

Análise de Itens do Desenho da Figura Humana: Aplicação de TRI¹

Carmen E. Flores-Mendoza²
Universidade Federal de Minas Gerais
Francisco J. Abad
Universidad Autónoma de Madrid
Álvaro José Lelé
Centro Universitário Newton Paiva

RESUMO — O Desenho da Figura Humana é um dos instrumentos mais divulgados e utilizados na prática de avaliação psicológica de crianças no Brasil. Recentemente um novo sistema de correção para avaliação cognitiva infantil foi publicado por Wechsler (2000, 2003). O presente trabalho constitui a primeira tentativa de análise psicométrica dos itens que compõem o desenho da figura masculina utilizando modelos matemáticos da Teoria de Resposta ao Item (TRI). A análise dos resultados obtidos em duas amostras de crianças (711 da cidade de Campo Grande e 564 da cidade de Belo Horizonte), aponta que a consistência interna do instrumento, considerando 53 itens, é adequada (0,87). Contudo, a análise de TRI, modelo politômico, mostra haver problemas de dificuldade e discriminação de um grupo de itens. Conclui-se que o instrumento não está bem calibrado e, portanto, há necessidade de ajustes quanto ao conjunto de itens que o compõe.

Palavras-chave: desenho da figura humana; TRI; análise de itens.

Item Analysis of Human Figure Drawings Test: Application of IRT

ABSTRACT – The Human Figure Drawings is one of the most divulged and used instruments in the practice of psychological evaluation of children in Brazil. Recently, Wechsler (2000, 2003) published a new correction system to cognitive evaluation. This work constitutes the first attempt of psychometric analysis of items related to male figure using mathematic models from Item Response Theory (IRT). The analysis of the results obtained in two samples of children (711 from the city of Campo Grande and 564 from the city of Belo Horizonte) with the Male Drawing points that the internal consistency of the instrument is adequate (.87) when 53 items are considered. Nevertheless, the IRT analysis, polytomous item responses, shows problems of difficulty and discrimination of a group of items. It's inferred that the instrument is not well calibrated and consequently there is a necessity of adjustments on the set of items that compound it.

Key words: human figure drawing; IRT; items analysis.

O desenho da figura humana é um dos temas preferidos pelas crianças e seu estudo remonta ao século XIX. Segundo Kamphaus e Pleiss (1991), já em 1885 com Ebenezer Cooke nos Estados Unidos e em 1887 com Conrado Ricci na Itália, encontram-se trabalhos sobre aspectos desenvolvimentais de desenhos de crianças. O ápice do interesse em grafismo infantil teria ocorrido entre os anos 1900 e 1915, período em que apareceram duas grandes investigações internacionais. A de Lambrecht em 1906, com desenhos coletados em diversos países e culturas e a de Claparède em 1907. A primeira não foi concluída e a segunda constituiu a tentativa de estabelecer estágios de desenvolvimento em paralelo à capacidade intelectual geral. As tentativas de estabelecer parâmetros do desenvolvimento psicológico infantil através do desenho baseiam no fato de que, no entender de Van Kolck (1981), o grafismo da criança constitui uma expressão do desenvolvimento em geral, isto é, pode-se estabelecer as fases de desenvolvimento de diversas dimensões psicológicas como

a percepção visual, a psicomotricidade, o desenvolvimento conceptual e outros.

O primeiro estudo bem sucedido e sistemático do desenho da figura humana para avaliação do desenvolvimento infantil, constitui, talvez, o trabalho de Florence Goodenough realizado em 1926 na Universidade de Minnesota, Estados Unidos (Abou-Jamra & Castilho, 1987). Conhecido como “*Draw-a-ManTest*” o teste de Goodenough visava analisar o aparecimento, no desenho, de partes do corpo humano conforme aumentava a idade cronológica. As diferenças que surgiam no desenho das crianças proporcionariam, assim, um instrumento de medida. No estudo de Goodenough foram escolhidos, catalogados e observados 4.000 desenhos de crianças da escola maternal e da escola primária. Objetivava-se assim identificar as diferenças existentes nas produções dos grupos superiores e observar as mudanças que ocorriam durante o desenvolvimento dos menores (Abou-Jamra & Castilho, 1987). A escala final de Goodenough constava de 51 itens cuja aplicação e avaliação estatística conferiu à autora a conclusão de que a criança desenha o que ela sabe e não o que ela vê. Posteriormente Harris, na década de 60, pretendeu estender a escala de Goodenough para uso em adolescentes e propor um sistema de pontuação para análise da figura feminina (no estudo de Goodenough empregou-se somente a figura masculina). Harris não conseguiu alcançar

1 Agradecimentos à psicóloga Andréia Pereira pela coleta de dados na cidade de Campo Grande, MS.

2 Endereço: Universidade Federal de Minas Gerais, Av. Antonio Carlos 6627, Laboratório de Avaliação das Diferenças Individuais, Gabinete 4042, Belo Horizonte, BH, Brasil 31270-901. E-mail: carmencita@fafich.ufmg.br

o primeiro objetivo. Com a adição da figura feminina, elaboração de novos itens para avaliação da figura masculina e a utilização do desenho de si mesmo como técnica projetiva, o sistema de Harris ficou conhecido como “*Goodenough-Harris Draw-A-Person*” (Alves, 1981). Segundo Abou-Jamra e Castilho (1987), os estudos visando obter informações das características psicométricas do instrumento, indicam que, no aspecto relacionado à estabilidade temporal, encontraram-se correlações entre 0,60 a 0,91 utilizando o sistema Goodenough e de 0,42 a 0,79 utilizando o sistema Harris. Por outro lado, apesar da aplicação duvidosa da técnica das metades para verificação da pertinência dos itens do Desenho da Figura Humana, os índices encontrados com essa técnica foram maiores a 0,75. Quanto à análise da precisão dos avaliadores, uma vez que o julgamento da pontuação depende da ótica do avaliador sobre o desenho, os índices encontrados foram maiores a 0,80, o que faz supor uma alta objetividade do instrumento.

Em 1988, surgiu nos Estados Unidos a proposta de Nagliari chamada *The Draw A Person: A Quantitative Scoring System* (DAP). Trata-se de uma revisão e atualização do trabalho de Goodenough e Harris. A nova proposta constitui de 64 itens e foi normatizada com base em uma amostra representativa de 2.622 crianças americanas. A análise de consistência interna apontou um coeficiente de 0,86 e uma estabilidade temporal de 0,74. Como em outros sistemas de correção, o sistema de Nagliari obteve uma correlação média entre avaliadores de 0,94 (Kamphaus & Pleiss, 1991). No Brasil, os estudos do Desenho da Figura Humana como forma de avaliação cognitiva podem ser encontrados em Almeida (1959), Alves (1981), Lessa (1953), Zausmer (1954) e, mais recentemente, Wechsler (2000, 2003).

Sistema Wechsler de Avaliação do Desenho da Figura Humana

A proposta de Wechsler (2000, 2003) revisa os trabalhos de Harris, Koppitz e Nagliari. Composto de 58 itens para avaliação da figura masculina e de 53 itens para a figura feminina, o sistema Wechsler foi normatizado com base a duas amostras, a primeira com 2.391 crianças de Brasília-DF e a segunda, com 588 crianças da cidade de Campinas-SP. O estudo de estabilidade temporal do instrumento, considerando a amostra total, apresentou um coeficiente de 0,75 para a figura feminina e um coeficiente de 0,69 para a figura masculina. A análise de consistência interna apontou um coeficiente *alpha* para o desenho feminino entre 0,77 e 0,89 (crianças do sexo feminino e masculino) e um coeficiente entre 0,76 a 0,86 para o desenho masculino (crianças do sexo feminino e masculino). Tais resultados indicam que o instrumento apresenta bons índices de precisão. Também se observaram diferenças em relação ao sexo da criança e ao sexo da figura desenhada, o que motivou a realização da normatização considerando, além da idade, as variáveis sexo da criança e sexo do desenho.

