

## Matrizes Progressivas Avançadas: análise de itens, precisão e normas

André Luiz Moraes Ramos e colaboradores<sup>1</sup>  
Núcleo Salesiano de Pesquisas Lorena/SP

Há poucas pesquisas sobre as Matrizes Progressivas Avançadas no Brasil e em outros países. Este estudo investigou a análise de itens, consistência interna e normas. A amostra consistiu de 200 universitários do Vale do Paraíba, com idade média de 25 anos e meio. Apesar de sua boa precisão (método da biparticipação = 0,860 e KR-20 = 0,835), os resultados indicaram que a ordem e discriminabilidade dos seus itens mostraram-se deficientes. São apresentadas normas para a conversão de escores brutos em percentis e escores T de McCall.

**Palavras-chave:** Matrizes progressivas, testes psicológicos, validação.

### Abstract

There are few researches about Advanced Progressive Matrices in Brazil and in other countries. This study investigated item-analysis, internal consistency, and norms. The sample consisted of 200 undergraduates from Vale do Paraíba with an average age of 25 and a half years. In spite of its good reliability (split half = .860 and KR-20 = .835), the results indicated that the order and the discriminability of their items appeared deficient. For the conversion of raw scores into standardized T scores and percentiles, there were presented norms.

**Key words:** Progressive Matrices, Psychological tests, validation.

Apesar de ser uma das mais tradicionais, a área de avaliação e medida em Psicologia tem sido negligenciada, no Brasil, em termos de pesquisa, principalmente no que se refere aos parâmetros psicométricos de precisão e validade.

Alguns estudos sobre padronização e normas têm sido realizados por editoras, como o Centro de Psicologia Aplicada (CEPA) e a Casa do Psicólogo.

A promoção de reuniões pelos profissionais que utilizam técnicas de exame psicológico, durante as Reuniões Anuais de Psicologia de Ribeirão Preto e o Encontro sobre Testes Psicológicos organizado pelo Conselho Regional de Psicologia - 7ª região, tem permitido a troca de

informações e a implementação de estratégias para combater estas deficiências. Entretanto, só recentemente, com a criação do Laboratório de Pesquisas em Avaliação e Medida da Universidade de Brasília (PAM-UnB), é que se tem investido de forma mais sistemática, tentando reverter este atraso tecnológico.

Um dos instrumentos validados pelo PAM foram as Matrizes Progressivas de J.C. Raven, em suas formas geral e infantil. Este teste se baseia na teoria de inteligência de Spearman, para medir a habilidade de educação de relações e correlatos, através de formas figurais abstratas (Raven, 1962), e para Anastasi (1982) é a melhor medida para esta finalidade.

As Matrizes Progressivas Avançadas, também chamadas de escala universitária (Van Kolk, 1981), representam a forma mais elevada do teste de Raven, e são indicadas para adultos com habilidade mental acima da média. Destacam-se, porém, por ser a forma menos pesquisada, sendo, portanto, escolhida como objeto de investigação do presente estudo.

<sup>1</sup> Auxiliaram na pesquisa: Michéle Ferreira de Lima, Valdete Aparecida Ferreira, Anelise Fontanezi Salazar, Sandra Antônia Kazuko Yazawa, Cynthia dos Santos Ragi, Sandra Mara dos Santos Milessi, Glaucine Prado da Silva, Euni Vieira e Silva Andrade Nunes, Maria Angela da Silva, Fábio Henrique Ribeiro dos Santos e Alda Patrícia Fernandes Nunes Rangel.

Endereço para correspondência: Rua Dom Bosco, 284, CEP 12600, Lorena/SP.

A versão original de 1947 foi revisada em 1962, tomando a forma atual com duas séries: I (12 itens) e II (36 itens). Todavia, o manual apresenta poucas informações sobre os aspectos técnicos do teste.

Os dados sobre a precisão teste-reteste da versão original (1947) indicam coeficientes entre 0,76 e 0,91. Estes números, de acordo com Vernon (1983), são inferidos para a versão de 1962. Alderton e Larson (1990) obtiveram, através da análise de consistência interna pela fórmula KR-20, coeficientes de 0,812 e 0,836, em uma amostra masculina de recrutas da marinha americana.

A maioria dos estudos sobre a validade do teste consideram-no altamente saturado em um grande fator, interpretado como fator geral de inteligência (Vernon, 1981; Alderton & Larson, 1990; Zimowski & Wortke, citados por Alderton & Larson, 1990). A exceção fica por conta de Dillon, Pohlmann e Lohman (1981), que identificaram dois fatores, considerados por eles como habilidade para deduzir figuras simples de formas mais complexas (similar à dimensão denominada por Guilford como cognição de transformação figurativa-visual) e um outro fator envolvendo aspectos ligados à habilidade mecânica.

