditivo negativo e positivo. Os valores de especificidade variaram de 0,79 a 0,95 e os de sensibilidade de 0,61 a 1,0, sendo observado que crianças que obtêm mais de 13 pontos de risco no MAI devem ser acompanhadas, em razão do risco aumentado para a paralisia cerebral.

Dando continuidade à análise de validade do MAI, no presente estudo foi adotado o modelo estatístico Rasch, que é específico para a análise de itens, e tem sido usado para examinar as qualidades

psicométricas de escalas usadas em reabilitação.¹⁴⁻¹⁷ Uma das vantagens do modelo Rasch é que ele oferece parâmetros objetivos para avaliar se os itens do teste medem um único constructo, no caso, desenvolvimento motor, ou se apresentam padrões erráticos de escore, o que compromete a validade de constructo do instrumento.¹⁸ Outra vantagem desse modelo estatístico é que ele permite a visualização da relação entre nível de habilidade dos sujeitos e índice de dificuldade dos itens, que é essencial para examinar a adequação do teste à amostra em estudo.¹⁹ Neste trabalho procuramos responder às seguintes perguntas: a) os itens do MAI se alinham para medir um único constructo unidimensional, desenvolvimento motor, definido por itens que apresentam padrões válidos de escore, como indicado pelo ajuste às expectativas do modelo Rasch?; b) a dificuldade do conjunto de itens está de acordo com o nível de habilidade da amostragem, aos quatro e aos oito meses de idade?; c) existem diferenças significativas no nível de dificuldade dos itens, quando o teste é aplicado aos quatro e aos oito meses de idade?

Métodos

Foram incluídas crianças com história de prematuridade, nascidas no Hospital das Clínicas da Universidade Federal de Minas Gerais (UFMG), em Belo Horizonte, Minas Gerais, Brasil. Tais crianças foram selecionadas no Ambulatório da Criança de Risco (ACRIAR) da UFMG, programa de acompanhamento do desenvolvimento de recém-nascidos pré-termo. Foram incluídas todas as crianças com registro completo no banco de dados do ACRIAR que se enquadravam nos seguintes critérios: a) peso ao nascimento menor ou igual a um quilo e meio; b) idade gestacional menor ou igual a 32 semanas e c) avaliação completa do MAI aos quatro e aos oito meses.

O MAI é composto por 65 itens que examinam o desenvolvimento neuromotor em quatro áreas: tônus muscular, reflexos primitivos, reações automáticas e movimentos voluntários.⁴ A área de tônus muscular avalia a consistência da musculatura e sua habilidade para responder à gravidade em várias posições; a área dos reflexos é avaliada de forma similar aos testes padronizados de reflexos, porém com escores qualitativos. A seção de reações automáticas avalia o desenvolvimento de reações de retificação e equilíbrio; por último, são avaliadas as respostas voluntárias à estimulação visual e auditiva, e as habilidades motoras grossas e finas. Cada área é pontuada

com critério de quatro ou de seis pontos, específico para cada item, baseada em critérios comportamentais estabelecidos no "Manual".⁴ A primeira área examinada, tônus muscular, tem escala bidirecional de seis pontos, na qual o valor três representa tônus normal; escores abaixo de três indicam sinais de hipotonia, e acima de três, sinais de hipertonia. Para as outras três áreas, a pontuação varia de um a quatro, e o escore um representa a resposta mais madura.

O teste não tem escore normativo, mas foram criados perfis de risco para as crianças aos quatro, seis e oito meses de idade.^{1,7} Quando o escore de um item difere do escore considerado normal, de acordo com os perfis traçados para cada idade, a criança recebe um ponto de risco. Os pontos de risco de cada área são somados para determinar um escore total. Apesar de haver variações, considera-se que crianças que apresentam mais de 10 pontos de risco têm maior chance de ter PC.⁷

Os dados para este estudo foram obtidos em prontuários e no banco de dados do ACRIAR da UFMG, onde estão registradas as informações clínicas e os resultados de avaliações das crianças em acompanhamento. Inicialmente as crianças foram selecionadas, no banco de dados, de acordo com os critérios de inclusão. Em seguida, todos os dados de prontuários foram conferidos, e foi dada entrada aos escores individuais de cada item do MAI, uma vez que no banco de dados constavam apenas os escores totais de risco. Crianças com prontuários incompletos foram excluídas do estudo. Foram encontradas 89 crianças que se enquadravam no critério; dessas, 16 foram excluídas por não apresentarem dados completos do MAI, resultando em 73 crianças na amostra final.

O teste foi administrado por fisioterapeutas ou terapeutas ocupacionais com grande experiência na avaliação de bebês, cuja confiabilidade entre examinadores é checada periodicamente, de acordo com as instruções contidas no "Manual".⁴ Cada criança foi avaliada por meio de observação e manuseio direto, sempre na presença dos pais ou dos responsáveis. O consentimento por escrito, para a utilização dos dados para a pesquisa, foi dado por todos os pais. O projeto foi aprovado pelo Comitê de Ética em Pesquisa da UFMG.

Neste estudo, usamos o programa estatístico WINSTEPS de análise Rasch²⁰ que gera valores que indicam a relação entre os escores obtidos e as expectativas do modelo estatístico. A análise Rasch parte do princípio de que só é possível derivar medidas verdadeiras de itens que medem o mesmo constructo.¹⁸ Quando os itens medem a mesma

habilidade, podemos ordená-los, de mais fáceis a mais difíceis, em um contínuo linear, análogo a uma régua.

