

Teste dos Relógios e R-1 – Forma B: Evidências de validade

Ana Paula Porto Noronha
Acácia Aparecida Angeli dos Santos
Fermino Fernandes Sisto
Universidade São Francisco (USF)

RESUMO

Considerando a necessidade de construção de testes brasileiros, com análises apropriadas de parâmetros psicométricos, o presente estudo objetivou analisar as evidências de validade do *R-1: Teste não-verbal de avaliação da inteligência*, por meio da correlação desse instrumento com o Teste dos Relógios. Foram estudados 68 alunos de cursos de Educação de Jovens e Adultos, com idade média de 34,12 e desvio padrão de 13,76. Os resultados indicaram que a correlação entre os dois instrumentos foi significativa e com uma magnitude que pode ser considerada boa (0,64), indicando que 41% da variância é comum a ambos os testes. Entretanto, não foi aceita a hipótese de que o Teste dos Relógios estaria mais correlacionado com o Fator 2 do R-1.

Palavras-chave: Avaliação psicológica, testes de inteligência, validade de construto.

ABSTRACT

Teste dos Relógios e R-1 – Forma B: Evidences of validity

Considering the necessity of Brazilian construction of tests with appropriated psychological parameters studies, the present study aimed to analyze evidences of validity about R-1: *Teste não-verbal de avaliação da inteligência* correlating it with the Teste dos Relógios. 68 students (mean age of 34.12 and standard deviation of 13.76) attending to the Educação de Jovens e Adultos Courses were studied. The results indicated that the correlation between the instruments was significant and with a magnitude that can be considered good (.64) indicating that 41% of the variance is common in both tests. However, the hypothesis of bigger correlation between the Teste dos Relógios and the Factor 2 than the Factor 1 of R-1 Test was not accepted.

Key words: Psychological assessment; intelligence tests; construct validity.

INTRODUÇÃO

No Brasil, a partir de 1990, foi identificada uma preocupação com os problemas epistemológicos e com os fundamentos da medida, assim como com a produção de instrumentos válidos e precisos (Pasquali, 2001). Tal preocupação reflete-se nos artigos de pesquisas publicados nos periódicos nacionais como, por exemplo, o de Noronha, Sartori, Freitas e Ottati (2001), no qual foram analisados 21 manuais de testes de inteligência. Cada instrumento foi estudado com base em critérios de um questionário para avaliar a qualidade dos testes usados na Espanha (Prieto e Muñiz, 2000). Os resultados revelaram que alguns testes não possuem nos seus manuais, informações como data de publicação, nome do autor ou bibliografia e que de uma maneira geral, a análise pertinente à aplicação e à avaliação recebeu uma pontuação maior do

que os dados de identificação, exceto no que se refere à bibliografia, que foi mal pontuada.

Os resultados foram corroborados pelo trabalho desenvolvido por Noronha e cols. (2003), cujo objetivo foi o de analisar os estudos de validade, precisão e padronização de 19 instrumentos de avaliação da inteligência. Lamentavelmente o trabalho permitiu constatar que parte dos instrumentos comercializados no Brasil necessita de fundamentação científica e não atende aos requisitos mínimos exigidos pela comunidade acadêmica, científica, bem como pelo Conselho Federal de Psicologia. No que se refere aos estudos de validade e precisão, os achados indicaram que aproximadamente 74% dos instrumentos apresentam os referidos estudos, embora não necessariamente conjuntamente. Esse dado, ainda que signifique um avanço em relação aos dados relatados por Sisto, Codenotti, Costa e Nascimento (1979), sugere que a

qualidade psicométrica dos testes ainda é um tema que permanentemente deve ser focalizado pelos estudiosos da avaliação psicológica.

Também nesse sentido, Noronha, Freitas e Ottati (2002) analisaram os parâmetros psicométricos 26 testes de inteligência. Os resultados obtidos demonstraram que, quase na totalidade, os manuais dos testes apresentavam estudos de validade e de precisão, embora muitos não trouxessem informações a respeito das amostras de normatização.

Tais considerações mostram a necessidade de estudos brasileiros na área de avaliação psicológica, incluindo a construção de materiais nacionais, com estudos apropriados de análise dos parâmetros psicométricos. Conseqüentemente, identifica-se a importância de pesquisas realizadas com amostras adequadas e que proponham a verificação dos construtos mais utilizados nos diferentes campos profissionais do psicólogo. Decorrente dessa preocupação, o presente trabalho visa analisar evidências de validade do *R-1: Teste não-verbal de avaliação da inteligência*.

