

USO DE COMPUTADORES NA AVALIAÇÃO PSICOLÓGICA: ESTUDO DE SUA INFLUÊNCIA SOBRE O DESEMPENHO INDIVIDUAL EM UM TESTE DE RACIOCÍNIO NUMÉRICO (RN)¹

WAGNER BANDEIRA ANDRIOLA

Professor Adjunto do Departamento de Fundamentos da Educação; Professor do Programa de Mestrado e Doutorado em Educação Brasileira, Universidade Federal do Ceará.

Resumo: O uso de testes computadorizados por psicólogos exige uma análise prévia de suas características métricas, assim como do desempenho dos sujeitos. Dessa forma, investigou-se a influência da informatização de um Teste de Raciocínio Numérico (RN) sobre o desempenho individual de estudantes do ensino médio. Utilizou-se uma amostra de 143 alunos, divididos randomicamente em dois grupos, sendo um controle (versão lápis-papel) e outro experimental (versão computadorizada). Os resultados indicaram a existência de diferenças estatisticamente significativas entre o desempenho dos grupos, considerando-se as variáveis *gênero* e *idade*. Sugere-se a elaboração de normas diferenciadas para os dois formatos do Teste RN (lápis-papel *versus* computadorizado), considerando-se as referidas variáveis.

Palavras-chave: psicometria; teste psicológico informatizado; avaliação psicológica; raciocínio numérico.

THE USE OF COMPUTERS ON PSYCHOLOGICAL ASSESSMENT: AN INVESTIGATION OF ITS INFLUENCE ABOUT INDIVIDUAL PERFORMANCE ON A NUMERIC REASONING TEST (RN)

Abstract: The use of computerized tests by psychologists demands a previous analysis of their metric characteristics, as well as that of the performance of the individual being tested. This study investigated the influence of a computerized Test of Numeric Reasoning (RN) concerning the individual performance of high school students. The sample was composed of 143 high school students randomly divided into two groups: a control (paper-pencil format) and an experimental (computerized format). The results indicated statistically significant differences between the performance of the subjects of the groups regarding the *gender* and

age variables. The results also indicate the necessity of elaborating different norms for both formats of the RN Test (paper-pencil *versus* computerized), taking into consideration the above mentioned variables.

Keywords: psychometry; computerized psychological test; psychological assessment; numeric reasoning.

1. Introdução

Nos Estados Unidos o uso de computadores na área da avaliação psicológica remonta aos anos 1950, com uma dupla função: calcular o somatório dos escores e efetivar a padronização dos resultados, ambos de maneira mais rápida e precisa (Lindquist, 1954; Merenda, 1987). Desde então, uma diversidade de *softwares* tem sido desenvolvida, visando a avaliação de amplos aspectos psicológicos, como personalidade, psicomotricidade, interesses e aptidões. Como exemplos, citamos os trabalhos de Barret, Alexander, Doverspike, Cellar e Thomas (1982), Elwood e Clark (1978), Fremer e Anastasio (1969), Harrell e Lombardo (1984), Klinge e Rodziewicz (1976), Lushene, O'Neil e Dunn (1974), O'Brien e Dugdale (1978), Sampson (1990) e Sapinkopf (1978).

Não obstante, essa é uma área que busca seu pleno desenvolvimento em países como Argentina (Adam e De Etchevers, 1993; Fogliato e Pérez, 1990, 1996), Portugal (Almeida e Silva e Sá, 1989; Taveira e Campos, 1989; Taveira 1992) e Espanha (Prieto, Cano, Orgaz e Pulido, 1993; Prieto, Cano, Pulido e Orgaz, 1996). No Brasil começam a surgir alguns testes computadorizados, como é o caso de uma Bateria de Testes destinada à Orientação Vocacional, e de um teste para avaliar as Condições e Hábitos de Estudo (CHE), ambos de autoria de Camara (1992a, 1992b). Mais recentemente têm surgido vários trabalhos objetivando a informatização de testes (Andriola, 1996a, 1996c; Castilho, Primi e Jonas, 1996; Thiers, Capovilla, Macedo e Duduchi, 1996).

Com o uso crescente de testes psicológicos computadorizados surgiram muitos questionamentos sobre essa nova prática (Roid, 1986; Zachary e Pope, 1983). Atualmente, percebe-se a existência de duas opiniões antagônicas: de um lado estão os que recusam todo e qualquer

tipo de utilização do computador na avaliação psicológica, e de outro aqueles que, influenciados pelos avanços e potencialidades tecnológicas, ressaltam suas vantagens (Andriola, 1997).

A partir dessas posições antagônicas, algum tipo de reflexão sobre essa nova prática deve ser feita. Devemos aclarar os objetivos da utilização de testes informatizados, por meio da caracterização de dois momentos distintos. Em uma etapa inicial, os computadores podem servir para dar maior rapidez e precisão à correção e normatização dos escores brutos. Os autores e investigadores desse grupo ressaltam algumas vantagens “visíveis”, como a possibilidade do aplicador realizar observações de comportamentos paralelamente à testagem, além de obter maior padronização na apresentação das instruções (Biskin e Koloktin, 1977; Cory, 1977). Mesmo assim, existem reservas quanto à utilização dos computadores durante a aplicação dos testes, visto que variáveis como formato do *software*, atitudes frente ao computador e familiaridade com o programa, podem influenciar o desempenho individual, e por conseguinte, as características métricas do instrumento (Murphy e Davidshofer, 1994).