O sistema Wechsler passou por três revisões antes da versão final. Para verificar a precisão dos avaliadores, utilizou-se o procedimento de juízes cegos. Da primeira versão à terceira versão, a análise de concordância entre os juízes apontou um aumento dos coeficientes de correlação. Assim, na versão

final, observam-se altas concordâncias entre os juízes que variaram entre 0,95 a 0,98 para a figura masculina e de 0,93 a 0,97 para a figura feminina. No que diz respeito à análise de itens, a autora seguiu os critérios de Koppitz, quais sejam a porcentagem de ocorrência em três faixas de amplitude (classificados como esperados, comuns e incomuns) de cada item em cada faixa etária, de acordo com o sexo da criança e o sexo da figura desenhada. Também se observou uma classificação chamada de “itens excepcionais” cujo significado não fica suficientemente claro no manual.

Em 2003, a autora apresentou novos dados atendendo a uma recomendação da *Internacional Testing Commission* sobre revalidação e repadronização de instrumentos psicológicos. O novo estudo refere-se a dados de crianças oriundas de cinco estados brasileiros, quais sejam Bahia, Minas Gerais, Rio de Janeiro, Rio Grande do Sul e São Paulo. A amostra final ficou conformada em 3.340 crianças, sendo 49,5% dela conformada por crianças de São Paulo. A amostra deste segundo estudo foi superior à primeira, não só pela quantidade de crianças envolvidas, mas também pela representação geográfica. O resultado da análise de variância mostrou haver efeito da idade, sexo e tipo de figura. A consistência interna apontou um coeficiente *alpha*, dependendo da faixa etária, entre 0,83 a 0,89 para a figura masculina e entre 0,77 a 0,85 para a figura feminina. A precisão entre cinco avaliadores ficou entre 0,94 a 0,99. Por outro lado, a comparação de médias entre as amostras do primeiro estudo (1996) e do segundo estudo (2003), mostrou diferenças significativas nos dois tipos de figuras, masculino e feminino. A amostra do segundo estudo apresentou um desempenho gráfico superior ao da primeira amostra, o que motivou a autora a realizar uma repadronização do instrumento.

Em geral, pode-se dizer que os estudos de Wechsler (2000 & 2003), mostram evidências de um efeito desenvolvimental no desempenho gráfico infantil, isto é, à medida que aumenta a idade cronológica aumenta também a pontuação total. Se esse efeito desenvolvimental está, ou não, relacionado à capacidade cognitiva deverá ser tema de outras investigações. No momento, focalizar-se-á uma especial atenção à ausência de análise de itens tanto no primeiro estudo quanto no segundo. Isto é, não se encontra informação sobre a capacidade dos itens em representar o traço que o instrumento pretende medir, como tampouco do grau de eficiência com que o fazem (índice de dificuldade e de discriminação). O estabelecimento de uma análise empírica dos itens permite que se obtenha informação sobre a eficiência do instrumento em discriminar diferenças individuais no desenvolvimento do traço. Assim por exemplo, um instrumento com itens muito fáceis, em que qualquer sujeito independente da sua capacidade acerte a resposta, produzirá uma curva assimétrica com distribuição do traço concentrada à direita. Inversamente, um instrumento com itens muito difíceis, em que qualquer sujeito independente da sua capacidade falhe na resposta, produzirá uma curva assimétrica concentrada à esquerda. Nesse sentido, a análise de itens torna-se necessária para saber o quão eficiente é um instrumento na identificação do grau de capacidade do indivíduo. Atualmente, existe uma forma moderna, porém complexa, de analisar o comportamento dos itens. Trata-se da Teoria de Resposta ao Item, e que a seguir será brevemente descrita.

TRI e análise de itens

A Teoria de Resposta ao Item (TRI), é uma teoria estatística geral em que por meio de diversos modelos matemáticos tenta-se investigar a probabilidade do sujeito responder corretamente a qualquer item do teste em função da sua habilidade. Assim os modelos trabalham na perspectiva de relacionar dados observáveis (desempenho no teste) a dados não observáveis (habilidade cognitiva). O relacionamento resultante pode ser observado através da função ou curva característica do item (*Item Characteristic Curve-ICC*), em que se registra os níveis de êxito para cada item e a proporção de pessoas de cada nível que supera cada item. Na estimativa do ICC freqüentemente considera-se três parâmetros: a discriminação (parâmetro a), a dificuldade (parâmetro b) e a adivinhação do item (parâmetro c). O item do teste pode ser contínuo ou discreto, ordenado ou não ordenado (Hamblenton & Jones, 1993). A utilização da Teoria de Resposta ao Item pressupõe duas exigências: unidimensionalidade e independência local.

A unidimensionalidade diz respeito ao pressuposto da TRI de que os itens do teste solicitam uma aptidão dominante, isto é, o teste seria sensível à expressão significativa de um domínio do comportamento humano. A independência local significa que as respostas a um conjunto de itens estão vinculadas à probabilidade da resposta correta do examinando a cada item individual. Num teste tradicional, a probabilidade de resposta correta a um determinado item depende do acerto ou erro nos itens anteriores. Isto não ocorre na análise de TRI. A resolução a cada item é independente da resolução dos outros itens.

Uma vez estabelecidos os parâmetros dos itens deve-se demonstrar a adequação do modelo escolhido (modelo de um, dois ou três parâmetros) aos dados observados. Para tanto se analisa os resíduos e a invariância dos parâmetros. A primeira, análise dos resíduos, diz respeito à diferença maior que zero entre o observado e o predito pelo modelo. A segunda, análise da invariância, diz respeito ao grau de semelhança entre as ICCs quando extraídas de duas amostras de sujeitos submetidos ao mesmo teste (Pasquali, 1999).

A maioria dos instrumentos de rendimento ótimo (ex. testes de inteligência) é composta de itens de resposta binária (0,1), razão pela qual tornaram-se mais conhecidos os modelos de TRI dicotômicos. O Desenho da Figura Humana, contudo, apresenta categorias ordenadas nas quais se supõe que para ter um desempenho ótimo em uma determinada categoria o sujeito deve responder bem a uma série de passos anteriores. Por exemplo, na categoria "Nariz", o sujeito ao obter um ponto em 4c (estrutura completa do nariz), o sujeito deve ter respondido 4b (duas dimensões) e, conseqüentemente, 4a (presença de nariz). A análise dos itens desse sistema de pontuação poderá ser feita, então, utilizando o "Modelo de Resposta Graduada". Esse modelo criado por Samejima (1997) pressupõe que codificando as respostas do sujeito entre 0 e $m-1$ (sendo m o número de respostas distintas), a probabilidade de ocorrer a escolha de uma categoria maior ou igual que k , em função do nível de habilidade θ , e é dada pela função:

$$P^*(x \geq k | \theta) = \frac{1}{1 + \exp[-1.7 * a(\theta - b_k)]}$$

A fórmula não está definida para $k = 0$ e $k = m$, porém, logicamente:

$$P^*(x \geq 0 | \theta) = 1 \quad P^*(x \geq m | \theta) = 0$$

Esta função é a mesma que se utiliza para o modelo logístico de dois parâmetros. Por tal motivo se afirma que o modelo de resposta graduada é uma generalização deste modelo ao caso politômico (Embretson & Reise, 2000).

Seguindo essa direção, o presente estudo visa apresentar a aplicação da Teoria de Resposta ao Item, modelo de resposta graduada, à análise do Desenho da Figura Humana (sistema Wechsler), numa tentativa de aperfeiçoar a identificação de suas propriedades psicométricas. Para tanto, realizaram-se dois estudos, o primeiro com uma amostra de Campo Grande (MS) e o segundo com uma amostra de Belo Horizonte (MG).

Estudo 1

Foi coletado no ano de 1996, um total de 711 desenhos da figura masculina de crianças sem déficits físicos ou sensoriais, sendo 360 desenhos realizados por crianças do sexo feminino e 351 desenhos realizados por crianças do sexo masculino, com faixa etária entre 6 anos e 5 meses até 12 anos, freqüentadoras de seis escolas públicas da região metropolitana de Campo Grande – Mato Grosso do Sul. As escolas foram sorteadas aleatoriamente de uma lista cedida pela Secretaria de Ensino. Cada escola participante autorizou, mediante a sua Direção, a realização da coleta de dados. O sistema de pontuação adotado para análise dos desenhos foi o proposto por Wechsler (2000, 2003), especificamente aquele relacionado ao desenho da figura masculina. Para verificar a precisão dos critérios de correção, dois conjuntos de 40 desenhos não pertencentes à amostra, foram avaliados por uma equipe de nove juízes.