Programas de análise Rasch emitem tabelas com a calibração dos itens, por nível de dificuldade, além do erro de medida e parâmetros associados a cada item, que permitem verificar se o teste se enquadra no princípio de unidimensionalidade.^{19,20} Um desses valores é o quadrado médio, MnSq (*infit* e *outfit*), cujo valor esperado é de 1 ($\pm 0,3$), sendo aceitável um desvio padrão (*t*) de até 2,0. Ou seja, itens com MnSq entre 0,7 e 1,3, associados a um $t \leq 2,0$, são considerados como se enquadrando nas expectativas do modelo. Valores de MnSq acima de 1,3 indicam itens com escore errático, nos quais pessoas com alta habilidade recebem pontuação inesperadamente baixa e pessoas com baixa habilidade recebem boa pontuação em itens difíceis. Esses itens ou estão mal formulados ou não se enquadram no mesmo constructo dos outros itens do teste. Valores baixos de MnSq apontam itens cujos padrão de escore é muito consistente, com pouca variação.¹⁵ As duas situações comprometem o teste, mas considera-se que o padrão errático de escores tem mais impacto na validade, já que escores imprevisíveis oferecem informações pouco claras sobre o nível de desempenho do indivíduo.¹⁵ Neste estudo, de acordo com outros trabalhos na área,²¹ consideramos que apenas os itens com $MnSq \geq 1,3$ e $t > 2,0$ contrariam as expectativas do modelo Rasch.

Considera-se aceitável que, no máximo, 5% dos itens de um teste não se enquadrem nos parâmetros do modelo Rasch. Como no presente estudo foram avaliados 53 itens, aos quatro meses, e 66 itens, aos oito meses, espera-se que não mais que dois itens e três itens, nas idades respectivas, apresentem valores fora do esperado. Deve-se observar que a diferença no número de itens, nas duas idades, deve-se ao fato de que alguns deles mais avançados, como sentar e engatinhar, não são aplicados aos quatro meses de idade. Ressaltamos, ainda, que incluímos na análise mais itens do que os 65 originais do MAI. Isso ocorreu porque, como a escala de pontuação da primeira seção (tônus muscular) do MAI é bidirecional (i.e., de um a seis pontos, em que o valor 3 representa tônus normal, escores abaixo de 3 indicam sinais de hipotonia, e acima de 3, sinais de hipertonia), ela foi transformada em duas: uma que varia de normal (escore 1) à hipotonia (escore 3), e outra que varia de normal (escore 1) à hipertonia (escore 3).

Para verificar a estabilidade dos itens, aos quatro e aos oito meses, foi comparada a calibração dos mesmos nas duas idades usando método gráfico, recomendado por Wright e Masters.²² Para tanto, foi

feito um gráfico de dispersão, da calibração dos itens aos quatro *versus* oito meses, e traçada uma linha de controle de qualidade, usando os erros-padrão associados à calibração de cada item. Itens localizados dentro dos limites das linhas de controle de qualidade não apresentam diferenças significativas de calibração.

Resultados

A amostra foi constituída de 73 crianças, das quais 38 eram do sexo feminino e 35 do sexo masculino (Tabela 1). Os fatores de risco gestacional mais comuns nas crianças da amostra foram: hipertensão materna (37,8%), uso de medicamentos durante a gestação (22,9), pré-eclâmpsia (20,2%), eclâmpsia (12,1%). Apenas 16,2% das gestações não apresentaram fatores de risco. As intercorrências neonatais mais frequentes incluíram: icterícia (89,4%), septicemia (63,5%), uso de ventilação mecânica e adaptação/dificuldade respiratória (45,9%) e outras (35,1%).

Tabela 1

Caracterização da amostragem quanto à idade gestacional, peso ao nascimento e sexo.

Idade gestacional (semanas)	Peso médio ao nascimento (gramas)	Sexo	
		Feminino	Masculino
≤ 28	958	11	7
29 e 30	1080	16	13
31 e 32	1283	11	15
Total: 29,75 ($\pm 1,58$)	1124,88 ($\pm 233,49$)	38	35

Os valores globais de MnSq, *Infit* e *Outfit*, respectivamente, aos quatro (MnSq=1,01; $t=-0,1$ /MnSq=0,95; $t=-0,3$) e aos oito meses (MnSq=1,01; $t=-0,1$ /MnSq=0,88; $t=-0,3$), estavam dentro dos parâmetros esperados, o que indica que, de maneira geral, os itens do MAI se enquadram nas expectativas do modelo Rasch. A consistência interna do teste também atingiu valores dentro do esperado, de - 0,94 aos quatro meses e 0,92 aos oito meses. Além disso, foi calculado o índice de separação, que indica que os itens do MAI dividem as crianças em aproximadamente três níveis diferentes de habilidade motora, aos quatro meses e, em apro-

Tabela 2

Calibração dos itens aos quatro meses, em ordem crescente de dificuldade, e estatísticas de enquadramento.

