

INDICAÇÕES DA VALIDADE CONVERGENTE DO TESTE NÃO-VERBAL DE INTELIGÊNCIA SON-R 2½-7[A]

Jacob A. Laros¹ - Universidade de Brasília, Brasília, Brasil
 Rafaela F. Reis - Universidade de Brasília, Brasília, Brasil
 Peter J. Tellegen - University of Groningen, Groningen, Holanda

RESUMO

O objetivo deste estudo foi avaliar a validade convergente do teste não-verbal de inteligência SON-R 2½-7[a] no Brasil. O teste em questão é a versão abreviada do SON-R 2½-7 que foi publicado na Holanda em 1998 e que foi normatizado e validado em vários países de Europa. A versão abreviada foi administrada no Brasil em 120 crianças com idade variando entre 5 e 7 anos junto com as Matrizes Progressivas Coloridas (MPC) do Raven e com a Escala de Maturidade Mental Colúmbia (EMMC). A correlação, corrigida para atenuação, do SON-R 2½-7[a] com estes dois testes foi respectivamente, 0,77, e 0,62. Foi observada uma relação particularmente forte entre o MPC e a escala de execução do SON-R 2½-7[a], e entre o EMMC e a escala de raciocínio do SON-R 2½-7[a]. Os resultados do presente estudo indicam uma boa validade convergente do teste para a faixa de idade investigada.

Palavras-chave: Teste não-verbal de inteligência; Validade convergente; SON-R 2½-7[a]; Correção para atenuação.

INDICATIONS OF THE CONVERGENT VALIDITY OF THE NON-VERBAL TEST OF INTELLIGENCE SON-R 2½-7[A]

ABSTRACT

This study aimed to investigate the convergent validity of the nonverbal test of intelligence SON-R 2½-7[a] in Brazil. This test is the abridged edition of the SON-R 2½-7 nonverbal test of intelligence that was published in 1998 in The Netherlands with normatization and validation studies realized in various European countries. The abridged edition was administered in Brazil to 120 children with ages varying between 5 and 7 years together with Raven's Colored Progressive Matrices (CPM) and the Columbia Mental Maturity Scale (EMMC). The correlation, corrected for attenuation, of the SON-R 2½-7[a] with these two tests was .77 and .62, respectively. A particularly strong relationship was observed between the CPM and the performance scale of the SON-R 2½-7[a], and between the EMMC and the SON-R reasoning scale. The results of this study indicate a fair convergent validity of the SON-R 2½-7[a] for the investigated age group.

Keywords: Nonverbal tests of intelligence; Convergent validity; SON-R 2½-7[a]; Correction for attenuation.

INTRODUÇÃO

Existe uma necessidade crescente no Brasil de testes psicológicos válidos, fidedignos e com normas brasileiras para a avaliação de habilidades cognitivas de crianças pré-escolares e escolares (Alves, 2002; Noronha & Alchieri, 2002; Noronha e colaboradores, 2003; Muñiz, Prieto, Almeida & Bartram, 1999; Oakland, Wechsler, Bensuan & Stafford, 1994). Testes com qualidades psicométricas comprovadas podem contribuir consideravelmente para diagnósticos clínicos de atrasos ou dificuldades específicas do desenvolvimento cognitivo de crianças. Os resultados de uma revisão recente sobre os testes psicológicos comercializados no Brasil

(Noronha, Primi & Alchieri, 2004), destacam ainda mais a necessidade de testes com qualidade no país. Esse estudo indicou que uma parte significativa dos instrumentos psicológicos existentes no Brasil carece de fundamentação teórica e não atende aos requisitos mínimos definidos nas perspectivas nacionais e internacionais. Essas constatações destacam a importância do presente estudo que objetivou demonstrar a validade convergente da versão reduzida do SON-R 2½-7, o SON-R 2½-7[a] (a letra "a" é da palavra inglesa *abbreviated*).

O SON-R 2½-7 é a última versão do teste não-verbal de inteligência Snijders-Oomen para a faixa etária de crianças de 2½ a 7 anos e foi normatizado e validado na Holanda em 1998 com base de uma amostra de 1.124 crianças (Tellegen, Winkel, Wijnberg-Williams & Laros, 1998). Os testes SON (Snijders-Oomen Não-verbal) devem seu nome à criadora do primeiro teste, Doutora N. Snijders-

¹ Contato:

Email: jalaros@gmail.com

Apoio CNPq (bolsa de produtividade) e Fundo SON do University of Groningen, The Netherlands

Oomen. Fazendo uso de tarefas novas e tarefas recém desenvolvidas na época, essa autora desenvolveu uma bateria de testes que incluiu diversas tarefas não-verbais relacionadas à habilidade espacial e ao raciocínio abstrato e concreto. A inteligência foi operacionalizada por essa autora em termos da habilidade de aprendizagem e das chances de sucesso na escola. O objetivo da primeira bateria SON era quebrar o viés dos testes de execução não-verbal empregados naquela época e permitir que funções como abstração, simbolismo e entendimento de situações comportamentais ficassem mais acessíveis para a avaliação não-verbal. Nenhum conceito de inteligência específico foi presumido como base para a primeira bateria SON. Porém, “forma”, “coerência concreta”, “abstração” e “memória de curto-prazo” foram vistos como sendo representações aceitáveis de funcionamento intelectual típico para sujeitos que estavam sofrendo de surdez precoce (Tellegen e colaboradores, 1998). Um requisito para a primeira bateria de testes SON foi que as vivências familiares e a educação formal deveriam influenciar os resultados do teste o mínimo possível. Além disso, as tarefas precisavam interessar à criança para evitar que ela ficasse entediada ou sem disposição para continuar.

