

Estudo de Validação da Bateria de Provas de Raciocínio (BPR-5)¹

Ricardo Primi²

Universidade São Francisco - Brasil

Leandro S. Almeida

Universidade do Minho - Portugal

RESUMO - Esse estudo investiga a validade e precisão da BPR-5, versão reduzida da Bateria de Provas de Raciocínio Diferencial (BPRD) de Almeida (1988), composta por cinco subtestes: Raciocínio Abstrato, Raciocínio Verbal, Raciocínio Numérico, Raciocínio Espacial e Raciocínio Mecânico. A BPR-5 foi aplicada em 1243 alunos brasileiros ($N=771$) e portugueses ($N=472$) da sexta série do ensino fundamental até a terceira série do ensino médio. Os coeficientes de consistência interna variaram de 0,62 a 0,84 e os de precisão pelo método das metades de 0,65 a 0,87. Encontrou-se um único fator explicando aproximadamente 55% da variância representando uma medida composta de inteligência fluida, cristalizada, processamento visual e habilidade quantitativa e conhecimento prático de mecânica. As correlações da BPR-5 com as notas escolares foram, no geral, positivas chegando a atingir 0,54 ($p < 0,01$). Conclui-se que a BPR-5 constitui um instrumento eficiente e rápido para avaliação simultânea do raciocínio geral e das aptidões podendo ser útil aos profissionais em suas diversas áreas de atuação.

Palavras-chave: avaliação da inteligência; estudo de validação; habilidades cognitivas

Validity of a Battery of Reasoning Tests (BPR-5)

ABSTRACT - This research investigates reliability and validity of the Reasoning Battery Tests (BPR-5), a reduced version of the Battery of Differential Reasoning Tests (BPRD; Almeida, 1988). It is comprised by five subtests: Abstract Reasoning, Verbal Reasoning, Numerical Reasoning, Spatial Reasoning and Mechanical Reasoning. The participants were 1243 sixth grade to high school students from Brazil ($N=771$) and Portugal ($N=472$). The internal consistency coefficients varied from 0,62 to 0,84 and split-half coefficients 0,65 to 0,87. The factor analysis indicated an unidimensional solution accounting for approximately 55% of the total variance representing a composite score based on fluid intelligence, crystallized intelligence, visual processing ability, quantitative ability and practical-mechanical knowledge. The criterion validity coefficients based on the correlation with school success were around 0,54 ($p < 0,01$). It was concluded that the BPR-5 constitutes an efficient and fast way to assess general and specific abilities and could be a valuable instrument for professionals in their assessment needs.

Key words: intelligence assessment; validity studies; psychological assessment.

Em muitas situações profissionais tais como na orientação profissional, avaliação das dificuldades de aprendizagem e seleção de pessoal o psicólogo necessita tomar decisões sustentadas em avaliação das aptidões. Obviamente, a qualidade dessas decisões está fortemente condicionada à qualidade dos instrumentos de medida disponíveis e à competência técnica e ética do profissional. Em relação ao primeiro aspecto observa-se no Brasil uma carência de instrumentos de avaliação das aptidões. Os testes disponíveis,

tradicionalmente usados, para este fim são a Bateria CEPA e o DAT (Bennett, Seashore & Wesman, 1947; Rainho, n.d.) obviamente desatualizados. Essa carência motivou este trabalho que apresenta o estudo de validação de um teste de aptidão chamado Bateria de Provas de Raciocínio (BPR-5) de Almeida e Primi (1998).

A BPR-5 originou-se da Bateria de Provas de Raciocínio Diferencial (BPRD) construída por Almeida (1986, 1988) que por sua vez originou-se dos Testes de Raciocínio Diferencial de Meuris (1969). Ela é composta por cinco subtestes: Raciocínio Abstrato (RA), Raciocínio Verbal (RV), Raciocínio Numérico (RN), Raciocínio Espacial (RE) e Raciocínio Mecânico (RM), em duas formas: Forma A, para alunos da sexta a oitava séries do Ensino Fundamental e Forma B, para alunos da primeira à terceira séries do Ensino Médio.

A BPR-5 fundamenta-se nas concepções fatoriais mais recentes da inteligência (Carroll, 1997) possibilitando a avaliação simultânea do fator g (Spearman, 1927) e de fatores mais específicos (Thurstone, 1938).

Com relação aos fatores específicos e adotando a terminologia e definições apresentadas por Carroll (1997) e Horn (1991) e Woodcock (1990), o subteste RA associa-se prin-

- 1 Este trabalho faz parte do Estudo de Validação e Padronização da Bateria de Provas de Raciocínio financiado pela Casa do Psicólogo e pelo projeto de colaboração interinstitucional das Universidade São Francisco e Universidade do Minho. Agradecemos a inestimável colaboração dos alunos: Cassio Ricardo Fares Riedo, Roseli Filizatti, Tristana Cesaretto, Claudinéia Aparecida Ferreira de Godoi, Romilda Simões de Queiroz, Cristiane Jardim Girioli, Rosangela Scrich, Kelly Fiorelli Ferro, todos do Curso de Psicologia do Centro de Ciências Humanas e Sociais da Universidade São Francisco e das alunas Teresa Oliveira Martins, Ana Antunes do Instituto de Educação e Psicologia da Universidade do Minho que participaram na coleta e organização dos dados.
- 2 Endereço: Laboratório de Avaliação Psicológica e Educacional - LabAPE, Universidade São Francisco, Mestrado em Psicologia, Rua Alexandre Rodrigues Barbosa, 45, CEP 13251-900, Itatiba - SP, E-mail: rprimi@uol.com.br.

principalmente à inteligência fluida (Gf) definida como capacidade de raciocinar em situações novas, criar conceitos e compreender implicações. O subteste RV associa-se à inteligência fluida e à inteligência cristalizada (Gc), definida como extensão e profundidade do conhecimento verbal vocabular, e à capacidade de raciocinar utilizando conceitos previamente aprendidos. O subteste RN associa-se à inteligência fluida e em parte à habilidade quantitativa (Gq) definida como a compreensão de conceitos quantitativos básicos como soma, subtração, multiplicação, divisão e manipulação de símbolos numéricos. O subteste RE associa-se em parte à inteligência fluida, mas, principalmente, à capacidade de processamento visual (Gv) definida como a habilidade de representar e manipular imagens mentais. O subteste RM associa-se em parte à inteligência fluida e aos conhecimentos práticos mecânicos.