Em uma segunda etapa, os computadores podem servir para a análise e interpretação dos resultados. Vários autores, como Millman e Outlaw (1978), questionam a possibilidade do computador atuar satisfatoriamente nesses dois momentos. Essa posição encontra fundamentação no fato dos computadores possuírem sérias limitações quanto à “capacidade analítica”, já que dependem sempre da eficácia do *software* desenvolvido. A situação que melhor exemplifica tal limitação dá-se na análise de resultados obtidos a partir da aplicação de uma bateria de testes ou de um instrumento multifatorial. Como ressalta Merenda (1987):

(...) o computador pode proporcionar perfis psicológicos e páginas e páginas de “*output*” para a interpretação dos resultados. O problema está no significado psicológico dessa informação. O computador fornece somente informação aditiva, cabendo ao psicólogo a integração dessa informação (p. 24).

Outros autores, como Watts, Baddley e Williams (1982), chamam a atenção para o uso das mesmas normas de apuração dos resultados em distintos formatos de testes. Defendem que no caso da transposição

do uso de normas elaboradas com testes tipo lápis-papel para os testes informatizados, deverá ter sido realizado algum estudo visando investigar a existência de diferenças quanto ao desempenho dos respondentes que se submeteram a uma e outra forma do mesmo teste. Como exemplo citamos o estudo efetivado por Andriola (1996a) com o Teste de Raciocínio Verbal (RV), que utilizou uma amostra de 120 estudantes do ensino médio, dividida randomicamente em dois grupos, submetida a dois formatos diferentes do mesmo teste: tradicional (lápis-papel) e informatizado. Os resultados indicaram a existência de diferenças significativas entre o desempenho de homens e mulheres nos formatos considerados. Como conclusão, o autor sugeriu a elaboração de normas diferenciadas para homens e mulheres em cada uma das formas do Teste RV.

De igual maneira, a realização do presente estudo visa esclarecer se as normas originais, elaboradas utilizando-se o formato tipo lápis-papel do Teste de Raciocínio Numérico (RN), desenvolvido por Andriola (1996b), podem ser transpostas para as situações em que esse instrumento psicológico seja utilizado no formato computadorizado. O objetivo é verificar a existência de diferenças significativas entre o desempenho dos sujeitos submetidos ao Teste RN em formatos diferenciados: lápis-papel *versus* informatizado.

As variáveis consideradas no estudo foram:

- *Independentes (VI's)*: formato do Teste RN (lápis-papel *versus* informatizado), gênero e idade. A presença das variáveis sexo e idade deveu-se ao fato destas terem sido utilizadas na elaboração das normas;
- *Dependente (VD)*: desempenho no Teste RN, expresso pelos escores individuais obtidos por meio do somatório dos acertos às 25 questões.

1.1. Informações técnicas sobre o Teste de Raciocínio Numérico²

O Teste de Raciocínio Numérico (RN) foi desenvolvido por Andriola (1994; 1996b), por meio de fases organizadas hierarquicamente:

- Escolha da base teórica para o teste e para a elaboração dos itens;
- Estudo-piloto com uma amostra de 275 estudantes, objetivando: (i) detectar problema quanto à compreensão das instruções; e (ii) determinar o índice de dificuldade das primeiras 40 seqüências numéricas elaboradas para o Teste RN;
- Estudo final com uma amostra de 545 estudantes, objetivando determinar: (i) os parâmetros métricos dos itens (dificuldade³, discriminação⁴ e homogeneidade⁵); (ii) a validade de construto⁶ e a precisão do instrumento⁷; (iii) as normas de apuração⁸.

Com base nos resultados obtidos, organizou-se um teste composto por 25 seqüências numéricas, com tempo de 15 minutos para sua resolução⁹.

2. Método

Segundo Campbell e Stanley (1979) e Vieira e Hoffmann (1989), trata-se de um delineamento experimental, visto que existem dois grupos: um controle (GC) e outro experimental (GE), aleatoriamente organizados. No GE houve a manipulação da variável experimental (ou independente) – no presente estudo caracterizada pelo fato dos sujeitos responderem ao teste no formato computadorizado – enquanto no GC utilizou-se a forma tradicional (lápiz-papel).

2.1. Amostra

Utilizou-se uma amostra de 143 estudantes do ensino médio de uma escola pública da cidade de Fortaleza (CE), sendo 59 alunos da 1ª série, 46 da 2ª série e 38 da 3ª série. Do total, 71 eram homens e 72 mulheres, com idade média de 17,28 anos (desvio-padrão = 2,19 anos).

O quadro 1 apresenta as estatísticas descritivas para ambos os grupos (GE e GC), considerando as variáveis *gênero*, *série escolar* e *idade*.

