

Temas em Psicologia

ISSN: 1413-389X

comissaoeditorial@sbponline.org.br

Sociedade Brasileira de Psicologia

Brasil

Arie Laros, Jacob; Nascimento de Almeida, Gabriel Olimpio; Valentini, Felipe; Feitosa de Lima, Renata Manuelly

Dimensionalidade e Evidências de Validade Convergente do SON-R 6-40

Temas em Psicologia, vol. 23, núm. 4, diciembre, 2015, pp. 929-945

Sociedade Brasileira de Psicologia

Ribeirão Preto, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=513751493011>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc

# Dimensionalidade e Evidências de Validade Convergente do SON-R 6-40

**Jacob Arie Laros<sup>1</sup>**

*Departamento de Psicologia Social e do Trabalho da Universidade de Brasília,  
Brasília, DF, Brasil*

**Gabriel Olimpio Nascimento de Almeida**

*Faculdade IESGO, Formosa, GO, Brasil*

**Felipe Valentini**

*Programa de Pós-Graduação em Psicologia da Universidade Salgado  
de Oliveira – Universo, Niterói, RJ, Brasil*

**Renata Manuelly Feitosa de Lima**

*Programa de Pós-Graduação em Psicologia Social, do Trabalho e Organizações  
da Universidade de Brasília, Brasília, DF, Brasil*

## Resumo

No presente artigo são fornecidas evidências da estrutura fatorial e da validade convergente do SON-R 6-40 (*Snijders-Oomen Non-Verbal Intelligence Test – Revised*), um instrumento não-verbal de inteligência para pessoas entre 6 e 40 anos. O teste foi aplicado em conjunto com dois outros testes de inteligência: a forma A da Bateria de Provas de Raciocínio (BPR-5) e o SON-R 2½-7[a] (para crianças de 2,5 até 7 anos). Participaram no estudo 112 alunos de escolas públicas do Distrito Federal. Os resultados da análise fatorial evidenciaram a unidimensionalidade dos escores do SON-R 6-40. A correlação entre o SON-R 6-40 e a BPR-5 foi de 0,74 ( $N = 68$ ) e entre o SON-R 6-40 e o SON-R 2½-7[a] foi igual a 0,80 ( $N = 44$ ). Estes resultados indicam evidências da validade convergente do SON-R 6-40 para a faixa de idade investigada.

**Palavras-chave:** Validade convergente, teste não-verbal de inteligência, SON-R 6-40.

## Dimensionality and Convergent Validity Evidence of the SON-R 6-40

## Abstract

This article presents evidence of the factor structure and of the convergent validity of the SON-R 6-40 (*Snijders-Oomen Non-Verbal Intelligence Test – Revised*), a nonverbal test of intelligence for persons with ages between 6 and 40 years. The test was administered together with two other tests of intelligence: form A of the *Bateria de Provas de Raciocínio* (BPR-5) and the SON-R 2½-7[a] (for children 2.5 to 7 years). In this study 112 students of public schools in the Distrito Federal of Brazil participated.

<sup>1</sup> Endereço para correspondência: Universidade de Brasília, Instituto de Psicologia, Instituto Central de Ciências (ICC)/Ala Sul, sala A1-061/4, Laboratório de Métodos e Técnicas de Avaliação, Brasília, DF, Brasil 70910-900. Fone: (61) 3107-6902. E-mail: jalaros@gmail.com, gabriel.nascimento1986@gmail.com, valentini.felipe@gmail.com e renatamanuelly@gmail.com

Trabalho realizado com apoio do Fundo SON da *Rijksuniversiteit Groningen* (RuG), Holanda.

Este artigo é resultado da dissertação de mestrado do segundo autor, sob a orientação do primeiro autor.

Os autores agradecem o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Factor analysis indicated the unidimensionality of the SON-R 6-40 scores. The correlation between the SON-R 6-40 and the BPR-5 was .74 ( $N=68$ ) and between the SON-R 6-40 and the SON-R 2½-7[a] was .80 ( $N=44$ ). These results indicate satisfactory evidence of the convergent validity of the SON-R 6-40 for the investigated age range.

**Keywords:** Nonverbal test of intelligence, convergent validity, SON-R 6-40.

## Dimensionalidad y Evidencia de Validez Convergente del SON-R 6-40

### Resumen

El presente artículo muestra evidencias de la estructura factorial y validez convergente del test SON-R 6-40 (*Snijders-Oomen Non-Verbal Intelligence Test – Revised*), instrumento no verbal de evaluación de inteligencia para personas de entre 6 y 40 años. El test fue aplicado junto con dos pruebas de inteligencia: forma A de la *Batería de Provas de Raciocínio* (BPR-5) y SON-R 2½-7[a] (para niños de 2.5 a 7 años). Participaron 112 alumnos de escuelas públicas del Distrito Federal. Los resultados del análisis factorial muestran la unidimensionalidad de los puntajes del SON-R 6-40. La correlación encontrada entre el SON-R 6-40 y la BPR-5 fue de .74 ( $N=68$ ) y entre el SON-R 6-40 y el SON-R 2½-7[a] fue de .80 ( $N=44$ ). Estos resultados muestran evidencias satisfactorias de validez convergente del test SON-R 6-40 para el intervalo etario investigado.

**Palabras clave:** Validez convergente, test no verbal de inteligencia, SON-R 6-40.

A inteligência é uma das variáveis latentes mais estudadas pela psicologia, e os resultados desses estudos têm contribuído de maneira significativa para a compreensão do desenvolvimento humano, aprendizagem e desempenho acadêmico (Berg, 2000; Chen & Siegler, 2000; Neisser et al., 1996). De maneira geral, o conceito de inteligência refere-se à capacidade cognitiva de processamento e está associado à aprendizagem e à adaptação ao meio (Almeida, 1994). O presente estudo insere-se neste contexto e busca discutir as evidências de validade de uma medida para a avaliação da inteligência.

As discussões sobre os modelos unidimensionais (principalmente de Spearman, 1904, 1927) versus os multidimensionais (principalmente de Thurstone, 1938) marcaram o início da história da inteligência. No entanto, modelos que integraram as duas perspectivas surgiram por volta da segunda metade do século XX. Nesse contexto, Cattell e Horn buscaram evidências de uma organização fatorial por meio de duas dimensões gerais: inteligência fluida ( $G_f$ ), que envolve as capacidades de raciocínio e de resolução de problemas para os quais a pessoa tem pouco

conhecimento prévio; e inteligência cristalizada ( $G_c$ ), que se refere à aquisição e à solidificação de conhecimentos formais e informais (Cattell, 1943, 1963; Davidson & Downing, 2000; Horn, 1994; Horn & Cattell, 1966). Posteriormente, foram incluídas outras dimensões gerais ao modelo, entre elas, o processamento visual ( $G_v$ ) e velocidade de processamento ( $G_s$ ; Horn, 1994; Horn & Blankson, 2012; Horn & Cattell, 1966). Tais trabalhos foram de suma importância para a psicologia, e algumas das dimensões propostas por Cattell e Horn são, ainda, utilizadas por grande parte dos instrumentos para avaliação da inteligência.

As dimensões propostas por Cattell e Horn foram, em grande parte, incorporadas pelas modernas teorias de Carroll (1993, 2005), bem como pelo modelo CHC (Cattell-Horn-Carroll; McGrew, 2005; McGrew & Flanagan, 1998). Carroll (1993), a partir de um estudo de meta-análise de mais de 400 artigos, evidenciou uma estrutura da inteligência em três níveis hierárquicos, variando do mais específico (fatores de primeira ordem) ao mais geral (um único fator de terceira ordem). Além disso, a maior par-

te das dimensões propostas por Cattell e Horn foi incorporada como fatores de segunda ordem do modelo de Carroll. Esse trabalho de Carroll é considerado por pesquisadores como um dos mais importantes para a história da psicologia (Burns, 1994; Schneider & McGrew, 2012), pois propõe, com base em dados empíricos, uma estrutura complexa para a organização fatorial da inteligência.