Itens	Escore bruto	Medida	Erro	Infit		Outfit	
				MnSq	t	MnSq	t
Respostas Auditivas (V)	73	5,63	1,82	Calibração Máxima Estimada			
Reflexo labirintico tônico (RTL) prono (R)	73	5,63	1,82	Calibração Máxima Estimada			
Reflexo labirintico tônico (RTL) supino (R)	79	2,54	0,42	0,95	-0,1	0,50	-1,0
Clonus (R)	83	1,98	0,34	1,37	1,1	2,13	1,9
Suspensão em prono (T hipo)	60	1,80	0,71	0,85	-0,2	0,44	-0,7
Acompanhamento visual (V)	87	1,58	0,29	1,09	0,3	1,17	0,4
Sumário de tônus de tronco (T hipo)	54	1,57	0,71	0,95	-0,1	0,63	-0,5
Assimetria de tônus (T)	88	1,41	0,28	1,46	1,7	1,59	1,5
Assimetria de movimentos voluntários (A)	86	1,38	0,28	1,57	2,0	1,16	0,4
Centralização da cabeça (V)	93	1,14	0,25	1,22	0,9	1,07	0,2
Reflexo de marcha automática (R)	93	1,14	0,25	1,40	1,6	1,22	0,7
Sumário de reflexos (R)	92	1,13	0,25	0,82	-0,9	0,52	-2,0
Postura em prono (T hiper)	56	1,08	0,52	0,68	-0,7	0,50	-0,9
Assimetria de movimentos automáticos (V)	93	1,01	0,24	1,46	1,9	1,11	0,4
Tremor (R)	97	0,91	0,23	1,72	2,9	1,95	2,8
Postura em supino (T hiper)	64	0,74	0,43	0,62	-1,1	0,35	-1,6
Sumário de reações automáticas (A)	99	0,67	0,22	0,82	-1,0	0,62	-1,8
Extensibilidade (T hipo)	64	0,63	0,42	1,15	0,4	1,24	0,4
Sumário movimentos voluntários (V)	100	0,56	0,22	0,65	-2,1	0,52	-2,5
Visão periférica (V)	101	0,55	0,22	1,88	3,8	1,74	2,6
R TCA evocado (R)	103	0,55	0,21	1,08	0,4	1,02	0,1
Variação na distribuição do tônus (T)	101	0,55	0,22	1,36	1,7	1,13	0,5
Assimetria de reflexos (R)	104	0,50	0,21	0,90	-0,6	0,79	-1,0
Reação positiva de suporte (R)	107	0,42	0,21	1,44	2,2	1,31	1,3
Consistência (T hiper)	67	0,34	0,37	0,69	-1,0	0,46	-1,5
Reflexo de preensão palmar (R)	109	0,34	0,20	1,03	0,2	0,87	-0,6
Endireita, com extensão cabeça (A)	116	0,06	0,19	0,89	-0,7	0,79	-1,2
Sumário tônus extremidades (T hiper)	73	0,02	0,31	0,88	-0,4	0,67	-1,0
Consistência (T hiper)	74	0,02	0,31	0,62	-1,6	0,50	-1,6
RTCA espontâneo (R)	119	-0,05	0,19	1,05	0,3	0,93	-0,4
Sumário de tônus de tronco (T hiper)	85	-0,14	0,27	0,79	-1,0	1,24	0,7
Extensibilidade (T hiper)	85	-0,15	0,27	0,78	-1,0	0,70	-1,0
Postura em supino (T hipo)	83	-0,19	0,27	0,80	-0,9	0,92	-0,2
Suspensão em prono (T hiper)	90	-0,20	0,26	1,07	0,3	1,03	0,1
Mãos abertas (V)	126	-0,29	0,18	1,17	1,1	1,10	0,6
Reflexo de Moro (R)	113	-0,33	0,19	0,61	-2,8	0,67	-2,2
Sumário tônus extremidades (T hipo)	78	-0,34	0,28	1,27	1,0	1,49	1,3
Colocação dos pés (A)	127	-0,39	0,18	1,07	0,4	1,07	0,4
Posição cabeça anterior/posterior (V)	135	-0,57	0,18	0,54	-3,8	0,55	-3,5
Reflexo de Galant (R)	133	-0,63	0,18	1,26	1,6	1,25	1,5
Postura em prono (T hiper)	94	-0,65	0,23	0,92	-0,4	0,87	-0,5
Suporta peso nos ombros (V)	136	-0,66	0,18	1,02	0,1	0,98	-0,1
Mãos na linha média (V)	140	-0,73	0,17	1,24	1,6	1,24	1,5
Equilíbrio de cabeça (V)	141	-0,76	0,17	0,75	-2,0	0,73	-2,0
Endireita, com flexão lateral cabeça (A)	143	-0,82	0,17	0,66	-2,8	0,69	-2,3
Endireita, com flexão cabeça (A)	144	-1,00	0,17	0,81	-1,4	0,88	-0,8
Colocação das mãos (A)	155	-1,25	0,17	1,14	1,0	1,22	1,4
Uso ativo dos quadris (V)	181	-1,95	0,18	1,01	0,1	1,02	0,1
Reflexo de Landau (A)	182	-1,98	0,18	0,85	-1,0	0,91	-0,6
Rotacão de tronco (A)	190	-2,24	0,18	0,79	-1,4	0,79	-1,3
Reflexo de Preensão plantar (R)	208	-2,87	0,19	0,91	-0,5	0,96	-0,2
Vocalização (V)	214	-3,10	0,20	0,95	-0,3	0,96	-0,2
Endireitamento de tronco (V)	220	-3,35	0,20	0,50	-3,4	0,53	-3,2

Os itens em negrito não se enquadram nos parâmetros do modelo Rasch. As letras entre parênteses indicam a qual seção do teste o item pertence. (T= tônus, R= reflexos primitivo, A= reações automáticas, V= movimentos voluntários). Observar que quanto maior a pontuação/escore bruto, pior o desempenho no item. MnSq = Mean Square (Quadrado -médio); t = Desvio padrão.

Tabela 3

Calibração dos itens aos oito meses, em ordem crescente de dificuldade, e estatísticas de enquadramento.