Variável	Grupo experimental (GE)	Grupo controle (GC)
Gênero	Masc.: n = 35; Fem.: n = 36	Masc.: n = 36; Fem.: n = 36
Série escolar	1ª: n = 39; 2ª: n = 14; 3ª: n = 18	1ª: n = 20; 2ª: n = 32; 3ª: n = 20
Idade	\bar{X} = 16,93; dp = 2,23	\bar{X} = 17,63; dp = 2,15

Quadro 1. Estatísticas descritivas dos grupos estudados.

Há uma distribuição aproximadamente equitativa nos dois grupos para as variáveis *gênero* e *idade*. Para a variável *série escolar* há diferenças entre a quantidade de respondentes nos dois grupos. Por exemplo, verifica-se maior quantidade de alunos da 1ª série no GE, enquanto no GC há maior quantidade da 2ª série.

2.2. Instrumento

Como destacam Andriola (1995a, 1995b, 1997, 1998, 1999a), Andriola e Cavalcante (1999), Barreto, Trompieri Filho e Andriola (1999), Andriola e Pasquali (1995), Friendenberg (1995) e Pasquali (1995, 1997a), um teste deve apresentar parâmetros métricos. Assim, a escolha do Teste de Raciocínio Numérico (RN) está fundamentada no fato de ser um instrumento de conteúdo recente, construído para estudantes secundaristas, possuidor de parâmetros métricos (validade de construto, fidedignidade e índices de dificuldade e discriminação dos itens) aceitáveis, além de possuir normas elaboradas a partir de uma amostra representativa da população à qual se destina (Andriola, 1996b). Trata-se de um instrumento de medida psicológica, composto por 25 questões em formato de seqüências numéricas. Nele, a tarefa do respondente será completar essas séries, por meio da escolha de uma dentre cinco alternativas propostas como respostas. O limite de tempo destinado à resolução dos itens é de 15 minutos.

No presente estudo foram utilizados dois formatos do Teste RN: a) tradicional (lápiz-papel) e b) informatizado (versão para computadores tipo PC's). A programação para computador foi feita em linguagem *Clipper*, respeitando as instruções existentes para a forma lápis-papel, e reduzindo os comandos ao uso das seguintes teclas: quatro teclas de direção; a tecla "*Enter*" e as cinco teclas com as letras correspondentes às alternativas propostas como respostas (a,b,c,d,e). O computador advertia ao respondente, por meio de um sinal sonoro ("*bip*"), quando alguma tecla distinta era escolhida. Os itens eram apresentados na tela do computador um de cada vez; ao término do tempo estabelecido de 15 minutos, o programa era bloqueado, impossibilitando dessa maneira a apresentação de respostas adicionais.

2.3. Procedimento

Antes da aplicação do Teste RN, os voluntários participantes eram informados sobre os objetivos da investigação, e solicitados a apresentar seus questionamentos ou dúvidas para que fossem esclarecidas pelo investigador. Feito isso, iniciava-se o processo de coleta de dados.

No GC as aplicações foram coletivas e realizadas por dois auxiliares de pesquisa. Inicialmente procedeu-se à distribuição das folhas de resposta para preenchimento por parte dos respondentes, sendo realizada em seguida a leitura das instruções. Após certificarem-se de que todos os respondentes haviam entendido a tarefa a ser executada, os aplicadores ordenavam o início do teste. No limite de 15 minutos os aplicadores informavam aos respondentes o término do tempo, os quais imediatamente paravam de responder ao Teste RN. Em seguida, as folhas de respostas eram recolhidas e apresentavam-se agradecimentos aos estudantes pela participação na pesquisa.

No GE as aplicações também foram coletivas, realizadas no Laboratório de Pesquisa em Avaliação e Medida Psico-Educacional (LABPAM) da Universidade Federal do Ceará (UFC). Os estudantes eram orientados a utilizar as teclas especificadas, e em seguida realizar a leitura das instruções, juntamente com os auxiliares de pesquisa. Não havendo dúvidas sobre a execução da tarefa, iniciavam o teste. Após o tempo limite de 15 minutos o computador travava o sistema, impossibilitando que o respondente fornecesse novas respostas. Ao término de cada sessão de aplicação, em um total de três, agradeciam-se aos participantes.

3. Resultados

Com as respostas dos sujeitos organizou-se um banco de dados e realizaram-se análises descritivas e inferenciais, por meio do uso do *software SPSS for Windows (7.5)*. A Tabela 1 apresenta os percentuais de acerto dos respondentes dos grupos (GC e GE) às 25 questões do Teste RN.

Questões	Acertos(%)		Respostas Fornecidas Não (%)	
	GE	GC	GE	GC
1	98,6	95,8	0,0	1,4
2	88,7	97,2	0,0	1,4
3	93,0	91,7	0,0	1,4
4	98,6	83,3	0,0	5,6
5	90,1	88,9	0,0	2,8
6	83,1	70,8	0,0	5,6
7	97,2	94,4	0,0	2,8
8	97,2	86,1	0,0	5,6
9	90,1	65,3	2,8	6,9
10	90,1	84,7	4,2	6,9
11	77,5	76,4	7,0	9,7
12	83,1	76,4	8,5	12,5
13	70,4	43,1	12,7	30,6
14	66,2	58,3	22,5	33,3
15	63,4	50,0	28,2	38,9
16	60,6	41,7	36,6	51,4
17	43,7	30,6	47,9	56,9
18	47,9	31,9	49,3	59,7
19	33,8	22,2	52,1	69,4
20	22,5	11,1	69,0	83,3
21	12,7	6,9	78,9	87,5
22	9,9	6,9	85,9	87,5
23	7,0	6,9	90,1	88,9
24	7,0	8,3	90,1	87,5
25	7,0	6,9	90,1	87,5

Tabela 1. Percentual de acerto às 25 questões do Teste RN.