No final da década de 1990, McGrew e Flanagan (1998) propuseram uma reorganização das dimensões evidenciadas por Cattell, Horn e Carroll (CHC). Na realidade, McGrew e Flanagan concentraram os seus trabalhos (inclusive os mais recentes) em discussões sobre os modelos teóricos discutidos na literatura sobre o tema. Nesse sentido, o modelo CHC também propõe uma organização hierárquica da inteligência em três níveis (ou estratos). O estrato mais amplo (fator de terceira ordem) é composto por uma capacidade cognitiva geral. O segundo estrato é composto por 16 dimensões amplas, entre elas a inteligência fluída (*Gf*) e a memória. Finalmente, o primeiro estrato é composto por aproximadamente 80 fatores de primeira ordem que representam habilidades específicas, tais como raciocínio indutivo, extensão de memória e tempo de reação, entre outras (McGrew, 2005; McGrew & Flanagan, 1998; McGrew & Wendling, 2010).

Por ser um modelo complexo, o CHC tornou difícil a tarefa de construir um único instrumento para avaliar todas as habilidades postuladas pela teoria. Por outro lado, por ser um modelo mais teórico do que empírico, a maior parte dos instrumentos para avaliação da inteligência pode ser interpretada, de alguma maneira, sob a perspectiva do CHC. Nesse contexto, é possível citar as baterias SON (*Snijders-Oomen Non-Verbal Intelligence Test*).

A primeira bateria SON foi criada na década de 1940 na Holanda pela pesquisadora N. Snijders-Oomen, que pretendia avaliar o potencial de aprendizagem de crianças com deficiência auditiva. Considerando que os testes de inteligência geral disponíveis naquela época não eram adequados para o grupo que esperava avaliar, devido à dependência de habilidades verbais, a autora criou o primeiro teste SON para avaliar

crianças surdas com idades compreendidas entre os 4 e 14 anos de idade (Snijders-Oomen, 1943).

As primeiras versões do SON são anteriores às concepções do modelo CHC. Contudo, a organização das dimensões desse instrumento pode ser interpretada com base no modelo CHC. Nesse sentido, os autores do SON construíram o instrumento de maneira a abranger uma bateria de subtestes de funções específicas, organizando-os numa dimensão geral relacionada à inteligência fluída. Os estudos de validade oferecem suporte à interpretação da estrutura hierárquica do SON-R 2½-7[a] (SON-Revisado para crianças entre 2,5 e 7 anos): uma dimensão geral que subjaz às dimensões específicas de raciocínio e de execução (Karino, Laros, & Jesus, 2011; Valentini, Laros, & Jesus, 2010).

Desde a década de 1940, foram realizadas várias atualizações das baterias SON e, consequentemente, outras versões surgiram. A última versão, o SON-R 6-40 (para pessoas de 6 até 40 anos), foi utilizado na presente pesquisa. Tal versão é indicada para a avaliação da inteligência de pessoas ouvintes e surdas entre 6 e 40 anos de idade. A bateria SON-R 6-40 é composta por quatro subtestes: Analogias, Mosaicos, Categorias e Padrões. Tais subtestes sustentam a predição de escores de inteligência associados ao raciocínio abstrato/indutivo, bem como à execução. Pesquisas recentes evidenciaram que os escores dos quatro subtestes estão explicados por uma dimensão geral, e esta pode ser interpretada como uma medida da inteligência fluida (*Gf*; Tellegen & Laros, 2014).

Ressalta-se que os itens dos subtestes do SON-R 6-40 foram desenvolvidos a partir de uma teoria de dificuldade; isto é, uma revisão analítica dos fatores mais importantes que explicam os níveis sucessíveis de dificuldade dos itens (Snijders, Tellegen, & Laros, 1989). Esse procedimento foi adotado para garantir uma ampla variedade de dificuldade dos itens, bem como facilitar uma aplicação adaptativa. Ademais, a teoria da dificuldade permite criar e testar hipóteses sobre os processos cognitivos envolvidos na resolução de itens. Nesse contexto, cada nível de dificuldade é associado às diferenças entre os estímulos dos itens. Por exemplo, para o subteste

Analogias, os itens mais difíceis estão relacionados, entre outros aspectos, à complexidade das transformações dos estímulos dos itens. As hipóteses que sustentam as teorias da dificuldade do SON foram testadas a partir da amostra de normatização do SON-R 5½-17 (para crianças de 5,5 a 17 anos;  $N = 1,350$ ), utilizando o modelo linear logístico Bayesiano (*Linear Logistic Test Model*). Os resultados desse estudo indicaram um bom ajuste entre as teorias da dificuldade dos subtestes Mosaicos e Padrões e as dificuldades dos itens empiricamente observadas (Geerlings, Laros, Tellegen, & Glas, 2013).

Os estudos de normatização e de validação do SON-R 6-40 na Holanda e Alemanha já foram concluídos (Tellegen & Laros, 2011, 2014; Tellegen, Laros, & Petermann, 2012). Dentre os estudos realizados com o SON-R 6-40 nos Países Baixos, foi conduzida uma pesquisa que buscava obter evidências de validade convergente com outros testes, tais como WAIS-III (*Wechsler Adult Intelligence Scale-III*) e WISC-III (*Wechsler Intelligence Scale for Children-III*). Os resultados encontrados mostram correlações altas de 0,83 entre os escores gerais do SON-R 6-40 e o WAIS-III. A correlação entre o escore geral do SON-R 6-40 e a Escala de Desempenho e Escala Verbal do WAIS-III foi 0,81 e 0,73, respectivamente. A correlação de 0,73 chama a atenção para o fato que um teste não-verbal de inteligência é capaz de explicar uma boa parte da variância de uma escala que pretende medir aspectos verbais da inteligência. Quanto ao teste WISC-III, o escore geral do SON-R 6-40 obteve um coeficiente de correlação de 0,80. Enquanto para os escores obtidos na escala de desempenho e na escala verbal do WISC-III, os resultados das correlações obtidas foram de 0,79 e 0,69, respectivamente.

No Brasil, os estudos de normatização e de validação deste instrumento estão em andamento. Nesse contexto, o presente estudo tem como objetivo investigar a estrutura fatorial e as evidências de validade convergente do teste SON-R 6-40 em uma amostra de participantes brasilienses.

Ressalta-se que os estudos apresentados a seguir apoiam-se nas concepções de validade e precisão defendidas pelo atual *Standards* da

*American Psychological Association* (APA), no qual a validade é compreendida como um conjunto de evidências que sustentam a interpretação dos escores para os fins propostos pelo instrumento (*American Educational Research Association* [AERA], APA, & *National Council on Measurement in Education* [NCME], 1999; Kane, 2001; Sireci & Parker, 2006; Zumbo, 2006). Destaca-se que os autores do presente artigo estão cientes de propostas contemporâneas que defendem a utilização do conceito de validade como uma propriedade do teste e de maneira dicotômica (teste válido ou teste não-válido). Cita-se o pesquisador holandês Denny Borsboom com um dos principais teóricos dessa última proposta. Borsboom entende que uma medida é válida se as variações nas respostas aos itens podem ser explicadas por um atributo psicológico, e se este atributo psicológico, de fato, existe (Borsboom, Cramer, Kievit, Zand Scholten, & Franic, 2009; Borsboom, Mellenbergh, & van Heerden, 2004). Obviamente, tal proposta restringe os estudos de validade a modelos causais. Por outro lado, autores também contemporâneos, como Bruno D. Zumbo, entendem que a utilização rígida de modelos causais como única forma de análise da validade é muito restrita e não condizente com os contextos amplos nos quais as medidas são utilizadas (Zumbo, 2006). Como proposta, Zumbo tem defendido que a interpretação dos escores é o centro do conceito de validade. Nesse sentido, a validade é um processo complexo e diz respeito às evidências que dão suporte às interpretações específicas realizadas a partir dos escores de uma medida. De maneira parecida, Kane (2001) entende que a validade refere-se à acurácia na qual os escores de um indivíduo são estimados. Ainda que Zumbo e Kane difiram em algumas concepções filosóficas, para ambos (assim como para a APA), os escores e a suas interpretações são centrais para o conceito de validade. Kane (2001) e Zumbo (2006) lembram, ainda, a proposta de Cronbach (1988) de separação dos estudos de validade em: *weak program* caracterizado pelos estudos de correlação entre os escores e qualquer outra medida; e *strong program*, na qual a validade dos escores é discutida e dirigida por uma teoria bem articula-

da. Portanto, as evidências de validade podem (e devem) ser obtidas por meio de diferentes fontes e por estudos que utilizam diversos métodos e delineamentos.