Como pode ser observado todos os subtestes associam-se, em maior ou menor grau, à inteligência fluida, habilidade que é mais semelhante ao fator *g* de Spearman (1927). O formato da maioria dos subtestes (analogias nas provas RA e RV e séries nas provas RN e RE) requer o raciocínio indutivo e por isso essas tarefas requerem conhecimentos de procedimentos gerais de raciocínio como a codificação de informação, ou seja, recuperação de atributos relevantes da memória de longo prazo para a solução do problema (como nas provas RV e RN) e/ou a seleção de atributos relevantes a partir da percepção (como nas provas RA e RE), a indução de relações abstratas entre essas informações e aplicação dessas relações para a criação de novas idéias. Também requer o gerenciamento metacognitivo composto por procedimentos de planejamento de estratégias de resolução, execução e controle da implementação dessas estratégias, monitoramento da implementação e revisão de estratégias quando necessário (Sternberg, 1984, 1986).

A adaptação desta bateria para uso no Brasil iniciou-se em 1995 quando desenvolvemos um estudo das propriedades psicométricas da BPRD o qual foi publicado na IV Conferência Internacional de Avaliação Psicológica (Primi, Almeida & Lucarelli, 1996). A BPRD é amplamente usada em Portugal desde a década de 80 a partir do que vários estudos se acumularam evidenciando sua precisão e validade (Almeida, 1987, 1989, 1992 e 1995; Almeida & Campos, 1986a e 1986b; Almeida & Costa, 1989; Almeida & Martins, 1996) e a versão para o Brasil também demonstrou possuir boas qualidades psicométricas (Primi e cols., 1996). A partir desses resultados promissores vislumbrou-se a possibilidade de diminuir o número de itens criando-se uma versão reduzida da bateria.

Não é pacífica, entre os pesquisadores, a questão do tamanho dos testes. Se por um lado se justifica a redução no número de itens por razões associadas ao menor tempo e recursos envolvidos, que também terá reflexos positivos nos alunos avaliados, por outro as propriedades psicométricas de uma prova psicológica demasiado curta podem estar aquém do desejado. Podemos pensar que testes mais longos possibilitam melhor discriminação dos traços e desempenhos avaliados (sensibilidade), assim como tendencial-

mente favorecem a obtenção de índices de precisão mais elevados. Neste caso, o recurso a adaptações na fórmula corretiva de Spearman-Brown, usada no cálculo da precisão pelo método das metades (Anastasi & Urbina, 1997), permite-nos estimar o número de itens para um teste em face de um índice de precisão que se espera atingir. Por último, podemos pensar que o número de itens está igualmente associado à validade do teste. A idéia assenta no pressuposto de que podemos ter uma melhor representação de um traço se um maior número de indicadores comportamentais do traço forem incluídos na avaliação.

Evidentemente não é correto centrar a discussão exclusivamente no número de itens, pois o mais importante para as propriedades psicométricas da prova passa pela qualidade dos mesmos. A Teoria de Resposta ao Item (Hambleton, Swaminatham & Rogers, 1991), e todas as suas aplicações como a avaliação adaptativa, ilustra como podemos fazer avaliações mais discriminativas, precisas e válidas recorrendo a um pequeno conjunto de itens tidos como os mais informativos em relação às características psicológicas do sujeito (Embretson, 1996). Obviamente essa metodologia, bem recente na construção e validação de testes psicológicos, pressupõe aplicações individuais e preferencialmente informatizadas.

Nas situações de aplicação coletiva de testes "papel-lápis" poder-se-á repensar o tamanho do teste à luz da questão seguinte: serão todos os itens igualmente importantes e necessários para uma determinada avaliação? Um teste planejado para avaliar alunos de uma ampla faixa de níveis escolares, como da sexta série do ensino fundamental até a terceira série do ensino médio deverá incluir desde itens mais adequados à avaliação dos alunos dos primeiros níveis dessa faixa escolar até itens mais adequados à avaliação dos últimos níveis. Entretanto os itens de um extremo, por exemplo, adequados à avaliação dos alunos do ensino médio, proporcionarão pouca ou nenhuma informação para a avaliação dos alunos de outro extremo, como, por exemplo, os alunos da sexta série do ensino fundamental. Portanto a manutenção de uma única prova, para esses seis níveis, implicará maior número de itens e, conseqüentemente, tempos de avaliação mais longos, mesmo que parte da informação não venha a ser relevante, sobretudo pensando-se nos alunos que se situam nos níveis extremos.

Nas décadas de 50 e 60 vários autores investigaram o impacto da redução do número de itens na precisão e validade de alguns testes de inteligência geral e de aptidões cognitivas (Davis & Fifer, 1959; Merenda, Hall, Clarke & Pascale, 1962; Clarke, 1960). Em alguns dos estudos comparou-se os parâmetros psicométricos (precisão e validade) do Teste de Aptidão Diferencial (DAT, Bennett e cols. 1947) com os parâmetros de uma segunda bateria formada por oito provas similares, no entanto com um número menor de itens e, conseqüentemente, com um tempo menor de aplicação (*Measurement of Skills Tests, MOS*; Clarke, 1960). As discrepâncias encontradas entre as duas baterias não foram relevantes (Merenda e cols., 1962; Merenda, Clarke & Jacobsen, 1965), observando-se inclusive melhor capa-

cidade preditiva dos resultados escolares ao longo de três anos letivos por parte da bateria abreviada (Merenda, Jacobsen & Clarke, 1969).