As normas contidas no manual são apresentadas em percentis (mas somente para 50, 75, 90 e 95), além de uma tabela de conversão para QI. Quanto ao Brasil, tem-se notícia de uma normatização realizada em São Paulo com 210 estudantes de Psicologia (Duarte, W.F. et al, 1983).

Neste artigo são apresentados dados relativos à análise de itens (ordem de dificuldade, distribuição por faixa percentilica de dificuldade e índice de discriminabilidade), consistência interna e normas para a interpretação dos resultados.

## Método

**Sujeitos.** Tomaram parte na pesquisa 200 universitários da região do Vale do Paraíba (Lorena, Guaratinguetá, Cruzeiro, Resende e Volta

Redonda), com idade média de 25 anos e meio (desvio-padrão de 7 anos e meio), entre 17 e 56 anos, maioria feminina (87,8%), dos cursos de Psicologia (84,3%), Pedagogia (8,1%), Administração e Economia (1,5%), Letras e Comunicação (1,5%) e Exatas (4,6%), de instituições particulares de ensino superior, curso noturno.

**Instrumento e procedimento.** Foi pesquisada a série II das Matrizes Progressivas de Raven, Escala Avançada, versão de 1962. O teste consiste em 36 itens, dispostos em ordem crescente de dificuldade, numa matriz figural 3x3, com o elemento do canto inferior direito incompleto. A tarefa do sujeito é determinar a regra das linhas e das colunas que definem a progressão figural e, então, selecionar a alternativa correta, entre 8 opções, que completa o modelo.

O teste foi aplicado coletivamente, seguindo-se as instruções do manual, iniciando-se com a série I (15 minutos) e depois passando à série II (40 minutos). Os sujeitos que não obtiveram 50% ou mais de acertos na série I tiveram seus resultados da série II desconsiderados para efeito desta pesquisa.

**Tratamento dos dados.** A análise dos itens investigou a ordem de dificuldade dos mesmos (Anastasi, 1982; Leary & Dorans, 1985; Wing, 1980), além da distribuição dos itens por faixa percentilica de dificuldade (Mollenkopf e Cerdá, citados por Almeida, 1988 e por Erthal, 1987, respectivamente). A discriminabilidade dos itens foi aferida a partir do teste de  $\chi^2$  entre grupos critérios constituídos por 30% dos sujeitos de maiores escores no teste (grupo superior) e pelos 30% com escores menores (grupo inferior), de acordo com o procedimento indicado por Guilford (1954), e pelo cálculo do poder discriminativo de cada item: IPD (Glass & Stanley, citados por Almeida, 1988; Erthal, 1987).

Quanto à precisão, foi calculada a consistência interna dos itens pelo método de biparticipação ("split-half"), através do estabelecimento de formas paralelas pela divi-

são dos itens, a partir do índice de dificuldade dos mesmos, usando-se a fórmula corretiva de Spearman - Brown (Anastasi, 1982), e pelo método de Kuder - Richardson, obtida pela fórmula KR-20 (Crombach, 1970).

Os parâmetros interpretativos foram determinados através de normas percentílicas e escore T de McCall.

## Resultados e Discussão

### Análise de itens

As Matrizes Progressivas, em suas três formas, foram concebidas como teste cujos itens estão dipostos em ordem crescente de dificuldade. Nas instruções transmitidas aos respondentes (CEPA, 1962), é mencionado que se deve tentar

Tabela 1. Distribuição dos itens por nível de dificuldade

ITEM	Índice de dificuldade (proporção de acertos)	Cerdá					Mollenkopf		
		Muito fácil	Fácil	Médio	Difícil	Muito difícil	Difícil	Médio	Fácil
		(1-.80)	(.79-.60)	(.59-.40)	(.39-.20)	(.19-0)	(>.65)	(.65-.35)	(<.35)
1	.925	*					*		
2	.935	*					*		
3	.940	*					*		
4	.860	*					*		
5	.735		*				*		
6	.795		*				*		
7	.860	*					*		
8	.800	*					*		
9	.785		*				*		
10	.770		*				*		
11	.815	*					*		
12	.760		*				*		
13	.645		*					*	
14	.715		*				*		
15	.720		*				*		
6	.635		*					*	
7	.640		*					*	
8	.450			*				*	
19	.630		*					*	
20	.575			*				*	
21	.425			*				*	
22	.400			*				*	
23	.420			*				*	
24	.330				*				*
25	.320				*				*
26	.275				*				*
27	.210				*				*
28	.210				*				*
29	.110					*			*
30	.175					*			*
31	.140					*			*
32	.145					*			*
33	.170					*			*
34	.095					*			*
35	.155					*			*
36	.050					*			*
Número de itens		7	11	5	5	8	14	9	13
Porcentagem		19	31	14	14	22	39	25	36

