



SON-R 6-40: Dimensionalidade e Validade Convergente com a CMMS, TONI-3 e GPM

Talita Alves¹

Centro Universitário Unieuro, Brasília-DF, Brasil

Jacob Laros

Universidade de Brasília, Brasília-DF, Brasil

RESUMO

O objetivo deste estudo foi avaliar a estrutura fatorial do teste não verbal de inteligência SON-R 6-40 e obter evidências de validade convergente deste com outros três instrumentos não verbais de inteligência: Matrizes Progressivas Coloridas de Raven, Escala de Maturidade Mental Colúmbia e TONI-3. Participaram do estudo 150 crianças de escolas públicas e privadas do Distrito Federal com idades entre 6 e 9 anos ($M=7,7$; $DP=1,0$). As análises fatoriais exploratórias indicaram uma estrutura com dois fatores para o SON-R 6-40 e uma estrutura unifatorial para os outros testes. Os dois fatores apresentaram uma correlação de 0,64 que explicam 68,3% da variância comum. Os coeficientes de correlação entre o SON-R 6-40 e os demais instrumentos variaram entre 0,62 e 0,73, com valor médio de 0,69. Os resultados indicam evidências adequadas da validade convergente do SON-R 6-40 com outros três testes não verbais de inteligência frequentemente utilizados no Brasil.

Palavras-chave: validade convergente; teste não verbal de inteligência; propriedades psicométricas; SON-R 6-40.

ABSTRACT – SON-R 6-40: Dimensionality and Convergent Validity with CMMS, TONI-3 and GPM

The aim of the study was to assess the factor structure of the non-verbal intelligence test SON-R 6-40 and obtain evidence of its convergent validity with three other non-verbal intelligence tests: Raven's Colored Progressive Matrices, Columbia Mental Maturity Scale and TONI-3. Study participants were 150 children aged 6 to 9 years ($M=7.7$, $SD=1.0$) from public and private schools of the Federal District. Exploratory factor analysis indicated a two-factor structure for the SON-R 6-40 and a unidimensional structure for the three other tests. The two factors showed a correlation of .64 and together explained 68.3% of the common variance. The correlation coefficients between the SON-R 6-40 and the other instruments ranged between .62 and .73, with a mean value of .69. The results provide evidence of adequate convergent validity for the SON-R 6-40 with three other non-verbal intelligence tests frequently used in Brazil.

Keywords: convergent validity; non-verbal intelligence test; psychometric properties; SON-R 6-40.

RESUMEN – SON-R 6-40: Dimensionalidad y Validez Convergente con CMMS, TONI-3 y GPM

El objetivo del estudio fue evaluar la estructura factorial del test de inteligencia no verbal SON-R 6-40 y obtener evidencias de su validez convergente con otros tres instrumentos de inteligencia no verbal: el Test de Matrices Progresivas de Color de Raven, la Escala de Madurez Mental de Columbia y el TONI-3. En el estudio participaron 150 niños de 6 a 9 años de escuelas públicas y privadas del Distrito Federal ($M=7.7$, $DS=1.0$). El Análisis Factorial Exploratorio indicó una estructura de dos factores para el SON-R 6-40, y una estructura unifactorial para los otros tres tests. La correlación entre los dos factores fue de 0,64 y, lo que explica el 68.3% de la varianza común. Los coeficientes de correlación entre el SON-R 6-40 y los otros instrumentos fluctuaron entre 0,62 y 0,73 con una media de 0,69. Los resultados muestran evidencias adecuadas de la validez convergente del test SON-R 6-40 con los otros tres tests de inteligencia no verbal de uso frecuente en Brasil.

Palabras clave: validez convergente; test de inteligencia no verbal; propiedades psicométricas; SON-R 6-40.

A validade é um aspecto imprescindível para determinar a utilidade e a qualidade dos escores obtidos em testes psicológicos. Os estudos sobre construção, adaptação e uso dos testes permitem atestar o rigor técnico e científico dessas medidas e a confiabilidade dos resultados obtidos por meio da testagem (Urbina, 2014). A área de avaliação psicológica no Brasil dispõe de um número

limitado de instrumentos para uso profissional. Esse campo também se caracteriza pela escassez de estudos sobre a validade e as propriedades psicométricas dos testes em diferentes amostras nacionais. Desse modo, pesquisas sobre novos instrumentos, e sobre aqueles que se encontram disponíveis, são importantes fontes de informação para o desenvolvimento de conhecimento científico

¹ Endereço para correspondência: Universidade de Brasília, Instituto de Psicologia – IP, Instituto Central de Ciências (ICC)/Ala Sul, sala A1-061/4, Laboratório de Métodos e Técnicas de Avaliação – META, 70910-900, Brasília, DF. E-mail: talitaalves.psi@gmail.com
Artigo derivado da dissertação de mestrado da autora Talita de Araújo Alves com orientação de Jacob Arie Laros, defendida em 2016 no programa de pós-graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília.

sobre as variáveis em psicologia e para o aprimoramento das práticas em avaliação psicológica.

O requisito da validade é um aspecto central para os testes psicológicos (*American Educational Research Association [AERA] et al., 2014*). É com base na validade dos testes que os escores são considerados adequados e representativos do construto que se objetiva mensurar (Pacico & Hutz, 2015). Em outras palavras, as inferências realizadas com base nos resultados dos testes estão diretamente associadas às evidências de validade, pois estas fornecem subsídios à interpretação e utilidade dos escores em testes psicológicos.