Resultados

Concordância entre avaliadores

O índice de concordância entre os juízes para as pontuações totais foi de 0,94. Contudo, considerando os 58 itens para avaliação da figura masculina, o índice de concordância, para cada item, variou entre 0,72 a 1,00. Os itens 8a (Rosto-proporção), 13e (Pernas-espaço entre as pernas), 16a (coordenação motora-linhas) e 16b (coordenação motora-união) foram os que obtiveram menor acordo entre os juízes (0,77; 0,75; 0,72 e 0,77 respectivamente). Para esses itens adicionou-se informação mais detalhada de avaliação. Um novo conjunto de 40 desenhos foi avaliado por oito juízes dos nove que compunha a equipe inicial, desta vez considerando as informações adicionais. O índice de concordância entre os juízes para os itens 8a e 13e aumentou para 0,88 e 0,80 respectivamente, porém os itens 16a e 16b mantiveram um índice de concordância um pouco abaixo (0,75 e 0,76) do tradicionalmente aceito (0,80).

Uma descrição da amostra quanto às variáveis sexo, idade, pontuação média e variabilidade é apresentada na Tabela 1.

Uma análise de variância foi realizada para verificar os efeitos das variáveis sexo e idade no desempenho da amostra.

Tabela 1. Médias e desvios padrões por sexo e grupo etário

Grupo Etário	Idade	Feminino			Masculino			Total		
		n	Média	DP	n	Média	DP	n	Média	DP
1	6,6 - 7,0	22	20,59	5,37	28	19,86	5,40	50	20,18	5,34
2	7,1 - 7,5	27	22,59	5,31	12	21,08	4,42	39	22,13	5,04
3	7,6 - 8,0	47	23,32	4,76	39	21,26	6,91	86	22,38	5,89
4	8,1 - 8,5	34	24,62	6,05	46	23,48	6,63	80	23,96	6,37
5	8,6 - 9,0	55	24,53	5,77	56	25,86	6,90	111	25,20	6,37
6	9,1 - 9,5	24	26,58	6,59	23	25,04	5,37	47	25,83	6,01
7	9,6 - 10,0	32	26,16	5,79	24	26,71	6,05	56	26,39	5,85
8	10,1 - 10,5	45	28,84	7,39	26	29,00	4,76	71	28,90	6,52
9	10,6 - 11,0	38	26,92	4,61	39	27,56	6,73	77	27,25	5,75
10	11,1 - 11,5	14	27,64	6,96	24	28,42	6,74	38	28,13	6,74
11	11,6 - 12,0	14	28,93	6,35	21	28,38	8,42	35	28,60	7,56
12	12,1 - 12,5	8	28,00	6,32	13	30,23	5,99	21	29,38	6,06
Total		360			351			711		

Os resultados mostraram haver efeito da idade [$F(11) = 11,318$, $p = 0,000$], o que reforça a hipótese do instrumento medir algum construto desenvolvimental. A proporção de variância na pontuação total explicada pela variável idade é de 15% (quadrado do valor Eta). Por outro lado, diferentemente do estudo de Wechsler (2000), não se encontrou efeito da variável sexo [$F(1) = 0,086$, $p = 0,769$], pelo menos no que diz respeito ao desenho da figura masculina. Tampouco foi significativa a interação sexo x idade [$F(11) = 0,615$, $p = 0,817$].

A Figura 1 mostra que conforme aumenta a idade, aumenta também a pontuação total no desenho da figura masculina. Tal ocorrência se observa tanto em crianças do sexo feminino quanto do sexo masculino.

Análise de itens: primeira exploração dos dados

Analysaram-se 57 dos 58 itens devido a que o item 1a (presença de cabeça) não apresentou variabilidade nas respostas. Para aplicar os modelos da teoria de resposta ao item comprovou-se em primeiro lugar o critério de unidimensionalidade. Para tanto, aplicou-se o NOHARM, um programa de análise fatorial. Encontrou-se quatro itens dos 57 com pesos negativos ou próximos a zero [8a (proporção do rosto), 16a (coord. motora-linha), 16b (coord. motora-união e 17 (presença de

perfil ou frente)]. Os pesos dos itens podem ser observados na Tabela 2. Conclui-se que esses itens não podem ser analisados mediante TRI (de fato correlacionam zero ou negativamente com a pontuação total, logo uma análise clássica deveria ter descartado esses itens). Inclusive, é preciso destacar que três desses itens (8a, 16a e 16b) são os mesmos que apresentaram baixa concordância entre juízes.

Considerando os 53 itens restantes encontrou-se que o primeiro fator explica 28% da variância. Logo, poderia-se afirmar que o teste é suficientemente unidimensional conforme alguns dos critérios clássicos. O programa NOHARM proporciona as covariâncias residuais (diferenças entre as covariâncias empíricas e as covariâncias teóricas segundo a solução unidimensional). A raiz quadrada média desses resíduos (*Root mean square of residuals* ou RMSR) para a solução unifatorial é de 0.010229, enquanto uma solução bifatorial exploratória somente produz um ligeiro decréscimo da RMSR a 0.009928 (um decréscimo inferior ao 10%). Portanto, um modelo unidimensional se ajusta moderadamente bem aos dados.

No entanto, se comprova que em geral sendo pequenos os resíduos, há pequenos subconjuntos de itens que mostram relações não explicadas pelo fator geral e que estão relacionados ao seu conteúdo específico (itens 2a e 2b; 6a e 6b; 7a e 7b; 9a e 9b; etc.). Tal situação constitui um problema posto que indicaria que o requisito de independência local, necessário para aplicar a TRI, não se cumpre. Tais subgrupos estão estritamente relacionados ao sistema de pontuação.

Segundo a TRI, para pessoas com um mesmo nível de traço, a probabilidade de acertar dois itens (ex. itens 2a e 2b) deveria ser o produto das probabilidades de acertar cada um desses itens em separado. Claramente no teste do Desenho da Figura Humana, as probabilidades não são independentes: se o cabelo segue a linha da cabeça (acerto no item 2b) necessariamente o cabelo está presente (acerto no item 2a). Tal situação se repete em numerosos itens. Portanto, o modelo de dois parâmetros deve ser aplicado com extrema precaução, uma vez que se viola um dos pressupostos de aplicação da TRI.

Tal problema não está relacionado à TRI. Deve-se recordar que os índices de consistência interna nessas situações geralmente se superestimam posto que a covariação entre

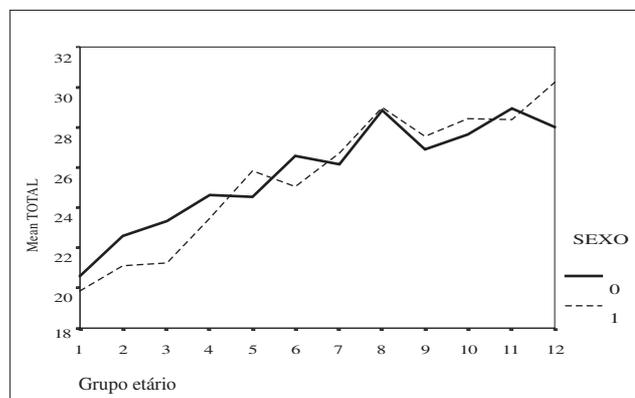


Figura 1. Pontuação média em cada faixa etária.