O SON-R 2½-7 é uma bateria que consiste em seis subtestes, a saber: Categorias, Analogias, Situações, Histórias, Mosaicos e Padrões. A bateria avalia um largo espectro de habilidades cognitivas que não exigem o uso da fala e da linguagem escrita. O teste é de aplicação individual e as instruções podem ser dadas tanto de maneira verbal quanto não-verbal, dependendo das possibilidades de comunicação da criança. Recentemente foram realizadas normatizações do teste na Alemanha (Tellegen, Laros & Petermann, 2007), na França (Tellegen, Laros & Kiat, 2009) e na República Tcheca (Tellegen, Laros & Heider, 2008). Além da Holanda, foram realizados estudos de validação do SON-R 2½-7 em Austrália, Estados Unidos, Reino Unido e Alemanha. No Brasil, a versão reduzida do teste, composta pelos subtestes Categorias, Situações, Mosaicos e Padrões, foi normatizada em 2008 (Laros, Tellegen, Jesus, & Karino, no prelo).

Além do SON-R 2½-7, existe o SON-R 5½-17 para sujeitos com idades entre 5½ anos e 17 anos (Laros & Tellegen, 1991). Esta bateria consiste em sete subtestes: Categorias, Mosaicos, Figuras Ocultas, Padrões, Situações, Analogias e Histórias. A bateria pode ser dividida em quatro tipos de subtestes: testes de raciocínio abstrato (Categorias e Analogias), testes de raciocínio concreto (Situações e Histórias),

testes espaciais (Mosaicos e Padrões) e um teste perceptual (Figuras Ocultas). A amostra de normatização na Holanda foi composta de 1.350 sujeitos ouvintes e 768 sujeitos surdos.

Uma similaridade entre o SON-R 2½-7 e outros testes de inteligência para crianças novas, como por exemplo, o BAS (Elliott, Murray & Pearson, 1979-82), o K-ABC (Kaufman & Kaufman, 1983), e o WPPSI-R (Wechsler, 1989) é que a inteligência é avaliada tendo como base o desempenho em tarefas bastante diversificadas. Entretanto, no SON-R 2½-7, itens verbais não são incluídos porque são geralmente dependentes do conhecimento e da experiência. Assim, o SON-R 2½-7 é mais focado na mensuração de “inteligência fluida” do que de “inteligência cristalizada” (Cattell, 1971) se comparado a esses outros testes. A “inteligência fluida” foi assim nomeada porque foi concebida como sendo capaz de permear muitos tipos de atividades mentais; a “inteligência cristalizada” foi assim chamada porque foi pensada como um tipo de produto final de experiências até um determinado ponto na vida de um indivíduo (Carroll, 1993). Uma característica interessante da Teoria Gf-Gc (Gf é o fator geral da inteligência fluida e Gc é o fator geral da inteligência cristalizada), desenvolvida por Horn (1994) é que as habilidades da inteligência fluida tendem a declinar com o avanço da idade, pelo menos em alguns indivíduos, enquanto as habilidades cristalizadas, como vocabulário, não (Carroll, 1993). McGrew e Flanagan (1998) elaboraram a teoria CHC (Cattell-Horn-Carroll) sobre a inteligência que sintetizava os modelos de Cattell, Horn e Carroll, enfatizando a presença de uma estrutura fatorial hierárquica de três níveis. O nível mais alto da estrutura fatorial corresponde ao fator “g” de Spearman indicando a existência de operações cognitivas comuns a todas as atividades mentais (Primi, 2002).

A independência dos testes SON das habilidades verbais das pessoas testadas os torna muito adequados para o Brasil, onde uma porcentagem substancial da população detém baixo nível de escolarização (IBGE, 2007). Para pessoas analfabetas ou semi-analfabetas, um desempenho fraco em um teste de inteligência que depende de habilidades verbais, pode evidenciar um conhecimento verbal deficiente antes que uma habilidade limitada de raciocínio ou de potencial de aprendizagem (Tellegen & Laros, 2005).

Os testes não-verbais de inteligência SON constituem opções vantajosas para pessoas com uma habilidade verbal subdesenvolvida, e particularmente

para pessoas mudas ou com problemas auditivos, pessoas com problemas de aprendizagem bem como para pessoas com problemas de expressão e linguagem. A este respeito, uma especialista na área de avaliação neuropsicológica, M. D. Lezak, autora da seminal obra intitulada “*Neuropsychological Assessment*”, comenta que os testes SON-R constituem ferramentas muito úteis na avaliação neuropsicológica (Lezak, Howieson & Loring, 2004).

Ambos os testes SON-R, o SON-R 2½-7 e o SON-R 5½-17 foram avaliados pela COTAN, Comissão Nacional de Avaliação de Testes dos Países Baixos, responsável pela avaliação de testes psicológicos. Todos os aspectos do teste (base teórica para a construção do teste, execução do manual e dos materiais do teste, normas, validade e fidedignidade) receberam a avaliação mais favorável possível (Evers, Van Vliet, Mulder & De Groot, 2000). Isto significa que na Holanda os testes SON-R são considerados testes de grande credibilidade.

Uma vantagem fundamental das baterias de inteligência SON é que os subtestes que as compõem não requerem tradução. Isto os torna ferramentas adequadas para estudos transculturais e internacionais. Cabe ressaltar que os profissionais brasileiros não dispõem de outros instrumentos psicológicos que reúnam todas as características apresentadas pelo SON-R 2½-7[a] na avaliação do funcionamento cognitivo de crianças na faixa etária de 2½ a 7 anos. Ainda não foram publicados artigos sobre a validade convergente do SON-R 2½-7[a] no contexto brasileiro. Assim, o objetivo geral do estudo atual foi verificar a validade convergente do teste não-verbal de inteligência SON-R 2½-7[a] na realidade brasileira. As evidências de validade do teste são fundamentais para a avaliação adequada e útil das habilidades cognitivas das crianças. Para conseguir evidências da validade convergente do SON-R 2½-7[a] foram aplicados dois outros testes de inteligência, a saber: Escala de Maturidade Mental Colúmbia (EMMC) e Matrizes Progressivas Coloridas (MPC) de Raven – Escala Especial. Considerando que o MPC e o EMMC visam medir o mesmo construto geral do que o SON-R 2½-7[a], a saber inteligência, podemos formular a hipótese que o SON-R 2½-7[a] mostrará correlações positivas e fortes com o MPC e o EMMC. Uma correlação pode ser considerada forte quando seu valor é maior do que 0,50 (Cohen, 1988).

