

ESTUDO EXPLORATÓRIO DAS PROPRIEDADES PSICOMÉTRICAS DO TESTE RAVEN-ESCALA AVANÇADA INFORMATIZADO

RICARDO PRIMI

Faculdade de Ciências Humanas - Universidade São Francisco

CARMEM E. FLORES-MENDOZA E ADAIL V. CASTEHO

Laboratório de Informática Aplicada à Psicologia - Instituto de Psicologia USP



RESUMO

Esse estudo investigou as propriedades psicométricas de uma versão experimental informatizada das Matrizes Progressivas de Raven - Escala Avançada. Participaram 360 universitários respondendo ao instrumento coletivamente sem limite de tempo. Para a Série II a dificuldade dos itens se correlacionou em 0,96 com a versão original. A precisão através da fórmula de Kuder-Richardson foi de 0,83 e o coeficiente de validade comparando as notas no Raven com o desempenho acadêmico foi de 0,43. Esses resultados indicaram que a versão informatizada foi, no que se refere as propriedades psicométricas, equivalente à original.

Descritores: Matrizes Progressivas de Raven - Escala Avançada, Testes Informatizados, Raciocínio Analítico, Testes de Inteligência.

ABSTRACT

EXPLORATORY STUDY OF THE COMPUTERIZED RAVEN ADVANCED PROGRESSIVE MATRICES PSYCHOMETRIC PROPERTIES.

This study has investigated the psychometric properties of a computerized experimental version of the Raven's Advanced Progressive Matrices (APM). The participants were 360 undergraduate students who answered collectively without time limit. For the Set II the item's difficulties correlated .96 with the original version. Internal consistency reliability coefficient was .83 and the correlation with academic achievement grades was .43. These results have indicated that the psychometric properties of the computerized version were equivalent to the original.

Key Words: Raven's Advanced Progressive Matrices (APM), Computerized Testing, Analytical Reasoning, Intelligence Tests.

Os autores agradecem Alicia M. S. Hernandez, Ana Lucia Jancovic, Carla C. Del Roy, Claudinéia A. E. de Godói, Fabiane C. Pedroso, Romilda S. de Queiroz que colaboraram na coleta de dados.

INTRODUÇÃO

No II Encontro de Técnicas de Exame Psicológico observou-se um maior interesse dos profissionais em informatizar testes psicológicos em comparação ao primeiro encontro. Neste último, dizia-se que, no trabalho de informatização, a primeira tarefa seria adequar o máximo possível o formato do teste à tela do computador e verificar no caso de populações especiais, a necessidade ou possibilidade da utilização de periféricos convenientes como, por exemplo, consoles especiais, telas sensíveis ao toque, etc., não esquecendo os custos envolvidos que possam limitar a utilização (Castilho e Flores-Mendoza, 1995). Afirmava-se também que uma providência importante, relacionada aos testes psicológicos informatizados, seria verificar se as propriedades psicométricas, tais como a precisão e a validade dos instrumentos tradicionalmente utilizados, se manteriam significativamente inalteradas quando os mesmos fossem aplicados via computador. Sabemos que o emprego de muitos dos testes tradicionalmente usados no Brasil - e de origem estrangeira - é freqüentemente alvo de críticas dadas a ausência de normas estabelecidas especificamente no país ou de estudos e pesquisas que assegurem a adequação do seu conteúdo e das características psicométricas referentes às nossas condições. Assim, também, quando se procede à informatização de um teste, seja brasileiro ou estrangeiro, no mínimo dever-se-ia proceder a estudos para estabelecer a validade de seus resultados comparando-os com os obtidos com a versão original tipo "lápiz e papel" ou de execução.

Poder-se-ia perguntar se um teste, quando informatizado, se manteria válido e, ainda, se haveria necessidade de estabelecimento de normas específicas para versões informatizadas. Essa questão, como muitas outras no campo acadêmico, não pode de pronto ser respondida. Entretanto, se um estudo de validação que utilize como critério externo o mesmo teste na versão tradicional mostrar que existe uma correlação significativa e substancial entre os dois, inclusive, próxima ao coeficiente de precisão teste-reteste, poderíamos dizer que a versão informatizada equivale à original. Se, alternativamente, um estudo de validação por outro método, indicasse que as propriedades psicométricas permanecessem significativamente inalteradas poderíamos, também, inferir a equivalência entre as formas tradicional e informatizada.