Desenho da figura humana

Tabela 2. Parâmetros estimados para os itens

Itens	Analise fatorial	Modelo 2p	Modelo de resposta graduada					
	peso	a	B	a	b1	b2	b3	
8/16a/16b/17	.04/-03/-03/-54	eliminados						
1b. Proporção (Cabeça)	.21	0.38	3.48	1b	0.37	3.54		
2a. Presença (Cabelo)	.59	0.8	-1.84					
2b. Adaptação (Cabelo)	.47	0.62	-1.56	2abc	0.63	-2.16	-1.54	3.23
2c. Penteado (Cabelo)	.42	0.64	3.02					
3a. Presença (olhos)	.62	0.92	-3.29					
3d. Proporção (olhos)	.64	0.96	1.02	3ad	0.99	-3.17	1.01	
3b. Sobrancelhas (olhos)	.34	0.4	0.78	3b	0.4	0.78		
3c. Pupilas (olhos)	.53	0.64	0.55	3c	0.64	0.55		
3e. Direção (olhos)	.22	0.26	0.06	3e	0.27	0.06		
4a. Presença (Nariz)	.34	0.47	-2.48					
4b. Duas dimensões (Nariz)	.54	0.69	0.57	4abc	0.59	-2.11	0.63	3.45
4c. Estrutura completa (Nariz)	.40	0.63	3.15					
4d. Fossas nasais (Nariz)	.49	0.66	1.96	4d	0.65	2.01		
5a. Presença (Boca)	.45	0.6	-2.99					
5b. Lábios (Boca)	.61	1	2.26	5abc	0.76	-2.54	2.89	3.09
5c. Nariz e lábios (Boca)	.69	1.18	2.27					
6a. Presença (Orelhas)	.39	0.51	1.6	6ab	0.51	1.61	2.91	
6b. Proporção e posição (Orelhas)	.47	0.66	2.31					
7a. Presença (pescoço)	.63	0.76	-0.59	7ab	0.66	-0.64	0.57	
7b. Duas dimensões (pescoço)	.55	0.68	0.48					
7c. Linha do pescoço (roupa)	.57	0.9	1.78	7c	0.88	1.83		
9a. Presença (Queixo)	.30	0.43	-2.76	9ab	0.35	-3.37	-2.91	
9b. Queixo e testa (Queixo)	.28	0.39	-2.65					
10a. Presença (Mãos)	.37	0.44	-0.07	10a	0.44	-0.07		
10b. Dedos (Mãos)	.53	0.62	-1.28	10bd	0.64	-1.26	1.67	
10d. Forma dos dedos (Mãos)	.51	0.71	1.57					
10c. Número de dedos (Mãos)	.59	0.82	1.3	10c	0.84	1.29		
10e. Posição do polegar (Mãos)	.54	0.9	2.8	10e	0.9	2.81		
11a. Presença (Braços)	.62	0.94	-2.65	11ab	0.43	-4.72	0.11	
11b. Junto ao corpo (Braços)	.32	0.4	0.07					
11c. Cotovelo (Braços)	.37	0.63	2.93	11c	0.6	3.05		
11d. Comprimento (Braços)	.37	0.49	1.78	11d	0.47	1.84		
11e. União (Braços)	.63	0.99	-2.22	11e	0.94	-2.29		
11f. União 2 (Braços)	.36	0.44	-1.32	11f	0.45	-1.3		
11g. Proporção (Braços)	.39	0.7	3.22	11g	0.69	3.28		
12a. Presença (ombros)	.59	0.72	0.44	12ab	0.73	0.42	1.68	
12b. Forma (ombros)	.63	0.93	1.42					
13a. Presença (Pernas)	.39	0.57	-3.96					
13b. Joelho (Pernas)	.14	0.5	4.19	13ab	0.43	-5.02	5.07	
13c. Extensão (Pernas)	.35	0.41	-0.11					
13d. Tornozelo ou pulso (Pernas)	.60	1.04	1.81	13d	0.98	1.88		
13e. Espaço entre pernas (Pernas)	.21	0.27	-1.05	13e	0.27	-1.05		
13f. Duas dimensões (Pernas)	.69	1.06	-1.28	13f	1.08	-1.27		
14a. Presença (Pés)	.79	0.97	-1.3					
14b. Tamanho (Pés)	.61	0.67	-0.54	14abc	0.74	-1.5	-0.49	2.51
14c. Detalhes (Pés)	.60	1.02	1.94					
15a. Presença (Tronco-corpo)	.43	0.75	-3.56					
15b. Duas dimensões (Tronco-corpo)	.65	0.86	-2.31	15abc	0.83	-3.33	-2.28	2.93
15c. Contorno do tronco (Tronco-corpo)	.53	0.87	2.7					
18c. Roupas 3 (Vestuário)	.77	1.44	0.17	18c	1.43	0.17		
18a. Roupas 1 (Vestuário)	.81	1.66	0.05					
18b. Roupas 2 (Vestuário)	.76	1.58	0.81	18abd	1.5	0.05	0.84	2
18d. Roupas 4 (Vestuário)	.72	1.48	1.97					

os itens está relacionada não apenas ao construto medido, mas também aos mais específicos (por exemplo, os erros dos itens se correlacionam). Portanto, se faz necessário aplicar modelos em unidades de análise nos quais seja possível o cumprimento do pressuposto de independência local. Assim sendo, a estratégia será ao invés de utilizar os itens como unidades de análise, utilizar grupos de itens.

A seguir são descritas duas análises sobre os 53 itens. Primeiro se aplica um modelo dicotômico (o modelo logístico de dois parâmetros) e depois se aplica um modelo politômico de TRI (o modelo de resposta graduada).

Ajuste do Modelo Logístico de Dois Parâmetros

Em geral, o ajuste dos 53 itens ao modelo é adequado. Para tanto, utilizou-se o programa PARSCALE (Muraki & Bock, 1997) estimando o modelo com métrica normal ($D = 1,7$). O programa proporciona uma estatística de ajuste para cada item mediante a qual se pode estabelecer o nível de significância estatística das discrepâncias entre as frequências de escolha das opções teóricas e as observadas para cada nível de habilidade (estatística X^2). Nas primeiras colunas da Tabela 3, se apresentam os X^2 , seus graus de liberdade, o nível de significância e a razão entre X^2 e os graus de liberdade. Observa-se que nenhuma das estatísticas é significativa e a razão $X^2/g.l$ é um valor inferior a 3 para todos os itens. Portanto, pode-se afirmar que todos os itens se ajustam razoavelmente ao modelo.

A distribuição dos parâmetros a (média = 0.76; $DP = 0.31$) e dos parâmetros b (média = 0.28; $DP = 2.07$), indica que os itens tendem a ser pouco discriminativos e que muitos itens apresentam parâmetros b extremos (ou são muito fáceis ou são muito difíceis). Itens claramente inadequados por sua baixa discriminação são: 1b (adaptação do cabelo), 3b (pupilas), 3e (direção dos olhos), 4a (presença nariz), 9a (presença queixo), 9b (queixo e testa), 10a (presença mãos), 11b (braços – tronco), 11d (cumprimento braços), 11f (união II), 13c (cumprimento pernas) e 13e (espaço entre pernas). Pressupondo a invariância de parâmetros pode-se dizer que esses itens serão poucos discriminativos independentemente do grupo de idade analisado. Observa-se, também, a discrepância entre a dificuldade dos itens e o nível de habilidade do grupo de idade analisado. Existem itens claramente inadequados por sua excessiva facilidade [3a (presença olhos), 4a (presença nariz), 5a (presença boca), 9a (presença queixo), 9b (queixo-testa), 11a (presença braços), 11e (união I), 13a (presença pernas), 15a (presença tronco) e o 15b (duas dimensões tronco)] ou por sua excessiva dificuldade [1B (proporção cabeça), 2c (penteado cabelo), 4c (estrutura completa nariz), 5b (duas dimensões lábios), 5c (nariz e lábios), 6b (proporção orelha), 10e (posição do polegar), 11c (cotovelo), 11g (proporção braços), 13b (joelho) e o 15c (contorno tronco)]. Os parâmetros a e b são apresentados na Figura 2 (ver também resumo na Tabela 2).

No gráfico seguinte (Figura 3) se mostra o erro padrão de medida em função do nível de habilidade estimado (segundo a TRI). Observa-se que os 53 itens apresentam um erro padrão de medida bastante aceitável em toda a amplitude da habilidade.