resolver os problemas na ordem em que se apresentam, sem omitir nenhum e sem voltar atrás, e que cada problema é um pouco mais difícil do que o anterior. A vantagem prática da hierarquização dos itens é a dos sujeitos não perderem o seu tempo com a tentativa de resolução de itens difíceis, deixando outros mais fáceis por responder (Almeida, 1988).

Entretanto, o coeficiente de correlação entre a posição do item no teste e a proporção de acertos do mesmo, apesar de ser elevado e negativo ( $r's < -0,978$ ,  $p < 0,001$ ), não é perfeito. Este dado concorda com Larson e Alderton (1990):  $r's < 0,967$ ,  $p < 0,001$ .

A distribuição dos itens por nível de dificuldade é apresentada na tabela 1, em que se observa que alguns itens se encontram fora de ordem. O item 5, por exemplo, apresentou maior dificuldade do que os sete que o sucedem, o que faz com que sua posição real seja a 12ª e não 5ª como se encontra no teste. O mesmo defeito de ordenação foi verificado, principalmente, com os itens 11, 29 e 35.

Diante destes dados, o teste fica vulnerável à seguinte situação descrita por Almeida (1988): dois sujeitos com o mesmo nível de capacidade podem apresentar resultados diferentes no teste: um, perante as dificuldades específicas de um item, persiste - por vezes em vão - na sua resolução, o outro perante essa dificuldade opta pela busca e realização de itens mais fáceis. O resultado é que o segundo sujeito será avaliado, através do instrumento, como tendo uma inteligência superior à do primeiro, quando na verdade, a diferença é devido a personalidades distintas e não ao potencial intelectual diferenciado, como alertam Leary e Dorans (1985) e Munz e Smouse (1968).

A tabela 1 apresenta, também, o grau de dificuldade dos itens de acordo com os critérios de Cerdá (citado por Erthal, 1987) e de Mollenkopf (citado por Almeida, 1988), indicando que o teste oferece um espectro de opções capaz de discriminar sujeitos em função dos níveis extremos de inteligência. Na distribuição proposta por Cerdá verifica-se que poucos são os itens com dificuldade

média (entre 0,40 e 0,59). A mesma situação se reflete quanto aos índices de dificuldade propostos por Mollenkopf, uma vez que apenas 25% dos itens têm dificuldade entre 0,35 e 0,65, comprometendo a discriminabilidade do instrumento, pois os itens que melhor discriminam o nível de inteligência do sujeito são os de dificuldade média, que nas Matrizes Progressivas Avançadas representam a minoria.

**Tabela 2.** Análise da discriminabilidade dos itens

ITEM	GRUPOS EXTREMOS		X <sup>2</sup>	IPD
	(proporção de acertos)			
	Superior	Inferior		
1	1.000	.833	10,909	.3013
2	.967	.867	3,927 *	.1644 *
3	.983	.850	6,928	.2835
4	.983	.667	20,837	.4099
5	.983	.433	43,926	.4463
6	.950	.600	21,075	.3570
7	1.000	.717	17,005	.3880
8	.900	.600	27,346	.3484
9	.983	.467	40,167	.4681
10	.983	.417	45,873	.5486
11	1.000	.567	33,191	.4436
12	.867	.500	18,639	.4170
13	.900	.417	31,158	.4041
14	.917	.417	33,750	.4496
15	.950	.483	32,175	.4552
16	.900	.300	45,000	.5331
17	.850	.483	18,150	.3172
18	.667	.200	19,286	.4312
19	.817	.417	20,306	.3312
20	.833	.300	34,751	.3940
21	.767	.117	51,400	.5299
22	.833	.033	78,190	.5726
23	.767	.183	40,936	.5125
24	.700	.133	39,634	.4512
25	.567	.083	31,947	.4677
26	.483	.067	26,123	.3851
27	.400	.117	12,570	.3269
28	.500	.017	36,578	.4464
29	.167	.067	2,911 *	.1620 *
30	.350	.117	9,130	.3165
31	.317	.017	19,440	.3607
32	.250	.067	7,566	.2706
33	.200	.117	1,563 *	14,53 *
34	.183	.033	6,988	.2570
35	.283	.050	11,760	.2337
36	.067	.017	1,878 *	13,29 *
Critérios de discriminabilidade:			$\chi^2 \geq 6,64$ ( $p < 0,01$ )	IPD $\geq 0,20$

\*Item não-discriminativo

Analisando-se especificamente a discriminabilidade de cada um dos 36 itens, através de dois procedimentos distintos, os resultados foram integralmente consistentes.