Com base nessas evidências buscou-se desenvolver uma versão reduzida da BPRD organizando-a em duas formas (Forma A, para alunos da sexta à oitava séries do Ensino Fundamental e Forma B, para alunos da 1ª à 3ª séries do Ensino Médio) contendo um conjunto de itens selecionados com base na sua eficácia em função da adequação da dificuldade e do poder discriminativo. Um relato mais detalhado da seleção dos itens e de alguns resultados preliminares das propriedades psicométricas da versão reduzida foi apresentado em Almeida, Antunes, Martins e Primi (1997). Esse conjunto de itens constituiu a BPR-5 a qual, portanto, é uma versão reduzida da BPRD. A versão completa também está sendo estudada no Brasil na avaliação das aptidões de estudantes universitários por Pasquali e Araújo (1998) que a denominaram Bateria de Raciocínio Diferencial (BRD). O propósito deste estudo, portanto, é apresentar os primeiros resultados sobre as evidências de validade e precisão da BPR-5 simultaneamente para as populações brasileira e portuguesa.

Método

Participantes

A amostra foi composta por 1243 alunos, portugueses e brasileiros, cuja distribuição segundo a forma do teste que responderam, o nível de escolaridade, o gênero e o país, é apresentada na Tabela 1. Com relação ao Brasil os alunos eram, na sua maioria, provenientes de escolas do sistema público estadual de ensino (83%) do período diurno (65%) das cidades de Indaiatuba, Itatiba, Jundiá e Mogi Guaçu do interior do Estado de São Paulo. Com relação a Portugal os alunos eram todos provenientes de escolas públicas do período diurno dos distritos de Braga e Viana do Castelo. No sistema português, no ensino secundário (10o. ao 12o.

ano), os alunos escolhem uma dentre quatro áreas: científico-natural, artes, econômico-social ou humanidades (Ministério da Educação, 1998). Para garantir a representatividade da amostra optou-se por escolher alunos de duas áreas distintas: científica e humanidades.

Materiais

A BPR-5 (Almeida & Primi, 1998) é constituída por duas formas (A e B), com cinco subtestes cada, com o mesmo número de itens descritos em seguida e exemplificados na Figura 1. A Forma A aplica-se aos estudantes da sexta à oitava série do ensino fundamental e a Forma B aos alunos da primeira à terceira série do ensino médio. As formas A e B compartilham itens para que no futuro possam ser feitos estudos de equiparação das notas.

Prova de Raciocínio Abstrato (Prova RA): Prova composta por 25 itens (19 itens são comuns às Formas A e B) de conteúdo abstrato, envolvendo analogia com figuras geométricas, ou seja, A:B // C: (A, B, C, D, E). É necessário que se descubra a relação existente entre os dois primeiros termos e aplicá-la ao terceiro, para se identificar o quarto termo entre as 5 alternativas de resposta. O tempo limite é de 5 minutos.

Prova de Raciocínio Verbal (Prova RV): Prova composta por 25 itens (18 itens são comuns às Formas A e B) envolvendo analogia, como a prova de raciocínio abstrato, contudo entre palavras. A relação analógica existente entre um primeiro par de palavras deverá ser descoberta e aplicada de forma a identificar a quarta palavra entre as cinco alternativas de resposta que mantém a mesma relação com uma terceira apresentada. O tempo limite é de 4 minutos.

Prova de Raciocínio Espacial (Prova RE): Prova composta de 20 itens (12 itens são comuns às Formas A e B) nos quais existem séries de cubos tridimensionais em diferentes posições que indicam movimento. Os movimentos podem ser constantes, por exemplo sempre para a direita, ou alternados, por exemplo para esquerda e para cima. Por meio da análise das diferentes faces pode-se descobrir o cubo que se seguiria se o movimento descoberto fosse aplicado ao último cubo da série. Existem cinco alternativas de resposta e o tempo limite é de 8 minutos.

Prova de Raciocínio Numérico (Prova RN): Prova composta por 20 itens (12 itens são comuns às Formas A e B) nos quais existem séries de números; o sujeito deve descobrir qual relação aritmética rege as progressões nas séries e aplicá-la respondendo quais seriam os dois últimos números que completariam a série. O tempo limite é de 9 minutos.

Prova de Raciocínio Mecânico (Prova RM): Prova composta por 25 itens (19 itens comuns às Formas A e B) constituídos por gravuras que retratam um problema e opções de resposta. As questões são compostas por problemas práticos que envolvem conteúdos físico-mecânicos. A resposta é dada escolhendo-se, dentre as alternativas, a resposta que melhor responde a questão proposta pelo problema. O tempo limite é de 7 minutos

Tabela 1 - Distribuição da Amostra Segundo Forma da BPR-5, Escolaridade, Sexo e País.

Forma	Escolaridade		Sexo	Portugal	Brasil
	Brasil	Portugal		N	N
A	6a. série	7o. ano	M	37	33
			F	30	41
	7a. série	8o. ano	M	19	36
			F	45	31
8a. série	9o. ano	M	23	25	
		F	47	28	
B	1a. série	10o. ano	M	51	55
			F	45	117
	2a. série	11o. ano	M	35	23
			F	60	83
3a. série	12o. ano	M	33	78	
		F	47	221	
Total				472	771

Procedimento

As aplicações foram feitas coletivamente após a permissão da diretoria das escolas e o aceite dos alunos em participar da pesquisa. Os autores contaram com a colaboração de oito alunos de graduação em Psicologia (dois em Portugal e seis no Brasil) que foram treinados para coordenar a aplicação da bateria. O procedimento de aplicação incluía uma breve introdução para a explicação dos objetivos do estudo, a leitura das instruções em voz alta e o esclarecimento de dúvidas. As provas foram aplicadas com tempo limitado conforme indicado anteriormente. Após o esclarecimento das dúvidas o aplicador instruíu os alunos a começarem e iniciava a contagem do tempo. Esgotando-se o tempo o aplicador instruíu os alunos a interromperem a resolução e passarem para o próximo subteste. O aplicador iniciava novamente a leitura das instruções e o esclarecimento de dúvidas e procedia da mesma maneira até que todos os subtestes fossem aplicados. A ordem de apresentação dos subtestes foi RA, RV, RE, RN e RM que eram aplicados em uma aula de aproximadamente 50 minutos.