A estimação da precisão por meio da TRI pode ser obtida mediante o cálculo do erro padrão de medida para uma distri-

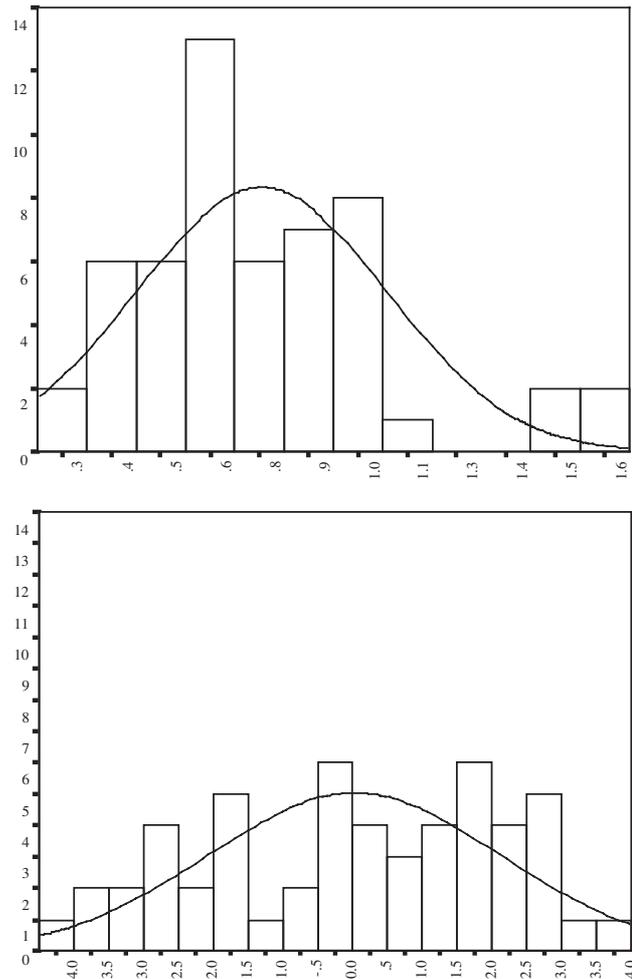


Figura 2. Distribuição dos parâmetros a e b apresentados respectivamente.

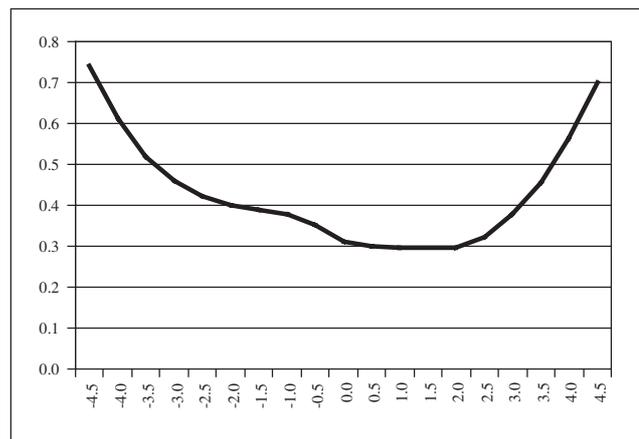


Figura 3. Erro padrão de medida dos 53 itens.

buição concreta de θ (ex. distribuição normal de θ com média 0 e desvio padrão de 1) e calculando o valor r_{xx} mediante a fórmula $r_{xx} = 1 - \text{Var}[e] / \text{Var}[\theta_{\text{est}}]$. Na presente análise, obteve-se um coeficiente de precisão de 0,87, portanto, ligeiramente superior à estimação clássica de consistência interna α (0,86).

Ajuste do modelo de resposta graduada de TRI

Em primeiro lugar, deve-se decidir quais conjuntos de itens são suscetíveis de tratamento politômico. O objetivo é *não perder informação*, isto é, ao seguir um sistema de pontuação politômica deve-se obter exatamente a mesma pontuação utilizada no sistema dicotômico. Por exemplo, se todos os sujeitos apresentam padrões de resposta para dois itens de 00 (falha-falha), 10 (acerto-falha) e 11 (acerto-acerto) então não haverá problemas em que esses dois itens dicotômicos possam ser pontuados como um único item politômico 0, 1 e 2. A pontuação politômica seria: 0 ponto se há falha no primeiro item, 1 ponto se há falha no segundo item e 2 pontos se não há falha em nenhum dos itens. O padrão 01 (falha-acerto) representa na verdade uma inconsistência lógica, posto que não se espera “penteadado” (2c) quando não houve “presença de cabelo” (2a). De qualquer forma se algum item permitir um padrão 01 (falha-acerto) os sujeitos serão pontuados em 0 pelo sistema de pontuação politômico. Nesse caso, mudarão as pontuações dos sujeitos dependendo do sistema de pontuação dicotômico ou

politômico e poderá haver perda de informação importante. Portanto, analisaram-se item a item quais eram suscetíveis de serem pontuados politômica. A análise identificou que os 53 itens podiam ser agrupados em 33 itens politômicos. Essa informação aparece resumida no Quadro 1.

Calculou-se uma nova pontuação para os novos 33 itens. A correlação entre as duas pontuações (dicotômica e politômica) foi de 0,997. Tal resultado permite inferir que se trabalhou com a mesma pontuação apesar da diferença entre os modelos.

Em geral o ajuste dos 33 itens ao modelo politômico é adequado. Novamente, utilizou-se o programa PARSCALE (Muraki & Bock, 1997) estimando o modelo de resposta graduada com métrica normal ($D = 1,7$).

Observa-se que os parâmetros a e b foram praticamente iguais para aqueles itens que se mantiveram na forma original sem agrupar (Tabela 2). Na Tabela 3, últimas colunas, se apresentam os X^2 , os graus de liberdade, o nível de significância e a razão entre X^2 e os graus de liberdade. Verifica-se que nenhuma estatística é significativa e que a razão X^2/df é, novamente, um valor inferior a 3 para todos os itens. Portanto,

Quadro 1. Suscetibilidade dos itens à análise politômica

Item	Suscetibilidade para agrupamento	Item análise politômica
1b	Não suscetível de ser agrupado	1
2abc	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença de cabelo, adaptação do cabelo, penteadado do cabelo.	2
3ad	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença, proporção dos olhos	3
3b	Sobrancelhas, não suscetível de ser agrupado	4
3c	Pupilas, não suscetível de ser agrupado (considerou-se 3ad)	5
3e	Direção do olhar, não suscetível de ser agrupado (considerou-se 3ad)	6
4abc	Suscetíveis de serem agrupados pela sua hierarquia lógica: Presença nariz, duas dimensões, estrutura completa	7
4d	Fossas nasais, não suscetível de ser agrupado (considerou-se 3abc)	8
5abc	Suscetíveis de serem agrupados pela sua hierarquia lógica: Presença, lábios, nariz e lábios.	9
6ab	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença, proporção/posição de orelhas	10
7ab	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença pescoço, duas dimensões	11
7c	Linha do pescoço, não suscetível de ser agrupado (considerou-se 7ab)	12
9ab	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença, queixo e testa	13
10a	Presença de mãos, não suscetível de ser agrupado	14
10bd	Suscetíveis de serem agrupados pela sua hierarquia lógica: dedos, forma dos dedos	15
10c	Número de dedos, não suscetível de ser agrupado (considerou-se 10bd)	16
10e	Posição do polegar, não suscetível de ser agrupado (considerou-se 10bd)	17
11ab	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença, braços junto ao tronco.	18
11c	Cotovelo, não suscetível de ser agrupado (considerou-se 11ab)	19
11d	Comprimento, não suscetível de ser agrupado (considerou-se 11ab)	20
11e	União I, não suscetível de ser agrupado (considerou-se 11ab)	21
11f	União II, não suscetível de ser agrupado (considerou-se 11ab)	22
11g	Proporção, não suscetível de ser agrupado (considerou-se 11ab)	23
12ab	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença, forma dos ombros.	24
13ab	Suscetíveis de serem agrupados pela sua hierarquia lógica: presença, joelho.	25
13c	Comprimento, não susceptível de ser agrupado (considerou-se 13ab)	26
13d	Tornozelo ou pulso, não suscetível de ser agrupado (considerou-se 13ab)	27
13e	Espaço entre as pernas, não suscetível de ser agrupado (considerou-se 13ab)	28
13f	Dois dimensões, não suscetível de ser agrupado (considerou-se 13ab)	29
14abc	Suscetíveis de serem agrupados pela sua hierarquia lógica: Presença, tamanho dos pés, detalhes.	30
15abc	Suscetíveis de serem agrupados pela sua hierarquia lógica: Presença, duas dimensões, contorno do tronco.	31
18c	Roupa 3, não suscetível de ser agrupado	32
18abd	Suscetíveis de serem agrupados pela sua hierarquia lógica: roupa 1, roupa 2, roupa 4	33