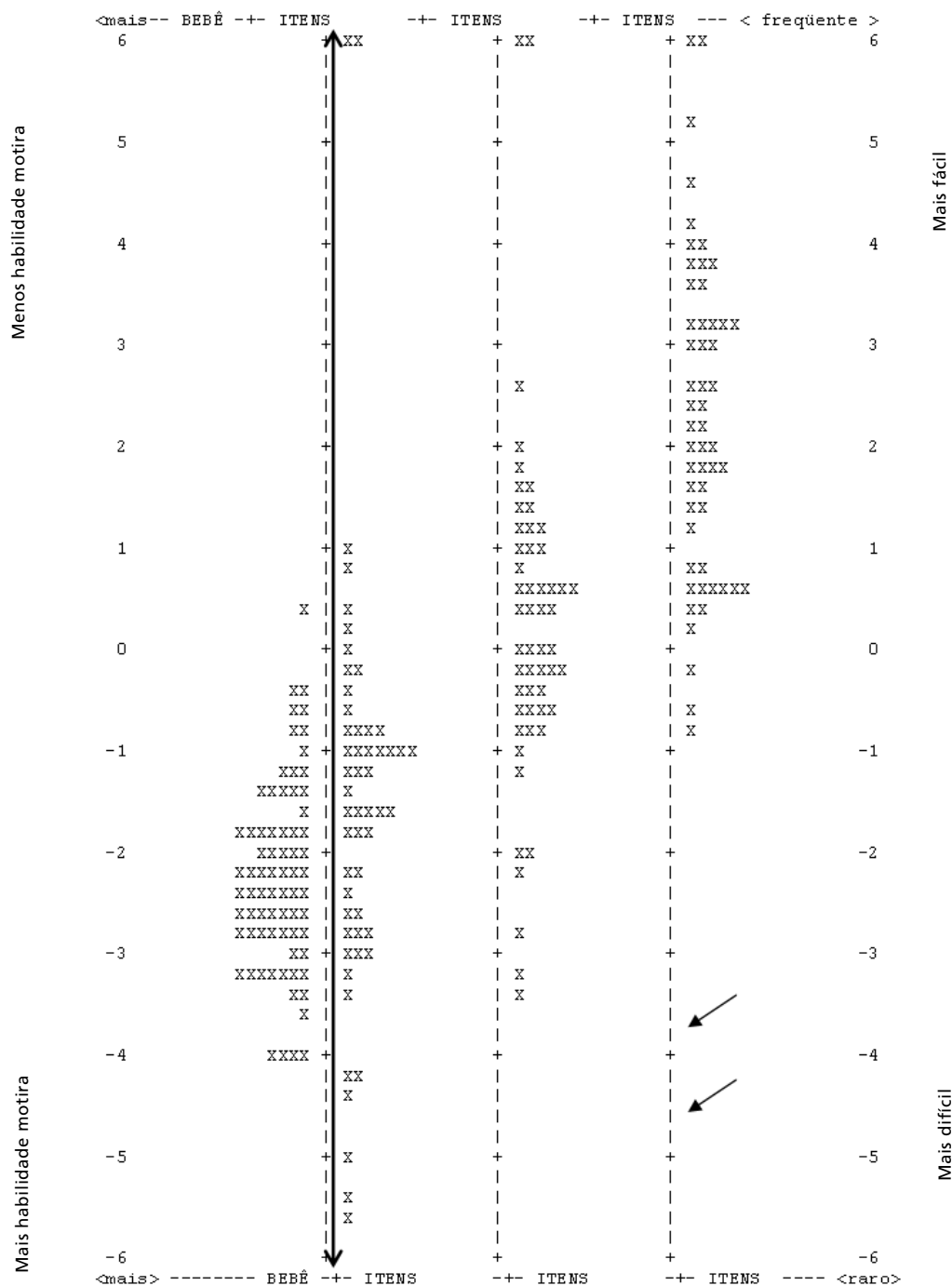
Itens	Escore bruto	Medida	Erro	Infit		Outfit	
				MnSq	t	MnSq	t
Reflexo labirintico tônico (RTL) prono (R)	72	4,64	1,82			Calibração Máxima Estimada	
Reflexo de Preensão palmar (R)	73	2,62	0,72	1,20	0,3	0,09	-1,1
Reflexo labirintico tônico (RTL) supino (R)	76	2,34	0,59	1,01	0,0	0,17	-1,1
Centralização cabeça (V)	77	2,03	0,52	0,71	-0,6	0,13	-1,4
RTCA evocado (R)	74	1,96	0,52	0,96	-0,1	0,69	-0,3
Reflexo de Marcha automática (R)	78	1,78	0,47	0,90	-0,2	0,64	-0,4
Visão periférica (V)	76	1,75	0,47	1,24	0,5	0,40	-0,8
Clônus (R)	78	1,58	0,43	1,00	0,0	1,29	0,3
Acompanhamento visual (V)	79	1,58	0,43	1,40	0,8	1,32	0,3
Sumário de tónus de tronco (T hipo)	58	1,38	1,00	0,86	-0,1	0,20	-0,8
Mãos abertas (V)	81	1,25	0,38	0,88	-0,3	0,26	-1,5
Postura em supino (T hiper)	70	1,21	0,60	0,43	-1,3	0,13	-1,3
Suspensão em prono (T hipo)	63	1,15	0,71	0,66	-0,5	0,75	-0,2
Postura em prono (T hiper)	64	1,14	0,61	0,50	-1,1	0,74	-0,2
Respostas auditivas (V)	82	1,11	0,36	1,87	1,9	1,58	0,7
Tremor (R)	82	1,11	0,36	0,67	-1,0	0,43	-1,1
Assimetria de tónus (T)	81	1,05	0,37	1,53	1,2	1,06	0,1
Assimetria de movimentos voluntários (V)	80	1,02	0,37	2,11	2,3	2,29	1,4
Sumário de reflexos (R)	83	0,98	0,35	0,80	-0,6	0,61	-0,7
Equilíbrio de cabeça (A)	84	0,86	0,33	0,76	-0,8	0,60	-0,8
Posição cabeça anterior/posterior (V)	84	0,86	0,33	0,75	-0,8	0,63	-0,7
RTCA espontâneo (R)	84	0,86	0,33	1,11	0,3	1,04	0,1
Alcançar (V)	85	0,76	0,32	1,21	0,6	0,78	-0,4
Mãos na linha média (V)	85	0,76	0,32	0,76	-0,8	0,39	-1,6
Assimetria de movimentos automáticos (A)	84	0,70	0,32	1,42	1,1	0,96	-0,1
Reflexo de Moro (R)	80	0,62	0,33	0,75	-0,8	0,91	-0,2
Reflexo de Galant (R)	84	0,60	0,31	1,28	0,8	2,09	1,6
Endireita, com extensão cabeça (A)	87	0,56	0,30	0,62	-1,4	0,44	-1,4
Preensão grossa (V)	89	0,39	0,28	0,96	-0,1	0,54	-1,2
Assimetria de reflexos (R)	89	0,39	0,28	0,98	-0,1	1,00	0,0
Colocação dos pés (A)	85	0,31	0,29	1,14	0,4	1,07	0,1
Variação na distribuição do tónus (T)	88	0,24	0,28	1,30	1,0	0,92	-0,2
Colocação das mãos (A)	87	0,15	0,27	1,34	1,1	1,03	0,1