Resultados

Como orientam os Parâmetros para Testes Educacionais e Psicológicos (American Psychological Association, 1985), a qualidade dos instrumentos de avaliação está condicionada à demonstração de evidências de precisão e validade. Seguindo essas orientações apresentaremos as evidências de precisão pelo método da consistência interna e pelo método das metades e as evidências de validade com referência ao construto e ao critério.

Na Tabela 2 apresenta-se as médias, desvios padrão, precisão pelo método das metades (*Split Half*), precisão por consistência interna calculado a partir dos coeficientes de correlação de *Pearson* (*Alfa 1*) e consistência interna calculada a partir dos coeficientes de correlação tetracórica (*Alfa 2*).

A análise dos itens (índice de dificuldade e poder discriminativo) indicou que, de um modo geral, as provas cumpriram os requisitos necessários para a composição de escalas precisas. Verificou-se que os itens apresentavam variância adequada e se correlacionavam com o escore to-

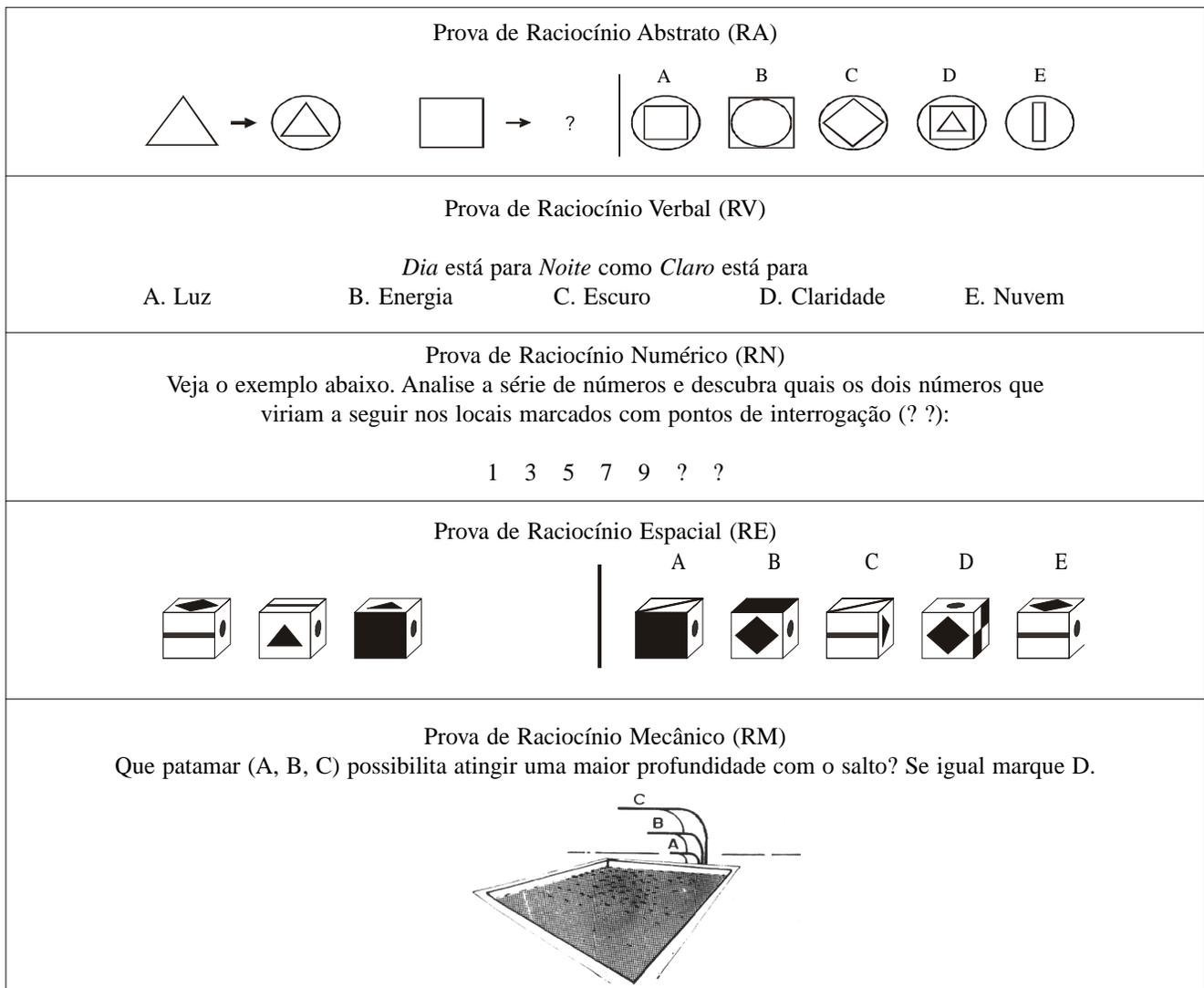


Figura 1 - Exemplos de itens dos cinco subtestes da BPR-5.