MÉTODO

Participantes

Nesse estudo participaram 120 crianças do Distrito Federal com idade variando entre 5;5 anos e 7;10 anos. Das 120 crianças, 68 eram meninos e 52 meninas. Todas as crianças foram testadas dentro da escola. As cinco escolas que participaram (duas escolas públicas e três particulares) foram selecionadas aleatoriamente com base no critério de ter crianças entre cinco e sete anos de idade. Foram enviados termos de consentimento juntamente com uma carta de explicação sobre os procedimentos aos pais, e aqueles que concordavam com a participação de seu filho a devolvia assinada.

Instrumentos

O SON-R 2½-7[a]

A versão abreviada do SON-R 2½-7 consiste de quatro subtestes: Categorias (15 itens), Situações (14 itens), Mosaicos (15 itens) e Padrões (16 itens). Exemplos dos itens dos quatro subtestes são oferecidos no *Manual and Research Report* do SON-R 2½-7 que está disponível no *web site* do testes SON (www.testresearch.nl). No presente estudo foram utilizadas as normas brasileiras do teste, que são baseadas em uma amostra nacional representativa de 1.200 crianças com idades que variam entre 3;3 anos a 7;9 anos. Cada grupo de idade foi representado por uma amostra de 120 crianças que foi estratificada segundo o sexo, situação educacional e variáveis demográficas. O processo de transformação dos escores brutos em escores distribuídos normalmente com média e desvio fixos (100,15) não foi efetuado separadamente para cada grupo de idade, no entanto, as distribuições dos escores para todos os grupos de idade foram ajustadas simultaneamente como uma função contínua da idade. Esse procedimento de ajuste foi usado a fim de minimizar as diferenças entre a distribuição observada e a distribuição populacional estimada. Na amostra de normatização brasileira o índice médio de consistência interna (Lambda 2 de Guttman) dos 4 subtestes do SON-R 2½-7[a] foi 0,79. O subteste Categorias mostrou a fidedignidade média mais alta (0,81) e Situações a mais baixa (0,77). Os subtestes do SON-R 2½-7[a] podem ser agrupados em dois tipos de teste: de raciocínio (Categorias e Situações) e de execução (Mosaicos e Padrões) com enfoque espacial. Assim, três escores são calculados para a bateria: para a escala total (SON-R ET), para a escala de execução (SON-R EE) e para a escala de raciocínio (SON-R ER). A fidedignidade média da escala total foi 0,92 e das escalas de execução e raciocínio 0,87.

Matrizes Progressivas Coloridas (MPC) do Raven

O teste Matrizes Progressivas Coloridas (MPC) foi normatizado e validado para o Brasil (Angelini, Alves, Custódio, Duarte & Duarte, 1999). A normatização brasileira do MPC foi realizada no município de São Paulo com base em uma amostra de 1.547 alunos (774 meninas e 773 meninos). Participaram 345 alunos (22,3%) de escolas particulares e 1.202 alunos (77,7%) de escolas públicas (municipais e estaduais). O critério empregado no sorteio das escolas permitiu a representatividade de todas as regiões da cidade de São Paulo (Angelini e colaboradores, 1999). O MPC do Raven foi desenvolvido a partir da teoria de Spearman sobre inteligência e tem como propósito fazer uma medida direta de “g” (inteligência geral). O fator “g” frequentemente funciona como um parâmetro para estudos de análise fatorial sobre a inteligência (Hogan, 2006). O teste pretende mensurar a dimensão fluida da inteligência geral ao invés da dimensão cristalizada (Carroll, 2003; Primi, 2002; Woodcock, 1998). O MPC é voltado para crianças de 5 a 11 anos e para pessoas adultas com retardo mental e outras situadas na extremidade inferior da distribuição de inteligência. O teste é dividido em três séries: A, Ab e B, sendo cada uma composta por 12 itens, impressos em cores vivas para aumentar o nível de interesse das crianças.

A Escala de Maturidade Mental Colúmbia (EMMC)

A *Escala de Maturidade Mental Colúmbia - EMMC* (Alves & Duarte, 2001) é um teste apropriado para crianças entre 3;6 anos e 9;11 anos. Seu nome refere-se à *University of Columbia* onde o teste foi desenvolvido. A versão brasileira do EMMC é constituída por 3 exemplos e 92 itens e é composto de 8 escalas ou níveis de idade (A, B, C, D, E, F, H e I). Com a divisão do teste completo em 8 níveis diferentes, apenas os itens mais adequados à faixa etária da criança são aplicados. O manual da adaptação brasileira do Colúmbia foi publicado quase 10 anos depois da coleta de dados: o estudo de normatização foi realizado no final do 2º semestre de 1992 e no início do 1º semestre de 1993 e o manual foi publicado em 2001.

A amostra de normatização no Brasil foi constituída de 1.535 crianças (777 meninas e 758 meninos). Participaram 539 alunos (35,1%) de escolas particulares e 996 alunos (64,9%) de escolas públicas (municipais e estaduais). Na faixa de 3;6 anos a 4;6 anos, as crianças foram testadas em creches de escolas privadas em função da dificuldade de se encontrar crianças nessa faixa nas pré-escolas

públicas. O estudo de normatização foi realizado em 50 escolas do município de São Paulo: 13 escolas privadas, 14 escolas estaduais e 23 escolas municipais.