Algumas questões fundamentais sobre a validade dos testes são encontradas nos trabalhos de autores importantes da psicometria, como em Cronbach e Meehl (1955), onde se discute que a validade de construto dos testes deve receber maior destaque, e que o processo de validação de um instrumento envolve critérios empíricos e teóricos, a fim de aprimorar a construção do conhecimento sobre os construtos. Para além das evidências empíricas e teóricas, Messick (1989) observa que a validade dos testes está diretamente associada à utilização dos escores, o que conduz os estudos da validade às consequências da testagem na vida dos sujeitos. As perspectivas contemporâneas de validade retomam discussões fundamentais sobre a mensuração e a utilização de métodos para a construção e validação dos testes, destacando a relação causal entre as variáveis psicológicas e os itens dos testes (Borsboom et al., 2004). Apesar das divergências, a ideia comum a todas as propostas é de que o processo de validação de um teste precisa assegurar que a medida é uma boa representação do construto psicológico que se busca avaliar.

A psicometria utiliza os principais critérios de levantamento de informações sobre a validade dos testes conforme adotados pelo atual *Standards for Educational and Psychological Testing*. O documento se refere à validade como um conjunto de evidências que subsidiam a interpretação dos escores obtidos por meio da testagem (AERA et al., 2014). Com base nesse documento, as evidências de validade são categorizadas de acordo com as diferentes fontes de informações, assim o delineamento contemporâneo sobre a validade é dividido da seguinte forma: (a) validade de conteúdo; (b) relações com outras variáveis (validade convergente e discriminante, validade teste-critério); (c) processos de resposta, (d) estrutura interna e (e) consequências da testagem.

O processo de validação de um instrumento é cumulativo. As pesquisas sobre a validade dos testes são fontes ilimitadas de informações para atestar a representatividade da medida. Nesse sentido, é fundamental que os testes apresentem evidências de validade provenientes de diferentes fontes de dados. O acúmulo de evidências de validade é indispensável, sobretudo para testes psicológicos que são novos em determinada população. Um exemplo disso é a pesquisa de validação e normatização

para a população brasileira do SON-R 6-40, um teste não verbal de inteligência, a qual foi concluída em 2019.

A primeira versão dos testes SON (Snijders-Oomen Não Verbal) foi desenvolvida na Holanda em 1943. Desde então diversas versões foram construídas para abranger pessoas em diferentes faixas etárias (Laros et al., 2015). O teste SON-R 6-40 é um teste não verbal de inteligência, o qual pode ser aplicado com ou sem o uso da fala, tornando o instrumento adequado para grupos especiais (Tellegen & Laros, 2014). Do ponto de vista teórico, o teste SON busca medir a inteligência fluida, que compõe o modelo CHC (Cattell-Horn-Carroll), e representa o raciocínio frente a problemas novos, cuja solução independe de conhecimento prévio, sendo o aspecto que mais se aproxima de um fator geral de inteligência (Flanagan & Harrison, 2012).

No Brasil, encontra-se disponível para uso a versão dos testes SON denominada SON-R 2½-7[a], voltada para crianças com idades entre 2 anos e 6 meses até 7 anos e 11 meses (Laros et al., 2015). A versão SON-R 6-40 apresenta maior abrangência etária, pois o instrumento foi desenvolvido para o público dos 6 aos 40 anos de idade. No que diz respeito aos estudos de validade, a versão SON-R 2½-7[a] incluiu estudos brasileiros sobre a validade convergente desse instrumento com outros testes de inteligência utilizados no Brasil. Dentre os estudos, destacam-se os estudos do SON-R 2½-7[a] com as Matrizes Progressivas Coloridas de Raven (CPM) e as Escalas de Maturidade Mental Colúmbia (CMMS) (Laros et al., 2010), além dos testes WISC-III e WPPSI-III (Karino et al., 2011).

Pesquisas sobre a validade do SON-R 6-40 incluem estudos internacionais que analisam as correlações desse com os testes: *Wechsler Intelligence Scale for Children – WISC III/IV* ($r=0,80$), *Wechsler Adult Intelligence Scale – WAIS-III* ($r=0,83$), *Wechsler Non-Verbal Scale of Ability, Dutch adaptation – WNV-NL* ($r=0,77$) e *Dutch Intelligence Test for Educational Level – NIO* ($r=0,79$) (Tellegen & Laros, 2014). O manual holandês do SON-R 6-40 também inclui análises para verificar a estrutura fatorial do instrumento e a adequação do teste para grupos especiais (Tellegen & Laros, 2014). No Brasil, foram analisadas as correlações entre o SON-R 6-40 e a Bateria de Provas de Raciocínio, forma A (BPR-5) e o SON-R 2½-7[a] (Laros et al., 2015). Entretanto, até o momento não são identificados estudos que analisam a relação entre o SON-R 6-40 e demais instrumentos disponíveis no contexto brasileiro para a avaliação da inteligência.

Nesse contexto, o presente estudo objetivou determinar a estrutura fatorial do SON-R 6-40 e reunir evidências da validade convergente do SON-R 6-40 com três outros testes não verbais de inteligência: Matrizes Progressivas Coloridas de Raven (CPM), a Escala de Maturidade Mental Colúmbia (CMMS) e o TONI-3, forma A (*Test of Nonverbal Intelligence, Form A, Third Edition*). Para poder obter evidências de validade

convergente do SON-R 6-40, a determinação da dimensionalidade dos outros três instrumentos também foi necessária, assim como o cálculo dos coeficientes de fidedignidade dos escores obtidos nas quatro medidas investigadas.

O teste das Matrizes Progressivas de Raven é utilizado em diversos países há mais de 50 anos. O instrumento apresenta três escalas: *Standard*, coloridas e avançadas (Angelini et al., 1999), que variam de acordo com o público-alvo. O Raven avalia o fator geral de inteligência e, de acordo com os autores do manual, este apresenta dois componentes: a capacidade educativa e de reprodução (Angelini et al., 1999). A educação representa a formação de novos significados com base em situações não estruturadas, e a reprodução está associada à memória, indicando a capacidade de replicar conteúdos com base cultural de conhecimentos.