Tabela 2 - Estatísticas Descritivas e Coeficientes de Precisão da BPR-5

Estatísticas Descr. e Precisão	Brasil									
	Prova RA		Prova RV		Prova RN		Prova RE		Prova RM	
	A	B	A	B	A	B	A	B	A	B
Média	8,53	8,64	9,64	10,89	4,59	6,01	6,01	5,76	6,16	5,76
Desvio padrão	4,14	3,54	3,58	3,52	2,60	3,04	3,04	2,93	2,85	2,68
Split Half	0,86	0,84	0,80	0,77	0,77	0,82	0,67	0,71	0,72	0,65
Alfa 1	0,82	0,79	0,74	0,72	0,75	0,79	0,62	0,66	0,68	0,62
Alfa 2	0,89	0,88	0,85	0,86	0,87	0,89	0,75	0,77	0,80	0,77

Estatísticas Descr. e Precisão	Portugal									
	Prova RA		Prova RV		Prova RN		Prova RE		Prova RM	
	A	B	A	B	A	B	A	B	A	B
Média	10,36	11,39	11,09	14,10	6,73	9,01	6,56	8,82	7,31	8,77
Desvio padrão	4,57	3,18	4,06	3,92	3,69	3,61	3,31	3,70	2,31	3,40
Split-Half	0,80	0,72	0,84	0,78	0,87	0,85	0,74	0,78	0,68	0,71
Alfa 1	0,82	0,72	0,77	0,77	0,84	0,82	0,69	0,77	0,68	0,70
Alfa 2	0,89	0,85	0,90	0,88	0,94	0,92	0,81	0,87	0,80	0,86

Alfa 1: baseado no coeficiente de correlação de Pearson; *Alfa 2*: baseado no coeficiente de correlação tetracórico.

tal. Isto indica que os itens estão consistentes dentro dos subtestes e contribuem coerentemente para a variância do escore produzindo coeficientes de precisão adequados como mostra a Tabela 2. Os valores mais baixos foram observados nas provas RM e RE.

Um ponto importante a ser comentado refere-se ao baixo número de acertos, principalmente nas provas RN, RE e RM entre os alunos brasileiros. O baixo número de acertos pode ter ocorrido pelo tempo reduzido impossibilitando a resposta à um número razoável de itens. Geralmente os últimos itens tinham proporções de acerto muito baixas o que gerou uma grande heterogeneidade nos índices de dificuldade. O problema, nestes casos, é que a amplitude de variação dos coeficientes de correlação é afetada pelo grau de heterogeneidade e assimetria da distribuição nas variáveis dicotômicas. Supondo que dois itens, de dificuldades diferentes e, portanto com distribuições heterogêneas, estejam medindo uma mesma variável espera-se que, se eles estiverem correlacionados ao máximo, o padrão de resposta seja tal que os sujeitos que acertam o item mais difícil acertam necessariamente o item mais fácil. Para um padrão como este se espera encontrar um coeficiente de correlação igual a um. No entanto usando-se o coeficiente Phi nesta situação o valor não será igual a um. O valor específico que ele poderá atingir será menor que um e dependerá do grau de heterogeneidade nas distribuições. Isto ocorre porque o coeficiente relaciona-se à previsibilidade de uma variável em função da outra e como as distribuições são diferentes, nunca o acerto em um item poderia ser perfeitamente previsto a partir do acerto no outro item. Mais detalhes sobre esta discussão podem ser encontrados em Ferguson (1959), Primi e Almeida (1998), Panter, Kimberly, Swygert e Dahlstrom, (1997); Waller, Tellegen, McDonald e Lykken (1996).

Por causa deste problema decidiu-se calcular, também os coeficientes de precisão com base nas correlações tetracóricas (Alfa 2) que não são afetados pela heterogenei-

dade de distribuição das variáveis (Primi & Almeida, 1998; Waller e cols., 1996). Observa-se que, por este método, todos os coeficientes são iguais ou maiores do que 0,75 atingindo, portanto níveis satisfatórios de precisão.

Como apresentamos na introdução, os cinco subtestes fazem exigência à inteligência fluida (Gf) e, portanto esperava-se encontrar um único fator explicando as correlações entre os subtestes. Para investigar a hipótese de que as diferentes provas estão medindo, através de conteúdos diferentes, uma única dimensão cognitiva efetuou-se a análise fatorial da matriz de correlação entre os cinco subtestes.

Na Tabela 3 apresentamos as matrizes de correlação entre as provas e os resultados da análise fatorial por componentes principais. Como pode ser observado há somente um fator com *eigenvalue*³ maior que um, responsável por pouco mais de 50% da variância dos escores confirmando a expectativa.

Na época da coleta de dados as notas escolares de parte da amostra foram obtidas e correlacionadas com a BPR-5 com o intuito de verificar a validade da bateria em prever o desempenho acadêmico. Esperava-se encontrar associações significativas entre a BPR-5 e as notas escolares uma vez que o desempenho acadêmico está relacionado ao raciocínio. As Tabelas 4, 5 e 6 apresentam os coeficientes encontrados.

Nota-se que, de um modo geral, os valores dos coeficientes são moderados e, na sua maioria significativos favorecendo a hipótese levantada.

3 Algebricamente, as matrizes de correlação podem ser decompostas em duas outras matrizes uma contendo os *eigenvalues* (autovalores) e outras os *eigenvectors* (autovetores). A primeira matriz terá tantos *eigenvalues* quantos forem os fatores extraídos da análise e a magnitude do *eigenvalue* está associada a variância que ele conseguiu condensar. Portanto o *eigenvalue* associa-se a importância do fator pois, dizendo de maneira não muito precisa, associa-se a quantas variáveis ele consegue agrupar.

Tabela 3 - Matriz de Correlações entre os Subtestes (Valores Acima e Abaixo da Diagonal respectivamente para Portugal e Brasil) e Resultados da Análise Fatorial.