Alves e Duarte (2001) publicaram coeficientes de fidedignidade das duas metades do Colúmbia baseados na pesquisa de normatização no Brasil variando entre 0,82 e 0,93, dependendo da idade da criança e um coeficiente mediano de 0,87. O coeficiente de fidedignidade obtido pelo método teste-reteste foi de 0,80 para a amostra total; de 0,90 para crianças de 6 anos e de 0,73 para crianças de 8 anos. Os dados psicométricos fornecidos pelos autores sugerem que o Colúmbia tem uma qualidade psicométrica melhor para crianças mais jovens em comparação com crianças mais velhas.

Procedimentos

Para a aplicação dos três testes foi realizado um treinamento prévio com os aplicadores, dez estudantes de psicologia da Universidade de Brasília (UnB). Os dados foram coletados entre novembro de 2007 e junho de 2008. As crianças eram chamadas dentro do horário de aula, e não durante o intervalo, o que poderia acarretar em uma falta de cooperação e motivação por parte da criança. Em um primeiro dia era aplicado o SON, de duração de aproximadamente 40 minutos, e, em um segundo, a mesma criança era chamada novamente para a aplicação dos dois outros instrumentos, que tinham duração de aproximadamente 15 minutos cada. Na metade dos participantes aplicou-se em primeiro lugar o MPC e na outra metade o EMMC foi aplicado em primeiro lugar. Após a coleta de dados, os pais receberam uma carta devolutiva com indicações do resultado de seu filho.

Análise de dados

Antes de realizar as análises principais foram realizadas análises exploratórias para verificar a ocorrência de erros de digitação e outros tipos de erros. Depois desta fase de controle, foram calculadas estatísticas descritivas para caracterizar os testes aplicados. No cálculo do índice de fidedignidade foi utilizado o coeficiente Lambda 2 de Guttman, uma vez que foi comprovada que Lambda 2 estima melhor a fidedignidade do que o coeficiente Alfa de Cronbach (Ten Berge & Zegers, 1978; Tellegen & Laros, 2004), principalmente quando os instrumentos contêm poucos itens ou quando a amostra é pequena. O cálculo das correlações entre os testes foi baseado em escores normatizados. Além das correlações brutas foram calculadas correlações

depois da aplicação da correção para atenuação, também chamada de correção para a não-fidedignidade (Hogan, 2006), usando a fórmula $r^*_{XY} = r_{XY} \div \sqrt{r_{YY}}$ (Miles & Shevlin, 2001; Cronbach, 1996; Nunnally & Bernstein, 1994; Osborne, 2003), onde r^*_{XY} é a correlação corrigida e r_{YY} é a fidedignidade do teste critério. Osborne (2003) afirma que neste contexto a mensuração não fidedigna causa a subestimação das relações reais entre as variáveis. O autor argumenta que pesquisadores deveriam corrigir para atenuação para obter uma estimativa melhor da relação verdadeira entre as variáveis na população. Para a investigação da validade convergente do SON-R 2½-7[a] foi utilizada, além da análise correlacional, também a técnica de análise fatorial exploratória. Uma análise fatorial exploratória PAF (*Principal Axis Factoring*)

com rotação Promax foi efetuada para verificar se o SON-R 2½-7[a] está medindo fatores semelhantes do que os testes MPC e EMMC. Para fundamentar a decisão quantos fatores precisavam ser extraídos o critério da Análise Paralela foi utilizado (Laros, 2005; Ledesma & Valero-Mora, 2007).

RESULTADOS E DISCUSSÃO

Como primeiro passo foram calculadas estatísticas descritivas para os três testes aplicados e para as escalas de execução (SON-R EE) e de raciocínio (SON-R ER) do SON-R 2½-7[a]. A Tabela 1 permite identificar os valores da média, desvio-padrão, erro-padrão da média, o intervalo de confiança de 95% das médias e o coeficiente de fidedignidade.

Tabela 1. Média, desvio-padrão, erro padrão da média e intervalo de confiança de 95% das médias obtidas e o coeficiente de fidedignidade (Lambda 2 de Guttman) para as escalas do SON-R 2½-7[a], o MPC e o EMMC [N = 120].

	Média	DP	EP	IC 95%	Fidedignidade
SON-R ET	113,10	12,08	1,10	109,9 – 114,3	0,91
SON-R EE	109,02	13,58	1,18	106,7 – 111,4	0,86
SON-R ER	114,71	11,78	1,19	112,3 – 117,1	0,84
MPC	114,03	12,03	1,10	111,8 – 116,2	0,79
EMMC	113,44	15,85	1,45	110,5 – 116,3	0,84

Notas: DP = desvio-padrão; EP = erro padrão da média; IC95% = intervalo de confiança de 95%; SON-R ET = Escala Total do SON-R; SON-R EE = Escala de Execução do SON-R; SON-R ER = Escala de Raciocínio do SON-R; MPC = Matrizes Progressivas Coloridas do Raven; EMMC = Escala de Maturidade Mental Colúmbia.

Verificando as médias dos testes na Tabela 1 torna-se evidente que os escores totais dos três testes (os escores de QI) mostram médias muito semelhantes. A inspeção dos intervalos de confiança de 95% aponta que não existem diferenças estatisticamente significativas em 5% entre as médias dos escores totais dos testes. Isso significa que podem ser utilizados os mesmos critérios de interpretação para os três testes.