A Escala de Maturidade Mental Colúmbia – CMMS foi originalmente publicada em 1954, diante da necessidade de se obter estimativas adequadas da habilidade intelectual (capacidade geral de raciocínio) de crianças com condições especiais (Burgemeister et al., 2001). Esse mesmo público motivou a construção do teste TONI, em 1982 (Brown et al., 1982). A proposta do teste TONI é avaliar a inteligência em amostras cujo desempenho é afetado por limitações linguísticas e/ou motoras. O TONI-3 é a terceira versão do teste, que foi publicada em 1997, apresentando duas formas paralelas: A e B. O teste TONI-3 objetiva avaliar a inteligência geral, com foco no raciocínio abstrato e na resolução de problemas. O desenvolvimento do TONI não adotou uma teoria particular sobre a inteligência. Segundo o manual (Brown et al., 2006), o teste é capaz de se adaptar frente aos importantes modelos teóricos da inteligência, tais como: o fator *g* de Spearman, a teoria das habilidades mentais primárias de Thurstone e a teoria de Cattell e Horn.

Método

Participantes

O estudo analisou dados de uma amostra composta por 150 estudantes (54,7 % do sexo feminino). A idade dos participantes variou entre seis e nove anos ($M=7,7$; $DP=1,0$). No momento da coleta, os estudantes cursavam do 1º ao 5º ano do Ensino Fundamental em quatro escolas do Distrito Federal, sendo três escolas da rede pública de ensino (74,7% da amostra) e uma escola da rede privada.

Instrumentos

Todos os participantes responderam aos quatro instrumentos de avaliação da inteligência: as Matrizes Progressivas Coloridas de Raven – CPM (Angelini et al., 1999), a Escala de Maturidade Mental Colúmbia – CMMS (Burgemeister et al., 2001), o TONI-3, forma A (Brown et al., 1982) e o SON-R 6-40 (Tellegen &

Laros, 2014). Todos os instrumentos apresentam tarefas não verbais, cuja resolução independe das habilidades de fala e escrita dos sujeitos. No período da coleta, todos os testes apresentavam parecer favorável para uso de acordo com o Conselho Federal de Psicologia – CFP.

O primeiro teste, as Matrizes Progressivas Coloridas de Raven – CPM, destina-se para crianças de 5 até 11 anos de idade. A escala é dividida em três séries: A, Ab e B, com 12 itens cada. O escore total da CPM apresentou coeficientes de fidedignidade variando de 0,52 a 0,93 para as faixas etárias pesquisadas da amostra de normatização brasileira (Angelini et al., 1999)

O segundo instrumento, a Escala de Maturidade Mental Colúmbia – CMMS (Burgemeister et al., 2001) é um teste de aplicação individual destinado para crianças com idades entre 3 anos e 6 meses e 9 anos e 11 meses. O teste é composto por 92 itens de classificação figurativa, que estão dispostos em oito níveis sobrepostos. Estudos de fidedignidade dos escores totais na CMMS indicaram coeficientes que variam de 0,82 a 0,93 (Alves & Duarte, 2001).

O teste TONI-3, forma A, é um teste de aplicação individual destinado para crianças na faixa etária de 6 até 10 anos. O teste consiste de 45 itens, organizados de acordo com o nível de dificuldade. Na pesquisa de normatização brasileira, foram encontrados coeficientes de fidedignidade (alfa de Cronbach) variando de 0,80 até 0,83 (Santos et al., 2006).

O SON-R 6-40 é destinado para pessoas na faixa etária de 6 até 40 anos de idade (Tellegen & Laros, 2014). No total, o SON-R 6-40 possui 124 itens divididos em quatro subtestes: Analogias e Categorias (subtestes de múltipla escolha); e Mosaicos e Padrões (subtestes de execução). Na pesquisa de validação e normatização na Holanda e Alemanha, os escores nos subtestes do SON-R 6-40 apresentam um coeficiente médio de fidedignidade de 0,87 e, para o escore total normatizado (QI-SON), foi encontrado um valor médio de 0,95 (Tellegen & Laros, 2014). A pesquisa brasileira de validação e normatização do SON-R 6-40 possui aprovação junto ao Comitê de Ética em Pesquisa do Instituto de Ciências Humanas da Universidade de Brasília – CEP/IIH – UNB desde 2011 e foi concluída no ano de 2019.

Procedimentos

Os alunos foram selecionados para a participação no estudo a partir do critério da idade. As aplicações ocorreram nas escolas, após o consentimento dos pais e/ou responsáveis. A aplicação dos instrumentos foi realizada em duas sessões com cada participante, com tempo médio de duração de 30 minutos para cada uma delas. Foi adotado um procedimento de alternância na ordem de aplicação dos testes, a fim de evitar vieses no processo de testagem (aprendizagem e fadiga). Dessa forma, metade dos estudantes iniciou respondendo ao SON-R 6-40 e depois os demais testes: CMMS, CPM e TONI-3, e o

procedimento inverso foi realizado com a metade restante da amostra.

Análise dos Dados

As análises exploratórias e descritivas foram viabilizadas pelo pacote estatístico SPSS. Os escores brutos dos quatro testes (SON-R 6-40, TONI-3, CMMS e CPM) foram transformados em escores normatizados (escores de $QI - \text{desvio}$), com média 100 e desvio padrão de 15.

A análise fatorial exploratória foi realizada com o programa FACTOR, versão 10.10.03 (Baglin, 2014; Lorenzo-Seva & Ferrando, 2013), e utilizou parcelas de itens para os testes. A construção de parcelas de itens foi adotada a fim de evitar o surgimento de fatores artificiais, relacionados à dificuldade dos itens dicotômicos (De Ayala, 2009; Little et al., 2002; Holt, 2004; Rocha & Chelladurai, 2012; Tran & Formann, 2009; Weng & Cheng, 2005). As análises não incluíram itens extremamente difíceis ($p < 0,10$) e itens extremamente fáceis ($p > 0,90$). Assim, cada um dos testes resultou na seguinte quantidade de parcelas: SON-R 6-40 (18 parcelas de itens), CMMS (17 parcelas de itens), CPM (nove parcelas de itens) e TONI-3 (cinco parcelas de itens).