	Forma A					Brasil		Portugal	
	RA	RV	RN	RE	RM	Cargas no Fator 1	h^2	Cargas no Fator 1	h^2
RA		0,44	0,50	0,49	0,51	0,74	0,55	0,81	0,66
RV	0,50		0,32	0,44	0,42	0,76	0,57	0,70	0,50
RN	0,40	0,36		0,46	0,34	0,66	0,44	0,71	0,50
RE	0,40	0,39	0,36		0,32	0,73	0,53	0,74	0,55
RM	0,31	0,40	0,27	0,43		0,67	0,44	0,69	0,49
					<i>Eigenvalue</i> (Prop. Expl.)	2,54 (50,8%)		2,71 (54,2%)	
	Forma B					Brasil		Portugal	
	RA	RV	RN	RE	RM	Cargas no Fator 1	h^2	Cargas no Fator 1	h^2
RA		0,44	0,49	0,48	0,43	0,76	0,58	0,75	0,57
RV	0,49		0,42	0,32	0,37	0,73	0,53	0,66	0,43
RN	0,45	0,45		0,56	0,49	0,79	0,62	0,79	0,63
RE	0,48	0,48	0,60		0,56	0,80	0,64	0,78	0,62
RM	0,44	0,35	0,43	0,41		0,69	0,47	0,76	0,58
					<i>Eigenvalue</i> (Prop. Expl.)	2,84 (56,9%)		2,83 (56,7%)	

Os valores das correlações acima da diagonal referem-se a amostra portuguesa e abaixo à amostra brasileira. Todos os coeficientes são significativos a $p < 0,01$. A notação h^2 refere-se à comunalidade.

Observa-se que as correlações significativas do escore geral (EG), calculado pela soma dos escores nos cinco subtestes, com as notas escolares variaram de 0,23 a 0,54, isto é, indicando que a variância compartilhada entre as duas variáveis atingiu até aproximadamente 29%.

Observa-se também indícios de que quando o conteúdo dos subtestes e da disciplina escolar se aproximam as correlações tendem a ser maiores. O subteste RV demonstrou possuir as correlações mais altas com as notas escolares, principalmente, com o desempenho na disciplina Português. Observe-se que na amostra brasileira (Tabela 6), as provas RN e RE parecem se aproximar de disciplinas de matemática, física e química.

Discussão

O presente artigo apresenta os resultados do estudo intercultural de validação Bateria de Provas de Raciocínio (BPR-5).

Com relação à precisão foram encontrados coeficientes aceitáveis principalmente nas provas RA, RV e RN. Nas provas RE e RM, embora razoáveis, os coeficientes tenderam a ser mais baixos. Na prova RM o resultado encontrado é coerente com estudos anteriores nos quais se tem obtido coeficientes semelhantes (Almeida, 1988). Este fato indica que os itens da prova RM possivelmente são mais heterogêneos, quanto aos conhecimentos práticos mecânicos que requerem, levando à baixa consistência interna.

Quanto ao subteste RE o coeficiente foi mais baixo do que se esperava. O subteste RN também geralmente apresenta coeficientes mais altos do que os encontrados nesse estudo (Almeida, 1988; Primi, Almeida, & Lucarelli, 1996).

Provavelmente esses resultados estão indicando que o tempo fixado para estes subtestes é muito reduzido. Nestes subtestes os alunos demoram mais para responder aos itens e, diante de um tempo reduzido, eles não respondem a uma quantidade razoável de itens necessária à uma estimação mais precisa de sua habilidade. De fato observou-se que uma proporção considerável dos alunos não chegou a responder os últimos itens. Os comentários dos alunos ao fim das provas eram de que o tempo era muito curto. Como a omissão foi tratada como erro os últimos itens tiveram índices de dificuldade mais baixos culminando em uma distribuição dos índices de dificuldade heterogênea limitando a extensão dos coeficientes de correlação de Pearson entre os itens e, conseqüentemente, os coeficientes de consistência interna. Quando esses coeficientes de correlação são calculados a partir de uma fórmula alternativa (coeficiente tetracórico), cuja extensão não é afetada pela heterogeneidade, os coeficientes de consistência interna são mais altos

Tabela 4 - Correlações entre a BPR-5 (Forma A) e as Notas Escolares a Partir da Amostra Portuguesa

Subtestes	Notas Escolares				
	Matemática N=172	Português N=200	Educação Visual N=201	Físico - Química N=201	Média das notas
RA	0,37**	0,34**	0,26**	0,39**	0,39**
RV	0,53**	0,49**	0,29**	0,52**	0,52**
RM	0,16*	0,10	0,10	0,17*	0,16*
RE	0,42**	0,45**	0,33**	0,47**	0,47**
RN	0,29**	0,41**	0,23**	0,39**	0,39**
EG	0,51**	0,50**	0,33**	0,23*	0,54**

EG escore composto pela soma dos subtestes. * $p < 0,05$; ** $p < 0,01$

Tabela 5 - Correlações entre a BPR-5 (Forma B) e as Notas Escolares a Partir da Amostra Portuguesa

Notas	Notas Escolares					
	Matemática N=177	Português N=264	Desenho Téc. Cient. N=21	Filosofia N=192	Informática N=67	Psicologia N=51
RA	0,35**	0,30**	0,37	0,36**	0,10	0,40**
RV	0,31**	0,35**	0,50*	0,31**	-0,03	0,21
RM	0,21**	0,30**	0,39	0,30**	0,21	0,37**
RE	0,27**	0,29**	0,66**	0,28**	0,27*	0,40**
RN	0,27**	0,31**	0,30	0,33**	0,32**	0,10
EG	0,36**	0,41**	0,54*	0,42**	0,26*	0,35*

EG score composto pela soma dos subtestes. * $p < 0,05$; ** $p < 0,01$

Tabela 6 - Correlações entre a BPR-5 (Forma B, N=97) e as Notas Escolares a Partir da Amostra Brasileira

Notas	Notas Escolares					
	NG	Português	Matemática	Física	Química	Biologia
RA	0,35**	0,23*	0,40**	0,39**	0,40**	0,32**
RV	0,51**	0,45**	0,38**	0,38**	0,49**	0,38**
RN	0,23*	0,17	0,37**	0,32**	0,21*	0,18
RE	0,20	0,08	0,26*	0,36**	0,17	0,15
RM	0,01	0,00	0,17	0,13	0,01	0,06
EG	0,38**	0,28**	0,44**	0,44**	0,38**	0,31**
Notas	Geografia	História	Filosofia	Inglês	Artes	
	RA	0,37**	0,17	0,03	0,24*	0,07
RV	0,40**	0,43**	0,32**	0,30**	0,28**	
RN	0,12	0,21*	-0,09	0,15	0,01	
RE	0,20*	0,10	-0,13	0,17	0,09	
RM	-0,01	-0,07	-0,17	-0,08	0,01	
EG	0,33**	0,25*	0,01	0,23*	0,14	

EG score composto pela soma dos subtestes e NG média das notas escolares. * $p < 0,05$; ** $p < 0,01$

confirmando a suspeita sobre o tempo reduzido (Primi & Almeida, 1998).