Tabela 3. Ajuste dos itens

	dicotômico					politômico			
	X ²	gl	P	X ² /gl		X ²	gl	P	X ² /gl
1b. Proporção (Cabeça)	18.20	20.00	0.58	0.91	1b	12.98	20.00	0.88	0.65
2a. Presença (Cabelo)	3.42	9.00	0.95	0.38					
2b. Adaptação (Cabelo)	8.04	13.00	0.84	0.62	2abc	34.03	36.00	0.56	0.95
2c. Penteado (Cabelo)	33.72	20.00	0.03	1.69					
3a. Presença (olhos)	0.00	0.00							
3d. Proporção (olhos)	15.41	17.00	0.57	0.91	3ad	11.24	17.00	0.85	0.66
3b. Sobrancelhas (olhos)	16.15	19.00	0.65	0.85	3b	15.62	19.00	0.68	0.82
3c. Pupilas (olhos)	41.48	17.00	0.00	2.44	3c	32.53	17.00	0.01	1.91
3e. Direção (olhos)	14.64	19.00	0.75	0.77	3e	10.02	19.00	0.95	0.53
4a. Presença (Nariz)	14.91	11.00	0.19	1.36					
4b. Duas dimensões (Nariz)	13.52	17.00	0.70	0.80	4abc	45.05	42.00	0.35	1.07
4c. Estrutura completa (Nariz)	16.89	20.00	0.66	0.84					
4d. Fossas nasais (Nariz)	19.85	20.00	0.47	0.99	4d	20.49	20.00	0.43	1.02
5a. Presença (Boca)	2.17	2.00	0.34	1.09					
5b. Lábios (Boca)	21.54	20.00	0.37	1.08	5abc	25.60	21.00	0.22	1.22
5c. Nariz e lábios (Boca)	19.19	20.00	0.51	0.96					
6a. Presença (Orelhas)	18.11	20.00	0.58	0.91					
6b. Proporção e posição (Orelhas)	17.30	20.00	0.63	0.87	6ab	40.45	29.00	0.08	1.39
7a. Presença (pescoço)	21.27	15.00	0.13	1.42					
7b. Duas dimensões (pescoço)	17.74	17.00	0.41	1.04	7ab	27.27	28.00	0.50	0.97
7c. Linha da roupa (pescoço)	7.85	20.00	0.99	0.39	7c	7.96	20.00	0.99	0.40
9a. Presença (Queixo)	7.02	10.00	0.73	0.70					
9b. Queixo e testa (Queixo)	12.86	13.00	0.46	0.99	9ab	15.68	13.00	0.27	1.21
10a. Presença (Mãos)	25.95	17.00	0.08	1.53	10a	19.18	17.00	0.32	1.13
10b. Dedos (Mãos)	17.14	14.00	0.25	1.22					
10d. Forma dos dedos (Mãos)	26.54	20.00	0.15	1.33	10bd	35.17	31.00	0.28	1.13
10c. Número de dedos (Mãos)	12.54	18.00	0.82	0.70	10c	11.02	18.00	0.89	0.61
10e. Posição do polegar (Mãos)	6.54	20.00	1.00	0.33	10e	12.89	20.00	0.88	0.64
11a. Presença (Braços)	0.23	1.00	0.63	0.23					
11b. Junto ao corpo (Braços)	15.29	17.00	0.58	0.90	11ab	9.41	17.00	0.93	0.55
11c. Cotovelo (Braços)	16.11	20.00	0.71	0.81	11c	9.40	20.00	0.98	0.47
11d. Extensão (Braços)	26.99	20.00	0.14	1.35	11d	31.40	20.00	0.05	1.57
11e. União (Braços)	0.08	2.00	0.95	0.04	11e	0.96	2.00	0.63	0.48
11f. União 2 (Braços)	15.89	16.00	0.46	0.99	11f	28.74	16.00	0.03	1.80
11g. Proporção (Braços)	21.76	20.00	0.35	1.09	11g	18.34	20.00	0.57	0.92
12a. Presença (ombros)	19.28	17.00	0.31	1.13					
12b. Forma (ombros)	22.17	18.00	0.22	1.23	12ab	26.80	30.00	0.63	0.89
13a. Presença (Pernas)	0.00	0.00	X	X	13ab	22.24	20.00	0.33	1.11
13b. Joelho (Pernas)	20.15	20.00	0.45	1.01					
13c. Extensão (Pernas)	13.19	17.00	0.72	0.78	13c	17.00	17.00	0.46	1.00
13d. Tornozelo ou pulso (Pernas)	18.97	20.00	0.52	0.95	13d	14.42	20.00	0.81	0.72
13e. Espaço entre pernas (Pernas)	10.47	16.00	0.84	0.65	13e	17.89	17.00	0.40	1.05
13f. Duas dimensões (Pernas)	9.16	11.00	0.61	0.83	13f	6.83	11.00	0.81	0.62
14a. Presença (Pés)	24.09	11.00	0.01	2.19					
14b. Tamanho (Pés)	26.09	16.00	0.05	1.63	14abc	42.22	42.00	0.46	1.01
14c. Detalhes (Pés)	14.21	20.00	0.82	0.71					
15a. Presença (Tronco-corpo)	0.00	0.00	X	X					
15b. Duas dimensões (Tronco-corpo)	1.67	3.00	0.65	0.56	15abc	7.83	24.00	1.00	0.33
15c. Contorno do tronco (Tronco-corpo)	14.01	20.00	0.83	0.70					
18c. Roupa 3 (Vestuário)	18.46	15.00	0.24	1.23	18c	24.42	15.00	0.06	1.63
18a. Roupa 1 (Vestuário)	16.23	14.00	0.30	1.16					
18b. Roupa 2 (Vestuário)	13.42	17.00	0.71	0.79	18abd	37.44	35.00	0.36	1.07
18d. Roupa 4 (Vestuário)	7.00	20.00	1.00	0.35					

pode-se afirmar que todos os itens se ajustam adequadamente ao modelo de resposta graduada.

A distribuição dos parâmetros *a* (média = 0,68; SD = 0,30) confirma o caráter pouco discriminativo de alguns itens. Assim, itens claramente inadequados por sua baixa discriminação são: 1b, 3b, 3e, 9ab, 10a, 11ab, 11d, 11f, 13ab, 13c e o item 13e.

Considerando o ajuste da dificuldade dos itens, observam-se itens claramente inadequados por sua excessiva facilidade (2abc, 3ad, 4abc, 5abc, 9ab, 11ab, 11e, 13ab e 15abc) ou por sua excessiva dificuldade (1b, 2abc, 4abc, 4d, 5abc, 6ab, 10e, 11c, 11g, 13ab, 14abc e 15abc). A interpretação de tais resultados é bastante similar à realizada para os modelos dicotômicos, à qual pode-se acrescentar:

- No item 9, a e b (queixo-testa) apresentam dificuldades muito próximas. Isto faz com que uma das pontuações seja muito pouco eleita. Poderia-se concluir que ao utilizar o item 9a, não faz falta o item 9b, ou vice-versa.
- No item 2abc, os critérios a e b (presença-adaptação de cabelo) possuem dificuldades muito próximas. Isto faz com que uma das pontuações seja muito pouco eleita. Poderia-se concluir que ao utilizar o item 2a, não faz falta o item 2b, ou vice-versa.
- No item 5abc, os critérios b e c (lábios-nariz) possuem dificuldades muito próximas. Isto faz com que uma das pontuações seja muito pouco eleita. Poderia-se concluir que ao utilizar o item 5b, não faz falta o item 5c, ou vice-versa.