Postura em supino (T hipo)	78	0,06	0,37	1,02	0,1	0,50	-1,0
Consistência (T hiper)	67	0,01	0,41	0,64	-1,0	0,42	-1,0
Sumário de reações automáticas (A)	92	-0,01	0,25	0,65	-1,6	0,44	-1,8
Endireita, com flexão lateral cabeça (A)	93	-0,03	0,25	0,77	-1,0	0,84	-0,4
Extensibilidade (T hiper)	73	-0,14	0,37	0,75	-0,7	0,67	-0,6
Suporte peso nos ombro (V)	96	-0,17	0,24	0,63	-1,8	0,46	-1,8
Postura em prono (T hipo)	77	-0,21	0,33	1,01	0,0	0,51	-0,9
Suspensão em prono (T hiper)	82	-0,25	0,32	0,94	-0,2	0,84	-0,3
Transferências (V)	98	-0,27	0,23	2,31	4,1	3,79	4,7
Extensibilidade (T.hipo)	72	-0,33	0,34	1,01	0,0	0,66	-0,7
Endireita, com flexão cabeça (A)	99	-0,34	0,23	0,82	-0,8	1,69	1,7
Sumário de movimentos voluntários (V)	98	-0,37	0,23	0,42	-3,2	0,34	-2,6
Sumário de tónus extremidades (T hiper)	73	-0,39	0,36	0,73	-0,8	0,61	-0,8
Rotação de tronco (A)	102	-0,48	0,22	0,91	-0,4	0,72	-0,9
Reação positiva de suporte (R)	104	-0,51	0,22	1,81	3,0	1,58	1,6
Sumário de tónus de tronco (T hiper)	91	-0,70	0,27	0,92	-3,0	0,86	-0,3
Sentar quando colocado (V)	108	-0,74	0,20	1,31	1,4	0,80	-0,7
Extensão protetora à frente (A)	104	-0,87	0,20	0,99	0,0	0,73	-1,1
Uso ativo quadris (V)	112	-0,89	0,20	1,02	0,1	0,82	-0,7
Sumário de tónus extremidades (T hipo)	78	-0,92	0,28	1,22	0,8	1,89	1,6
Consistência (T hipo)	82	-0,93	0,27	0,88	-0,5	0,59	-1,1
Rolar (V)	117	-1,14	0,19	1,77	3,4	1,60	2,1
Endireitamento de tronco (V)	122	-1,25	0,18	0,87	-0,7	0,90	-0,5
Extensão protetora para o lado (A)	121	-1,28	0,18	1,14	0,7	0,91	-0,4
Equilíbrio em suspensão ventral (A)	139	-1,79	0,16	0,38	-5,2	0,44	-3,7
Reação de Landau (A)	139	-1,79	0,16	0,98	-0,1	1,16	0,8
Reação de Pára-queda (A)	144	-1,95	0,16	0,95	-0,3	1,07	0,4
Equilíbrio sentado (V)	128	-2,00	0,17	0,39	-4,8	0,47	-3,4
Vocalização (V)	163	-2,43	0,16	1,20	1,3	1,36	1,9
Progressão em prono (V)	177	-2,98	0,16	1,23	1,4	1,12	0,7
Preensão fina (V)	193	-3,07	0,16	0,71	-2,1	0,81	-1,2
Reflexo de preensão plantar (R)	207	-3,36	0,16	0,97	-0,2	1,50	2,4
Passar para sentado (V)	209	-3,57	0,16	1,34	1,9	1,17	0,9