O presente estudo constitui um primeiro passo, uma aproximação inicial, nessa direção com relação à versão informatizada do teste "Teste das Matrizes Progressivas Raven - Escala Avançada", revisada em 1962, (Raven, 1965), desenvolvida no Laboratório de Informática Aplicada à Psicologia, (Castilho e Primi, 1996) no Instituto de Psicologia da Universidade de São Paulo. Trata-se de um estudo preliminar, de precisão interna do teste, do nível de dificuldade dos itens, da comparação dessas propriedades com as da versão ori-

ginal e de correlação dos resultados obtidos via versão informatizada com as notas de alunos de uma instituição de Ensino de Nível Superior do Estado de São Paulo. A partir dos resultados da análise desses dados poderemos decidir pela realização de um estudo de validade utilizando como critério a versão original "lápiz e papel" do mesmo. Esse instrumento, bem como as Escalas Geral e Especial do mesmo teste, têm sido freqüentemente mencionadas na literatura nacional. A validade intercultural desse instrumento tem sido provada, obtendo-se altas correlações (0,98 a 1,00) entre o índice de dificuldade dos itens e amostras de diferentes culturas, por J. Raven, J. C. Raven e Court (1998a). Um levantamento efetuado num banco de dados, mantido pela American Psychological Association, o Psyclit, revelou mais de 160 referências nos últimos 5 anos a esse teste. Esses estudos foram realizados dentro de contextos de cunho psicossocial, clínico, neuropsicológico, cognitivo e teórico; a maioria orientada por princípios psicométricos tradicionais. Dentre esses apenas cinco estudos se relacionavam à informatização do teste ou envolviam novas estratégias de análise e mensuração como a Teoria de Resposta ao Item (Kubinger, Formann e Farkas, 1991; Styles e Andrich, 1993), além de um trabalho psicométrico tradicional.

No Laboratório de Informática Aplicada à Psicologia do IP/USP inicia-se pois uma série de estudos psicométricos relacionados a instrumentos psicológicos informatizados. Um dos primeiros, o do "Teste das Matrizes Progressivas de Raven - Escala Avançada" será seguido pelo estudo das outras escalas do teste e de outros instrumentos de avaliação intelectual, assim como de inventários de interesses e de personalidade já adaptados pelo Laboratório, para aplicação e avaliação informatizada.

MÉTODO

Participantes

A amostra foi composta por 360 alunos universitários provenientes dos cursos de Engenharia, Análise de Sistemas, Administração e Psicologia, de uma Universidade particular do Estado de São Paulo, sendo 49,3% provenientes do 1º ano de Psicologia. Dentre os sujeitos 71,9% pertenciam ao sexo feminino e 28,1% ao sexo masculino.

Materiais

Utilizou-se o "Teste das Matrizes Progressivas de Raven - Escala Avançada" (Raven, 1965), versão informatizada em ambiente "Windows" (Castilho e Primi, 1996).

A coleta de dados foi realizada em salas com 20 microcomputadores PC-486, IBM compatíveis, com monitores SVGA de 14" e mouse configurados em rede local.

Para o tratamento estatístico dos resultados empregou-se o software "SPSS" (Statistical Package for the Social Sciences) - Versão 6.0, 1996.

Procedimento

As aplicações foram coletivas com grupos de 20 sujeitos. O software apresentava as questões, tela a tela, uma de cada vez. Manteve-se um "lay-out" de tela similar à página do teste. E como nesse, o sujeito deveria escolher uma das alternativas de resposta.

A escolha da alternativa julgada correta era feita colocando o cursor do mouse sobre uma das alternativas, e, pressionando o botão esquerdo do mouse o software emoldurava essa alternativa fornecendo ao sujeito uma indicação visual da alternativa que estava escolhendo. Nesse ponto o aluno poderia então passar para o próximo item ou refazer sua escolha caso julgasse necessário. Semelhante à aplicação tradicional, o programa não forneceu "feedback" ao aluno, isto é, o programa não o informava sobre o acerto ou não da sua escolha.

A cada item respondido, o programa procedia ao registro do tempo de reação (em milissegundos) e do resultado da escolha: 1 ponto para acerto e 0 para erro. As instruções foram semelhantes às usadas na forma tradicional. Não houve limite de tempo. Nessas sessões seis monitores atuaram como auxiliares.

O programa dispunha de uma base de dados que registrava automaticamente os resultados em um arquivo de formato compatível com o exigido pelo programa SPSS (Statistical Package for the Social Science, 1996). As análises estatísticas foram realizadas e apresentadas com a utilização desses software.

RESULTADOS

Uma primeira preocupação foi com o tempo total de realização do teste. Isto porque observou-se que alguns alunos respondiam rapidamente evidenciando, por algum motivo, que não apresentavam motivação suficiente para responder adequadamente ao teste. E, como nessas condições os escores poderiam não refletir estimativas confiáveis das habilidades desses alunos, decidiu-se eliminar os escores de 42 sujeitos.