Para discussões filosóficas mais aprofundadas sobre o conceito de validade sugere-se a leitura dos textos de Borsboom, Kane, Sireci e Zumbo. No entanto, o presente artigo tem como objetivo apresentar as evidências empíricas, apoiadas na definição de validade do *Standard* da APA (AERA et al., 1999), que corroboram a interpretação dos escores do SON-R 6-40, com base na validade convergente (*weak program*) e na dimensionalidade associada à inteligência fluida (*strong program*).

## Método

Considerando que em estudos de validação de testes psicológicos é recomendável aplicar mais de uma medida que avalia o mesmo construto teórico optou-se pela realização de dois estudos: no primeiro estudo o SON-R 6-40 (Tellegen & Laros, 2011, 2014) foi aplicado junto com a Bateria de Provas de Raciocínio (BPR-5; Primi & Almeida, 2000), enquanto que no segundo estudo aplicou-se o SON-R 6-40 junto com o SON-R 2½-7[a] (Laros, Tellegen, Jesus, & Karino, 2015).

## Participantes

No primeiro estudo, participaram 68 estudantes entre o sétimo e o nono ano do ensino básico, sendo 36 (52,9%) do sexo masculino e 32 (47,1%) do sexo feminino. A idade dos participantes neste estudo variou entre 11,7 anos e 16,1 anos ( $M = 14,2$ ;  $DP = 1,17$ ). Os participantes eram alunos de escolas da rede pública de Brasília/DF. No segundo estudo de validade, participaram 44 alunos do primeiro e segundo ano do ensino básico de escolas públicas de Brasília/DF, sendo 23 do sexo masculino (52,3%) e 21 do sexo feminino (47,7%) com idades variando entre 5,9 e 7,8 anos ( $M = 6,9$ ;  $DP = 0,47$ ) no momento da aplicação do teste. As amostras dos dois estudos de validade foram amostras de conveniência. A pesquisa foi realizada entre os meses finais de 2012 e os primeiros meses de 2013.

## Instrumentos

1. O SON-R 6-40 é uma bateria não-verbal que busca avaliar a inteligência fluida como descrita pela teoria de Carroll (2005) e pela teoria de Carroll-Horn-Cattell (Schneider & McGrew, 2012). O teste é de aplicação individual com duração média de 60 minutos e pode ser administrado para pessoas na faixa etária entre 6 e 40 anos com ou sem uso da linguagem falada. A bateria é composta de quatro subtestes: dois subtestes de múltipla escolha com 36 itens (Analogias e Categorias), e dois subtestes de execução com 26 itens (Mosaicos e Padrões). Os subtestes de múltipla escolha são divididos em três séries paralelas (itens 1a-12a, itens 1b-12b e itens 1c-12c) e os subtestes de execução (Mosaicos e Padrões) são divididos em duas séries paralelas (itens 1a-13a e itens 1b-13b). A divisão dos subtestes em séries paralelas foi feita para possibilitar o procedimento de testagem adaptativa. Com o procedimento de testagem adaptativa apenas os itens mais adequados são aplicados a cada participante, minimizando a quantidade de itens aplicados que são fáceis demais ou difíceis demais para o respondente. O procedimento adaptativo é importante para manter a motivação do sujeito durante a aplicação dos subtestes. No subteste Analogias a tarefa consiste em descobrir qual transformação ocorreu na primeira figura e aplicar a mesma transformação na segunda figura. No subteste Categorias o sujeito precisa descobrir o conceito subjacente aos três desenhos apresentados. No subteste Mosaicos, diversas figuras de mosaicos precisam ser compiados numa moldura utilizando quadrados vermelhos, brancos e vermelho-brancos. No subteste Padrões o sujeito deve completar um padrão formado por uma ou duas linhas em que uma parte está faltando. Os subtestes Mosaicos e Padrões tem um tempo limite na aplicação. Os escores totais são normatizados do SON-R 6-40 na escala de QI ( $M=100$ ,  $DP=15$ ); os escores dos subtestes são normatizados na escala de subtestes de Wechsler ( $M=10$ ,  $DP=3$ ). Uma vez que

escores normatizados para Brasil ainda não estavam disponíveis foram utilizados as normas da Holanda /Alemanha. Maiores informações sobre o teste SON-R 6-40 podem ser encontradas no site criado pelos autores ([www.testresearch.nl](http://www.testresearch.nl)).

2. A Bateria de Provas de Raciocínio (BPR-5; Almeida & Primi, 2000) avalia os seguintes tipos de raciocínio: Raciocínio Verbal (RV), Raciocínio Abstrato (RA), Raciocínio Espacial (RE), Raciocínio Numérico (RN) e Raciocínio Mecânico (RM). O último subteste não foi aplicado devido ao alto grau de dificuldade para a amostra estudada. A BPR-5 possui duas formas: forma A indicada para estudantes cursando as séries finais do ensino fundamental e forma B indicada para alunos do ensino médio. A forma A foi aplicada na atual pesquisa considerando as características da amostra. A BPR-5 foi normatizada para o Brasil em 2000: a pesquisa de normatização foi realizada com uma amostra de 1.722 alunos (603 para a forma A e 1119 para a forma B). Em um estudo de validação a bateria BPR-5 foi aplicada junto com as escalas Matrizes Progressivas Coloridas de Raven e o G-36 (Teste Não Verbal de Inteligência; Almeida et al., 2010). O estudo foi realizado com uma amostra de 330 estudantes universitários cursando o segundo ano de Psicologia. Análise fatorial exploratória dos dados desse estudo forneceu evidências que os subtestes Raciocínio Mecânico, Raciocínio Espacial e Raciocínio Abstrato medem a inteligência fluida (*Gf*), enquanto os subtestes Raciocínio Verbal e Raciocínio Numérico medem a inteligência cristalizada (*Gc*). Os escores da BPR-5 são normatizados na escala padrão (Escore Padrão Normalizado [EPN]), que se assemelha à escala de QI ( $M=100$ ;  $DP=15$ ).
3. O SON-R 2½-7[a] é um teste não-verbal de inteligência para crianças com idades entre 2,5 anos e 7 anos composto por quatro subtestes: Mosaicos, Categorias, Situações e Padrões. O subtestes Categorias e Situações compõem a Escala de Raciocínio e os sub-