Os itens em negrito não se enquadram nos parâmetros do modelo Rasch. As letras entre parênteses indicam a qual seção do teste o item pertence (T = tónus, R= reflexos primitivo, A= reações automáticas, V = movimentos voluntários). Observar que quanto maior a pontuação/escore bruto, pior o desempenho no item. MnSq = mean square (quadrado médio; t = desvio padrão).

Figura 1

Relação entre o nível de habilidade dos bebês e o nível de dificuldade dos itens para o *movement assessment of infants* (MAI) aos quatro meses.

Mapa de sujeitos e itens aos quatro meses.

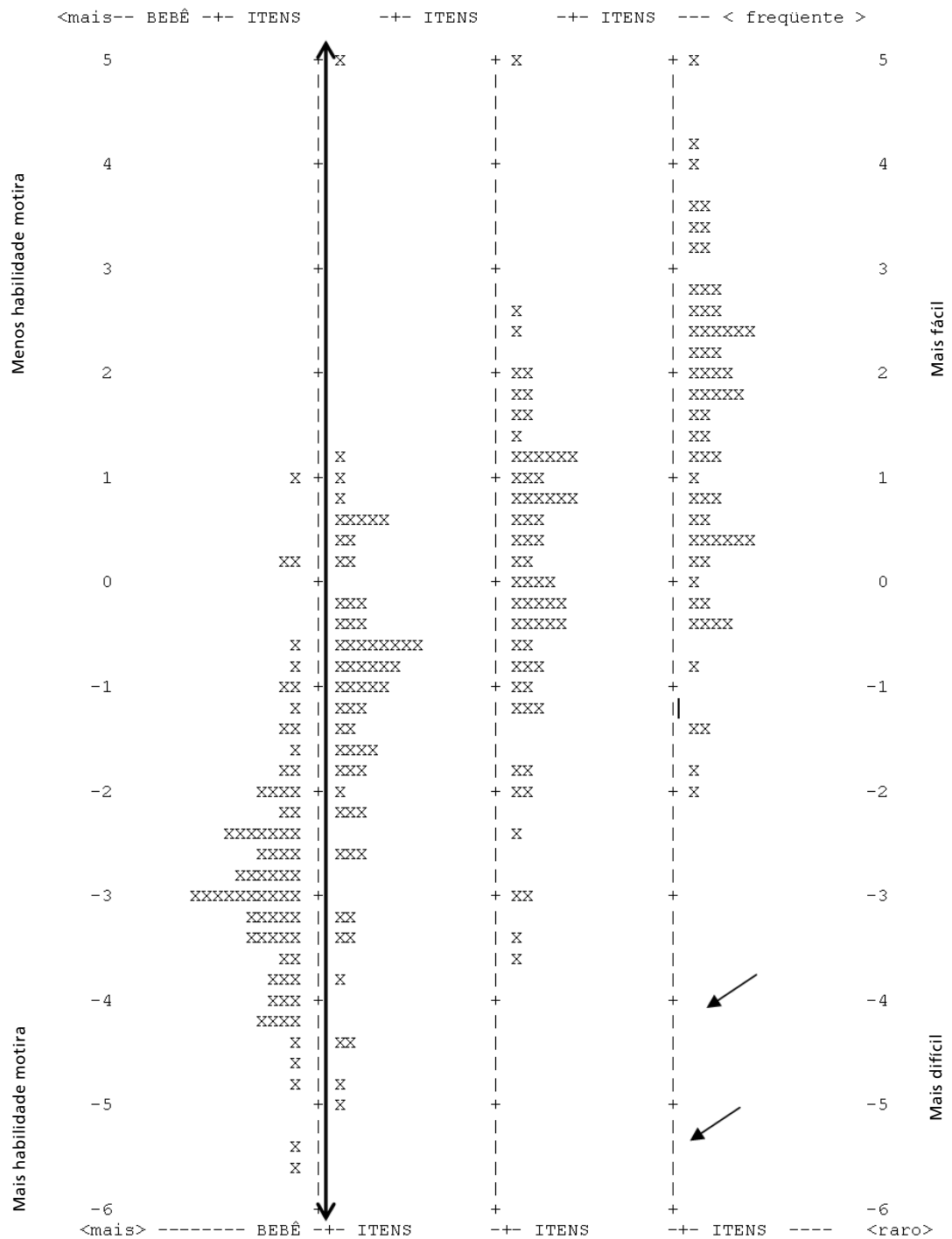


O contínuo de habilidade motora é ilustrado pela linha contínua vertical, as medidas de habilidade motora das crianças são representadas pelos "X", à esquerda da linha, e a calibração dos itens por nível de dificuldade pelos "X", à direita da linha. As linhas pontilhadas verticais indicam incrementos de dificuldade relacionados à pontuação dos itens. As setas indicam lacunas na área de cobertura dos itens.

Figura 2

Relação entre o nível de habilidade dos bebês e o nível de dificuldade dos itens para o *movement assessment of infants* (MAI) aos oito meses.

Mapa de sujeitos e itens aos oito meses.



O contínuo de habilidade motora é ilustrado pela linha contínua vertical, as medidas de habilidade motora das crianças são representadas pelos "X", à esquerda da linha, e a calibração dos itens por nível de dificuldade pelos "+", à direita da linha. As linhas pontilhadas verticais indicam incrementos de dificuldade relacionados à pontuação dos itens. As setas indicam lacunas na área de cobertura dos itens.

ximadamente 3,5 níveis, aos oito meses.

A calibração da dificuldade dos itens, aos quatro e aos oito meses, é apresentada nas Tabelas 2 e 3. Examinando a calibração dos itens, observa-se que ela segue a progressão esperada do desenvolvimento, mas o número de itens que não se encaixam no modelo Rasch ($infit/outfit \geq 1,3$; $t \geq 2$ pontos) ultrapassa o limite de 5%. Esses itens estão assinalados, em negrito, nas Tabelas 2 e 3.

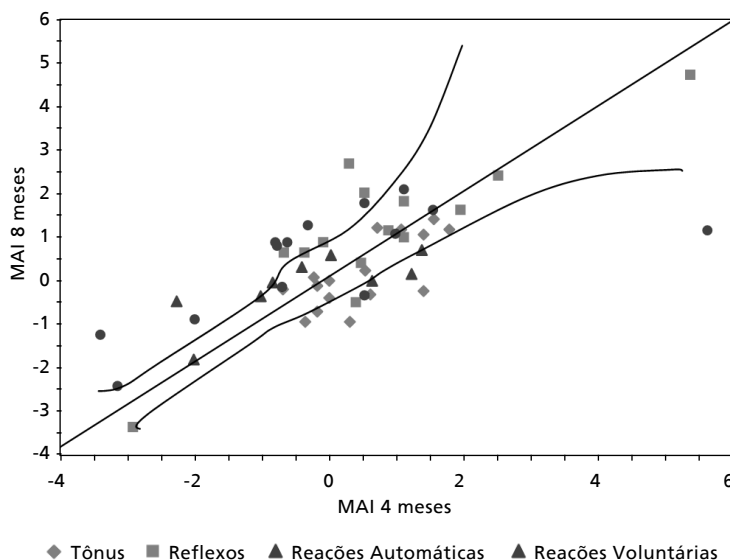
Aos quatro meses, como indicado pela Tabela 2, os itens mais fáceis foram "Respostas Auditivas" e o "Reflexo Tônico Labiríntico" em prono (RTL em prono), nos quais todos os bebês obtiveram escore total, e o "Reflexo Tônico Labiríntico em supino (RTL em supino)". Os itens mais difíceis foram referentes à "Vocalização e ao Endireitamento de Tronco". Um grande número de bebês obteve escores entre 1 e 2, indicando predomínio de crianças com pontuação dentro da faixa de normalidade. Como pode ser observado na Figura 1, que representa a relação entre a dificuldade dos itens e o nível de habilidade, todos os bebês estão localizados na metade inferior da figura, o que indica que o teste foi fácil e muitos bebês apresentaram habilidades que não puderam ser diferenciadas pela dificuldade dos itens. Além disso, muitos itens se agrupam para medir o mesmo nível de habilidade e, como indicado pelas setas, há algumas lacunas na distribuição dos itens, indicando níveis nos quais há poucos itens para avaliar ou diferenciar as habilidades das crianças.

Aos oito meses (Tabela 3), o "Reflexo Tônico Labiríntico em prono (RTL em prono)" continua sendo o item de menor dificuldade, ao lado de outros como o "Reflexo de Preensão Palmar", o "RTL em supino, Centralização da Cabeça e Reflexo Tônico Cervical Assimétrico evocado (RTCA evocado)". Os itens de maior dificuldade foram: "Reflexo de preensão plantar e o Passar para sentado". Na Figura 2, observa-se melhor distribuição na calibração dos itens e das medidas dos bebês, em relação aos quatro meses. Entretanto, ainda existem itens muito fáceis para o nível de habilidade dos bebês; alguns deles se acumulam no mesmo nível de habilidade, e persistem lacunas na calibração. Observa-se, também, maior concentração de bebês com nível de habilidade acima do nível de dificuldade dos itens, ou seja, a habilidade de alguns bebês de oito meses é maior do que a que pode ser medida pelo teste.