Merece destaque na Tabela 1 o desvio-padrão relativamente baixo para o MPC e as escalas do SON-R 2½-7[a], sendo 12,03 para o MPC e entre 12,52 e 13,01 para as escalas da bateria SON-R. Uma vez que os escores totais do MPC e da bateria SON-R foram expressos numa escala de M = 100 e DP = 15 seria esperado se encontrar um desvio-padrão ao redor de 15. Aparentemente, na amostra pesquisada nem todos os níveis de habilidade cognitiva foram

representados. O valor mais alto do desvio-padrão do EMMC em comparação com os outros testes aplicados pode, pelo menos parcialmente, ser explicado levando em consideração que o escore total no EMMC foi padronizado utilizando uma escala de M = 100 e DP = 16 ao invés da escala usual de M = 100, DP = 15. Considerando os coeficientes de fidedignidade (Lambda 2 de Guttman) da Tabela 1 torna-se evidente que todos os testes mostram um coeficiente de fidedignidade satisfatório, variando entre 0,79 e 0,91 sendo que o valor para a Escala Total do SON-R 2½-7[a] é o valor mais elevado.

A validade convergente do SON-R 2½-7[a] foi verificada neste estudo através da correlação com o MPC e o EMMC. A Tabela 2 mostra as correlações brutas e as correlações depois da aplicação da correção para atenuação.

Tabela 2. Correlação, antes e depois da aplicação da correção para atenuação, das escalas do SON-R 2½-7[a] com o MPC e o EMMC e o intervalo de confiança de 95% da correlação corrigida das escalas (N = 120).

	SON-R ET	IC – 95%	SON-R EE	IC – 95%	SON-R ER	IC – 95%
MPC	0,77 (0,68)	0,69 – 0,84	0,68 (0,61)	0,58 – 0,77	0,63 (0,56)	0,50 – 0,73
EMMC	0,62 (0,56)	0,51 – 0,72	0,47 (0,43)	0,32 – 0,60	0,59 (0,54)	0,47 – 0,70

Notas: SON-R ET = Escala Total do SON-R; SON-R EE = Escala de Execução do SON-R; SON-R ER = Escala de Raciocínio do SON-R; as correlações são corrigidas para atenuação: a correlação bruta é mostrada entre parêntesis. Os intervalos de confiança de 95% (IC 95%) foram calculados para as correlações corrigidas.

Os resultados da Tabela 2 apontam que a correlação bruta da escala total do SON-R 2½-7[a] com o MPC foi de 0,68. Depois da aplicação da correção para atenuação, a correlação aumentou para 0,77. Para obter a correlação corrigida, a correlação bruta precisa ser dividida pela raiz quadrada da fidedignidade do teste critério ($r_{XY}^* = r_{XY} \div \sqrt{r_{YY}} = 0,68 \div \sqrt{0,79} = 0,68 \div 0,89 = 0,77$). O valor da correlação corrigida de 0,77 é bem próximo ao valor de 0,74, encontrado para a correlação entre o SON-R 5½-17 e o Raven em um estudo feito na Holanda com 35 crianças de um hospital universitário (Tellegen, 1993). A correlação bruta do SON-R 2½-7[a] com o EMMC foi 0,56 e a correlação corrigida foi 0,62. A inspeção dos intervalos de confiança de 95% da correlação corrigida da bateria SON-R com o MPC e EMMC justifica a conclusão que a bateria SON-R 2½-7[a] mostra correlações positivas e fortes com o MPC e o EMMC uma vez que o limite inferior dos intervalos é maior do que 0,50. Este conclusão

corroborar a hipótese formulada no início deste estudo.

É importante informar que a correlação corrigida encontrada neste estudo de 0,59 entre o MPC e o EMMC é praticamente igual à correlação de 0,58 para crianças de 8 anos entre o MPC e o EMMC relatada pelos autores do manual brasileiro do EMMC (Alves & Duarte, 2001). Merece destaque os resultados apresentados na Tabela 2 que sugerem uma relação mais forte entre o MPC e a escala de execução da bateria SON-R, e entre o teste EMMC e a escala de raciocínio da bateria SON-R.

Na Tabela 3 as correlações brutas das escalas do SON-R 2½-7[a] com o MPC e o EMMC são mostradas para três grupos de idade: de 5;6 a 6;5 anos, de 6;6 a 6;11 anos e 7;0 a 7;11 anos. Na Tabela 3 as correlações brutas das escalas do SON-R 2½-7[a] com o MPC e o EMMC são mostradas para três grupos de idade: de 5;6 a 6;5 anos, de 6;6 a 6;11 anos e 7;0 a 7;11 anos.

Tabela 3. Correlações brutas das escalas do SON-R 2½-7[a] com o MPC (Raven) e EMMC (Colúmbia) para três grupos de idade.

Idade	Escala	MPC	EMMC
5;6 – 6;5 (N = 45)	SON-R ET	0,57	0,43
	SON-R EE	0,48	0,31
	SON-R ER	0,49	0,44
6;6 – 6;11 (N = 45)	SON-R ET	0,72	0,61
	SON-R EE	0,65	0,51
	SON-R ER	0,57	0,55
7;0 – 7;11 (N = 30)	SON-R ET	0,83	0,63
	SON-R EE	0,77	0,45
	SON-R ER	0,63	0,65

Notas: SON-R ET = Escala Total do SON-R; SON-R EE = Escala de Execução do SON-R; SON-R ER = Escala de Raciocínio do SON-R; MPC = Matrizes Progressivas Coloridas do Raven; EMMC = Escala de Maturidade Mental Colúmbia

Os resultados da Tabela 3 indicam que a correlação da escala completa do SON-R 2½-7[a] com o MPC aumenta de 0,57 na faixa etária de 5;6 a 6;5 anos para 0,83 na faixa etária de 7;0 a 7;11 anos. A correlação da escala completa do SON-R 2½-7[a] com o EMMC aumenta de 0,43 na faixa etária de 5;6 a 6;5 anos para 0,63 na faixa etária de 7;0 a 7;11

anos. Aparentemente, existe uma tendência de acréscimo da correlação do SON-R 2½-7[a] com os outros dois testes de inteligência com o aumento da idade das crianças. A Tabela 4 mostra os resultados de uma análise fatorial. Utilizou-se para a extração dos fatores o método PAF (*Principal Axis Factoring*) e para a rotação dos fatores o método Promax.