O procedimento utilizado para determinar o número de fatores a serem retidos foi a Análise Paralela de Horn (Horn, 1965) conforme implementada no

programa FACTOR 10.10.03. Para os quatro instrumentos utilizados, o método de extração de fatores foi o *Minimum Rank Factor Analysis* (MRFA) (Ten Berge & Kiers, 1991; Timmerman & Lorenzo-Seva, 2011), tendo por base a matriz de correlação policórica (Muthén & Kaplan, 1992) e a matriz de correlação de Pearson em caso de falta de convergência. A fidedignidade dos escores dos quatro testes foi estimada por meio do cálculo do alfa de Cronbach.

Para as análises de validade convergente entre os escores normatizados do SON-R 6-40 e os escores normatizados nos demais testes, utilizou-se o coeficiente de correlação bivariada de Pearson (r). A correlação bruta entre os escores normatizados foi corrigida por meio da fórmula de correção dupla para atenuação (Osborne, 2003; Thompson, 2003).

Resultados

Inicialmente foi conduzida uma análise fatorial exploratória com o programa FACTOR a fim de verificar a estrutura fatorial dos quatro instrumentos: SON-R 6-40, CMMS, CPM e TONI-3. Posteriormente, foi analisada a estrutura fatorial de todos os testes em conjunto. A Tabela 1 apresenta os resultados da análise fatorial exploratória para o SON-R 6-40.

Tabela 1

Média e Amplitude das Cargas Fatoriais e Comunalidade da Solução Fatorial do SON-R 6-40

Subteste	Carga Fatorial Fator 1		Carga Fatorial Fator 2		h^2
	Média	Amplitude	Média	Amplitude	
Mosaicos	0,68	0,61 – 0,74	0,11	0,08 – 0,15	0,77
Padrões	0,83	0,77 – 0,90	-0,04	(-0,07) – (-0,03)	0,84
Analogias	0,14	0,06 – 0,21	0,52	0,39 – 0,63	0,63
Categorias	-0,06	(-0,17) – 0,06	0,69	0,55 – 0,79	0,70

Nota. h^2 =comunalidade; método de extração de fatores: *Minimum Rank Factor Analysis* (MRFA); método de rotação: Promin; percentagem da variância comum explicada=68,3%; correlação entre os fatores=0,64

O índice de Kaiser- Meyer- Olkin (KMO) apresentou um valor considerado muito bom (0,90), indicando que a matriz de correlações é passível de fatoração. Na Análise Paralela de Horn, os *eigenvalues* empíricos e aleatórios (percentil 95) apresentaram os seguintes valores, (a) empíricos: 7,7; 1,7 e 1,2; (b) aleatórios: 1,8; 1,6 e 1,5. Esses resultados indicam uma solução de dois fatores para o SON-R 6-40, explicando 68,3% da variância comum. Os subtestes de execução Mosaicos e Padrões, que

medem raciocínio espacial, têm cargas altas no primeiro fator, enquanto os subtestes Analogias e Categorias, que medem raciocínio abstrato, mostram cargas altas no segundo fator. O primeiro fator será identificado como a Escala Espacial (SON-EE) e o segundo fator como a Escala de Raciocínio (SON-ER).

A seguir, os resultados da análise fatorial exploratória para os testes CMMS, CPM e TONI-3 são resumidos na Tabela 2.

Tabela 2

Média e Amplitude das Cargas Fatoriais e Percentagem da Variância Comum Explicada da Solução Fatorial do CMMS, CPM e TONI-3

Teste	Carga Fatorial		VCE
	Média	Amplitude	
CMMS	0,52	0,23 – 0,71	40,1%
CPM	0,67	0,51 – 0,78	68,9%
TONI-3	0,82	0,62 – 0,94	86,4%

Nota. VCEv=percentagem da variância comum explicada; método de extração de fatores=Minimum Rank Factor Analysis (MRFA)

O método da Análise Paralela de Horn indicou uma estrutura unifatorial para os testes CMMS, CPM e TONI-3. Para os três instrumentos, os índices de KMO variaram de 0,72 até 0,86. Em média, as cargas fatoriais para os três instrumentos apresentaram valores satisfatórios. Entretanto, os dados da Tabela 2 permitem verificar que a Escala de Maturidade Mental Colúmbia (CMMS) apresentou valores insatisfatórios das cargas fatoriais (<0,30). A média das cargas fatoriais do CMMS foi a mais baixa, enquanto a média das cargas fatoriais do TONI-3 foi a mais alta. O teste que apresentou a maior proporção de variância comum explicada foi o TONI-3 (86,4%), seguido pelos testes CPM (68,9%) e CMMS (40,1%).

A análise fatorial conjunta dos quatro instrumentos resultou em uma solução com um fator geral de

segunda ordem e dois fatores de primeira ordem. Para tanto, foi utilizada a solução Schmid-Leiman, proposta por Thompson (2005). A solução Schmid-Leiman é um método elegante para expressar os fatores de primeira e segunda ordem em termos das variáveis mensuradas. A remoção de toda a variância nos fatores de primeira ordem, e que também está presente nos fatores de segunda ordem, permite ao pesquisador determinar qual parte da variância é exclusiva para um determinado nível de análise (Thompson, 2005). O procedimento Schmid-Leiman é o modelo *bifactor* exploratório mais frequentemente usado (Canivez, 2016).