O construto "inteligência" tem sido explicado por uma diversidade de teorias e segundo Primi (2002) os testes de avaliação da inteligência estão cada vez mais sofisticados. No entanto, sabe-se que ainda hoje não é possível encontrar uma única bateria de testes capaz de avaliar todos os fatores amplos da inteligência. Vale destacar que os testes de inteligência, assim como os de personalidade, são os instrumentos mais utilizados na prática do psicólogo e isto se justifica em razão da exigência presente nos próprios contextos de atuação profissional com fortes demandas para a avaliação da inteligência (Vendramini e Noronha, 2002). Aliado a isso há maior número de testes para avaliação deste construto do que dos demais, tais como orientação vocacional, personalidade, dentre outros.

Os primeiros instrumentos de avaliação da inteligência tinham sustentação psicofísica. Por acreditarem que o construto estava basicamente relacionado a uma base sensorial, Os pesquisadores construíram medidas pautadas nas características da sensibilidade e da discriminação sensorial dos indivíduos. Dentre eles, Galton recebeu destaque, considerando que, por volta de 1880, criou instrumentos para aferir a discriminação sensorial, por meio de barras para medir percepção de comprimento e apito para percepção de altura do tom (Anastasi e Urbina, 2000).

No que se refere às concepções teóricas de inteligência, a perspectiva mais generalista e tradicional dentro da psicologia foi proposta por Spearman em 1927. Este autor considerou que toda a atividade intelectual está expressa em um fator geral (*g*), comum a toda atividade mental e possuidor de uma índole fisiológica, e um fator específico (*s*) associado a tarefas

individualizadas. Mais tarde acrescentou os fatores de grupo, constituindo uma teoria hierárquica da inteligência. Nesse sentido, Almeida, (1988) resalta que a inteligência tem sido concebida como entidade simples ou complexa de fatores, traços ou aptidões mentais, que compõem a mente humana.

Ainda no que se refere às várias definições de inteligência, Almeida (1988) enfatiza que Thurstone e Guilford assumiram uma posição contrária à de Spearman. O primeiro autor, por meio de uma análise fatorial, propôs sete fatores, entendidos como aptidões mentais diferenciadas, a saber, compreensão verbal, fluência verbal, aptidão numérica, velocidade perceptiva, aptidão espacial, memória, raciocínio, dedução e indução.

Posteriormente Cattell (1971) assumiu a existência de dois fatores gerais, os quais batizou com os nomes de Inteligência Fluida (*Gf*) e Inteligência Cristalizada (*Gc*). A fluida pode ser entendida como a capacidade para raciocinar em situações novas ou inesperadas, enquanto a cristalizada, como representante da profundidade e quantidade de conhecimentos e experiências adquiridos (Schelini, 2000). À concepção *Gf-Gc*, Horn (1991) adicionou outras quatro capacidades cognitivas, quais sejam, processamento visual, memória de curto prazo, armazenamento e recuperação a longo prazo e velocidade de processamento. Mais tarde, o autor acrescentou outras duas capacidades às anteriores (rapidez para a decisão correta e processamento auditivo).

Carroll (1993) analisou as principais teorias da inteligência. Por meio de vários estudos de meta análise diferenciou as capacidades em três diferentes camadas ou estratos, a saber, I) camadas específicas, II) camadas amplas ou gerais e, III) relativa à uma única camada geral.

Estudos mais recentes (Flanagan, McGrew e Ortiz, 2000) propõem a integração entre os modelos de Cattell e Horn (teoria da inteligência fluida e cristalizada) e o modelo de Carroll (teoria dos três estratos), denominada Teoria Psicométrica das Habilidades Cognitivas de Cattell-Horn-Carroll (CHC). Esse modelo enfatiza uma natureza multidimensional da inteligência e vem gradualmente sendo considerado um modelo consensual sobre a organização estrutural das capacidades à medida que fornece uma taxonomia para classificação dos fatores medidos pelos testes de inteligência (McGrew e Flanagan, 1998).

É nesse contexto teórico que o manual da Forma B do *'R-1: Teste não-verbal de inteligência'* (Sisto, Noronha e Santos, no prelo) propõe sua fundamentação, com base em dois tipos de estudos. Em um deles seus itens foram analisados por meio de análise fatorial e dois fatores foram constatados. Em um se-

gundo estudo, objetivando conhecer as habilidades cognitivas envolvidas, os itens foram analisados qualitativamente com base nas leis que regem o fator *g* (Spearman, 1927) e na Teoria de *Gf-Gc* de Cattell (1992).