De um modo geral os coeficientes atingiram níveis aceitáveis e poderão ser melhorados se o limite de tempo for aumentado. Portanto, sugere-se uma nova definição desses limites antes de prosseguir com o estudo de padronização da BPR-5.

Com relação à validade de construto, a BPR-5 mantém-se coerente com seu propósito, ou seja, avaliar uma única dimensão a partir de diferentes conteúdos. Há um único fator responsável pela maior parte da variação entre os escores nas cinco provas e este fator associa-se a uma medida composta pelas inteligências fluida, cristalizada, processamento visual e habilidade quantitativa. A dimensão central responsável pelas correlações entre essas provas refere-se a capacidade de raciocinar em situações novas criar conceitos e compreender implicações (Horn, 1991; Woodcock, 1990).

Com relação à validade de critério verificou-se que os subtestes têm correlações moderadas com as notas escolares. Observou-se que em alguns momentos 29% da variância das notas é compartilhada com as diferenças de raciocínio medidas pela BPR-5. Podemos interpretar que a covariância existente entre os subtestes da BPR-5 e as notas escolares

seria aquela correspondente aos processos de raciocínio necessários igualmente nas provas e nas atividades de aprendizagem escolar. Nesse sentido não devemos esperar altos coeficientes de correlação entre as notas da BPR-5 e as notas escolares já que, embora relacionadas ao raciocínio, a variância das notas escolares não indica exclusivamente diferenças nessa capacidade. A nota escolar pode indicar, por exemplo, que um determinado conteúdo foi memorizado, que o aluno se esforçou durante o ano, que o aluno tem interesse em determinada disciplina à ainda outros fatores desconhecidos.

O subteste que demonstrou estar mais associado com as notas foi o RV. Esse resultado é coerente, uma vez que o componente específico avaliado por ele é a inteligência cristalizada referindo-se à extensão e profundidade do conhecimento conceitual-verbal adquirido, principalmente, nas experiências educacionais. Portanto, entre os subtestes da BPR-5 o subteste RV é o que possui conteúdo mais próximo às tarefas escolares.

Em síntese, mesmo com alguns coeficientes de precisão mais baixos do que a expectativa inicial, as evidências de validade da BPR-5 demonstram que os coeficientes atingiram níveis aproximadamente idênticos, aos encontrados com a BPRD (Almeida, 1995; Almeida & Campos, 1986b).

Embora, em alguns subtestes, a parcela confiável da variância dos escores observados da BPR-5 seja menor à parcela dos mesmos subtestes da BPRD, desta parcela confiável, o montante que se correlaciona com critérios externos é aproximadamente o mesmo nas duas baterias. Considerando que a BPR-5 consome cerca de 54% do tempo consumido pela BPRD sem, no entanto, reduzir a qualidade da avaliação efetuada, pode-se concluir pela superioridade da primeira.

Conclui-se, portanto, que as evidências apoiam a precisão e validade da BPR-5. Pensamos que ela será de grande utilidade aos profissionais em suas atividades de psicodiagnóstico, orientação profissional, orientação escolar e seleção uma vez que consiste um instrumento eficiente e rápido para avaliação simultânea do raciocínio geral e das aptidões.