A correlação obtida entre o tempo total gasto na realização do teste e o escore final foi de 0,49 ($p < 0,01$, veja Figura 1). Portanto há uma tendência de obtenção de melhores resultados por aqueles sujeitos que permanecem mais tempo respondendo ao teste. Isto, também pode estar indicando uma relação entre o comportamento de reflexão (e por isso um gasto maior de tempo na resolução do teste) e a capacidade de raciocínio (altos escores). Houve exceções, como o caso de um sujeito que obteve 29 pontos em 5 minutos de realização do teste.

A regra adotada na identificação e exclusão dos casos foi: ter realizado o teste em menos de 30 minutos e ter uma quantidade de acertos menor do que 10 pontos. Para 19 alunos essa condição foi verdadeira, assim eles foram eliminados e, portanto, restaram 341 alunos para as análises subsequentes. De um modo global o tempo médio de realização das duas séries conjuntamente foi de 44,1 minutos, com um desvio-padrão de 19,5.

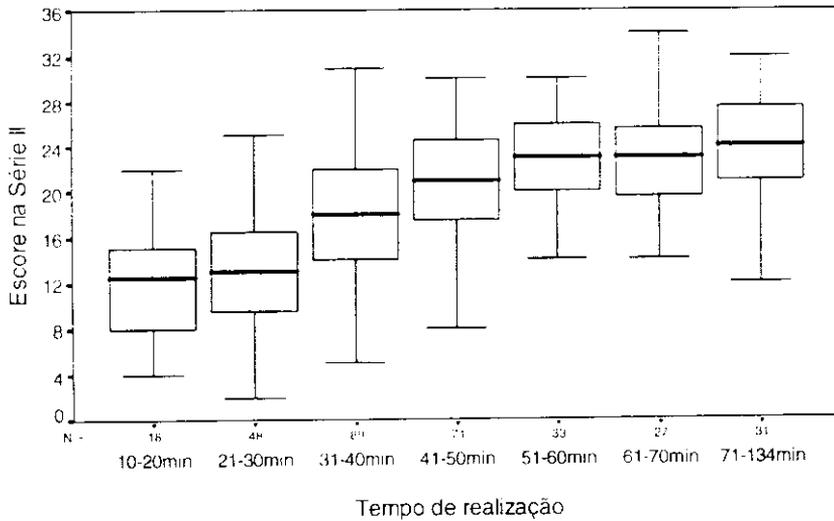


Figura 1 - Distribuição dos escores da Série II em função do tempo.

As retas horizontais no interior dos retângulos correspondem às medianas, os extremos aos percentis 25 e 75, e as retas horizontais extremas aos valores mínimos e máximos.

Como pode ser observado na Tabela 1 os índices de dificuldade estiveram bem dispersos e os índices de discriminação altos. Os valores mais baixos do índice de discriminação ocorreram nos casos em que a variância foi muito baixa, ou seja, em itens muito fáceis (como o Item 1, Série I) ou itens muito difíceis (como o Item 36, Série II). Os índices de dificuldade obtidos aqui foram comparados com os índices apresentados no manual (J. Raven, J. C. Raven e Court, 1998b; p. APM9) e as correlações foram de 0,93 para a Série I e de 0,96 para a Série II ($p < 0,001$).

Na Tabela 2 apresentou-se as estatísticas gerais incluindo a consistência interna do teste que foi determinada pelo método de Kuder-Richardson, obtendo-se o coeficiente de precisão de 0,83 para a Série II. O índice médio de dificuldade para a mesma série foi de 0,52. Na versão tradicional o coeficiente de precisão obtido foi 0,87 (J. Raven, J. C. Raven e Court, 1998b).