Tabela 4. Cargas fatoriais das escalas SON-R EE, SON-R ER, EMMC e MPC no fator geral e depois da extração de 2 fatores com *Principal Axis Factoring* e rotação PROMAX (N=120).

	Fator "g"	Fator 1	Fator 2
CPM	0,82	0,70	0,17
SON-R EE	0,68	0,74	-0,10
SON-R ER	0,71	0,08	0,68
EMMC	0,68	0,03	0,69

Notas: 52,3 % de variância foi explicada com a solução de um fator: com a solução de 2 fatores 57,7% da variância foi explicada. A correlação entre fator 1 e fator 2 foi de 0,80.

Os resultados da análise fatorial sugerem que, depois da rotação Promax, o primeiro fator pode ser interpretado como um fator espacial, que está medindo a capacidade de mentalmente fazer rotações de figuras geométricas (Thurstone, 1938). Este fator também já foi chamado de raciocínio visual (Neisser & colaboradores, 1996) e inteligência visual (McGrew, 1997; McGrew & Flanagan, 1998). O segundo fator da Tabela 4 é mais voltado para o raciocínio abstrato e está medindo a capacidade de encontrar uma regra ou princípio para resolver um problema, como em classificações de figuras ou analogias entre padrões. Thurstone chamou este segundo fator um fator de raciocínio indutivo (Hogan, 2006). Na classificação do McGrew (1997) onde são diferenciados 10 fatores de habilidades amplas e que tem seus fundamentos no modelo adaptado Gf-Gc, o fator é identificado como um fator de inteligência fluida. Considerando que se encontrou uma correlação de 0,80 entre os dois fatores da Tabela 4, foi rodada também uma solução fatorial com um único fator. Este fator pode ser interpretado como o fator "g". Neste fator geral, o teste MPC tem a carga mais alta. Isto está de acordo com a afirmação do Neisser (1998) que os testes de Raven são os que medem melhor o fator "g". A escala de raciocínio do SON-R 2½-7[a] apresenta uma carga mais alta no fator geral que a escala de execução. Entretanto, a diferença nas cargas fatoriais não é estatisticamente significativa no nível de 5%.

CONSIDERAÇÕES FINAIS

Os resultados obtidos no presente estudo mostram que existem evidências positivas da validade convergente do SON-R 2½-7[a] no Brasil na faixa etária pesquisada entre 5 e 7 anos de idade. Foram encontradas correlações altas com dois outros testes de inteligência, o MPC e o EMMC. A fidedignidade do SON-R 2½-7[a] se mostrou muito satisfatória neste estudo: o coeficiente de fidedignidade da escala total foi de 0,91, e de 0,86 e de 0,84 para as escalas de execução e de raciocínio respectivamente. Outro resultado positivo foi a indicação que no Brasil os mesmos critérios de interpretação do escore total podem ser utilizados para o SON-R 2½-7[a] e para as MPC e o EMMC, uma vez que não foram encontradas diferenças estatisticamente significativas entre as médias dos respectivos testes. Este resultado não era esperado devido ao efeito Flynn (Flynn, 2000; Neisser, 1998). Flynn demonstrou em uma série de estudos científicos (Wicherts & colaboradores, 2004) que os escores brutos em testes de inteligência aumentam com o tempo para cada faixa etária, não somente nos Estados Unidos, mas em muitos outros países. O aumento é de 3 pontos na escala de QI (100,15) em 10 anos para baterias amplas de inteligência, como o WISC-III e o WAIS. Para testes com uma carga alta no fator "g", como os testes de Raven, o aumento pode chegar a 7 pontos em 10 anos (Neisser, 1998).

Com uma nova normatização de um teste de inteligência, o escore médio é colocado novamente em 100, de modo que o aumento desaparece de vista. Neste contexto, o efeito Flynn significa que, se passados 10 anos desde a última vez em que o teste foi normatizado, as pessoas que agora obtêm um escore de 100 na nova versão, provavelmente obteriam um escore cerca de 103 na versão antiga (Neisser, 1998). Uma vez que a normatização do EMMC aconteceu há 16 anos e a das MPC há 10 anos, seriam esperados escores normatizados de 5 a 7 pontos mais altos.

A razão dos testes MPC e EMMC terem médias iguais ao SON-R 2½-7[a], apesar do efeito Flynn foi possivelmente o fato que as normas de ambos os testes são fundamentadas em amostras nas quais crianças do nível socioeconômico mais baixo não foram incluídas. Neste contexto, os autores do EMMC (Alves & Duarte, 2001) afirmam que na faixa de 3;6 a 4;6 anos, as crianças foram testadas em creches de escolas privadas em função da dificuldade de se encontrar crianças nessa faixa etária nas pré-escolas públicas. Além disso, nas amostras de normatização dos dois testes não foram incluídas crianças que estavam fora da escola. É bem razoável presumir que crianças que não frequentam uma escola, em geral, têm um nível socioeconômico mais baixo. Por causa destes razões os escores normatizados no EMMC e MPC são mais baixos do que deveriam ser. O efeito Flynn pode ter neutralizado a falta de representatividade de crianças do nível socioeconômico mais baixo nas amostras de normatização.

Outro resultado do presente estudo que merece destaque é que existem indicações fundamentadas na análise fatorial que o EMMC está medindo um construto de inteligência semelhante ao construto de raciocínio mensurado pela escala SON-R ER e que o MPC avalia um construto parecido com o construto que a escala SON-R EE mensura. O primeiro construto é também identificado como inteligência visual e o segundo como inteligência fluída.