GE = Grupo experimental (submetido ao formato computadorizado do Teste RN).

GC = Grupo controle (submetido ao formato lápis-papel do Teste RN).

Pelos dados da Tabela 1 constatamos que o percentual de acerto às 25 questões, em ambos os grupos, possui a mesma tendência: à medida que os respondentes avançam no Teste RN, há uma diminuição do índice de acerto, revelado pelo decréscimo dos respectivos valores percentuais. Por outro lado, ao serem considerados os percentuais de respostas não fornecidas, constata-se que nas oito primeiras questões o valor obtido pelo GE foi zero, isto é, os respondentes desse grupo apresentaram respostas às oito questões.

No GC a repetição dessa tendência não foi observada, isto é, os percentuais de respostas não fornecidas variou de 1,4% a 5,6%. Tal dado, adicionado ao fato de que os percentuais de respostas não fornecidas pelo GC são sempre superiores aos do GE, exceção feita às questões 23, 24 e 25, pode indicar que a tendência dos sujeitos do GE é apresentar, sistematicamente, respostas às questões, aumentando desse modo as respostas aleatórias.

Para a análise do desempenho dos respondentes segundo a variável *idade*, foram organizadas três faixas etárias, com sua distribuição empírica: a) até 16 anos; b) entre 17 e 18 anos; e c) acima de 18 anos. Os resultados revelaram decréscimo no desempenho dos sujeitos à medida que as idades avançavam. A Tabela 2 apresenta as médias de acertos nas respectivas faixas etárias.

Faixas etárias	Média	DP	Amplitude	N	%
Até 16 anos	15,24	4,57	6-25	55	38,5
17 - 18 anos	14,79	4,86	0-25	56	39,2
+ de 18 anos	12,13	5,09	3-24	31	21,7

Tabela 2. Desempenho médio dos respondentes de acordo com as faixas etárias.

Os dados da Tabela 2 indicam a tendência de decréscimo no desempenho dos respondentes, na medida em que as faixas etárias aumentam. Com o objetivo de detectar diferenças significativas quanto ao desempenho dos respondentes nas distintas faixas etárias, procedeu-se ao Teste da Análise de Variância (ANOVA). O resultado indicou que a variável *idade* é uma fonte de variação significativa [$F_{(2,141)} = 5,684$; $p < 0,05$] para explicar o desempenho dos estudantes. Esse resultado corrobora outros similares, encontrados por Almeida (1988), Andriola e Pasquali (1995) e Andriola (1996a), que observaram tendência de decréscimo no desempenho dos sujeitos em testes de raciocínio na medida em que a idade avança.

Nesse contexto, representamos graficamente o desempenho dos sujeitos componentes das três faixas etárias, no âmbito do GE e do

GC, para termos maior clareza na interpretação dos resultados do Teste ANOVA (ver Figura 1).

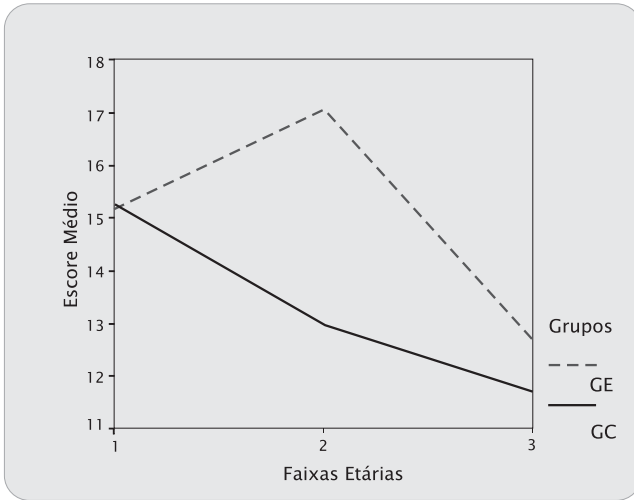


Figura 1. Desempenho dos respondentes das três faixas etárias consideradas.

A Figura 1 apresenta informações sobre possíveis diferenças entre as faixas etárias, no tocante ao rendimento médio no Teste RN, quando consideramos os grupos estudados (GE e GC). Para as Faixas 2 e 3, observa-se um aumento da pontuação média para os sujeitos do GE (teste informatizado). Não obstante, observamos não haver diferença entre os escores médios dos indivíduos mais jovens (Faixa 1) em ambos os grupos.