Referências

- Almeida, L.S. (1986). *Bateria de Provas de Raciocínio Diferencial (BPRD)*. Porto: Faculdade de Psicologia e de Ciências da Educação.
- Almeida, L.S. (1987). O impacto das experiências educativas na diferenciação cognitiva dos alunos: análise dos resultados em provas de raciocínio diferencial. *Revista Portuguesa de Psicologia*, 24, 131-157.
- Almeida, L.S. (1988). *O raciocínio diferencial de jovens*. Porto: INIC.
- Almeida, L.S. (1989). Gender and social class effects on differential reasoning tasks performance with portuguese secondary students. *Personality and Individual Differences*, 10 (5), 565-572.
- Almeida, L.S. (1992). *Baterias de Provas de Raciocínio Diferencial (BPRD) - Manual*. Braga, Portugal: Universidade do Minho.
- Almeida, L.S. (1995). Bateria de Provas de Raciocínio Diferencial (BPRD). Em L.S. Almeida, M.R. Simões & M.M. Gonçalves (Orgs.), *Provas Psicológicas em Portugal* (pp. 19-28), Braga: Apport.
- Almeida, L.S. & Campos, B.P. (1986a). Aferição de uma bateria de provas de raciocínio diferencial para utilização nas atividades de orientação vocacional de alunos do Ensino Secundário. *Psicologia: Teoria e Pesquisa*, 2 (3), 201-212.
- Almeida, L.S. & Campos, B.P. (1986b). Validade preditiva dos testes de raciocínio diferencial. *Cadernos de Consulta Psicológica*, 2, 105-118.
- Almeida, L.S. & Costa, A.R. (1989). A diferenciação cognitiva progressiva com a idade: Resultados na Bateria de Provas de Raciocínio Diferencial. *Psicologia*, 7 (3), 231-236.
- Almeida, L.S. & Martins, T. (1996). A validade de critério em provas cognitivas: Contributos para um banco de dados na população portuguesa. *Avaliação Psicológica: Formas e Contextos*, IV, 577-584.
- Almeida, L.S. & Primi, R. (1998). *Baterias de Provas de Raciocínio - BPR-5*. São Paulo: Casa do Psicólogo.
- Almeida, L.S., Antunes, A.M., Martins, T.B.O. & Primi, R. (1997). Bateria de provas de raciocínio (BPR-5): apresentação de procedimentos na sua construção. Em: I Congresso Luso-espanhol de psicologia e educação. Coimbra. *Actas*, 295-298.
- American Psychological Association (1985). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.
- Anastasi, A. & Urbina, S. (1997). *Psychological Testing*. Upper Saddle River, New Jersey: Prentice Hall.
- Bennett, G.K., Seashore, H.G. & Wesman, A.G. (1947). *Testes de Aptidões Específicas DAT*. Rio de Janeiro: Centro de Psicologia Aplicada (CEPA).
- Carrol, J.B. (1997). The three-stratum theory of cognitive abilities. Em D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Orgs.), *Contemporary intellectual assessment: theories, tests, and issues*. (pp. 122-130). New York: The Guilford press.
- Clarke, W.V. (1960). *Measurement of Skills Tests*. Providence, RI: Walter Clarke Associates, Inc.
- Davis, F.B. & Fifer, G. (1959). The effect on test reliability and validity of scoring aptitude and achievement tests with weights for each choice. *Educational and Psychological Measurement*, 19, 159-170.
- Embretson, S. (1996). The new rules of measurement. *Psychological Assessment*, 8 (4), 341-349
- Ferguson, G.A. (1959). *Statistical analysis in psychology and education*. New York: McGraw-Hill.
- Hambleton, H.K., Swaminatham, H. & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Horn, J.H. (1991) Measurement of intellectual capabilities: a review of theory. Em K.S. McGrew, J.K. Werder, & R.W. Woodcock WJ-R *Technical Manual* (pp. 197-245). Allen, TX: DLM.
- Merenda, P.F., Hall, Ch. E., Clarke, W.V., & Pascale, A.C. (1962). Relative predictive efficiency of the DAT and a short multi-factor battery of tests. *Psychological Reports*, 16, 151-155.
- Merenda, P.F., Clarke, W.V., & Jacobsen, G. (1965). Relative predictive validities of the MOS and DAT batteries for junior high school students. *Psychological Reports*, 16, 151-155.
- Merenda, P.F., Jacobsen, G. & Clarke, W.V. (1969). Further cross-validities of the MOS and DAT batteries. *Psychological Reports*, 24, 541-542.
- Meuris, G. (1969). *Tests de Raisonement Différentiel*. Bruxelles: Editest.
- Ministério da Educação (1998). *Ensino secundário: ofertas educativas e formativas*. Lisboa, Portugal: Eurodois Artes Gráficas.
- Panter, A.T., Kimberly, A., Swygert A. & Dahlstrom, W.G. (1997). Factor analytic approaches to personality item-level data. *Journal of Personality Assessment*, 68 (3), 561-589.
- Pasquali, L. & Araújo, R.M. (1998). Validação da bateria de raciocínio diferencial - BRD. [Resumo] Em: Sociedade Brasileira de Psicologia (Org.), *Resumos de Comunicações Científicas, XXVIII Reunião Anual de Psicologia* (pp. 59). Ribeirão Preto: SBP.
- Primi, R., Almeida, L.S. (1998). Considerações sobre a análise factorial de itens com resposta dicotômica. *Psicologia: Teoria, Investigação e Prática*, 3, 225-234.
- Primi, R., Almeida, L.S. & Lucarelli, M.D.M. (1996). Bateria de Provas de Raciocínio Diferencial (BPRD): Resultados numa amostra de adolescentes de São Paulo. *Avaliação Psicológica: Formas e Contextos*, IV, 385-392.

- Rainho, O. (n.d.). *Bateria CEPA Testes de Aptidões Específicas*. Rio de Janeiro: Centro de Psicologia Aplicada (CEPA).
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.
- Sternberg, R.J. (1984). Toward a triarchic theory of human intelligence. *The Behavioral and Brain Sciences*, 7, 269-315.
- Sternberg, R. J. (1986). Toward a unified theory of human reasoning. *Intelligence*, 10, 281-314.
- Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: University Chicago Press.
- Waller, N. G., Tellegen, A., McDonald, R. P. & Lykken, D. T. (1996). Exploring nonlinear models in personality assessment: development and preliminary validation of a negative emotionality scale. *Journal of Personality*, 64 (3), 545-576.
- Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment*, 8, 231-258.

Recebido em 27.08.1999

Primeira decisão editorial em 31.03.2000

Versão final em 17.11.2000

Aceito em 18.11.2000 ■

Princípios Normativos Para Publicação em *Psicologia: Teoria e Pesquisa*

Confidencialidade na revisão por pares Anonimato do autor

Psicologia: Teoria e Pesquisa adota os princípios de confidencialidade na revisão por pares.

É sabido que os autores podem ser identificados por seus trabalhos e os revisores, pelo teor dos comentários que acompanham seus pareceres. Não obstante — presumindo que há vantagens na revisão cega e do anonimato da consultoria — procura-se manter, nesta Revista, procedimentos que facilitem resguardar o caráter confidencial do processo. Alguns dos procedimentos visam proteger o anonimato do autor, minimizam constrangimentos na revisão e facilitam a tarefa do editor.

A identificação e indicações de autoria — incluindo nome completo e afiliação institucional de todos os autores, e notas informativas com agradecimentos e endereço para correspondência — devem constar apenas em folha personalizada, que é retirada pelo editor quando o manuscrito é encaminhado para exame por consultores *ad hoc*.

O anonimato dos autores não pode ser assegurado se o corpo do texto contém indicações da autoria, explícitas ou implícitas (local de realização do trabalho e referências a outros trabalhos do autor, p. ex.). Embora não seja uma exigência, recomenda-se aos autores que tomem cuidado para preparar o manuscrito sem indicação de autoria.

Em sua maior parte, o conteúdo deste texto foi traduzido e adaptado do *Manual de Publicação da APA* (APA, 1994).