Os autores do manual informam que o Fator 1 engloba, em maior quantidade, itens que podem depender de aprendizagem cultural e de peculiaridades do meio ambiente, como também outros, que exigem como resposta igualdade, similaridade ou continuidade gestáltica e que podem ser interpretados como de aprendizagem cultural. A esse respeito recuperaram a informação de que Spearman (1927) afirmou que itens com essas características perceptuais mantêm relação com a capacidade de educação. Além de itens com essas características, o Fator 1 também congrega itens que demandam educação de relação e de correlatos, ainda que em bem menor número. Em razão dessas peculiaridades, concluíram que o Fator 1 se assemelha à denominada inteligência cristalizada de Cattell (1971). Por sua vez, os itens do Fator 2 estão mais claramente definidos, por mais específicos. Segundo os autores do manual tratam fundamentalmente de educação de relações e de correlatos, o que estaria em consonância ao fator *g* de Spearman.

O Teste dos Relógios (Souza e Cambraia, 2002) apresenta uma fundamentação teórica com base na teoria de Spearman (1927) ainda que não tenha sido feita uma análise qualitativa dos itens para ver se correspondem às três leis noegenéticas. Não é realizado também análise fatorial ou qualquer outro estudo sobre a consistência interna dos itens. Apresentam como evidência de validade uma correlação de 0,68 com o *R-1 Teste não-Verbal de Inteligência* (Oliveira, 1963).

Com base nas informações de ambos os manuais, é possível supor que os instrumentos estejam medindo os mecanismos psicológicos propostos por Spearman (1927) como essenciais para resolver problemas que englobam o fator *g*. Além disso, pela fundamentação teórica contida no manual do *‘Teste dos Relógios’* é possível supor uma correlação maior com o Fator 2 da Forma B do *‘R-1: Teste não-verbal de inteligência’*.

MÉTODO

Participantes

Foram estudados 68 alunos de cursos de Educação de Jovens e Adultos, 50% deles freqüentando o telecurso de segundo grau e os outros 50% o primeiro grau. Desses alunos, 33 (48,5%) eram do sexo feminino e 35 (51,5%) do gênero masculino. A idade mínima foi de 16 anos e a máxima de 65, com uma média de 34,12 e um desvio padrão de 13,76.

Instrumentos

O manual da Forma B do *‘R-1: Teste não-verbal de inteligência’* foi elaborado por Sisto, Noronha e Santos (no prelo). Esse teste foi desenvolvido por Rynaldo de Oliveira, publicado pela Editora Vetor em 1973. Foi proposto para ser empregado no exame psicotécnico de motoristas, bem como para pessoas com baixa escolaridade ou mesmo analfabetas, como também para aquelas com dificuldades específicas para a compreensão do português.

Por meio da análise fatorial foi realizado o estudo de sua validade de construto. Os resultados apontaram para a existência de dois fatores, sendo que os itens do Fator 2 tratam basicamente de educação de relações e de correlatos. Já os itens do Fator 1, mesmo possuindo elementos de educação de relação e de correlatos, possui também itens que podem depender de aprendizagem cultural e de peculiaridades do meio ambiente, igualdade ou similaridade gestáltica e continuidade gestáltica. Para o estabelecimento das normas foram estudados 752 alunos de Cursos de Educação de Jovens e Adultos, com média de idade de 32,11 anos, variando entre 15 e 76 anos, com desvio padrão de 12,16 anos.

O instrumento é constituído por 40 itens apresentados em um caderno, com folha de resposta à parte. Para cada questão há de 6 a 8 possibilidades de respostas, sendo que os itens iniciais caracterizam-se por figuras de objetos concretos e os demais são figuras geométricas. Sua correção implica em atribuir um ponto a cada resposta correta, sendo 40 o máximo de pontos possíveis.

O Teste dos Relógios está na sua 4ª revisão (Souza e Cambraia, 2002). A primeira forma foi publicada em 1967 com o objetivo de minimizar a influência da escolaridade sobre os resultados de um teste de inteligência. A versão atual (Souza e Cambraia, 2002) é composta por 40 problemas (além dos 4 exemplos), sendo que cada problema é formado por 9 relógios.

A validade do teste foi determinada pela comparação dos resultados do Relógio – Forma B com o R-1 Teste não-Verbal de Inteligência. A correlação de Pearson determinou coeficiente de 0,68 (significativo a 0,01). Já a precisão foi obtida pelo método das metas, o que gerou coeficiente de 0,93.