Em resumo, a aplicação dos modelos de TRI à análise dos itens do Desenho da Figura Humana, no Estudo 1, aponta as seguintes evidências: 1) unidimensionalidade do instrumento; 2) adequada precisão; 3) possibilidade de analisar 33 itens com o modelo politômico de TRI.

Estudo 2

Participaram 559 crianças de cinco escolas da região de Belo Horizonte, a maioria (77,7%) proveniente de escolas públicas. A coleta de dados realizou-se em 2003, após o devido consentimento informado por parte das instituições participantes. A estatística descritiva é apresentada logo a seguir.

Uma análise de variância foi realizada para verificar os efeitos das variáveis sexo e idade no desempenho da amostra. De forma semelhante ao Estudo 1, os resultados mostram haver efeito da idade [$F(5) = 28,796, p = 0,000$], o que reforça a hipótese do instrumento medir algum construto desenvolvimental. Contudo, novamente, diferentemente do estudo de

Tabela 4. Estatística descritiva da segunda amostra

Idade	Masculino			Feminino		
	<i>n</i>	média	dp	<i>n</i>	média	dp
7	60	25,25	5,66	62	28,11	5,31
8	45	28,89	6,57	57	29,95	5,08
9	66	32,18	7,12	51	32,55	5,62
10	71	33,54	6,65	39	34,51	6,32
11	38	35,47	7,53	47	34,94	5,37
12	15	36,20	7,88	08	35,38	6,21
Total	295			264		

Wechsler (2000), não se encontrou efeito da variável sexo [$F(1) = 0,567, p = 0,568$], pelo menos no que diz respeito ao desenho da figura masculina.

TRI e análise de itens

Para efeito de verificar a validação da calibração dos itens obtida no Estudo 1, aplicou-se o modelo de resposta graduada considerando os 33 itens selecionados no estudo anterior. Observou-se que um dos itens (item 23-proporção dos braços) não se ajustou à nova amostra ($p = 0,01$). A correlação entre os parâmetros estimados nas duas amostras para os 32 itens restantes foi de 0,66 para o parâmetro *a* e de 0,96 para o parâmetro *b*. Posteriormente, procedeu-se à calibração dos 32 itens restantes utilizando ambas as amostras simultanea-

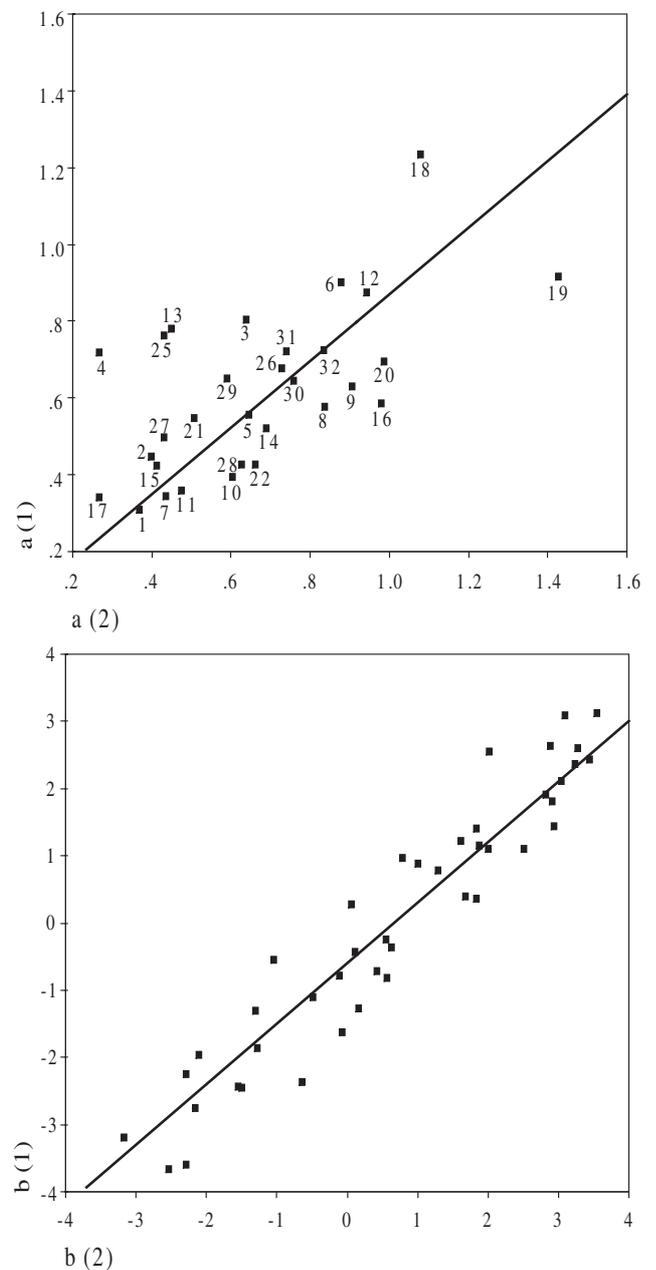


Figura 4. Dispersão dos parâmetros a e b.

mente, isto para efeito de uma maior qualidade da estimação. A nova calibração identificou o desajuste de dois novos itens (item 24-presença e forma de ombros e item 33-roupa I, II e IV). Isto quer dizer que embora cada um desses itens se ajustava bem ao modelo de TRI, ao se considerar ambas as amostras o ajuste não ocorria. Assim sendo, procedeu-se à correlação entre os parâmetros estimados para ambas as

amostras considerando os 30 itens restantes. Obteve-se uma correlação de 0,63 para o parâmetro *a* e 0,95 para o parâmetro *b*. A Figura 4 mostra os gráficos de dispersão para o parâmetro *a* e para o parâmetro *b* respectivamente, estimados em ambas as amostras.

Observa-se que para alguns itens (ex. 4-sobrancelha ou 19-cotovelo) os resultados no parâmetro *a* diferem bastante

Tabela 5. Calibração final dos 30 itens politômicos com a amostra total

Item	Item politômico	a	b1	b2	b3
13a. presença (Pernas)/ 13b. joelho (Pernas)	13ab	.47	-4.78	4.27	
11a. presença (Braços)/ 11b. junto ao corpo (Braços)	11ab	.59	-3.78	-.17	
3a. presença (olhos)/ 3d. proporção (olhos)	3ad	.74	-3.43	1.04	
15a. presença (Tronco-corpo)/ 15b. duas dimensões (Tronco-corpo)/ 15c. contorno do tronco (Tronco-corpo)	15abc	.93	-3.36	-2.45	1.87
5a. presença (Boca)/ 5b. lábios (Boca)/ 5c. nariz e lábios (Boca)	5abc	.74	-2.87	2.72	3.03
11e. união (Braços)	11e	.80	-2.47		
2a. presença (Cabelo)/ 2b. adaptação (Cabelo)/ 2c. penteado (Cabelo)	2abc	.52	-2.43	-1.92	2.76
4a. presença (Nariz)/ 4b. duas dimensões (Nariz)/ 4c. estrutura completa (Nariz)	4abc	.65	-2.00	.15	2.85
14a. presença (Pés)/ 14b. tamanho (Pés)/ 14c. detalhes (Pés)	14abc	.78	-1.79	-.73	1.68
13f. duas dimensões (Pernas)	13f	1.22	-1.47		
11f. união 2 (Braços)	11f	.59	-1.29		
13e. espaço entre pernas (Pernas)	13e	.23	-1.04		
7a. presença (PESCOÇO)/ 7b. duas dimensões (PESCOÇO)	7ab	.68	-1.04	.06	
10a. presença (Mãos)	10a	.40	-.65		
18c. roupa 3 (Vestuário)	18c	.97	-.39		
13c. extensão (Pernas)	13c	.43	-.39		
12a. presença (Ombros)/ 12b. forma (Ombros)	12ab	.89	-.06	.92	
3c. pupilas (olhos)	3c	.70	.17		
3e. direção (olhos)	3e	.33	.25		
7c. linhas do pescoço (Roupa)	7c	.97	1.01		
3b. sobrancelhas (olhos)	3b	.33	1.06		
10c. número de dedos (Mãos)	10c	.67	1.13		
6a. presença (Orelhas)/ 6b. proporção e posição (Orelhas)	6ab	.50	1.47	2.42	
13d. tornozelo ou pulso (Pernas)	13d	.78	1.54		
11d. comprimento (Braços)	11d	.42	1.66		
10e. posição do polegar (Mãos)	10e	.78	2.27		
11c. cotovelo (Braços)	11c	.55	2.36		
4d. fossas nasais (Nariz)	4d	.48	2.67		
11g. proporção (Braços)	11g	.67	2.69		
1b. proporção (Cabeça)	1b	.33	3.48		

Quadro 2. Itens eliminados da análise politômica na nova amostra

Itens eliminados
9a. presença (Queixo)+9b. Queixo e testa
10b. dedos (Mãos)+10d. forma dos dedos (Mãos)
18a. roupa 1 (vestuário)+18b. roupa 2 (vestuário)+18d. roupa 4 (vestuário)

de um grupo para outro. Os resultados da calibração final dos 30 itens com toda a amostra aparecem na Tabela 5, ordenados pela dificuldade do primeiro sub-item, assim como os itens eliminados (Quadro 2).