testes Mosaicos e Padrões a Escala de Execução. O subteste Categorias possui 15 itens e um exemplo. Nos primeiros sete itens quatro ou seis cartões precisam ser postos na categoria correta. Os itens seguintes são de múltipla escolha, nos quais a criança deve indicar, entre os cinco objetos, os dois adequados à categoria em análise. O subteste Situações possui 14 itens e um exemplo. Na primeira parte do subteste são apresentadas metades de quatro figuras. Depois de ter recebido os cartões com as metades faltantes, a criança precisa colocar a metade certa no desenho correto. A segunda parte de Situações é de múltipla escolha. Cada item consiste em uma situação desenhada em que uma ou duas peças estão faltando. As peças corretas precisam ser escolhidas entre várias alternativas a fim de deixar a situação consistente. O subteste Mosaicos é composto por 15 itens e um exemplo. Na primeira parte, a criança precisa copiar padrões simples de mosaicos numa moldura utilizando 3 a 5 quadrados vermelhos. Na segunda parte, requer-se da criança que padrões de mosaicos sejam copiados com quadrados vermelhos, amarelos e vermelhos / amarelos. O subteste Padrões é formado por 16 itens. Solicita-se à criança que copie uma forma utilizando um lápis. Os itens inicialmente são demonstrados pelo examinador e, no decorrer da aplicação, a criança precisa copiar a figura sem ajuda. Exemplos dos itens dos quatro subtestes podem ser encontrados no *Manual and Research Report* do SON-R 2½-7 que está disponível no *website* dos testes SON (<http://www.testresearch.nl/sonr/sonr257manual.pdf>). Os escores totais normatizados do SON-R 2½-7[a] e os escores normatizados das duas escalas são na escala de QI ( $M=100$ ,  $DP=15$ ). Foram utilizadas as normas brasileiras na determinação dos escores normatizados.

### Procedimento

Depois de obter aprovação do Comitê de Ética, os estudos de validade do SON-R 6-40

começaram a ser realizadas com o apoio de uma equipe de alunos de graduação em psicologia. Antes do inicio da pesquisa a equipe recebeu um treinamento na aplicação dos testes. No contato inicial com as escolas foram entregues a carta de apresentação da pesquisa, o documento de aprovação no Comitê de Ética em Pesquisa e os documentos de Consentimento Livre e Esclarecidos (TCLE) para os alunos e seus responsáveis. No total, quatro escolas públicas do Plano Piloto de Brasília participaram da coleta. Para possibilitar a aplicação dos testes, os alunos levaram um Termo de Consentimento Livre e Esclarecido para os seus pais. Assim, os pais podiam optar pela participação ou não participação de seus filhos na pesquisa.

A administração do SON-R 6-40 e do SON-R 2½-7[a] foi realizada individualmente. A instrução dos testes foi feita de maneira padronizada de acordo com os manuais dos testes. A ordem de aplicação dos testes foi alternada: 50% recebeu o SON-R 6-40 como primeiro teste e 50% recebeu o SON-R 2½-7[a] como primeiro teste.

A administração da BPR-5 foi realizada em grupo. A instrução da BPR-5 foi feita de maneira padronizada de acordo com o manual da bateria. A BPR-5 foi o primeiro teste aplicado, o SON-R 6-40 foi aplicado como segundo teste.

*Análise dos Dados.* Para as análises exploratórias e para as análises descritivas foi utilizado o pacote estatístico SPSS (*Statistical Package for the Social Sciences*) versão 18.0. Para a verificação da estrutura fatorial do SON-R 6-40 utilizou-se o software FACTOR versão 9.2 (Lorenzo-Seva & Ferrando, 2006).

A estrutura fatorial dos itens do SON-R 6-40 foi analisada juntando os dados de Estudo 1 ( $N=68$ ) e Estudo 2 ( $N=44$ ). Para determinar o número de fatores a extrair utilizou-se um tipo de análise Paralela – *Optimal Implementation of Parallel Analysis* (Timmerman & Lorenzo-Seva, 2011), em função da sua robustez para avaliar o número de fatores a ser retido (Baglin, 2014; Damásio, 2012; Timmerman & Lorenzo-Seva, 2011). A análise fatorial foi realizada com o método *Minimum Rank Factor Analysis* (Ten

Berge & Kiers, 1991), com base em correlações politómicas (Baglin, 2014).

Os itens muito fáceis ( $p > 0,90$ ) e os itens muito difíceis ( $p < 0,10$ ) não foram considerados na verificação da estrutura fatorial do SON-R 6-40. Assim, 43 dos 112 itens não foram incluídos na análise fatorial. Para evitar o surgimento de fatores artificiais de dificuldade a análise fatorial foi realizada com parcelas de itens em vez de itens individuais. O uso de parcelas de itens reduz a possibilidade de surgimento de fatores artificiais de dificuldade na análise fatorial com itens dicotómicos (Holt, 2004) e gera soluções mais estáveis (Little, Cunningham, Shahar, & Widaman, 2002; Rocha & Chelladurai, 2012). Assim, os 69 itens remanescentes foram distribuídos em 17 parcelas de itens. Cada parcela de itens consiste de quatro ou cinco itens: a distribuição dos itens nas parcelas de itens foi realizada na sequencia dos itens. Assim, por exemplo, a primeira parcela de itens consiste dos primeiros quatro itens.

No que diz respeito à análise de consistência interna, foram utilizados três índices: o coeficiente Lambda 2 de Guttman, a correlação item-resto (correlação item-total corrigida) e a correlação média entre os itens. Utilizou-se como índice de consistência interna o coeficiente Lambda 2 de Guttman, uma vez que estudos apontam que o esse coeficiente é uma estimativa melhor da fidedignidade, principalmente quando a amostra é pequena (Sijtsma, 2009, 2012; Ten Berge & Zegers, 1978).

Para a realização das análises de validade convergente do SON-R 6-40, utilizou-se da correlação não paramétrica (rho) de Spearman. Além das correlações brutas, foram calculadas as correlações corrigidas. Para tanto, utilizou-se a correção dupla para atenuação (Osborne, 2003; Thompson, 2003), cuja equação é:  $r_{xy} = r_{xy} \div \sqrt{r_{xx}r_{yy}}$ , na qual,  $r_{xy}$  é a correlação depois da correção dupla para atenuação,  $r_{xy}$  = correlação bruta e  $r_{xx}$  e  $r_{yy}$  são os coeficientes de fidedignidade dos dois testes. Osborne (2003) e Thompson (2003) argumentam que as relações reais entre variáveis podem ser subestimadas caso uma parte da variância é variância de erro. Segundo esses autores,

pesquisadores precisam corrigir para atenuação no intuído de obter uma estimativa melhor da relação verdadeira entre as variáveis na população. Outro índice calculado da relação entre os testes foi o índice de determinação, obtido por meio da correlação elevada ao quadrado. Esse índice é utilizado para mostrar a variância em comum ou a variância compartilhada entre duas variáveis.

## Resultados

### Estrutura Fatorial do SON-R 6-40

Inicialmente foi realizada uma Análise Paralela (AP) de 17 parcelas dos itens do SON-R 6-40 para avaliar a quantidade de fatores a ex-

trair. Usando o programa FACTOR 9.2 foram analisados 500 matrizes de correlações aleatórias e foi escolhida a opção de permutação com os dados observados (Buja & Eyuboglu, 1992). Os *eigenvalues* empíricos e aleatórios (percentil 95) apresentaram os seguintes valores, em sequência por dimensão sugerida: (a) empíricos 61,0 e 7,9; (b) aleatórios 14,3 e 12,3. Tais resultados indicaram a presença de um fator único.

Após a AP, foi realizada uma análise fatorial com o método *Minimum Rank Factor Analysis* (MRFA; Ten Berge & Kiers, 1991), indicada por Timmerman e Lorenzo-Seva (2011). Os resultados da análise fatorial com MRFA podem ser observados na Tabela 1.