A Figura 3 compara a calibração dos itens aos quatro e aos oito meses. Quando a dificuldade dos itens é estável, a relação entre as duas calibrações é estatisticamente linear. Nota-se, no entanto, que alguns itens se encontram fora das linhas de con-

Figura 3

Comparação da calibração dos itens do *Movement Assessment of Infants (MAI)* aos quatro e oito meses.



trole, indicando que as escalas apresentam 23 itens com níveis diferenciados de dificuldade, nas duas idades. Itens acima da linha diagonal foram mais fáceis aos oito meses, e itens abaixo da linha, aos quatro meses.

Discussão

Apesar de os resultados globais indicarem adequação do MAI ao modelo Rasch, uma análise mais minuciosa dos itens aponta alguns problemas. Vários itens se mostraram muito fáceis, tanto para bebês de quatro quanto de oito meses, embora a dificuldade de alguns itens varie de uma idade para a outra. Quando os itens são muito fáceis, não é possível medir a progressão da habilidade dos bebês, o que compromete o uso do MAI para medir resultados de intervenção terapêutica. Tais itens podem ser desnecessários para avaliar bebês nessas faixas etárias e poderiam ser eliminados, numa tentativa de simplificar o teste. Por outro lado, a presença de itens fáceis indica que o teste poderia ser aplicado em bebês com idade inferior a quatro meses. Pesquisas adicionais poderiam comprovar essa hipótese, sendo então necessário traçar o perfil de desempenho para a idade. Observou-se, também, que aos oito meses, parte significativa dos bebês apresentou habilidades superiores às medidas pelo teste, indicando que esse teste não é sensível para avaliar

crianças aos oito meses de idade corrigida. Isso já era esperado e se justifica pelo fato de que o MAI avalia especialmente reflexos e reações, com pouca ênfase na movimentação voluntária. À medida que o bebê se torna mais maduro, a atividade motora voluntária aumenta, mas o MAI tem poucos itens para medir essas habilidades mais avançadas.

Nota-se, ainda, que há algumas lacunas no contínuo de calibração dos itens que poderiam ser preenchidas pela adição de novos itens, o que melhoraria a precisão do teste. Além de preencher as lacunas, tal adição poderia também ser feita na extremidade inferior do mapa de calibração dos itens. A inclusão de itens mais difíceis melhoraria o poder de discriminação da escala e tornaria o instrumento mais útil para medir o desempenho de bebês acima de oito meses ou que apresentam maior habilidade motora.

Com relação à análise dos itens individuais, cerca de 7,5% deles não se enquadraram no modelo Rasch, tanto aos quatro quanto aos oito meses, o que ultrapassa os 5% aceitáveis para se confirmar a validade de constructo do teste. Aos quatro meses, quatro itens apresentaram escores erráticos (assimetria, tremor, visão periférica e suporte positivo) e aos oito meses, cinco itens (assimetria, transferência de objetos, reação positiva de suporte, rolar e preensão plantar).

Em um estudo sobre os perfis de desenvolvimento do MAI, Schneider *et al.*¹⁰ compararam o perfil de risco de bebês normais aos quatro meses com o perfil definido pelos autores do teste. Eles encontraram vários itens classificados como de risco, diferentes do perfil original. Dentre eles, dois itens - assimetria e visão periférica - coincidem com os presentes resultados, indicando que esses itens vêm apresentando problema em outros estudos.¹⁰ Os autores acreditam que a assimetria possa ser um componente normal do desenvolvimento aos quatro meses, e assinalam que esse item tem história de pobre confiabilidade, como demonstrado por Haley *et al.*⁹. Schneider *et al.*¹⁰ sugerem a exclusão de alguns itens e a adição de outros, além da criação de outro perfil de risco, para melhorar a confiabilidade e a validade do MAI. Os autores recomputaram os escores de risco da amostra, após as modificações sugeridas, e observaram especificidade de 100%, usando ponto de corte ≥ 10 pontos de risco, o que indica que a versão modificada resulta em definição mais precisa do desempenho motor de bebês aos quatro meses de idade.¹⁰

Harris²³ analisou, retrospectivamente, quais itens específicos do MAI foram discriminativos de PC em uma amostra de 229 bebês de alto risco. Esse

trabalho mostrou que 15 itens foram classificados como não-significativos; dentre eles estão "tremor", "reação positiva de suporte", "visão periférica e assimetria" (de reflexos, reações automáticas e movimentos voluntários). Esses itens também aparecem como sujeitos a escore errático em nosso estudo. Harris²³ indica que os resultados de seu estudo poderiam ser usados para criar uma versão mais curta do MAI, o que melhoraria sua validade. No presente estudo os dados corroboram essa conclusão em uma população culturalmente diferente.

A análise da estabilidade dos itens aos quatro e aos oito meses ressalta a importância dos perfis de desempenho específicos para cada idade. Como esperado, não se observam grandes diferenças nos itens relacionados ao tônus postural, uma vez que essa é uma característica relativamente estável. No entanto, os itens de reflexos ficam gradativamente mais fáceis, assim como as respostas voluntárias. Os resultados ilustram essas diferenças e mostram que os itens de reflexos, reações automáticas e movimentação voluntária foram relativamente mais fáceis aos oito meses que aos quatro meses. Nota-se, ainda, que o erro de medida aumenta substancialmente à medida que os itens ficam mais fáceis, ou seja, quanto mais fácil o item, menor sua contribuição para estimar o nível de habilidade da criança, em razão da grande margem de erro associado à medida.

Finalmente, considerando que a amostragem estudada foi de bebês pré-termo, geralmente mais debilitados, causa surpresa que os itens do MAI não sejam mais desafiantes para as suas habilidades motoras. Isso reforça a necessidade de reformulação do teste, caso ele seja usado para monitorizar progresso em terapia. Nesta pesquisa, os dados, apesar de colocarem em questão a validade de constructo do MAI, não invalidam, necessariamente, sua eficiência na detecção de PC. É verdade que alguns itens que contribuem para os escores de risco estão sujeitos a escore errático. No entanto, o que mais chama a atenção na análise é que o MAI parece ter mais itens do que o necessário. Portanto, usando os resultados deste estudo, bem como outros trabalhos publicados,^{10,23} seria possível tornar o MAI mais curto e menos desgastante para as crianças avaliadas. Os itens que descrevem os reflexos primitivos parecem ser úteis apenas aos quatro meses, uma vez que aos oito meses eles não são mais discriminativos das habilidades do bebê. Todavia, aos oito meses, há necessidade de maior número de itens para avaliação da movimentação voluntária do bebê.