Os resultados do procedimento Schmid-Leiman para os quatro instrumentos avaliados (o SON-R 6-40 formado por duas escalas, o TONI-3, CPM e CMMS) encontram-se na Tabela 3.

Tabela 3

Média de Cargas Fatoriais no Fator Geral e nos Fatores de Primeira Ordem Usando o Procedimento Schmid-Leiman

Escala	Segunda ordem		Primeira ordem	
	Fator Geral		Fator 1	Fator 2
	CF Média		CF Média	CF Média
SON-EE	0,64		0,30	0,06
SON-ER	0,56		-0,09	0,24
TONI-3	0,57		0,17	0,10
CPM	0,49		0,27	0,03
CMMS	0,38		0,23	-0,03
Média	0,53		0,21	0,09

Nota. CF Média=a média das cargas fatoriais; SON-EE=Escala Espacial do SON-R 6-40; SON-ER=Escala de Raciocínio do SON-R 6-40; os dois fatores de primeira ordem explicaram 42,0% da variância comum

A partir das informações da Tabela 3, observa-se que as cargas fatoriais dos fatores de primeira ordem são bem inferiores às cargas fatoriais no fator de segunda ordem. Isso ocorre devido à remoção da variância nos fatores de primeira ordem explicada pelo fator de segunda ordem. Os resultados da Tabela 3 mostram que a contribuição do CMMS no Fator Geral é a mais baixa e a contribuição do SON-EE no Fator Geral é a mais alta. No Fator 1, o SON-EE e o CPM mostram as cargas fatoriais mais

altas. No Fator 2, o SON-ER obteve a carga fatorial mais alta, e o CPM e o CMMS a carga fatorial mais baixa. Dos quatro testes avaliados, o SON-R 6-40 é o que parece mensurar melhor o raciocínio abstrato.

A Tabela 4 apresenta os escores normatizados nos quatro instrumentos. Os escores dos subtestes do SON-R 6-40 estão na escala de subtestes Wechsler ($M=10$, $DP=3$), e os escores totais dos quatro instrumentos estão na escala de QI ($M=100$, $DP=15$).

Tabela 4

Média, Desvio Padrão, EP e IC 95% das Médias dos Escores Normalizados nos Quatro Instrumentos

Subtestes SON-R 6-40	Média	DP	EP	IC 95%
Analogias	10,53	2,74	0,22	10,10 – 10,96
Mosaicos	10,19	2,66	0,22	9,76 – 10,62
Categorias	10,77	3,19	0,26	10,26 – 11,28
Padrões	9,99	2,96	0,24	9,52 – 10,46
Testes	Média	DP	EP	IC 95%
SON-R 6-40	102,3	12,99	1,06	100,2 - 104,4
CPM	106,5	13,11	1,07	104,4 - 108,6
CMMS	105,9	12,33	1,01	103,9 – 107,9
TONI-3	87,3	15,83	1,29	84,8 - 89,8

Nota. DP=desvio-padrão; EP=erro padrão da média; IC 95%=intervalo de confiança de 95% da média.

Os valores da Tabela 4 indicam que os escores normalizados mais altos foram obtidos nos testes CPM e CMMS (106,5 e 105,9 respectivamente). Os intervalos de confiança de 95% para os dois testes se sobrepõem, indicando que não existem diferenças significativas entre as médias para os dois instrumentos. O instrumento que apresentou o menor valor para a média, e maior desvio padrão e erro padrão da média foi o TONI-3. Na pesquisa de normalização brasileira do TONI-3, a média dos escores brutos totais foi de 16,57 e desvio padrão de 6,32 (Santos et al., 2006). O valor da média para os escores brutos totais no presente estudo foi de 12,11, com desvio padrão de 5,3. O cálculo do tamanho de efeito da diferença das médias apresentou um d de Cohen de 0,76, sendo um efeito forte (Vacha-Haase & Thompson, 2004). A diferença substancial entre as médias nas duas pesquisas indica que, na pesquisa de

normalização do TONI-3, o nível cognitivo da amostra foi bem mais alto em comparação com o nível cognitivo dos participantes no presente estudo. Isso leva a questionar a representatividade da amostra de normalização do TONI-3. O manual brasileiro do teste não apresenta a descrição detalhada da amostra, o que limita a interpretação das diferenças entre a média dos escores do TONI-3 e as médias dos demais instrumentos (Brown et al., 2006).

Foram calculados os coeficientes de fidedignidade dos escores nos subtestes do SON-R 6-40 e dos escores totais normalizados dos quatro testes. De acordo com Hogan (2006), os coeficientes de fidedignidade obtidos por meio da análise da consistência interna dos itens podem ser considerados como pré-requisitos para a validade dos testes. Os resultados encontram-se na Tabela 5, a seguir.

Tabela 5

Propriedades Psicométricas dos Subtestes do SON-R 6-40 e dos Escores Totais dos Quatro Testes Avaliados

Propriedades psicométricas	Subtestes do SON-R 6-40			
	Analogias	Mosaicos	Categorias	Padrões
consistência interna (α)	0,85	0,87	0,85	0,89
correlação item-total (r_{it})	0,40	0,40	0,39	0,52
correlação entre os itens	0,18	0,19	0,17	0,23
proporção de acertos	0,47	0,43	0,41	0,41
número de itens	36	26	36	26
Propriedades psicométricos	Escore totais normalizados			
	SON-R 6-40	CMMS	TONI-3	CPM
consistência interna (α)	0,94	0,85	0,90	0,84
correlação item-total (r_{it})	0,37	0,25	0,45	0,31
correlação entre os itens	0,14	0,08	0,24	0,14
proporção de acertos	0,43	0,58	0,35	0,60
número de itens	124	92	45	36

Os coeficientes de fidedignidade (alfa de Cronbach) dos escores dos subtestes do SON-R 6-40 variaram entre 0,85 e 0,89. O subteste que apresentou o maior coeficiente de fidedignidade foi Padrões (0,89). O teste com o maior coeficiente de fidedignidade do escore total foi o SON-R 6-40 (0,94), seguido do TONI-3 (0,90). Os coeficientes menores foram observados para o teste CPM (0,84) e CMMS (0,85). Em relação à proporção de acertos, o teste TONI-3 apresentou o menor valor (35%), indicando que, para a faixa etária

investigada, este foi o teste mais difícil. A proporção maior de acertos foi identificada para os testes CPM (60%) e CMMS (58%).