**Tabela 1**  
**Cargas Fatoriais (CF) e Comunalidades ( $h^2$ ) Análise Fatorial Exploratória do SON-R 6-40 (Estudo 1 & 2;  $n = 68$ )**

Parcela de itens	<i>n</i> itens	CF	$h^2$
Analogias 1	4	0,79	0,62
Analogias 2	4	0,71	0,50
Analogias 3	4	0,70	0,49
Analogias 4	4	0,71	0,50
Analogias 5	4	0,66	0,44
Mosaicos 1	4	0,75	0,56
Mosaicos 2	4	0,69	0,48
Mosaicos 3	5	0,84	0,71
Categorias 1	4	0,67	0,45
Categorias 2	4	0,64	0,41
Categorias 3	4	0,65	0,42
Categorias 4	4	0,63	0,40
Categorias 5	4	0,72	0,56
Padrões 1	4	0,85	0,72
Padrões 2	4	0,82	0,67
Padrões 3	4	0,76	0,58
Padrões 4	4	0,86	0,74
Média	----	0,73	0,54

*Notas.* Percentagem explicada da variância comum = 66,2%; RMSR (*Root Mean Square of Residuals*) = 0,0755; Valor máximo do RMSR esperado para um modelo aceitável = 0,0949.

Os resultados da análise fatorial indicaram que o fator único explicou 66,2% da variância comum. Segundo Baglin (2014) a percentagem da variância comum explicada pode ser considerada uma medida de ajuste do modelo aos dados (*goodness of fit*). Além disso, estimou-se o indicador de ajuste RMSR (*Root Mean Square of Residuals*), cujo valor foi igual a 0,075. Considerando que o ponto de corte do RMSR, calculado para esta análise fatorial, é de no máximo de 0,095, o resultado empírico obtido (abaixo do ponto de corte) indica que o resíduo do modelo é aceitável. Tal resultado também aponta para a adequação do modelo unidimensional do SON-R 6-40.

Ademais, conforme apresentado na Tabela 1, as cargas fatoriais, para o modelo unidimen-

sional, foram todas acima de 0,60 (cargas fatoriais variaram de 0,63 a 0,86). Reise, Waller e Comrey (2000) afirmam que se as cargas fatoriais são altas ( $>0,60$ ) e o fator consiste de um grande número de variáveis, uma amostra de 100 participantes é suficiente para usar a análise fatorial. Nesse sentido, ressalta-se a adequação da análise com MRFA, cujos resultados indicaram que o SON-R 6-40 tem uma estrutura unidimensional.

### **Evidências da Validade Convergente do SON-R 6-40 com a BPR-5**

Depois de analisar a estrutura fatorial do SON-R 6-40 foram analisadas as características psicométricas dos escores no SON-R 6-40 e da BPR-5, aplicados no primeiro estudo. A Tabela 2 descreve os resultados dessas análises.

**Tabela 2**  
**Características Psicométricas do SON-R 6-40 e da BPR-5 (Estudo 1; N = 68)**

Parâmetros psicométricos	SON-R 6-40: Escala Geral e subtestes				
	Escala Geral	Analogias	Mosaicos	Categorias	Padrões
Consistência Interna ( $\lambda_2$ )	0,95	0,87	0,89	0,90	0,91
Correlação item-total ( $r_{it}$ )	0,37	0,36	0,44	0,38	0,47
Correlação entre os itens	0,15	0,15	0,22	0,17	0,24
Proporção de acertos	0,60	0,61	0,62	0,55	0,64
Número de itens	124	36	26	36	26

  

Parâmetros psicométricos	BPR-5: Escala Geral e subtestes				
	Escala Geral	RV	RA	RE	RN
Consistência Interna ( $\lambda_2$ )	0,98	0,95	0,95	0,92	0,94
Correlação item-total ( $r_{it}$ )	0,57	0,60	0,62	0,57	0,59
Correlação entre os itens	0,33	0,38	0,40	0,35	0,38
Proporção de acertos	0,33	0,37	0,39	0,29	0,24
Número de itens	90	25	25	20	20

*Notas.* RV = Raciocínio Verbal; RA = Raciocínio Abstrato; RE = Raciocínio Espacial e RN = Raciocínio Numérico.

A Tabela 2 indica valores altos para a fidedignidade dos escores gerais dos dois testes: 0,95 para o SON-R 6-40 e 0,98 para a BPR-5. Os escores nos subtestes também mostram coe-

ficientes de fidedignidade elevados: para o SON-R 6-40 os valores ficaram entre 0,87 e 0,91 e para a BPR-5 entre 0,92 e 0,95. Em comparação com o SON-R 6-40 a BPR-5 é um teste mais di-

fícl: a porcentagem de acertos média na BPR-5 é 33% e no SON-R 6-40 é 60%.

Para obter evidências de validade convergente foram calculadas as correlações Spearman

( $r_s$ ) entre o SON-R 6-40 e a BPR-5 e os quatro subtestes da BPR-5. As correlações encontram-se na Tabela 3.

**Tabela 3**

**Coeficientes de Correlação de Spearman ( $r_s$ ) e Coeficientes de Determinação ( $r_s^2$ ) entre o SON-R 6-40 e a BPR-5 (Estudo 1; N=68)**

Coeficientes de correlação ( $r_s$ ) e coeficientes de determinação ( $r_s^2$ )						
	SON-R 6-40	RA	RV	RE	RN	BPR-5
SON-R 6-40	----	0,50 (0,46)	0,45 (0,41)	0,46 (0,41)	0,32 (0,29)	0,55 (0,50)
RA	0,71 (0,68)	----	0,45 (0,40)	0,64 (0,56)	0,31 (0,28)	0,81 (0,76)
RV	0,66 (0,63)	0,67 (0,63)	----	0,52 (0,46)	0,53 (0,48)	0,79 (0,71)
RE	0,68 (0,64)	0,80 (0,75)	0,72 (0,68)	----	0,42 (0,36)	0,85 (0,76)
RN	0,57 (0,54)	0,56 (0,53)	0,73 (0,69)	0,65 (0,60)	----	0,71 (0,66)
BPR-5	0,74 (0,71)	0,90 (0,87)	0,89 (0,84)	0,92 (0,87)	0,84 (0,81)	----

*Notas.* RV = Raciocínio Verbal; RA = Raciocínio Abstrato; RE = Raciocínio Espacial; RN = Raciocínio Numérico. Os coeficientes de correlação encontram-se no triângulo inferior da tabela, e os coeficientes de determinação (os coeficientes de correlação elevados ao quadrado) encontram-se no triângulo superior. Os primeiros valores são os coeficientes corrigidos para atenuação e os valores entre parênteses são os coeficientes brutos.

O exame do triângulo inferior da Tabela 3 revela coeficientes de correlação altos entre o SON-R 6-40 e o BPR-5 e seus subtestes, cujas correlações foram acima de 0,60, com exceção da correlação bruta entre o SON-R 6-40 e o subteste Raciocínio Numérico ( $r_s = 0,54$ ). O coeficiente mais importante no triângulo inferior para avaliar a validade convergente do SON-R 6-40 é a correlação corrigida entre os escores totais do SON-R 6-40 e do BPR-5 ( $r_s = 0,74$ ). Uma correlação de 0,74 significa que os dois testes têm 54,8% da sua variância em comum. Como esperado, as correlações mais altas foram encontradas entre o SON-R 6-40 e os subtestes de Raciocínio Abstrato e Raciocínio Espacial da BPR-5 (valores da correlação corrigida de 0,71 e 0,68 respectivamente) uma vez que estes subtestes medem o mesmo construto de inteligência que o SON-R 6-40 visa medir, a saber, a inteligência fluida.

#### *Evidências da Validade Convergente do SON-R 6-40 com o SON-R 2½-7[a]*

Antes de calcular as correlações entre os escores do SON-R 6-40 e do SON-R 2½-7[a] foram realizadas análises das características psicométricas dos escores dos ambos os testes. A Tabela 4 mostra os resultados dessas análises.