A correlação, considerando os 30 itens, entre os novos parâmetros e os parâmetros estimados na amostra inicial é de 0,86 (parâmetro *a*) e de 0,99 (parâmetro *b*), o que sugere que quando se considera a calibração completa não se observam mudanças substanciais nos parâmetros ao adicionar a nova amostra. Contudo, deve-se destacar diferenças de magnitude no desempenho de ambas amostras. A primeira amostra (obtida em 1996 na cidade de Campo Grande-MS) obteve uma média de -0,32 e um desvio padrão de 0,81. A segunda amostra (obtida em 2003 na cidade de Belo Horizonte-MG) obteve uma média de 0,40 e um desvio padrão de 0,81. Isto significa que a primeira amostra apresentou um nível mais baixo do traço subjacente à realização do teste do que a segunda amostra (Figura 5). A interpretação deste resultado deverá perpassar pela análise das condições sócio-educativas e econômicas de cada Estado assim como a possibilidade de um efeito temporal no desenvolvimento infantil.

Considerações Finais

A análise do Desenho da Figura Humana utilizando o sistema de correção Wechsler (2000,2003) mostrou que o instrumento em geral, descartando-se cinco itens, apresenta razoável unidimensionalidade. Isto é, os itens que compõem o instrumento provavelmente medem um mesmo construto (o primeiro fator explicou 28% da variância). A definição sobre qual construto está subjacente ao instrumento deverá ser estabelecida através de análises de validade convergente e divergente. Diferentemente dos estudos de Wechsler, na presente investigação não se observaram diferenças de sexo no desenho da figura masculina em nenhuma das duas amos-

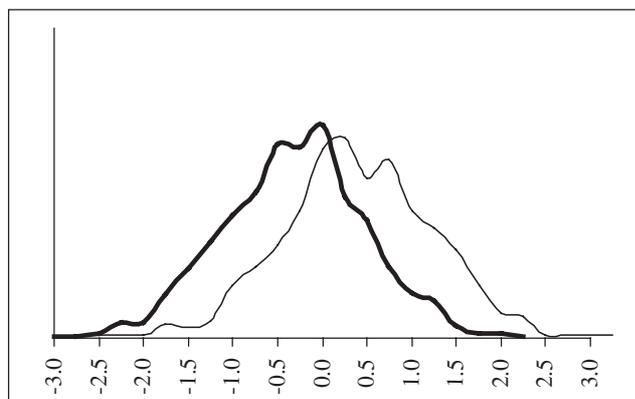


Figura 5. Distribuição do traço nas duas amostras: Campo Grande (linha grossa) e Belo Horizonte (linha fina).

tras. Observou-se sim, como já era esperado, efeito da idade. No que diz respeito à sua consistência interna, o instrumento apresentou um alto coeficiente de fidedignidade (0,87) bastante próximo dos encontrados nos estudos de Wechsler.

Embora deva-se ter cuidado com a aplicação do modelo dicotômico da TRI à análise do Desenho da Figura Humana, pela não independência local de alguns de seus itens, tanto esse quanto o modelo politômico de TRI (correlação entre os modelos de 0,86 para o parâmetro *a* e de 0,99 para o parâmetro *b*) mostra claramente problemas de discriminação de 12 itens. Também observaram-se problemas de dificuldade dos itens, seja por serem muito fáceis (10 itens) ou por serem muito difíceis (11 itens). Tais resultados sugerem uma revisão técnica dos itens que compõem o instrumento.

Um outro tipo de resultado que merece uma atenção cuidadosa da comunidade acadêmica refere-se às diferenças de desempenho entre as amostras. Nos estudos de Wechsler, observou-se que o desempenho da amostra de 2003 (São Paulo, Minas Gerais, Bahia, Rio de Janeiro e Rio Grande do Sul) foi superior ao da amostra de 1996 (Brasília e Campinas), por tanto encontraram-se diferenças em um intervalo de sete anos. Na presente investigação, a amostra de 2003 (Belo Horizonte) também foi superior à amostra de 1996 (Campo Grande), portanto encontrou-se o mesmo fenômeno no mesmo intervalo de tempo. Apesar das análises estatísticas serem muito diferentes entre os estudos (teste *t* de *Student* no estudo de Wechsler e TRI no presente estudo), e da variedade das amostras, as evidências são fortes o suficiente para iniciar um processo amplo de investigação sobre um possível aumento do desenvolvimento gráfico infantil na população infantil brasileira. Para tanto, se faz necessário primeiro aperfeiçoar a qualidade psicométrica do instrumento. O presente estudo pretendeu caminhar nessa direção.

Referências

- Abou-Jamra, C. C. & Castilho, M. T. U. R. (1987). Testes de Colúmbia, Raven-escala especial, INV e Goodenough. Em M. Ancona-Lopez (Org.), *Avaliação da inteligência* (pp. 1-61). Vol 20-II. São Paulo: EPU.
- Almeida, R. M. (1959). Um estudo do status mental em um grupo de crianças nordestinas de idade escolar. *Boletim de Psicologia*, 38, 35-55.
- Alves, I. C. B. (1981). O teste Goodenough-Harris em pré-escolares paulistanos. *Boletim de Psicologia*, 33(80), 40-52.
- Embretson, S. E. & Reise, S. (2000). *Item response theory for psychologists*. NJ: Erlbaum Publisher.
- Hambleton, R. K. & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 12(3), 38-47.
- Kamphaus, R. W. & Pleiss, K. L. (1991). Draw-a-Person techniques: Tests in search of a construct. *Journal of School Psychology*, 29(2), 395-401.
- Lessa, L. M. P. (1953). Estudo da escala de Goodenough. *Arquivos Brasileiros de Psicotécnica*, 5(4), 107-122.
- Muraki, E. & Bock, R. D. (1997). *PARSCALE: Parametric scaling of rating data*. Chicago: Scientific Software International.
- Pasquali, L. (1999). *Instrumentos psicológicos: Manual prático de elaboração*. Brasília: LabPAM/IBAPP.

- Samejima, F. (1997). Graded response model. Em W. J. Van der Linden & R. K. Hambleton (Orgs.), *Handbook of Modern Item Response Theory* (pp. 85-100). Nova Iorque: Springer.
- Van Kolck, O. L. (1981). *Técnicas de exame psicológico e suas aplicações no Brasil: Testes de aptidão*. Petrópolis: Vozes.
- Wechsler, S. (2000). *O desenho da figura humana. Avaliação do desenvolvimento cognitivo infantil*. (2ª ed.) Campinas: Livro Pleno.
- Wechsler, S. (2003). *DFH III: O desenho da figura humana: Avaliação do desenvolvimento cognitivo de crianças brasileiras* (3ª ed.). Campinas: LAMP/PUC.
- Zausmer, A. (1954). Um estudo sobre o teste de Goodenough num grupo de pré-escolares de São Paulo. *Arquivos Brasileiros de Psicotécnica*, 6(4), 37-55.

Recebido em 09.07.2004

Primeira decisão editorial em 14.02.2005

Versão final em 05.04.2005

Aceito em 25.04.2005 ■