Na tabela 2, estão indicados os índices de dificuldade e de discriminação para os 36 itens avaliados, tomando-se como limiar de aceitação um IPD igual ou superior a 0,20 (Garret, citado por Almeida, 1988).

Nota-se que quatro itens não satisfazem o critério de IPD, a saber 2, 29, 33 e 36. Da mesma forma, o teste de  $\chi^2$  entre grupos extremos reflete que os quatro itens supracitados não apresentam diferença estatisticamente significativa, a nível de  $p < 0,01$ , entre os grupos superior e inferior.

O item 2 é tão fácil que mesmo as pessoas de inteligência inferior acertam, enquanto que os outros (29, 33 e 36) são tão difíceis que, por maior que seja a capacidade intelectual dos testandos, este potencial não consegue ser captado pelos itens.

### Precisão

A consistência interna das Matrizes Progressivas Avançadas foi avaliada satisfatoriamente, seja quando os itens foram tomados em grupo (método da bipartição), aferindo-se a homogeneidade dos mesmos, seja quando considerados individualmente (KR-20), medindo-se o seu grau de uniformidade ou heterogeneidade dos itens entre si, sendo obtidos, respectivamente, os coeficientes de 0,860 e 0,835. Resultados semelhantes foram obtidos por Alderton e Larson (1990).

### Normas

O estabelecimento de critérios interpretativos adequados exige um número elevado de sujeitos e representatividade da população a que o teste se destina. Esta é uma tarefa custosa.

Entretanto, a prudência aconselha que, para se lançar neste empreendimento, é necessário que o instrumento possua comprovadamente os parâmetros psicométricos já mencionados, o que envolve análise de itens e precisão, elementos constitutivos da validade da medida.

Neste sentido, como o interesse primordial deste trabalho foi analisar os itens e aferir a preci-

são das Matrizes Progressivas Avançadas, não se teve preocupação em compor uma amostra numerosa, uma vez que, para tanto, são importantes dados sobre os aspectos aqui estudados.

Mesmo assim, devido às limitações técnicas do manual no que se refere à interpretação dos resultados, são apresentadas na tabela 3 as normas percentilicas relativas a 165 alunos do curso de Psicologia da Faculdade Salesiana de Filosofia, Ciências e Letras de Lorena.

**Tabela 3.** Norma percentilica

Curso de Psicologia da Faculdade Salesiana de Filosofia, Ciências e Letras de Lorena (N = 165)

PERCENTIL	ESCORE
95	28-36
90	26-27
80	24-25
75	23
70	22
60	-
50	20-21
40	18-19
30	16-17
20	14-15
10	11-13
5	10

Apesar de ser um dos sistemas mais utilizados, o percentil é uma escala que representa apenas a posição relativa ou ordinal do indivíduo dentro do grupo normativo, não permitindo o cálculo de média, do desvio-padrão e outras medidas estatísticas deles decorrentes, além da marcante desigualdade sobretudo nos extremos da distribuição (Erthal, 1987).

Com o intuito de apresentar os resultados em escala intervalar e de indicar de forma mais precisa a extensão das diferenças nos extremos da distribuição, são apresentados na tabela 4 os escores-T de McCall (média = 50 e desvio-padrão = 10) para a amostra de estudantes de Psicologia.

**Tabela 4.** Escores Padronizados T De McCall

	ESCORE T	ESCORE BRUTO
	77	31
	72	30
<b>+2<math>\sigma</math></b>		
	69	29
	67	28
	65	27
	63	26
	61	25
<b>+1<math>\sigma</math></b>		
	59	24
	57	23
	55	22
	54	21
	51	20
	50	19
<b>X</b>		
	48	18
	46	17
	45	16
	43	15
	41	14
	40	13
<b>-1<math>\sigma</math></b>		
	39	12
	36	11
	34	10
	31	9
<b>-2<math>\sigma</math></b>		
	26	8
	25	7
	23	6

Estas normas permitem a interpretação bem diferente daquela oferecida pelo manual. Para exemplificar a distância entre estas normas e a que consta no manual, basta ver que uma pessoa com 30 anos que tivesse acertado 21 itens, pelo manual estaria no percentil 95, enquanto que pela tabela 3 seu percentil seria 50.