Tabela 1. Propriedades Psicométricas dos Itens do Teste Raven - Escala Avançada

Nº do item	Série	ID	Índice de discr.	r_{pbi}
1	I	0,98	0,04	0,14
2	I	0,95	0,11	0,36
3	I	0,89	0,18	0,28
4	I	0,96	0,08	0,38
5	I	0,87	0,23	0,46
6	I	0,81	0,31	0,45
7	I	0,93	0,14	0,43
8	I	0,75	0,46	0,61
9	I	0,48	0,64	0,56
10	I	0,83	0,30	0,48
11	I	0,40	0,71	0,60
12	I	0,71	0,49	0,58
1	II	0,88	0,24	0,35
2	II	0,87	0,36	0,46
3	II	0,90	0,31	0,49
4	II	0,77	0,47	0,48
5	II	0,82	0,42	0,46
6	II	0,84	0,30	0,35
7	II	0,81	0,39	0,45
8	II	0,79	0,42	0,46
9	II	0,83	0,43	0,46
10	II	0,73	0,62	0,56
11	II	0,86	0,44	0,52
12	II	0,78	0,43	0,46
13	II	0,57	0,50	0,44
14	II	0,69	0,64	0,55
15	II	0,74	0,50	0,48
16	II	0,51	0,56	0,42
17	II	0,68	0,57	0,49
18	II	0,37	0,47	0,41
19	II	0,65	0,60	0,51
20	II	0,62	0,57	0,47
21	II	0,37	0,63	0,54
22	II	0,40	0,46	0,38
23	II	0,46	0,56	0,45
24	II	0,26	0,48	0,41
25	II	0,29	0,44	0,39
26	II	0,39	0,43	0,39
27	II	0,15	0,31	0,37
28	II	0,19	0,25	0,28
29	II	0,09	0,12	0,23
30	II	0,26	0,52	0,45
31	II	0,21	0,31	0,30
32	II	0,19	0,24	0,26
33	II	0,26	0,42	0,41
34	II	0,24	0,57	0,53
35	II	0,24	0,46	0,43
36	II	0,07	0,07	0,12

Tabela 2. Estatísticas Finais da Série I e II do Teste Raven - Escala Avançada

	Série I	Série II
X	9,7	19,9
S	1,8	5,8
S ²	3,3	33,6
Simetria	-0,96	-0,12
Curtose	1,2	-0,56
KR - 20	0,61	0,83
EPM	1,12	2,4
ID	0,79	0,52
<i>r_{pbi}</i>	0,44	0,42
Nº de itens	12	36

Para o estabelecimento da validade do teste foram utilizadas as notas escolares de uma parte da amostra (alunos do curso de Psicologia). Essas notas correspondiam às avaliações bimestrais da disciplina de TEP (Técnicas de Exame Psicológico). Antes de efetuar a correlação, verificou-se que as avaliações de TEP solicitassem dos alunos o raciocínio analítico. A correlação entre as notas e os escores da Série I foi de 0,34 ($n = 86$; $p < 0,001$) e com os escores da Série II de 0,41 ($n = 108$; $p < 0,001$). Na versão tradicional, o critério externo utilizado foi o emprego de notas de amostras de estudantes de Escolas de Mecânica, Técnicas e Comerciais e Universitárias. A comparação foi feita apenas mediante a distribuição dos resultados.

CONCLUSÃO

Os resultados obtidos nos cálculos dos índices de dificuldade dos itens, da consistência interna do teste e da validade concorrente parecem indicar que a versão informatizada do Teste das Matrizes Progressivas de Raven - Escala Avançada mantém propriedades psicométricas similares às da versão tradicional, e, nesse sentido, a primeira versão poderá ser utilizada como instrumento de avaliação. Assim, deveremos dar continuidade ao presente trabalho fazendo estudos comparativos com as outras escalas do teste, ou seja, Escala Especial e Geral.

REFERÊNCIAS BIBLIOGRÁFICAS

- Castilho, A. V. & Flores-Mendoza, C.E. (1995). Informática Aplicada a Técnicas de Exame Psicológico. *Boletim de Psicologia*, XLV, 102, 97-104.
- Castilho, A. V. & Primi, R. (1996). Sistema geral para informatização de testes. *Programas e Resumos do II Encontro de Técnicas de Exame Psicológico: Ensino, Pesquisa e Aplicações*, p.7.
- Kubinger, K. D.; Forman, A.K. & Farkas, M. G. (1991). Psychometric shortcomings of Raven's Standard Progressive Matrices, in particular for computerized Testing. *Revue Européenne de Psychologie Appliquée*, 41, 4, 295-300. . .
- Styles, I. & Andrich, D. (1993). Linking the standard and advanced forms of the Raven's Progressives Matrices in both the pencil-and-paper and computer-adaptive-testing formats. *Educational and Psychological Measurement*, 53, 4, 905-925.
- Raven, J., Raven J. C. & Court, J. H. (1998a). *Manual for Raven 's Progressive Matrices and Vocabulary Scales: Section 1. General Overview*. Oxford: Oxford Psychologists Press.
- Raven, J., Raven J. C. & Court, J. H. (1998b). *Manual for Raven 's Progressive Matrices and Vocabulary Scales: Section 4. Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.
- Raven, J. C. (1965). *Matrizes Progressivas - Escala Avançada*. Tradução e Adaptação do original em inglês por Francisco Campos. Rio de Janeiro: Centro Editor de Psicologia Aplicada.