Em síntese, o estudo atual oferece evidências sólidas sobre a validade convergente do SON-R 2½-7[a] no Brasil para a faixa de idade investigada. Assim, esperamos ter contribuído para o campo de conhecimento da avaliação psicológica em Brasil através de um estudo sobre um novo instrumento de inteligência com normas brasileiras que apresenta evidências positivas de validade e fidedignidade.

REFERÊNCIAS

- Alves, I. C. B. & Duarte, J. L. M. (2001). *Escala de maturidade mental Colúmbia: Manual para aplicação e interpretação*. 1ª Edição Brasileira. São Paulo: Casa do Psicólogo Livraria e Editora Ltda.
- Alves, I. C. B. (2002). Instrumentos disponíveis no Brasil para avaliação da inteligência. Em R. Primi (Org.), *Temas em avaliação psicológica* (pp. 80-102). Campinas, SP: Impressão Digital do Brasil Gráfica e Editora Ltda.
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F. & Duarte, J. L. M. (1999). *Manual Matrizes Progressivas Coloridas de Raven: escala especial*. São Paulo: CETEPP - Centro Editor de Testes e Pesquisas em Psicologia.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second Edition. New Jersey: Lawrence Erlbaum.
- Cronbach, L. J. (1996). *Fundamentos da testagem psicológica*. 5ª Edição. Porto Alegre: Artes Médicas.
- Elliott, C. D., Murray, D. J. & Pearson, L. S. (1979-82). *British Ability Scales: Manuals*. Windsor: National Foundation for Educational Research.
- Evers, A., Vliet-Mulder, J. C. & De Groot, J. (2000). *Documentation of tests and test research in The Netherlands*. Assen: Van Gorcum – NIP.
- Flynn, J. R. (2000). IQ trends over time: intelligence, race, and meritocracy. Em K. Arrow, S. Bowles & S. Durlauf. M. (Orgs.), *Meritocracy and economic inequality*. Princeton NJ: Princeton University Press.
- IBGE (2007). *Síntese de indicadores sociais: uma análise das condições de vida da população brasileira*. Estudos & Pesquisas: Informação demográfica e socioeconômica (21). Rio de Janeiro: Centro de Documentação e Disseminação de Informações.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro: LTC – Livros e Técnicos e Científicos Editores S.A.
- Horn, J. L. (1994). Theory of fluid and crystallized intelligence. Em R. J. Sternberg (Org.), *Encyclopedia of human intelligence* (pp. 443-451). New York: Macmillan.
- Jacquemin, A. & Xavier, M. A. (1984). O teste das Matrizes Progressivas de Raven – Escala Especial. Estudo de sua fidedignidade. *Ciência e Cultura*, 36, 77-79.

- Jesus, G. R., de (2009). *Normatização e validação do teste não-verbal de inteligência SON-R 2½-7[a] para o Brasil*. Tese de Doutorado, Universidade de Brasília, Brasília.
- Kaufman, A. S. & Kaufman, N. L. (1983). *K-ABC Kaufman Assessment Battery for Children. Interpretive Manual*. Circle Pines, MN: American Guidance Service.
- Laros, J. A. (2005). O uso de análise fatorial: algumas diretrizes para pesquisadores. Em L. Pasquali (Org.), *Análise fatorial para pesquisadores* (pp. 163-184). Brasília: LabPAM.
- Laros, J. A. & Tellegen, P. J. (1991). *Construction and validation of the SON-R 5½-17, the Snijders-Oomen non-verbal intelligence test*. Groningen, The Netherlands: Wolters-Noordhoff.
- Laros, J. A., Tellegen, P. J., Jesus, G. R., de & Karino, C. A. (no prelo). *SON-R 2½-7[a], Teste não-verbal de inteligência. Manual com normatização e validação brasileira*.
- Ledesma, R. D. & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: an easy-to use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation*, 12,1-10.
- Lezak, M. D., Howieson, D. B. & Loring, D. W. (2004). *Neuropsychological Assessment*. Fourth Edition. New York, NY: Oxford University Press.
- McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. Em D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Orgs.). *Contemporary intellectual assessment: theories, tests, and issues*. New York: The Guilford Press.
- McGrew, K. S. & Flanagan, D. P. (1998). *The intelligent test desk reference (ITDR): Gf-Gc cross-battery assessment*. Needham Heights: Allyn & Bacon.
- Miles, R. & Shevlin, M. (2001). *Applying regression & correlation: A guide for students and researchers*. London: Sage Publications.
- Muñiz, J., Prieto, G., Almeida, L. & Bartram, D. (1999). Test use in Spain, Portugal and Latin American Countries. *European Journal of Psychological Assessment*, 15, 151-157.
- Neisser, U. (1998). *The rising curve: long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.
- Neisser, U., Boodoo, G., Bouchard Jr., T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J. & Urbina, S. (1996). Intelligence: Knowns and Unknowns. *American Psychologist*, 51, 77-101.
- Noronha, A. P. P. & Alchieri, J. C. (2002). Reflexões sobre os instrumentos de avaliação psicológica. Em R. Primi (Org.), *Temas em avaliação psicológica* (pp. 7-16). Campinas, SP: Imprensa Digital do Brasil Gráfica e Editora Ltda.
- Noronha, A. P. P., Vendramini, C. M. M., Canguçu, C., Souza, C. V. R. de, Cobêro, C., Paula, L. M. de, Franco M. O. de, Lima, O. M. P. de, Guerra, P. B. C. de & Filizatti, R. (2003). Propriedades psicométricas em manuais de testes de inteligência. *Psicologia em Estudo*, 8, 93-99.
- Noronha, A. P. P., Primi, R. & Alchieri, J. C. (2004). Parâmetros psicométricos: uma análise dos testes psicológicos comercializados no Brasil. *Psicologia: Ciência e Profissão*, 24, 88-99.
- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory*. Third Edition. New York: McGraw-Hill Series in Psychology.
- Primi, R. (2002). Inteligência fluida: definição fatorial, cognitiva e neuropsicológica. *Padeia*, 12, 57-77.
- Oakland, T., Wechsler, S., Bensuan, E. & Stafford, M. (1994). The construct of intelligence among Brazilian children: An exploratory study. *School Psychology International*, 15, 361-370.
- Osborne, J. W. (2003). Effect sizes and the disattenuation of correlation and regression coefficients: lessons from educational psychology. *Practical Assessment, Research & Evaluation*, 8 (11), 1-5. Retirado em 24/07/2008 do www.PAREonline.net.
- Tellegen, P. J. (1993). *A nonverbal alternative to the Wechsler scales: The Snijders-Oomen Nonverbal Intelligence Tests*. Trabalho apresentado na Primeira Reunião Anual da South Padre Island Conference on cognitive assessment of children and youth in school and clinical settings.
- Tellegen, P.J., Winkel, M., Wijnberg-Williams, B. J. & Laros, J. A. (1998). *Snijders-Oomen Nonverbal Intelligence Test, SON-R 2½-7, Manual & Research Report*. Lisse: Swets & Zeitlinger.
- Tellegen, P. J. & Laros, J. A. (2004). Cultural bias in the SON-R test: Comparative study of Brazilian and Dutch children. *Psicologia: Teoria e Pesquisa*, 20 (2), 103-111.
- Tellegen, P. J. & Laros, J. A. (2005). Fair assessment of children from cultural minorities: a description of the SON-R non-verbal intelligence tests. Em D. Kopcanová (Org.), *Quality Education for*