A superioridade dos escores aqui apresentados também se sobressai quando se comparam a média da amostra lorenense com os dados americanos do estudo de Alderton e Larson (1990): média de 18,86 no Brasil e de 17,83 e 16,49 nas duas amostras por eles pesquisadas.

## Conclusão

Esta pesquisa, ao investigar criteriosamente, aspectos relevantes às medidas psicológicas, abordou uma questão fundamental dos testes psicológicos: toda crítica pode ser construtiva se calçada em dados minuciosamente analisados.

As Matrizes Progressivas Avançadas não tinham sido, até então, no Brasil, objeto de tal análise. Sabe-se, no entanto, que particularmente nas empresas do Vale do Paraíba, no Estado de São Paulo, este teste tem sido muito utilizado em processo seletivo para cargos de nível superior. Daí pergunta-se: que garantia se tem nos resultados por ele fornecidos?

De acordo com esta pesquisa, o teste possui uma deficiência quanto à ordenação dos itens, sendo que, na ordem em que se encontram dispostos, os escores aferidos podem ser uma função de características de personalidade, por exemplo persistência, ao invés de uma medida adequada de inteligência.

Há, também, a distribuição insatisfatória de itens por nível de dificuldade, o que limita o poder discriminativo do instrumento. Além disso, há quatro itens que se mostraram incapazes de diferenciar o potencial intelectual dos indivíduos pesquisados, sendo a manutenção dos mesmos desnecessária.

Por outro lado, o teste apresentou consistência interna satisfatória, o que sugere, mas não garante, que os itens meçam um mesmo aspecto, uma mesma dimensão.

Para se pesquisar a dimensionalidade do teste é prudente que se proceda ao estudo de validade das Matrizes Progressivas Avançadas através da análise escalar de Guttman, bem como

da análise fatorial, pelo método dos componentes principais.

Em virtude das limitações quanto ao tamanho e a diversidade da amostra, os resultados obtidos neste trabalho sugerem cautela quanto ao uso das Matrizes Progressivas Avançadas, até que se realizem investigações complementares sobre este importante teste psicológico.

## Referências

- ALDERTON, D.L. & LARSON, G.E. Dimensionality of Raven's Advanced Progressive Matrices items. *Educational and Psychological Measurement*, 1990, **50**, 887-900.
- ALMEIDA, L.S. **O raciocínio diferencial dos jovens**. Porto: Instituto Nacional de Investigação Científica, 1988.
- ANASTASI, A. **Psychological testing**. New York: Macmillan, 1982.
- CEPA. **Manual do teste Matrizes Progressivas - Escala Avançada**. Rio de Janeiro: CEPA, 1965.
- CROMBACH, L.J. **Essentials of psychological testing**. New York: Harper & Row, 1970.
- DILLON, R. F.; POHLMANN, J.T. & LOHMAN, D.F. A factor analysis of Raven's Advanced Progressive Matrices freed of difficulty factors. *Educational and Psychological Measurement*, 1981, **41**, 1295-1302.
- DUARTE, W.F. et al. Um estudo do teste Raven-Escala Avançada numa amostra de estudantes de Psicologia. **Resumos da 35ª Reunião Anual da Sociedade Brasileira para o Progresso da Ciência**, 1983, 804.
- ERTHAL, T.C. **Manual de psicometria**. Rio de Janeiro: Zahar, 1987.
- GUILFORD, J. P. **Psychometric methods**. New York: McGraw-Hill, 1954.
- LEARY, L.F. & DORANS, N.J. Implications for altering the context in which test items appear: A historical perspective on an immediate concern. *Review of Educational Research*, 1985, **55**, 387-413.
- MUNZ, D.C. & SMOUSE, A.D. Interactions effects of item-difficulty sequence and achievement-anxiety reaction on academic performance. *Journal of Educational Psychology*, 1968, **59**, 370-374.
- RAVEN, J.C. **Advanced Progressive Matrices, sets I and II**. London: H.K. Lewis, 1962.
- VAN KOLCK, O.L. **Técnicas de exame psicológico e suas aplicações no Brasil**. Petrópolis: Vozes, 1981.
- VERNON, P.A. Speed of information processing and general intelligence. *Intelligence*, 1983, **7**, 53-70.
- VERNON, P.A. Advanced Progressive Matrices. In KEYSER, D.J. & SWEETLAND, R.C. **Test critiques**. Kansas City, Mis: Test Corporation of America, 1984, 47-50.
- WING, H. Practice effects with traditional mental items. *Applied Psychological Measurement*, **4**, 141-155.