Os coeficientes de consistência interna Lambda 2 ( $\lambda_2$ ) dos escores nos subtestes do SON-R 6-40 mostram valores entre 0,83 (Mosaicos) e 0,92 (Analogias) nessa amostra. Para a bateria inteira foi encontrada um coeficiente de consistência interna de 0,95 (Tabela 4).

A proporção média de acertos de 0,36 do SON-R 6-40 mostra que o teste foi difícil para esse grupo de crianças com idades entre 5,9 e 7,8 anos. O subteste Padrões do SON-R 6-40 foi o subteste mais difícil para essa amostra com uma proporção de acertos de 0,32 e o subteste Analogias foi o subteste mais fácil com uma proporção de acertos de 0,42.

**Tabela 4****Características Psicométricas dos Escores no SON-R 6-40 e no SON-R 2½-7[a] (Estudo 2; N = 44)**

Parâmetros psicométricos	SON-R 6-40: Escala Geral e subtestes				
	Escala Geral	Analogias	Mosaicos	Categorias	Padrões
Consistência Interna ( $\lambda_2$ )	0,95	0,92	0,83	0,89	0,90
Correlação item-total ( $r_{it}$ )	0,38	0,47	0,45	0,42	0,51
Correlação entre os itens	0,16	0,24	0,26	0,21	0,29
Proporção de acertos	0,36	0,42	0,35	0,35	0,32
Número de itens	124	36	26	36	26

  

Parâmetros psicométricos	SON-R 2½-7[a]: Escala Geral e subtestes				
	Escala Geral	Mosaicos	Categorias	Situações	Padrões
Consistência Interna ( $\lambda_2$ )	0,83	0,78	0,78	0,60	0,62
Correlação item-total ( $r_{it}$ )	0,29	0,49	0,43	0,27	0,30
Correlação entre os itens	0,11	0,32	0,25	0,14	0,15
Proporção de acertos	0,79	0,75	0,80	0,78	0,83
Número de itens	60	15	15	14	16

Escala de Execução:  $\lambda_2 = 0,80$ ; correlação entre os itens = 0,20; proporção de acertos = 0,79; n itens = 31.

Escala de Raciocínio:  $\lambda_2 = 0,71$ ; correlação entre os itens = 0,10; proporção de acertos = 0,79; n itens = 29.

Para a consistência interna dos subtestes do SON-R 2½-7[a] foram encontrados valores entre 0,60 (Situações) e 0,78 (Mosaicos e Categorias). Para a bateria inteira foi encontrado um coeficiente de consistência interna de 0,83. O SON-R 2½-7[a] foi muito fácil para esse grupo de crianças com uma proporção média de acertos de 0,79. Esse fato provavelmente afetou negativamente a consistência interna dos escores dos subtestes e da bateria inteira.

As análises das características psicométricas indicam que para este grupo de crianças com idades entre 5,9 e 7,8 anos a qualidade psicométrica do SON-R 6-40 é melhor em comparação com o SON-R 2½-7[a].

Depois de analisar as características psicométricas foram calculadas as correlações entre o SON-R 6-40 e o SON-R 2½-7[a] e as suas duas escalas. Os resultados dessa análise podem ser observados na Tabela 5.

Na inspeção do triângulo inferior da Tabela 5, nota-se que o coeficiente de correlação corrigida para atenuação entre o SON-R 6-40 e o SON-R 2½-7[a] é 0,80. A correlação de 0,80 entre os dois testes pode ser interpretada como coeficiente de validade do SON-R 6-40. No triângulo superior da Tabela 5 podemos ver o valor do coeficiente de determinação de 0,64 entre o SON-R 6-40 e o SON-R 2½-7[a] o que indica que os testes têm 64% da sua variância em comum. Os resultados encontrados sugerem que a bateria SON-R 6-40 tem uma relação mais forte com a escala de Escala de Raciocínio do SON-R 2½-7[a] do que com a Escala de Execução ( $r_s$  corrigida de 0,79 versus  $r_s$  corrigida de 0,66).

## Discussão

A presente pesquisa buscou reunir evidências da dimensionalidade, precisão e validade

**Tabela 5**

**Coeficientes de Correlação de Spearman ( $r_s$ ) e Coeficientes de Determinação ( $r_s^2$ ) entre o SON-R 6-40 e SON-R 2½-7[a] (Estudo 2; N=44)**

Coeficientes de correlação ( $r_s$ ) e coeficientes de determinação ( $r_s^2$ )				
	SON-R 6-40	SON-R EE	SON-R ER	SON-R 2½-7[a]
SON-R 6-40	----	0,44 (0,34)	0,62 (0,42)	0,64 (0,50)
SON-R EE	0,66 (0,58)	----	0,36 (0,20)	0,86 (0,69)
SON-R ER	0,79 (0,65)	0,60 (0,45)	----	0,90 (0,72)
SON-R 2½-7[a]	0,80 (0,71)	0,93 (0,83)	0,95 (0,85)	----

*Notas.* SON-R EE = SON-R Escala de Execução; SON-R ER = SON-R Escala de Raciocínio. Os coeficientes de correlação encontram-se no triângulo inferior da tabela, e os coeficientes de determinação (os coeficientes de correlação elevados ao quadrado) encontram-se no triângulo superior. Os primeiros valores são os coeficientes corrigidos para atenuação e os valores entre parênteses são os coeficientes brutos.

convergente da bateria SON-R 6-40, por meio da aplicação desse instrumento em conjunto com outras medidas de inteligência: o SON-R 2½-7[a] e a BPR-5. Salienta-se que a bateria SON-R 6-40 está em processo de adaptação para o Brasil. Contudo, até a presente data, nenhum outro estudo sobre a validade desse instrumento foi realizado neste país. Portanto, ressalta-se a importância do estudo atual uma vez que é o primeiro estudo nesse país.

Em relação às análises da estrutura interna do SON-R 6-40, os resultados, de maneira geral, evidenciam a unidimensionalidade de escala. Esse resultado tem uma importante aplicação prática, pois permite a predição de um escore geral. Ademais, as evidências da unidimensionalidade são coerentes com os resultados do estudo realizado na Holanda, no qual também se encontrou evidência de uma dimensão única (Tellegen & Laros, 2014).

A dimensão única também é coerente com pesquisas recentes e teorias fatoriais da inteligência (Carroll, 1993, 2005; McGrew & Flanagan, 1998; McGrew & Wendling, 2010). Por exemplo, Primi, Silva, Santana, Muniz e Almeida (2013) avaliaram a adequação de um modelo *bi-factor* para a estrutura da BPR-5. Nessa pesquisa, foi modelada uma dimensão geral explicando todos os itens do instrumento, bem como fatores específicos associados aos subtestes. Ainda que

a BPR-5 apresente subtestes associados tanto à inteligência fluida quanto à cristalizada, a adequação do modelo *bi-factor* respaldou a modelagem uma dimensão geral. Além disso, a teoria de Carroll (1993, 2005) e o modelo CHC (McGrew & Flanagan, 1998; McGrew & Wendling, 2010) também advogam em favor da existência de uma dimensão geral, embora dimensões específicas também sejam estruturadas por estes dois últimos modelos.

Cabe salientar que a evidência de uma dimensão geral não invalida a estruturação de dimensões específicas. Em outras palavras, as evidências da unidimensionalidade indicam que as tarefas avaliadas pelo SON-R 6-40 podem ser explicadas por uma dimensão geral e que tais escores podem ser preditos; mas, ainda assim, pesquisas futuras podem testar e evidenciar estruturas alternativas, modelando dimensões específicas, além do fator geral. Nesse contexto, ressalta-se que o tamanho da amostra desta pesquisa não é grande. Portanto, os resultados apresentados tratam de evidências iniciais da estrutura do SON-R 6-40 e que tais evidências precisam ser testadas em pesquisas futuras com amostras maiores.