Para o estabelecimento das normas do instrumento foi aplicado em 348 sujeitos, do sexo masculino, que participavam de processos seletivos realizados pela Editora Vetor, nos estados de São Paulo, Minas Gerais e Rio de Janeiro. A idade dos sujeitos variou de 29 a 69 anos, sendo a idade média 39,48 anos e o DP 6,07. O grau de escolaridade dos sujeitos variou do Ensino Fundamental incompleto ao Ensino Médio completo.

Procedimento

Após a devida autorização do Comitê de Ética e das escolas, deu-se início à coleta de dados. A aplicação do instrumento foi coletiva, em grupos de 15 a 20 sujeitos, sendo que em cada sala havia pelo menos um aplicador e um auxiliar. O tempo médio de aplicação foi de 45 minutos para cada instrumento. Houve um intervalo de aproximadamente 30 minutos entre as aplicações dos dois testes.

Para efeitos de análise foram consideradas algumas medidas possibilitadas pelos instrumentos, assim como faixas etárias. Com relação ao R1 – Forma B foram usadas sete medidas, quais sejam, a pontuação total, do Fator 1, do Fator 2, dos itens de 1 a 10, dos itens de 11 a 20, dos itens de 21 a 30 e finalmente dos itens de 31 a 40. Concernente ao Teste dos Relógios, foram levadas em consideração cinco medidas, ou seja, a pontuação total, dos itens de 1 a 10, dos itens de 11 a 20, dos itens de 21 a 30 e finalmente dos itens de 31 a 40.

As faixas etárias usadas para estudo foram as que o manual do R-1 indica como significativas. São elas: faixa etária 1 de 15 a 26 anos, faixa etária 2 de 27 a 37 anos, faixa etária 3 de 38 a 49 anos e, finalmente, faixa etária 4 de 50 a 76 anos.

Critérios de correção

A correção foi feita de acordo com os critérios fornecidos pelos manuais. Foi atribuído um ponto para cada resposta correta e de cada instrumentos foram retiradas algumas medidas, conforme destacado a seguir.

RESULTADOS

A Tabela 1 apresenta os dados de estatística descritiva de ambos os instrumentos, para as várias medidas. Como pode ser observada, nas medidas dos itens 21 a 30 e de 31 a 40 do R-1 – Forma B nenhum sujeito obteve a pontuação máxima possível nesse intervalo, diferentemente das outras duas medidas em que todas tiveram pontuação máxima possível (10 pontos). Em relação ao Teste dos Relógios, fato semelhante pôde ser notado nos segmentos compostos pelos itens de 1 a 10 e itens de 31 a 40, pois a pontuação máxima possível por segmento também seria de 10 pontos em todos os casos.

Tanto no R-1 quanto no Teste dos Relógios é possível observar que a média de acerto é inversamente proporcional aos grupos de itens, ou seja, enquanto nos itens de 1 a 10 no R-1 a média é 7,66, nos itens de 31 a 40 é de 1,29. Ainda no que se refere às médias, o Teste dos Relógios teve menos acertos do que o R-1.

Em seguida foram analisados os dados com vistas a constatar a existência de relações entre idade e gênero e as medidas em estudo. Para tanto, foram analisadas as diferenças de média entre os grupos masculino e feminino, por meio da prova *t* de Student e correlacionadas as várias medidas obtidas por meio de ambos os instrumentos com a idade dos participantes. Os resultados estão na Tabela 2.

As análises indicaram que os valores de *t* não alcançaram significância em nível de 0,05, em quaisquer das medidas estudadas. Assim, não se averiguou diferenças entre gênero em relação às pontuações do R1-Forma B e do Teste dos Relógios.

TABELA 1
Dados descritivos para o R-1 – Forma B e o Teste dos Relógios nas várias medidas.

	Valor mínimo	Valor máximo	Média	Média do erro padrão	Desvio padrão
<i>R1 – Forma B</i>					
Itens de 1 a 10	0	10	7,66	0,35	2,91
Itens de 11 a 20	0	10	6,09	0,40	3,29
Itens de 21 a 30	0	9	3,97	0,30	2,48
Itens de 31 a 40	0	4	1,29	0,16	1,32
Fator 1	1	19	13,72	0,72	5,94
Fator 2	0	12	5,29	0,42	3,49
Pontuação total	1	31	19,02	1,06	8,78
<i>Teste dos Relógios</i>					
Itens de 1 a 10	0	9	5,41	0,32	2,60
Itens de 11 a 20	0	10	4,44	0,40	3,30
Itens de 21 a 30	0	10	3,25	0,36	2,97
Itens de 31 a 40	0	9	1,68	0,27	2,24
Pontuação total	0	37	14,78	1,08	8,90

TABELA 2
Valores de *t* de Student para gênero e as correlações (*r*) para idade por bloco de itens nas medidas do R-1 – Forma B e o Teste dos Relógios.