Também verificou-se que a variável *gênero* foi uma fonte de variação relevante, visto que obteve valor $F_{(1, 141)} = 1,873$; $p < 0,05$. Os homens obtiveram escore médio de 14,59 ($dp = 4,90$) e as mulheres de 13,19 ($dp = 4,96$), podendo-se afirmar que os homens foram superiores às mulheres no Teste RN. No âmbito dos grupos, esses valores sofreram pequenas alterações (ver Figura 2).

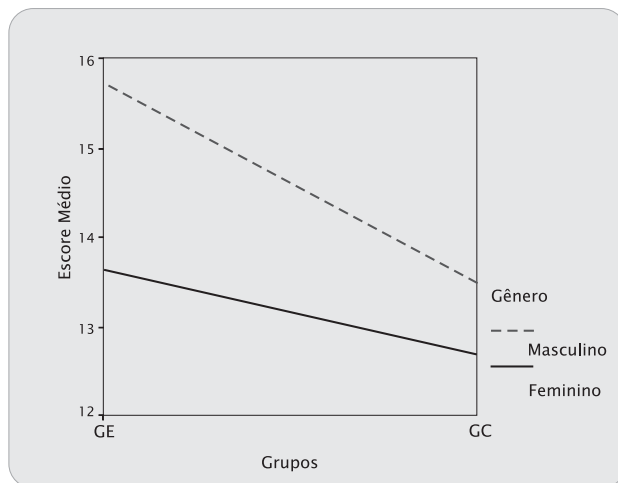


Figura 2. Desempenho dos respondentes segundo o gênero.

Segundo os dados da Figura 2, os homens e as mulheres obtiveram melhores rendimentos na forma computadorizada do Teste RN, com pontuações médias de 15,72 ($dp = 4,70$) e 13,64 ($dp = 5,37$). Na forma lápis-papel verificou-se uma diminuição das pontuações médias de ambos: homens com $\bar{X} = 13,48$ ($dp = 4,87$) e mulheres com $\bar{X} = 12,70$ ($dp = 4,69$). Observamos que tanto no âmbito do GE quanto no GC, os homens obtiveram rendimentos superiores às mulheres.

Por último, a comparação da possível diferença entre *os grupos* (GC e GE), quanto ao desempenho dos alunos, revelou ser insignificante [$F_{(1, 141)} = 1,183$; n.s.]. Observou-se que os respondentes do GC obtiveram média de desempenho no valor de 13,36 ($dp = 4,79$), enquanto os respondentes do GE obtiveram média 15,39 ($dp = 4,83$), podendo-se afirmar que os respondentes do formato computadorizado (GE) não se diferenciam dos sujeitos que se submeteram ao formato tradicional do mesmo teste (GC), no tocante ao desempenho no Teste RN.

A Figura 3 apresenta a distribuição da freqüência de acertos dos sujeitos do GE e GC.

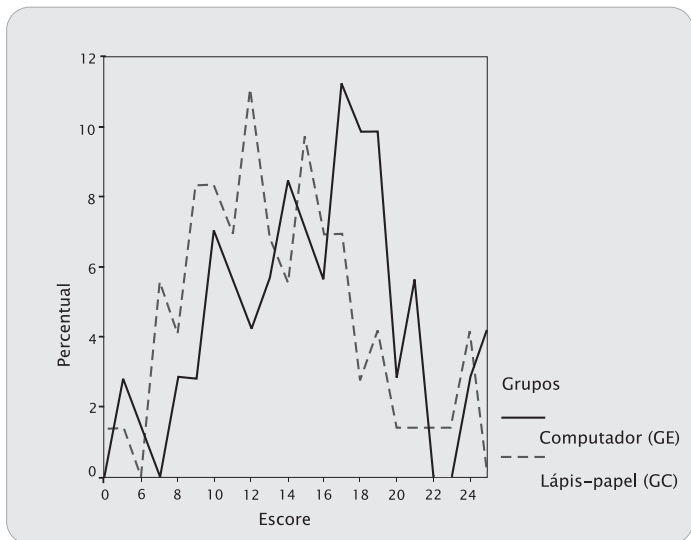


Figura 3. Frequência de acerto às questões segundo o grupo (GE ou GC).

As informações contidas na Figura 3 são importantíssimas para compreendermos porque o GE obteve desempenho médio superior ao GC, ainda que estatisticamente insignificante. Observando o gráfico de linhas, damos-nos conta de existir um maior percentual de sujeitos do GC com escores entre 7 e 13. Por outro lado, para escores compreendidos entre 16 e 22, aproximadamente, há mais sujeitos do GE. Com essas informações, e baseados no princípio de que a média aritmética é um parâmetro de tendência central muito suscetível a sofrer influências de valores extremos, podemos inferir que:

- A média do GC é inferior ao valor verificado no GE porque sofre influência da maior quantidade de respondentes com escores entre 7 e 13: aproximadamente 28% do total de indivíduos do grupo;
- A média do GE é superior ao valor verificado no GC porque sofre influência da maior quantidade de respondentes com escores entre 16 e 22: aproximadamente 25% do total de sujeitos do grupo.

Como a ANOVA compara as médias dentro dos grupos (*within groups*) e entre os grupos (*between groups*), observamos que a variabilidade dentro dos grupos foi mais significativa que entre os grupos, acarretando a inexistência de diferença relevante entre as pontuações médias do GE e do GC.