Para obter os coeficientes de validade convergente do SON-R 6-40 com os demais instrumentos, foram calculadas as correlações de Pearson entre o escore normatizado do SON-R 6-40 e os escores normatizados dos outros testes. As correlações brutas e corrigidas por meio da fórmula de correção dupla para atenuação encontram-se na Tabela 6.

Tabela 6

Coefficientes de Correlação de Pearson e Coeficientes de Determinação para os Escores Totais Normatizados

	Coeficientes de correlação (r_s) e coeficientes de determinação (r_s^2)			
	SON-R 6-40	CPM	CMMS	TONI-3
SON-R 6-40	-	0,52 (0,42)	0,45 (0,37)	0,37 (0,32)
CPM	0,73 (0,65)	-	0,61 (0,44)	0,28 (0,24)
CMMS	0,72 (0,64)	0,78 (0,66)	-	0,32 (0,26)
TONI-3	0,62 (0,57)	0,53 (0,49)	0,57 (0,51)	-

Nota. No triângulo inferior, são apresentados os coeficientes de correlação e, no triângulo superior, encontram-se os coeficientes de correlação elevados ao quadrado; os valores que aparecem inicialmente são os coeficientes corrigidos e entre parênteses estão os coeficientes brutos

A inspeção do triângulo inferior da Tabela 6 permite concluir que todas as correlações corrigidas do SON-R 6-40 com os outros testes são altas, entre 0,62 e 0,73. O coeficiente de validade convergente mais alto pôde ser observado entre o SON-R 6-40 e o teste CPM ($r=0,73$), indicando que os dois testes têm 73% da variância em comum. A correlação entre os dois testes (CPM e SON-R 6-40) é semelhante aos resultados

encontrados por Laros et al. (2010), onde foi encontrada uma correlação corrigida alta ($r=0,77$) entre o SON-R 2½-7[a] e o teste CPM.

Por último, foram analisadas as correlações entre os testes CPM, CMMS e TONI-3 e as duas Escalas do SON-R 6-40, a Escala Espacial (SON-EE) e a Escala de Raciocínio (SON-ER). Os resultados encontram-se na Tabela 7.

Tabela 7

Correlações entre as Duas Escalas do SON-R 6-40 e os Testes CPM, CMMS e TONI-3

Testes	SON-EE	IC 95%	SON-ER	IC 95%
MPC	0,71 (0,63)	(0,52 – 0,72)	0,58 (0,51)	(0,37 – 0,63)
EMMC	0,65 (0,59)	(0,47 – 0,69)	0,59 (0,53)	(0,41 – 0,63)
TONI-3	0,58 (0,54)	(0,39 – 0,65)	0,50 (0,46)	(0,30 – 0,60)

Nota. SON-EE=Escala Espacial; SON-ER=Escala de Raciocínio; IC 95%=Intervalo de confiança de 95% das correlações brutas entre os escores normatizados. Os valores em parênteses se referem às correlações brutas

De acordo com as informações da Tabela 7, verifica-se que a correlação mais alta foi encontrada entre a Escala Espacial do SON-R 6-40 e o teste CPM ($r=0,71$). Esse achado é semelhante aos resultados obtidos por Laros et al. (2010), onde foi encontrada uma correlação de 0,68 entre a Escala de Execução do SON-R 2½-7[a] e o teste CPM. Os testes CMMS e o TONI-3 apresentaram maior correlação com a Escala Espacial do SON-R 6-40 (SON-EE) do que com a Escala de Raciocínio (SON-ER). Esses resultados estão em concordância com

os obtidos com a análise fatorial, usando o procedimento Schmid-Leiman que indicam que a Escala de Raciocínio do SON-R 6-40 está medindo um aspecto da inteligência que os três outros testes não medem.

Discussão

A análise fatorial exploratória do SON-R 6-40 indicou a existência de dois fatores para a amostra do presente estudo, formada por um grupo de crianças com

idades entre 6 e 9 anos. Os subtestes Mosaicos e Padrões formam o primeiro fator, que apresentou cargas fatoriais altas, com valor médio de 0,76. O primeiro fator pode ser interpretado como um fator espacial, denominado SON-EE, e envolve a habilidade de construção visuoespacial, com tarefas que demandam a manipulação ativa de figuras e padrões de desenho. O segundo fator é composto pelos subtestes Analogias e Categorias, que apresentaram cargas fatoriais altas, com valor médio de 0,61. Esse fator pode ser interpretado como um fator de raciocínio do SON-R 6-40 (SON-ER) e está associado à habilidade de raciocínio abstrato. Ambos os fatores avaliam aspectos da inteligência fluida (Gf), a qual pode ser definida como a capacidade de solucionar problemas novos (Schneider & McGrew, 2018).

A solução de dois fatores encontrada para o SON-R 6-40 é semelhante aos resultados da pesquisa de normatização brasileira da versão SON-R 2½-7[a], onde também se identificou uma solução com dois fatores para o instrumento (Laros et al., 2015). O SON-R 2½-7[a] é composto por quatro subtestes: Mosaicos, Categorias, Situações e Padrões. A solução fatorial do SON-R 2½-7[a] indicou uma Escala de Execução (SON-EE), composta pelos subtestes Mosaicos e Padrões (subtestes visuoespaciais de execução) e uma Escala de Raciocínio (SON-ER), formada pelos subtestes Situações e Categorias (subtestes de raciocínio).