- children from socially disadvantaged settings (pp. 50-73). Bratislava: Commission for UNESCO.
- Tellegen, P. J., Laros, J. A. & Petermann, F. (2007). *SON-R 2½-7, Non-verbaler intelligenz-test: Testmanual mit deutscher normierung und validierung*, Göttingen, Bern, Wien, Paris, Oxford: Hogrefe.
- Tellegen, P. J., Laros, J. A. & Heider, D. (2008). *SON-R 2½-7, Neverbalni inteligencni test: Prirucka*. [SON-R 2½-7, non-verbal test of intelligence: Manual]. Praha: Testcentrum-Hogrefe
- Tellegen, P. J., Laros, J. A. & Kiat, M. (2009). *SON-R 2½-7. Test Non-Verbal d'Intelligence. Adaption Française. Manuel*. Paris: Editions Hogrefe France.
- Ten Berge, J. M. F. & Zegers, F. E. (1978). A serie of lower bounds to the reliability. *Psychometrika*, 43 (4), 575-579.
- Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Wechsler, D. (1989). *WPPSI-R, Wechsler Preschool and Primary Scale of Intelligence – Revised. Manual*. San Antonio, TX: The Psychological Corporation.
- Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., Baal, G. C. M. van, Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, 32, 509–537.
- Woodcock, R. W. (1998). Extending Gf-Gc theory into practice. Em J. J. McArdle & R. W. Woodcock (Orgs.), *Human cognitive abilities in theory and practice* (pp. 137-156). Mahwah: Erlbaum.

Recebido em julho de 2009
1ª revisão em outubro de 2009
2ª revisão em novembro de 2009
Aprovado em dezembro de 2009

SOBRE OS AUTORES:

Jacob Arie Laros: Doutor PhD em *Personality and Educational Psychology* pela *University of Groningen* (1991), pós-doutorado pela *University of Groningen* (2001) e pela *Universitat de València* (2009). O professor Laros é autor de dois testes não-verbais de inteligência, o SON-R para crianças de 2,5 até 7 anos e o SON-R para a faixa etária de 5 até 17 anos. Atualmente é professor Adjunto IV do curso da pós-graduação em Psicologia Social, do Trabalho e das Organizações (PSTO) da Universidade de Brasília (UnB) e bolsista de produtividade em pesquisa do CNPq. Suas áreas de expertise tanto para ensino como para pesquisa são: avaliação educacional, avaliação de programas educacionais e sociais, elaboração de testes de habilidades cognitivas, análise fatorial e análise multinível. *E-mail*: jalaros@gmail.com

Rafaela Frade Reis: aluna de graduação do curso de Psicologia da Universidade de Brasília. Ex-bolsista de Projeto de Iniciação Científica do CNPq. Concluiu o curso de extensão Universitária em Psicologia Cultural e Reabilitação, com duração de 424 horas, promovido pela Rede Sarah de Hospitais de Reabilitação em parceria com a Universidade da Califórnia e com a Federação Mundial de Neuroreabilitação. Atualmente atua como estagiária de psicologia da saúde no Hospital de Apoio de Brasília. *E-mail*: rafaelafreis@gmail.com

Peter J. Tellegen: Doutor PhD em *Personality and Educational Psychology* (1991) pela *University of Groningen*. Após a conclusão dos seus estudos de Psicologia na *University of Utrecht* obtendo o título de Psicólogo, Peter Tellegen lecionou e realizou pesquisas na *University of Groningen*. O professor Tellegen é autor de dois testes não-verbais de inteligência, o SON-R para crianças de 2,5 até 7 anos e o SON-R para a faixa etária de 5 até 17 anos e é responsável pelo desenvolvimento futuro destes testes. Além disso, ele é autor de um teste de inteligência de uso educacional e de um teste computerizado. O professor Tellegen é responsável do site www.testresearch.nl que oferece informações e uma avaliação crítica de testes de inteligência e o uso destes testes. *E-mail*: p.j.tellegen@rug.nl