Cabe ressaltar, por último, que não foram identificadas interações entre as três variáveis consideradas na presente investigação.

4. À guisa de conclusão

Segundo Almeida, a variável *idade* é importante para a explicação do desempenho cognitivo, visto que “a partir dos 17 anos verifica-se já uma relativa estabilidade e, na generalidade dos casos, inicia-se já um certo declínio na média dos resultados” (1988, p. 168). A tendência de declínio no desempenho dos sujeitos, à medida que a idade aumenta, foi observada no presente estudo, já tendo sido também verificada pelo próprio Almeida (1988), por Andriola e Pasquali (1995) e Andriola (1996a). Como decorrência, há necessidade de elaboração de normas distintas para o formato computadorizado do Teste RN, considerando-se a variável *idade*.

A variável *gênero* também mostrou-se relevante para a explicação das diferenças entre os desempenhos dos sujeitos. Nesse contexto, foram corroborados os resultados dos estudos de Almeida (1988), Andriola e Pasquali (1995) e Andriola (1996a), que observaram diferenças significativas entre homens e mulheres. Por conseguinte, há necessidade de elaboração de normas diferenciadas para os dois formatos do Teste RN, considerando essa variável.

Quanto ao formato do teste RN (computadorizado *versus* lápis-papel), os resultados obtidos indicaram a inexistência de diferença significativa entre ambos os grupos. Portanto, não há necessidade de elaboração de normas de apuração diferenciadas para os formatos estudados (computadorizado e lápis-papel).

Também deve ser mencionado que os coeficientes de precisão (*alpha* de Cronbach) de ambas as formas do Teste RN têm valor muito

aproximados. No formato lápis-papel a precisão obteve valor 0,91, em uma amostra de 545 sujeitos (Andriola, 1996b), enquanto no formato computadorizado obteve valor 0,88, utilizando-se uma amostra de 71 respondentes. Quanto à validade de construto da forma computadorizada, há necessidade de estudos posteriores para sua determinação, utilizando-se uma amostra de maior tamanho, que seja representativa dos estudantes do ensino médio.

Deve ser destacado ainda que, segundo Pasquali:

(...) o rápido avanço da informática está permitindo o desenvolvimento de tecnologias novas de coleta de dados em Psicologia, que se baseiam na própria capacidade do computador e da viabilidade de trabalhar modelos multivariados (1997b, p. 45).

Com a citação queremos ressaltar o caminho sem retorno pelo qual passa atualmente a avaliação psicológica. O uso do computador nessa atividade é um fato consumado e corriqueiro em países mais avançados. No Brasil, penso que se deva buscar: (i) discutir os problemas éticos subjacentes a esses novos procedimentos; (ii) reestruturar os currículos dos cursos de graduação em Psicologia; (iii) discutir o papel da psicometria clássica na era da informática; (iv) debater o papel dos Conselhos Regionais de Psicologia e do Conselho Federal de Psicologia no controle do uso desses “novos” instrumentos psicológicos (testes informatizados). Em suma, devemos enxergar as vantagens e desvantagens no uso do computador para a avaliação psicológica, além de observar a adequação dos instrumentos psicológicos e controlar seu uso. Também não podemos esquecer de discutir a qualidade da formação dos profissionais que atuarão nesse novo e promissor mercado de trabalho. Nesse âmbito, não devemos olvidar um antigo e inteligente adágio latino: *non omnes qui habent citharam sunt citharaédi* (nem todos os que têm cítara sabem tocar cítara).

Notas

- ¹ Financiada pela Fundação Cearense de Amparo à Pesquisa (FUNCAP, Proc. 036/97-P&D).
- ² Estudo realizado com financiamento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Proc. 300.919/93, Bolsa de Desenvolvimento Científico Regional – DCR).
- ³ Com base nos percentuais de acerto, observou-se que 34% dos itens foram considerados fáceis e 64% de dificuldade média.
- ⁴ Todos os 25 itens possuem adequado poder de discriminação, verificado pela comparação de grupos-critério: os 27% de sujeitos com maiores escores *versus* os 27% com menores escores.
- ⁵ Os 25 itens obtiveram cargas fatoriais superiores a 0,30 em algum dos dois fatores extraídos.
- ⁶ Extração de dois fatores pelo Método de Extração dos Componentes Principais, explicando ambos 46,1% da variância total dos escores, com auto-valores de 16,72 e 5,88.
- ⁷ Alpha de Cronbach = 0,91 para a forma de 25 itens.
- ⁸ Foram estabelecidas normas percentílicas para o formato de 25 itens.
- ⁹ Baseado no tempo levado pelos primeiros 20% de respondentes para a sua resolução.