Conclusões

Este estudo apresentou evidências importantes sobre o Movement Assessment of Infants. Embora os resultados gerais indiquem adequação do teste às expectativas do modelo Rasch, uma análise mais detalhada revela que um número estatisticamente significativo de itens não se encaixa no modelo, o que compromete a validade de constructo do teste. Muitos desses itens já vinham sendo apontados como problemáticos por outros autores, que dão sugestões para se amenizar essas falhas. Com base nos resultados da análise e em outros estudos sobre o MAI, os itens erráticos devem ser revisados ou mesmo eliminados, por não apresentarem informações consistentes a respeito do desenvolvimento motor nas idades corrigidas de quatro e oito meses.

Apesar das limitações, o MAI tem aspectos posi-

tivos. A calibração dos itens, nas duas idades, segue a progressão esperada do desenvolvimento e, de certa forma, observa-se o padrão previsto na literatura, com itens que avaliam reflexos posturais apresentando mais estabilidade, e itens de respostas voluntárias mostrando maior variabilidade aos quatro e aos oito meses. Índices diferentes de dificuldade de itens para as idades de quatro e oito meses corroboram a necessidade de perfis diferenciados por idade, como já é feito para o MAI. O grande número de itens fáceis para a amostragem indica a possibilidade de aplicação do teste em bebês ainda mais jovens, sendo também possível incluir itens mais difíceis para crianças mais velhas. Os resultados aqui apresentados podem ser úteis para guiar a análise de outros testes infantis e para pesquisadores interessados em desenvolver instrumentação específica para crianças brasileiras.

Referências

1. Washington K, Deitz JC. Performance of full-term 6-month-old infants on the movement assessment of infants. *Pediatr Phys Ther.* 1995; 7: 65-74.
2. Dubowitz L, Dubowitz V. The neurological assessment of the full-term newborn infant. London: Spastics International Medicinal Publication; 1891.
3. Galscoe FP, Byrne KE, Ashford LG, Johnson KL, Chang B, Strickland B. Accuracy of the Denver-II in developmental screening. *Pediatrics.* 1992; 89: 1221-4.
4. Chandler LS, Andrews MS, Swanson MW. Movement assessment of infants: a manual. Rolling Bay (WA): 1980.
5. 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.
6. Campbell SK, Kolobe THA, Osten ET, Girolani GL, Lemke M. Construct validity of the test of Infant Movement Performance. *Phys Ther.* 1995; 75: 585-96.
7. Swanson MW, Bennett FC, Shy KK, Whitfield MF. Identification of neurodevelopmental abnormality at four and eight months by the movement assessment of infants. *Dev Med Child Neurol.* 1992; 34: 321-37.
8. Harris SR, Haley SM, Tada WL, Swanson MW. Reliability of observational measures of the movement assessment of infants. *Phys Ther.* 1984; 64: 471-6.
9. Haley SM, Harris SR, Tada WL, Swanson MW. Item reliability of the movement assessment of infants. *Phys Occup Ther Pediatr.* 1986; 6: 21-39.
10. Schneider J, Lee W, Chasnoff IJ. Field testing of the movement assessment of infants. *Phys Ther.* 1988; 68: 321-7.
11. Harris SR. Early detection of cerebral palsy: sensitivity and specificity of the motor assessment tools (Abstract). *J Perinatol.* 1987; 7: 11-5.
12. Magalhães LC, Amorim FP, Paixão ML, Mancini MC, Barbosa VM. Influência de fatores de risco biológico nos escores de um teste para detecção de paralisia cerebral em crianças pré-termo. *Temas Desenvol.* 2001; 10: 5-12.
13. Cardoso AM, Magalhães LC, Paixão ML, Mancini MC, Amorim RHC, Rossi LDF. Validade preditiva do MAI para crianças pré-termo brasileiras. *Arq Neuro-Psiquiatr.* 2004; 62:1052-7.
14. Ficher AG. The assessment of IADL motor skills: an application of many faceted Rasch analysis. *Am J Occup Ther.* 1993; 47: 319-29.
15. Mancini MC, Fiuza PM, Rebelo JM, Magalhães LC, Coelho ZC, Paixão ML, Gontijo AP, Fonseca ST. Comparação do desenvolvimento de atividades funcionais em crianças com desenvolvimento normal e crianças com paralisia cerebral. *Arq Neuro-Psiquiatr.* 2002; 60: 446-52.
16. Velozo CA, Magalhães LC, Pan AW, Leiter P. Functional scale discrimination at admission and discharge: Rasch analysis of the level of rehabilitation scale-III. *Arch Phys Med Rehabil.* 1995; 76: 705-12.
17. Haley SM, Ludlow LH, Coster WJ. Pediatric evaluation of disability inventory. *Phys Med Rehabil Clin North Am.* 1993; 4: 529-40.
18. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil.* 1989; 70: 857-60.
19. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human science. New Jersey: Lawrence Erlbaum; 2001.
20. Linacre MJ. WINSTEPS: Rasch-model [computer program]. Chicago: MESA Press; 2000.
21. Lai JS, Fischer AG, Magalhães LC, Bundy AC. Construct validity of the sensory integration and praxis tests. *Occup Ther J Res.* 1996; 16: 75-97.

22. Wright BD, Masters GN. Rating scale analysis. Chicago: MESA Press; 1982.

23. Harris SR. Early neuromotor predictors of cerebral palsy in low-birthweight infants. Dev Med Child Neurol. 1987; 29: 508-19.

Recebido em 10 de outubro de 2005
Versão final apresentada em 27 de julho de 2006
Aprovado em 8 de agosto de 2006