	<i>t</i>	<i>r</i>
<i>R-1 – Forma B</i>		
Itens de 1 a 10	1,765	-0,440**
Itens de 11 a 20	0,507	-0,379**
Itens de 21 a 30	1,586	-0,574**
Itens de 31 a 40	0,496	-0,491**
Fator 1	1,220	-0,429**
Fator 2	1,163	-0,586**
Pontuação total	1,292	-0,524**
<i>Teste dos Relógios</i>		
Itens de 1 a 9	0,612	-0,297*
Itens de 10 a 18	0,996	-0,309*
Itens de 19 a 27	0,426	-0,247*
Itens de 28 a 36	-1,498	-0,013
Pontuação total	0,317	-0,287*

* Significativa em nível de 0,05.

** Significativa em nível de 0,01.

Optou-se, ainda, por utilizar a prova de correlação momento-produto (r) de Pearson para investigar o grau de associação entre as pontuações do R-1 – Forma B e do Teste dos Relógios com relação às idades. Vale destacar que essa prova paramétrica foi escolhida por se tratar de escores que provêm de uma população normal bivariada (Siegel, 1975). Neste caso, em apenas uma situação não foi encontrado um coeficiente de correlação significativo a 0,05, qual seja, o segmento contendo os itens de 28 a 36. Em todas as outras situações as correlações foram significativas e negativas, indicando que conforme aumentavam as idades dos respondentes, diminuam suas pontuações.

Os coeficientes de correlação de Pearson entre as pontuações do R-1 – Forma B (Fatores 1 e 2 e total) e a pontuação do Teste dos Relógios estão na Tabela 3. Os valores foram calculados para as quatro faixas etárias e independentemente delas. São apresentados (Figura 1), também, três gráficos de dispersão em relação às correlações entre os dados, independentemente das faixas de idade.

Como pode ser observado, a correlação entre os dois instrumentos foi significativa e com uma magnitude que pode ser considerada boa (0,64), indicando que 41% da variância é comum a ambos os testes. Com relação aos fatores 1 e 2, o Teste dos Relógios está

levemente mais correlacionado com o Fator 2 (0,62) que com o Fator 1 (0,58).

No que tange às faixas etárias, os coeficientes de correlações variaram e o menor foi o coeficiente produto do Fator 1 na faixa etária 1. De fato, nessa faixa, a correlação média foi a menor das quatro faixas (0,44). Nas outras duas situações dessa faixa etária as correlações foram significativas, mesmo não podendo ser consideradas altas. Esses dados sugerem que nessas idades mesmo existindo uma comunalidade entre os mecanismos avaliados pelos dois instrumentos, há uma importante parte da variância que indica a presença de outro ou outros mecanismos sendo solicitados para a resolução dos itens dos dois instrumentos.

Nas faixas etárias 2 e 3 as médias de correlação encontradas (0,67 e 0,60, respectivamente) facilitam a possibilidade de se falar em bons níveis e uma comunalidade maior dos mecanismos exigidos para a resolução dos itens de ambos instrumentos. Entretanto, há uma parte substancial da variância não explicada pelo que há de comum nos dois instrumentos, sugerindo também a presença de outro ou outros mecanismos para resolver os itens propostos nos testes. Ao lado disso, na faixa etária 4, a média de correlação encontradas (0,81) sugere quase que formas paralelas avaliando o desempenho cognitivo.

TABELA 3
Coeficientes de correlação de Pearson entre as pontuações do R-1 – Forma B (Fatores 1 e 2 e total) e a pontuação do Teste dos Relógios para as faixas etárias e independentemente das faixas etárias.

	Todas as idades (n=68)	Faixa 1 (n=25)	Faixa 2 (n=18)	Faixa 3 (n=13)	Faixa 4 (n=12)
Pontuação total R1-Forma B	0,64**	0,46*	0,70**	0,65*	0,84**
Fator 1 – R-1 – Forma B	0,58**	0,28	0,68**	0,64*	0,81**
Fator 2 – R-1 – Forma B	0,62**	0,59**	0,63**	0,53	0,80**

* Significativa em nível de 0,05.
** Significativa em nível de 0,01.