Em pesquisa com o SON-R 6-40 para adolescentes com idades entre 11 e 16 anos, Laros, Almeida, et al. (2015) encontraram uma solução unifatorial para o instrumento. Na amostra de normatização holandesa e alemã, também foi encontrada uma solução unifatorial para o instrumento, considerando a faixa etária dos 6 aos 40 anos de idade (Tellegen & Laros, 2014). No manual técnico do SON-R 6-40 (Tellegen & Laros, 2014), observa-se que a estrutura fatorial de dois fatores dos testes SON tende a desaparecer com o avanço da idade dos sujeitos. Existe uma alta probabilidade de que a estrutura de dois fatores encontrada no presente estudo esteja associada à idade dos participantes, considerando que eles pertencem à faixa etária mais jovem indicada para o SON-R 6-40.

A análise fatorial hierárquica usando o procedimento Schmid-Leiman mostrou que, no fator de segunda ordem, que pode ser interpretado como o fator geral de inteligência, o SON-EE apresentou a maior carga fatorial (0,64) e o CMMS apresentou a menor. Esses resultados sugerem que, dos quatro testes não verbais investigados, o fator espacial do SON-R 6-40 está mensurando melhor o fator geral da inteligência, enquanto o CMMS é o teste mais fraco nesse aspecto. Depois da remoção da variância nos dois fatores de primeira ordem, a qual também está presente no fator de segunda ordem (o fator geral), todos os testes (com exceção da SON-ER) permanecem com cargas fatoriais positivas e significativas no primeiro fator de primeira ordem. No segundo fator de primeira ordem, entretanto, apenas o SON-ER ficou

com uma carga fatorial acima de 0,20. Os resultados obtidos com o procedimento Schmid-Leiman indicam que os testes CPM, CMMS e TONI-3 avaliam a mesma variável latente (construto) mensurada pelo fator espacial do SON-R 6-40. Os resultados de análise fatorial hierárquica também indicaram que os testes CPM, CMMS e TONI-3 não avaliam a mesma variável latente mensurada pelo fator de raciocínio do SON-R 6-40. Esses resultados corroboram o objetivo inicial do teste SON de romper a unilateralidade dos testes não verbais de inteligência da época e de tornar habilidades como o raciocínio abstrato mais acessíveis para a testagem não verbal de inteligência (Tellegen & Laros, 2014).

Considerações Finais

O estudo teve por objetivo obter informações sobre a validade convergente do teste não verbal de inteligência SON-R 6-40 com três outros testes não verbais de inteligência: as Matrizes Progressivas Coloridas de Raven - CPM, a Escala de Maturidade Mental Colúmbia - CMMS e o TONI-3, forma A (*Test of Nonverbal Intelligence, Form A*). A escolha dos instrumentos foi feita a partir dos critérios: tipo de tarefa (itens não verbais), faixa etária e relevância do teste no contexto da avaliação psicológica no Brasil. O principal resultado esperado foi um alto grau de associação entre os escores obtidos no SON-R 6-40 e os escores nos demais testes utilizados. O estudo também avaliou a estrutura fatorial dos instrumentos por meio da análise fatorial exploratória e foram obtidos coeficientes de fidedignidade dos escores totais dos quatro testes investigados.

Os resultados forneceram indícios satisfatórios da validade convergente entre o SON-R 6-40 e os instrumentos: CPM, CMMS e TONI-3. O coeficiente de validade mais alto foi encontrado entre o SON-R 6-40 e o CPM ($r=0,73$). Em relação à dimensionalidade das medidas, foi encontrada uma estrutura com dois fatores para o teste SON-R 6-40, a saber: um fator espacial (subtestes Mosaicos e Padrões) e um fator de raciocínio (subtestes Analogias e Categorias). Esse resultado divergiu da análise fatorial exploratória dos demais instrumentos, que indicou uma estrutura unifatorial para os testes CPM, CMMS e TONI-3.

A solução Schmid-Leiman demonstrou que o teste SON-R 6-40 mensurou aspectos mais amplos das habilidades cognitivas em comparação com os outros três testes investigados na amostra avaliada. Em relação aos testes CPM, CMMS e TONI-3, as correlações indicaram a relação mais forte entre esses e o fator espacial do SON-R 6-40. Esse resultado indica que os três instrumentos mensuram aspectos mais específicos, associados às habilidades visuoespaciais requeridas na solução dos itens. Os índices de fidedignidade obtidos para os escores totais normatizados nos quatro instrumentos são considerados adequados, variando entre $\alpha=0,84$ para

as Matrizes Progressivas Coloridas de Raven e $\alpha = 0,94$ para o SON-R 6-40. Esses resultados reforçam a confiabilidade dos escores obtidos nas medidas, e consequentemente certificam a qualidade técnico-científica dos instrumentos.

Para pesquisas futuras da validade convergente do SON-R 6-40, sugere-se ampliar a faixa etária dos participantes do estudo, a fim de verificar o efeito da idade sobre a estrutura fatorial e sobre as correlações entre os escores dos instrumentos. No presente estudo, destaca-se que a utilização de uma faixa etária mais restrita limitou a extensão da interpretação dos resultados, uma vez que os dados obtidos apresentaram menor variabilidade de respostas.

Destaca-se ainda a importância da realização de estudos com o teste SON-R 6-40 e fontes diversas de validade no contexto brasileiro. Sugere-se a inclusão de amostras especiais e medidas de critério a fim de subsidiar a utilização dos escores do SON-R 6-40 em diferentes contextos de avaliação da inteligência.

Agradecimentos

Não há menções.