Referências Bibliográficas

- ADAM, G.; De ETCHEVERS, A.P. (1993). Diagnostico psicológico computadorizado [Resumo]. In: *XXIV Congresso Interamericano de Psicologia*. Santiago do Chile: Sociedade Interamericana de Psicologia.
- ALMEIDA, L.S. (1988). *O raciocínio diferencial dos jovens*. Porto: Instituto Nacional de Investigação Científica.
- ALMEIDA, L.S.; SILVA e SÁ, J.M. (1989). Uso dos computadores na avaliação psicológica: estudo em testes de aptidão intelectual. In: CRUZ, J. F.; GONÇALVES, R. A.; MACHADO, P. P. (eds). *Psicologia e educação: investigação e intervenção*. Porto: Associação dos Psicólogos Portugueses. p. 395-404.
- ANDRIOLA, W. B. (1994). Apresentação dos resultados parciais relativos às análises dos itens do Teste de raciocínio Numérico (RN). *Educação em Debate*. 27-28 (1): 76-80.

_____. (1995a). Os testes psicológicos no Brasil: problemas, pesquisas e perspectivas. In: ALMEIDA, L. S.; RIBEIRO, I. S. (org.). *Avaliação psicológica: formas e contextos*. Braga: Associação dos Psicólogos Portugueses. p. 77-82.

_____. (1995b). Problemas e perspectivas quanto ao uso dos testes psicológicos no Brasil. *Psique*. 6: 46-57.

_____. (1996a). O uso de computadores na avaliação psicológica: verificação de sua influência sobre o desempenho individual num Teste de Raciocínio Verbal (RV). *Psicologia: Reflexão e Crítica*. 9(2): 257-268.

_____. (1996b). A construção de um Teste de raciocínio Numérico (RN) para estudantes do 2º grau. In: ALMEIDA, L. S. et al. (org.). *Avaliação psicológica: formas e contextos*. Braga: Associação dos Psicólogos Portugueses. p. 149-156.

_____. (1996c). Estudo comparativo do desempenho de estudantes do ensino médio em dois formatos (lápis-papel e computador) de um teste de Raciocínio Verbal (RV) [Resumo]. In: INSTITUTO DE PSICOLOGIA/USP (org.). *II Encontro de Técnicas de Exame Psicológico*. São Paulo: Universidade de São Paulo (USP).

_____. (1997). A influência da informática na área da avaliação psicológica. *Psique*. 7(11): 102-110.

_____. (1998). Utilização da Teoria da resposta ao Item (TRI) para a organização de um banco de itens destinados à avaliação do raciocínio verbal. *Psicologia: Reflexão e Crítica*. 11(2): 295-308.

ANDRIOLA, W. B.; CAVALCANTE, L.R. (1999). Avaliação da depressão em alunos da pré-escola. *Psicologia: Reflexão e Crítica*. 12(2): 419-428.

ANDRIOLA, W. B. (1999). Avaliação do raciocínio abstrato em estudantes do ensino médio. *Estudos de Psicologia*. 4(1): 23-37.

ANDRIOLA, W. B.; PASQUALI, L. (1995). A construção de um Teste de Raciocínio Verbal (RV). *Psicologia: Reflexão e Crítica*. 8: 51-72.

BARRET, G.V., ALEXANDER, R.A., DOVERSPIKE, D., CELLAR, D.; THOMAS, J.C. (1982). The development and application of a computerized information-processing test battery. *Applied Psychological Measurement*. 6: 13-29.

- BARRETO, J. A. E.; TROMPIERI FILHO, N.; ANDRIOLA, W. B. (1999). Desenvolvimento da estrutura cognitiva de alunos da 4ª e 8ª séries. *Educação em Debate*. 37(1): 101-113.
- BISKIN, B.H.; KOLOKTIN, R.L. (1977). Effects of computerized administration on scores on the Minnesota Multiphasic Personality Inventory. *Applied Psychological Measurement*. 1: 543-549.
- CAMARA, A. E. (1992a). *Bateria de Testes para Orientação Vocacional*. Manual de aplicação. Brasília: Centro de Educação Unificada de Brasília (CEUB).
- _____. (1992b). *Teste para Avaliação das Condições e Hábitos de Estudos de estudantes*. Manual de aplicação. Brasília: Centro de Educação Unificada de Brasília (CEUB).
- CAMPBELL, D.T.; STANLEY, J.C. (1979). *Delimitamentos experimentais e quase-experimentais de pesquisa*. São Paulo: EPU-EDUSP.
- CASTILHO, A. V., PRIMI, R.; JONAS, A. L. (1996). Informatização do Inventário de Personalidade MMPI [Resumo]. In: PUC-RS. (org). *VII Encontro Nacional sobre Testes Psicológicos. I Congresso Ibero-Americano de Avaliação Psicológica*. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).
- CORY, C.H. (1977). Relative utility of computerized *versus* paper-and-pencil tests for predicting job performance. *Applied Psychological Measurement*. 1: 551-564.
- ELWOOD, D.L.; CLARK, C.L. (1978). Computer administration of the Peabody Picture Vocabulary Test to young children. *Behavior Research Methods and Instrumentation*. 10: 43-46.
- FOGLIATO, H.M.; PÉREZ, E. (1990). Cuestionario de intereses profesionales: desarrollo y descripción. *Revista Interamericana de Psicología*. 24(1): 69-82.
- _____. (1996). La informática en la evaluación del comportamiento vocacional [Resumo]. In: PUC-RS. (Org), *VII Encontro Nacional sobre Testes Psicológicos. I Congresso Ibero-Americano de Avaliação Psicológica*. Porto Alegre: PUC-RS.
- FREMER, J.; ANASTASIO, E.J. (1969). Computer-assisted test writing. *Journal of Educational Measurement*. 6: 69-74.