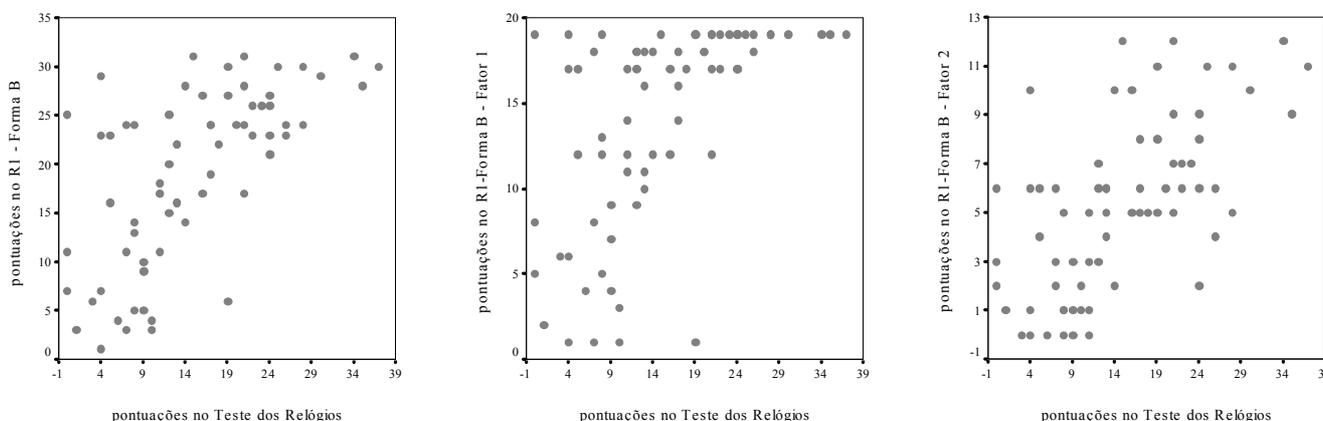


Figura 1 – Gráficos de dispersão das relações entre as pontuações do R-1 – Forma B (Fatores 1 e 2 e total) e a pontuação do Teste dos Relógios

No geral, a pontuação global do Teste R-1 – Forma B teve, nas cinco situações estudadas, apenas em uma delas um valor inferior a 0,60, mas todos eles significativos. Em relação ao Fator 1, ainda que com um coeficiente bastante baixo entre as cinco situações estudadas, os coeficientes são todos significativos e três deles acima de 0,60. Por sua vez, no que tange ao Fator 2, todos os coeficientes foram significativos e dois deles abaixo de 0,60. Esses resultados facilitam a interpretação de que essas informações podem ser usadas como evidência de validade.

Isso significa que a boa parte do Fator 1 e do Fator 2 e da pontuação geral pode ser explicada pelo Teste dos Relógios. De certa forma, ambas as medidas possuem em comum parte substantiva dos mecanismos implicados na resolução dos itens, ainda que uma outra parte da variância, de forma alguma desprezível, indique a presença de outro mecanismo como necessário para resolver esses itens.

Apesar disso, não se pode afirmar a comprovação da hipótese inicialmente colocada. De fato, os dados parecem sugerir que o Teste dos Relógios relaciona-se em intensidade semelhante, dificultando concluir que meça apenas o fator *g* como proposto em seu manual. Com vistas a aprofundar um pouco mais essa questão, na Tabela 4 são apresentados os coeficientes de correlação entre os vários segmentos de ambos os testes.

TABELA 4

Coeficientes de correlação de Pearson entre as pontuações dos segmentos do R-1 – Forma B e dos segmentos do Teste dos Relógios.

		<i>Teste dos Relógios</i>			
		<i>Itens de 1 a 10</i>	<i>Itens de 11 a 20</i>	<i>Itens de 21 a 30</i>	<i>Itens de 31 a 40</i>
<i>R-1 – Forma B</i>	<i>Itens de 1 a 10</i>	0,56**	0,46**	0,40**	0,21
	<i>Itens de 11 a 20</i>	0,61**	0,58**	0,42**	0,21
	<i>Itens de 21 a 30</i>	0,57**	0,61**	0,48**	0,24
	<i>Itens de 31 a 40</i>	0,38**	0,43**	0,47**	0,25*

* Significativa em nível de 0,05.

** Significativa em nível de 0,01.

Por esses dados pode-se observar que os segmentos dos itens de 1 a 10 e de 11 a 20 do Teste dos Relógios foram os que forneceram as médias de correlações mais altas em relação aos outros dois segmentos. Também, que do primeiro ao último segmento houve uma diminuição gradativa e constante das médias dos coeficientes de correlação.

Por sua vez, os segmentos dos itens 11 a 20 e 21 a 30 do R-1 – Forma B tiveram médias de correlações

mais altas que os outros dois segmentos desse mesmo teste. Esses dados permitem levantar a hipótese de que o Teste dos Relógios pode também não ser unidimensional.