Os parâmetros dos itens, estimados por meio da Teoria Clássica dos Testes (TCT), apresentados nas Tabelas 2 e 4, indicam que a bateria SON-R 6-40 é mais fácil do que a BPR-5,

para crianças e adolescentes entre 6 e 11 anos; bem como o SON-R 6-40 é mais difícil do que o SON-R 2½-7[a], para crianças entre 6 e 7 anos. Tais resultados eram esperados uma vez que a faixa etária das crianças avaliadas no Estudo 2 representa o limite superior de idade do SON-R 2½-7[a] e o limite inferior de idade do SON-R 6-40. Para o Estudo 1, no que se refere à BPR-5, a dificuldade estimada na presente amostra é semelhante aos resultados apresentados por Primi e Almeida (2000), estudo no qual indicou proporções de acerto dos subtestes entre 0,24 e 0,39 para uma amostra de participantes de Minas Gerais.

Salienta-se, contudo, que os parâmetros de dificuldade dos itens do SON-R 6-40 do Estudo 1 não podem ser comparados, diretamente, com os do Estudo 2, uma vez que a proporção de acertos, na TCT, é dependente do parâmetro de habilidade da amostra. Em estudos posteriores, a avaliação da dificuldade do SON-R 6-40 por meio da Teoria de Resposta ao Item (TRI), poderia apresentar estimações invariantes desse parâmetro (b). Contudo, na presente pesquisa, não foi possível estimar os parâmetros dos itens por meio da TRI em função do tamanho da amostra dos Estudos 1 e 2.

Portanto, o número reduzido de participantes nos dois estudos deve ser apontado como uma limitação da pesquisa. O aumento do tamanho da amostra proporcionaria maior estabilidade na estimação dos parâmetros dos itens, bem como dos escores dos participantes. Ademais, amostras maiores permitiram a estimação dos parâmetros por meio dos modelos de dois e três parâmetros logísticos (2-PL e 3-PL) da TRI. No entanto, é necessário mencionar que a aplicação dos testes utilizados nos Estudos 1 e 2 é longa e deve ser realizada individualmente (com exceção da aplicação da BPR-5). Portanto, todos os participantes foram avaliados individualmente, ao menos, uma vez. Nesse sentido, o aumento da amostra representa um custo financeiro bastante significativo, bem como um acréscimo considerável no período de coleta de dados.

No que se refere ao parâmetro de discriminação dos itens, por meio da correlação item-

-total é possível avaliar a qualidade do instrumento em diferenciar participantes com escores semelhantes. Para os Estudos 1 e 2, as médias das correlações item-total foram acima de 0,30, indicando que o instrumento discrimina adequadamente os escores dos participantes. Cabe salientar, ainda, que alguns itens são muito fáceis ou muito difíceis e apresentam, portanto, proporções de acerto muito baixa ou muito alta. Essa restrição da variabilidade tende a diminuir o valor de correlação, subestimando o parâmetro de discriminação dos itens. Neste contexto, realizou-se uma reanálise dos dados, na qual os itens com proporções de acerto muito baixa (proporção < 0,10) ou muito alta (proporção > 0,90) foram desconsiderados. Nessa reanálise, as médias das correlações item-total aumentaram aproximadamente 0,10, sendo todas acima de 0,40. Ressalta-se, contudo, que os itens muito fáceis ou muito difíceis não podem ser excluídos do teste, pois são úteis para as avaliações de pessoas em faixas etárias muito distintas (entre 6 e 40 anos), bem como para as avaliações de pessoas com habilidades cognitivas muito baixas ou muito altas. Portanto, a reanálise, aqui discutida, busca apenas indicar que, provavelmente, a discriminação média dos itens do teste SON-R 6-40 é superior a 0,40. Estudos futuros, com amostras mais heterogêneas em relação às capacidades cognitivas, poderiam aumentar a variabilidade dos dados e, consequentemente, tenderiam a aumentar o valor do parâmetro de discriminação.

Os valores adequados de discriminação dos itens também estão relacionados à precisão dos escores. Sabe-se que quanto maior é a capacidade discriminativa do item, mais informação útil ele agrega ao instrumento e, consequentemente, maior é a consistência interna. Nos dois estudos, o SON-R 6-40 apresentou coeficientes de precisão igual a 0,95, para a escala geral, bem como coeficientes acima de 0,83 para todos os subtestes. Tais resultados indicam que SON-R 6-40 é capaz de avaliar a inteligência de crianças e adolescente com boa precisão (associada à consistência interna), minimizando o erro de medida relacionado ao conteúdo dos itens.

A estimação da precisão dos escores também foi importante para o estudo de validade convergente. Uma vez que as medidas não são perfeitamente precisas, era esperado que as correlações entre os escores dos testes fossem subestimadas. Buscando evitar esse problema, foram calculadas as correlações atenuadas para a ausência de fidedignidade perfeita, que são interpretadas como a estimativa da correlação caso ambos os testes produzissem escores livres de erro de medida. Entre os resultados do presente estudo, destacam-se as altas correlações encontradas entre os escores gerais do SON-R 6-40, SON-R 2½-7[a] e BPR-5, superiores a 0,70. Tais resultados indicam que se os escores fossem mensurados sem erro de medida, os testes compartilhariam mais de 50% da variância. Resultados de um estudo internacional apontaram para efeitos semelhantes. Tellegen e Laros (2011) avaliaram uma amostra de 35 crianças alemãs e evidenciaram um coeficiente de correlação igual a 0,77 entre os escores das baterias SON-R 6-40 e o WISC-IV. Portanto, apresentam-se evidências da validade convergente do SON-R 6-40.

No que se refere especificamente à relação entre os escores do SON-R 6-40 e os escores dos subtestes da BPR-5 também foram observadas correlações entre moderadas e fortes. Nesse sentido, foram encontradas correlações mais elevadas entre o SON-R 6-40 e os subtestes Raciocínio Abstrato e Raciocínio Espacial da BPR-5 ( $r = 0,68$  para ambos os subtestes) do que com os subtestes Raciocínio Verbal e Raciocínio Numérico ( $r = 0,62$  e  $r = 0,44$ , respectivamente). Almeida et al. (2010) estudaram a relação entre o BPR-5, o Raven e o G-36 e indicaram que os subtestes Raciocínio Verbal e Raciocínio Numérico estão mais associados à inteligência cristalizada e os subtestes Raciocínio Abstrato e Raciocínio Espacial se relacionam mais à inteligência fluida. Assim, os resultados da presente pesquisa indicam que os escores do SON-R 6-40 estão mais associados à inteligência fluida do que à cristalizada. Portanto, os escores do SON-R 6-40 estão associados à habilidade para apreender e solucionar proble-

mas novos, para os quais a pessoa tem pouco conhecimento prévio.

Os resultados do Estudo 2 indicaram uma pequena diferença no tamanho do efeito para os subtestes: os escores do SON-R 6-40 associaram-se mais fortemente à Escala de Raciocínio do SON-R 2½-7[a] do que à Escala de Execução. No entanto, essa diferença é pequena (respectivamente,  $r = 0,72$  e  $r = 0,62$ ) e ambas as correlações oferecem suporte à validade convergente. Portanto, os escores do SON-R 6-40 associam-se às habilidades exigidas para a solução de problemas simples e complexos relacionadas ao raciocínio abstrato e concreto (Escala de Raciocínio), bem como estão relacionadas às habilidades visuomotoras e espaciais (Escala de Execução). Tais habilidades estão relacionadas à inteligência fluida, dentro do modelo CHC (McGrew, 2005; McGrew & Flanagan, 1998).