- FRIENDENBERG, L. (1995). *Psychological Testing. Design, Analysis and Use*. Massachusetts: Simon; Schuster Company.
- HARRELL, T.H.; LOMBARDO, T.A. (1984). Validation of an automated 16PF administration procedure. *Journal of Personality Assessment*. 48: 638-642.
- KLINGE, V.; RODZIEWICZ, T. (1976). Automated and manual intelligence testing of the Peabody Picture Vocabulary Test on a psychiatric adolescent population. *International Journal of Man-Machine Studies*. 8: 243-246.
- LINDQUIST, E.F. (1954). *The Iowa electronic test processing equipment*. New Jersey: Educational Testing Service.
- LUSHENE, R.E., O'NEIL, H.F.; DUNN, T. (1974). Equivalent validity of a completely computerized MMPI. *Journal of Personality Assessment*. 38: 353-361.
- MERENDA, P.F. (1987). O uso do computador na avaliação psicológica. *Jornal de Psicologia*. 6(2): 24.
- MILLMAN, J.; OUTLAW, W.S. (1978). Testing by computer. *Association for Educational Data Systems (AEDS) Journal*. 11: 57-72.
- MURPHY, K.R.; DAVIDSHOFER, C.O. (1994). *Psychological testing. Principles and Applications*. New Jersey: Prentice-Hall International.
- O'BRIEN, T.; DUGDALE, V. (1978). Questionnaire administration by computer. *Journal of the Market Research Society*. 20: 228-237.
- PASQUALI, L. (1995). O problema dos parâmetros psicométricos dos testes. In: ALMEIDA, L.S.; RIBEIRO, I.S. (org.). *Avaliação psicológica: formas e contextos*. Braga: Associação dos Psicólogos Portugueses. p. 5-14.
- _____. (1997a). *Psicometria: teoria e aplicações*. Brasília: Editora Universidade de Brasília.
- _____. (1997b). A elaboração de Testes Psicológicos [Resumo]. In: *VII Encontro Nacional sobre Testes Psicológicos. I Congresso Ibero-Americano de Avaliação Psicológica*. Porto Alegre: PUC-RS.

PRIETO, G., CANO, J., ORGAZ, B.; PULIDO, R. (1993). Uso del Hypercard para la construcción de tests informatizados de aptitudes espaciales. *Psicología*. 14(2): 229-237.

_____. (1996). Medición de la visualización espacial mediante tests informatizados. *Estudios de Psicología*. 55: 41-59.

ROID, G.H. (1986). Computer technology in testing. In: PLAKE, B.S.; WITT, J.C. (org.). *The Future of Testing*. New Jersey: Lawrence Erlbaum Associates. p. 29-69.

SAMPSON Jr., J.P. (1990). Evaluating computer-assisted career guidance systems: synthesis and implications. *Journal of Career Development*. 17(2): 143-149.

SAPINKOPF, R.G. (1978). A computer adaptative testing approach to the measurement of personality variables. *Dissertation Abstracts International*. 38(10-B): 4993.

TAVEIRA, M.C. (1992). Sistema Projectos: Orientação Vocacional apoiada por computador. *Jornal de Psicologia*. 10(1): 21-25.

TAVEIRA, M.C.; CAMPOS, B.P. (1989). Computer assisted career guidance for secondary education. *European Conference on Computers in Careers Guidance*. Cambridge: Cambridge University.

THIERS, V.O., CAPOVILLA, F.C., MACEDO, E.C.; DUDUCHI, M. (1996). Versão computadorizada: subteste Arranjo de Figuras do WISC III-R [Resumo]. In: INSTITUTO DE PSICOLOGIA/USP (org.). *II Encontro de Técnicas de Exame Psicológico*. São Paulo: Universidade de São Paulo (USP).

VIEIRA, S.; HOFFMANN, R. (1989). *Estatística experimental*. São Paulo: Atlas.

WATTS, K., BADDLEY, A.; WILLIAMS, M. (1982). Automated tailored testing using Raven's Matrices and the Mill Hill Vocabulary Tests: a comparison with manual administration. *International Journal of Man-Machine Studies*. 17: 331-344.

ZACHARY, R.A.; POPE, K.S. (1983). Legal and ethical issues in the clinical use of computerized testing. In: SCHWARTZ, M. D. (ed.). *Using Computers in Clinical Practice*. New York: Haworth Press. p. 37-68.

WAGNER BANDEIRA ANDRIOLA

Universidade Federal do Ceará – Campus do Benfica – Faculdade de Educação – Deptº
Fundamentos da Educação

Rua Waldery Uchoa, 1 – 60020-110 – Benfica – Fortaleza/CE

tel/fax: (85) 288-7677

e-mail: w_andriola@yahoo.com

- recebido em 06/09/02 –
- aprovado em 26/06/03 –