DISCUSSÃO E CONCLUSÕES

O presente trabalho versou sobre a análise de evidências de validade convergente do R-1 por meio do estudo de correlação entre os resultados obtidos com sua aplicação e a do Teste dos Relógios no mesmo grupo de alunos. Estudos que proponham esse tipo de verificação, utilizando desse ou de outro procedimento de análise são necessários no cenário nacional, considerando as deficiências na construção, normatização e na validação dos instrumentos de avaliação da inteligência apontadas desde 1979 por Sisto, Codenotti, Costa e Nascimento, e mais recentemente, por Noronha, Sartori, Freitas e Ottati (2001), Noronha, Freitas e Ottati (2002), Vendramini e Noronha (2002), Noronha e cols. (2003), dentre outros.

A validação convergente propõe a relação entre os resultados do teste e outras medidas que avaliam um construto diferente ou similar. Quando os construtos são similares, constitui-se a validade convergente e, quando são distintos, divergente ou discriminante. (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999).

O R-1 Teste não-Verbal Forma B utiliza-se do referencial proposto por Cattell (1971) para explicar o Fator 1 e das três leis noegenéticas de Spearman (1927) para explicar o Fator 2. Assim, o Fator 1 contém essencialmente solicitações que dependem de aprendizagem cultural e de peculiaridades do meio ambiente, além de solicitações que demandam resposta de igualdade, similaridade ou continuidade gestáltica como também, em menor quantidade e em forma mais elementar, de educação de relação e de correlatos. Ao lado disso, o Fator 2 aborda principalmente educação de relações e de correlatos, o que estaria em consonância ao fator *g* de Spearman.

Já o Teste dos Relógios foi concebido tendo como base que a atividade intelectual está expressa em um fator geral (*g*), comum a toda atividade mental e possuidor de uma índole fisiológica, a fatores de grupo e um fator específico (*s*) associado a tarefas individualizadas. Nesse sentido, a inteligência é concebida como entidade simples ou complexa de fatores, traços ou aptidões mentais, que compõem a mente humana (Spearman, 1927).

Com base nessas duas informações dos construtos dos testes em estudo foi aventada a hipótese de que o Teste dos Relógios estaria mais correlacionado com o

Fator 2 do R-1 do que com o Fator 1. Entretanto, observou-se que o Teste dos Relógios está apenas levemente mais correlacionado com o Fator 2 do R-1, que por sua vez, trata basicamente de educação de relações e de correlatos, o que estaria em consonância ao fator g de Spearman (1927).

De fato, considerando os resultados encontrados nesse estudo, pôde-se inferir que os valores das correlações indicaram que os fatores em questão foram significativos e satisfatórios o suficiente para justificar a similaridade entre eles e o Teste dos Relógios. Ainda, a pontuação total do R-1 a correlação foi bastante parecida com as encontradas para os dois fatores separadamente. Nesse contexto, os dados desta pesquisa facilitam a interpretação de que possivelmente o Teste dos Relógios não seja unidimensional.

Ainda no que se refere à validação de testes, Anastasi e Urbina (2000) apontam que o nível desenvolvimental pode ser considerado como evidência de validade. Nesse sentido, as correlações entre R-1 – Forma B e Relógios com as idades foram significativas e negativas, indicando que conforme aumentavam as idades dos respondentes, diminuía suas pontuações. O resultado está em concordância com o pressuposto por Spearman (1927) no sentido de que o fator (g) possui uma índole fisiológica, que tende a declinar com a idade.

Esse trabalho não teve a intenção de realizar todas as análises possíveis, mas que de alguma forma pode vir a amenizar a necessidade da área de avaliação psicológica, à medida que apresenta evidências de validade para o R-1, instrumento bastante utilizado pelos psicólogos. Sugere-se que novos estudos sejam desenvolvidos tanto no que se refere ao R-1 como no que tange ao Teste dos Relógios. Este último mereceria uma análise quanto à questão da unidimensionalidade. Também, outros tipos de evidência poderiam ser buscados o que ajudaria a entender um pouco mais os construtos que parecem subjazer aos dois testes.

REFERÊNCIAS

Almeida, L. S. (1988). *Teorias da inteligência*. Porto: Edições Jornal de Psicologia.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington: American Educational Research Association.

Anastasi, A. & Urbina, S. (2000). *Testagem psicológica*. Porto Alegre: Artes Médicas.

Carroll, J.B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cattell, R. B. (1992). The relevance of fluid and crystallized intelligence concepts to nature-nurture investigation. *The Manking Quarterly*, 23, 359-375.