Em que pese às limitações da amostra da presente pesquisa, os resultados obtidos forneceram evidências fortes da validade e da precisão dos escores do SON-R 6-40 no contexto brasileiro. Tais evidências oferecem suporte à utilização do SON-R 6-40 para avaliação, com boa precisão, da inteligência fluida de crianças e adolescentes. Ademais, o SON-R 6-40 pode contribuir com futuras pesquisas e avaliações clínicas, fornecendo uma medida ampla e precisa das habilidades cognitivas.

## Referências

- Almeida, L. S. (1994). *Inteligência: Definição e medida*. Aveiro, Portugal: CIDInE.
- Almeida, L. S., Nascimento, E. do, Lima, A. O. F., Vasconcelos, A. G., Akama, C. T., & Santos, M. T. (2010). Bateria de Provas de Raciocínio (BPR-5): Estudo exploratório em alunos universitários. *Avaliação Psicológica*, 9(2), 155-162.
- Almeida, L. S., & Primi, R. (2000). *BPR-5: Bateria de provas de raciocínio: Manual técnico*. São Paulo, SP: Casa do Psicólogo.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

- Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research & Evaluation*, 19(5), 1-14.
- Berg, C. A. (2000). Intellectual development in adulthood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 117-137). Cambridge, UK: Cambridge University Press.
- Borsboom, D., Cramer, A., Kievit, R., Zand Scholten, A., & Franic, S. (2009). The end of construct validity. In R. Lissitz (Ed.), *The concept of validity* (pp. 135-170). Charlotte, NC: Information Age.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111, 1061-1071. doi:10.1037/0033-295X.111.4.1061
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27(4), 509-540. doi:10.1207/s15327906mbr2704\_2
- Burns, R. B. (1994). Surveying the cognitive terrain [Book reviews]. *Educational Researcher*, 23(3), 35-37. doi:10.3102/0013189X023003035
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (2<sup>nd</sup> ed., pp. 69-76). New York: The Guilford Press.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40(3), 153-193.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22. doi:10.1037/h0046743
- Chen, Z., & Siegler, R. S. (2000). Intellectual development in childhood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 92-116). Cambridge, UK: Cambridge University Press.
- Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 3-17). Hillsdale, NJ: Lawrence Erlbaum.
- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213-228.
- Davidson, J. E., & Downing, C. L. (2000). Contemporary models of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 34-52). New York: Cambridge University Press.
- Geerlings, H., Laros, J. A., Tellegen, P. J., & Glas, C. A. W. (2013). Testing the difficulty theory of the SON-R 5½-17, a non-verbal test of intelligence. *British Journal of Mathematical and Statistical Psychology*, 67(2), 248-265. doi:10.1111/bmsp.12017
- Holt, J. K. (2004). *Item parceling in Structural Equation Modeling for optimum solutions*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, Ohio, USA.
- Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 443-451). New York: Macmillan.
- Horn, J. L., & Blankson, A. N. (2012). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3<sup>rd</sup> ed., pp. 73-98). New York: The Guilford Press.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57(5), 253-270. doi:10.1037/h0023816
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319-342. doi:10.1111/j.1745-3984.2001.tb01130.x
- Karino, C. A., Laros, J. A., & Jesus, G. R. (2011). Evidências de validade convergente do SON-R 2 ½ - 7 [a] com o WPPSI-III e WISC-III. *Psicologia: Reflexão e Crítica*, 24(4), 621-629. doi:10.1590/S0102-79722011000400001
- Laros, J. A., Tellegen, P. J., Jesus, G. R., & Karino, C. A. (2015). *SON-R 2½-7[a]. Manual - Teste não-verbal de inteligência*. São Paulo, SP: Hogrefe.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9(2), 151-173. doi:10.1207/S15328007SEM0902\_1
- Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, 38(1), 88-91.

- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (2<sup>nd</sup> ed., pp. 136-182). New York: The Guilford Press.
- McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR) - Gc-Gf cross battery assessment*. Boston, MA: Allyn and Bacon.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the past 20 years of research [Special issue]. *Psychology in the Schools*, 47(7), 651-675. doi:10.1002/pits.20497
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77-101.
- 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. Retrieved from <http://pareonline.net>
- Primi, R., & Almeida, L. S. (2000). Estudo de validação da Bateria de Provas de Raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa*, 16(2), 165-173. doi:10.1590/S0102-37722000000200009
- Primi, R., Silva, M. C. R., Santana, P. R., Muniz, M., & Almeida, L. S. (2013). The use of the bi-factor model to test the uni-dimensionality of a battery of reasoning tests. *Psicothema*, 25(1), 115-122. doi:10.7334/psicothema2011.393
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12(3), 287-297. doi:10.1037//1040-3590.12.3.287
- Rocha, C. M., & Chelladurai, P. (2012). Item parcels in structural equation modeling: An applied study in sport management. *International Journal of Psychology and Behavioral Sciences*, 2(1), 46-53. doi:10.5923/j.ijpbs.20120201.07
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, Tests, and Issues* (3<sup>rd</sup> ed., pp. 99-144). New York: The Guilford Press.
- Sijtsma, K. (2009). On the use, misuse and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107-120. doi:10.1007/s11336-008-9101-0
- Sijtsma, K. (2012). Future of psychometrics: Ask what psychometrics can do for psychology. *Psychometrika*, 77(1), 4-20. doi:10.1007/S11336-011-9242-4
- Sireci, S. G., & Parker, P. (2006). Validity on trial: Psychometric and legal conceptualizations of validity. *Educational Measurement: Issues and Practice*, 25(3), 27-34. doi:10.1111/j.1745-3992.2006.00065.x
- Snijders, J. T., Tellegen, P. J., & Laros, J. A. (1989). *Snijders-Oomen Non-verbal intelligence test. SON-R 5½-17. Manual and research report*. Groningen, Holland: Wolters-Noordhoff.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen* [The examination of intelligence with deaf-mute children]. Nijmegen, Holland: Berkhouwt.
- Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York, NY: Macmillan.
- Tellegen, P. J., & Laros, J. A. (2011). *SON-R 6-40: Snijders-Oomen Niet-verbale intelligentietest. Verantwoording, instructies & normen* [SON-R 6-40: Snijders-Oomen Nonverbal Intelligence Test. Research report, instructions & norms]. Göttingen, Germany: Hogrefe.
- Tellegen, P. J., & Laros, J. A. (2014). *SON-R 6-40. Snijders-Oomen Non-verbal intelligence test. Research report, instructions & norms*. Göttingen, Germany: Hogrefe.
- Tellegen, P. J., Laros, J. A., & Petermann, F. (2012). *Non-verbaler intelligentztest. Technisches Manual mit deutscher normierung und validierungsstudien*. Göttingen, Germany: Hogrefe.
- Ten Berge, J. M. F., & Kiers, H. A. L. (1991). A numerical approach to the exact and the approximate minimum rank of a covariance matrix. *Psychometrika*, 56(2), 309-315. doi:10.1007/BF02294464
- Ten Berge, J. M. F., & Zegers, F. E. (1978). A serie of lower bounds to the reliability. *Psychometrika*, 43(4), 575-579. doi:10.1007/BF02293815
- Thompson, B. (2003). Understanding reliability and coefficient alpha, really. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 1-24). Thousand Oaks, CA: Sage.

- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, CA: University of Chicago Press.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209-220. doi:10.1037/a0023353
- Valentini, F., Laros, J. A., & Jesus, G. R. (2010). *Invariância da estrutura fatorial do teste de inteligência SON-R 2½7[a] por meio da análise fatorial confirmatória*. Trabalho apresentado no IX Encontro Mineiro de Avaliação Psicológica, Belo Horizonte, MG, Brasil.
- Zumbo, B. D. (2006). Validity: Foundational issues and statistical methodology. In C. S. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26* (pp. 45-80). Amsterdam, Netherlands: Elsevier B.V.

Recebido: 27/03/2014

1<sup>a</sup> revisão: 21/08/2014

Aceite final: 30/10/2014