Flanagan, D. P., McGrew, K. S. & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales and CHC Theory: A contemporary approach to interpretation*. Boston: Allyn & Bacon.

Horn, J. L. (1991). Measurement of Intellectual Capabilities: a review of theory. In K. S. McGrew, J. K. Werder & R. W. Woodcock. *Woodcock-Johnson technical manual* (pp. 197-232). Chicago: Riverside.

McGrew, K. S. & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR) – Gf-Gc cross battery assessment*. Boston: Allyn and Bacon.

Noronha, A. P. P., Vendramini, C. M. M., Canguçu, C., Souza, C. V. R., Cobero, C., Paula, L. M., Lima, O. M. P., Guerra, P. B. C., Filizatti, R. & Franco, M. O. (2003). Propriedades psicométricas apresentadas em manuais de testes de inteligência. *Psicologia em Estudo*, 8, 1, 93-99.

Noronha, A. P. P., Freitas, F. A. & Ottati, F. (2002). Parâmetros Psicométricos de Testes de Inteligência. *Interação em Psicologia*, 6, 2, 195-202.

Noronha, A. P. P., Sartori, F. A., Freitas, F. A. & Ottati, F. (2001). Informações contidas nos manuais dos testes: um estudo com testes de inteligência. *Psicologia em Estudo*, 6, 2, 101-106.

Oliveira, R. (1963). *R1: Teste não verbal de inteligência*. Manual. São Paulo: Vetor Editora Psico-Pedagógica.

Pasquali, L. (2001). *Técnicas de exame psicológico – TEP*. Vol. I: Fundamentos das técnicas psicológicas. São Paulo: Casa do Psicólogo.

Prieto, G. & Muñoz, J. (2000). *Um modelo para evaluar la calidad de los tests utilizados em España*. [On line] <http://www.cop.es/tests/modelo.htm> (Acessado em 04/12/2000).

Primi, R. (2002). Inteligência fluida: definição fatorial, cognitiva e neuropsicológica. *Paidéia*, 12, 2, 57-77.

Primi, R. & Almeida, L. S. (2000). Estudo de Validação da Bateria de Provas de Raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa*, Brasília, 16, 2, 165-173.

Schelini, P. W. (2000). *Bateria multidimensional de inteligência infantil: Proposta de instrumento*. [Tese de Doutorado]. PUC-Campinas, Campinas-SP.

Siegel, S. (1975). *Estatística não-paramétrica*. São Paulo: McGraw-Hill do Brasil.

Sisto, F. F., Codenotti, N., Costa, C. A. J. & Nascimento, T. C. N. (1979). Testes Psicológicos no Brasil: que medem realmente? *Educação e Sociedade*, 1, 2, 152-165.

Sisto, F. F., Noronha, A. P. P. & Santos, A. A. A. (no prelo). *R-1: Teste não-verbal de inteligência – Forma B*. Manual. São Paulo: Vetor Editora Psico-Pedagógica.

Souza, F. V. S. & Cambraia, S. V. (2002). *Teste dos Relógios. B e C*. Manual, (4ª ed.). São Paulo: Vetor Editora Psico-Pedagógica.

Spearman, C. (1927). *Las Habilidades del Hombre: su naturaleza y medición*. Buenos Aires: Paidós.

Vendramini, C. M. M. & Noronha, A. P. P. (2002). Estudo comparativo entre testes de inteligência e de personalidade. *Psico*, 33, 2, 413-426.

Recebido em: 27/10/2004. Aceito em: 22/11/2005.

Autores:

Ana Paula Porto Noronha – Doutora em Psicologia: Ciência e Profissão pela PUC-Campinas. Docente em Psicologia da Universidade São Francisco – Itatiba-SP. Bolsista Produtividade do CNPq.

Acácia Aparecida Angeli dos Santos – Doutora em Psicologia Escolar e do Desenvolvimento Humano pela USP. Docente em Psicologia da Universidade São Francisco – Itatiba-SP. Bolsista Produtividade do CNPq.

Fermino Fernandes Sisto – Doutor em Pedagogia pela Universidad Complutense de Madrid. Docente em Psicologia da Universidade São Francisco – Itatiba-SP. Bolsista Produtividade do CNPq.

Endereço para correspondência:

ANA PAULA PORTO NORONHA
Rua Alexandre Rodrigues Barbosa, 45 – Centro
CEP 13251-900, Itatiba, SP, Brasil
Fone: (11) 4534-8040 – E-mail: ana.noronha@saofrancisco.edu.br