Financiamento

Todas as fontes de financiamento para elaboração e

produção do estudo (coleta, análise e interpretação dos dados, bem como, escrita dos resultados no presente no manuscrito) foram fornecidas pelo projeto de pesquisa 'Fundo SON de pesquisa da Universidade de Groningen – Holanda.

Contribuições dos autores

Declaramos que todos os autores participaram da elaboração do manuscrito. Especificamente, o(s) autor(es) Talita de Araújo Alves participou(ram) da redação inicial do estudo – conceitualização, investigação, visualização, o(s) autor(es) Jacob Arie Laros participou(ram) da análise dos dados, da redação final do trabalho – revisão e edição.

Disponibilidade dos dados e materiais

Todos os dados e sintaxes gerados e analisados durante esta pesquisa serão tratados com total sigilo devido às exigências do Comitê de Ética em Pesquisa em Seres Humanos. Porém, o conjunto de dados e sintaxes que apoiam as conclusões deste artigo estão disponíveis mediante razoável solicitação ao autor principal do estudo.

Conflito de interesses

Os autores declaram que não há conflitos de interesses.

Referências

- Alves, I. C. B., & Duarte, J. L. M. (2001). Padronização brasileira da Escala de Maturidade Mental Colúmbia. Em B. Burgemeister, L. H. Blum & I. Lorge (Eds.), *Escala de Maturidade Mental Colúmbia – CMMS. Manual para aplicação e interpretação* (1ª edição brasileira) (pp. 31- 54). Casa do Psicólogo.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven – Escala Especial. Manual técnico*. CETEPP – Centro Editor de Testes e Pesquisas em Psicologia.
- Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research & Evaluation, 19*(5), 1-15. <https://doi.org/10.7275/dsep-4220>
- Borsboom, D., Mellenbergh, G. J., & Heerden, V. J. (2004). The concept of validity. *Psychological Review, 111*(4), 1061-1071. <https://doi.org/10.1037/0033-295X.111.4.1061>
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1982). *Test of nonverbal intelligence*. PRO-ED.
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2006). *TONI-3 (forma A): Teste de Inteligência Não Verbal. Manual do examinador*. Vetor Editora.
- Burgemeister, B., Blum, L. H., & Lorge, I. (2001). *Escala de Maturidade Mental Colúmbia – CMMS. Manual para aplicação e interpretação* (1ª ed.). Casa do Psicólogo.
- Canivez, G. L. (2016). Bifactor modeling in construct validation of multifaceted tests: Implications for multidimensionality and test interpretation. Em K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advances* (pp. 247-271). Hogrefe.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281- 302. <https://doi.org/10.1037/h0040957>
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford.
- Flanagan, D. P., & Harrison, P. L. (2012). *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed.). Guilford Press.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. LTC - Livros Técnicos e Científicos Editora.
- Holt, J. K. (2004, October). *Item parceling in structural equation modeling for optimum solutions* [Paper presentation]. Annual Meeting of the mid-Western Educational Research Association, Columbus, OH, United States.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179- 185. <https://doi.org/10.1007/BF02289447>
- Karino, C. A., Laros, J. A., & Jesus, G. R. (2011). Evidências de validade convergente do SON-R 2½-7 [a] com o WPPSI-III e WISC-III. *Psicologia: Reflexão e Crítica, 24*(4), 621-629. <https://doi.org/10.1590/S0102-79722011000400001>

- Laros, J. A., Almeida, G. O. N., Valentini, F., & Lima, R. M. F. (2015). Dimensionalidade e evidências de validade convergente do SON-R 6-40. *Temas em Psicologia*, 23(4), 929 – 945. <http://dx.doi.org/10.9788/TP2015.4-10>
- Laros, J. A., Reis, R. F., & Tellegen, P. T. (2010). Indicações da validade convergente do teste não verbal de inteligência SON-R 2½-7[a]. *Avaliação Psicológica*, 9(1), 43-52. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712010000100006&lng=pt&nr=iso. ISSN 1677-0471
- Laros, J. A., Tellegen, P. J., Jesus, G. R., & Karino, C. A. (2015). SON-R 2½-7[a]. *Manual – Teste não-verbal de inteligência*. Hogrefe-CETEP.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighting the merits. *Structural Equation Modeling*, 9(2), 151-173. https://doi.org/10.1207/S15328007SEM0902_1
- Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, 38(1), 88-91. <https://doi.org/10.3758/BF03192753>
- Messick, S. J. (1989). Validity. Em R. L. Linn (Ed.), *Educational measurement* (3ª ed., pp. 13-103). Macmillan.
- Muthén, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology*, 45(1), 19-30. <https://doi.org/10.1111/j.2044-8317.1992.tb00975.x>
- 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. <https://doi.org/10.7275/0k9h-tq64>
- Pacico, J. C., & Hutz, C. S. (2015). Validade. Em C. S. Hutz, D. R. Bandeira, & C. M. Trentini (Eds.), *Psicometria* (pp. 71 - 84). Artmed Editora Ltda.
- 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. <https://doi.org/10.5923/j.ijpbs.20120201.07>
- Santos, A. A. A., Noronha, A. P. P., & Sisto, F. F. (2006). Estudos psicométricos no Brasil. Em L. Brown, R. J. Sherbenou & S. K. Johnsen (Eds.), *TONI-3 Forma A – Teste de Inteligência Não-Verbal* (pp. 67-120). Vetor Editora.
- Schneider, W. J., & McGrew, K. S. (2018). The Cattell - Horn - Carroll theory of cognitive abilities. Em D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (4ª ed., pp. 71-163). The Guilford Press.
- Tellegen, P. J., & Laros, J. A. (2014). *SON-R 6-40. Non-verbal intelligence test: Research report*. Hogrefe Verlag.
- 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. <https://doi.org/10.1007/BF02294464>
- Thompson, B. (2003). Understanding reliability and coefficient alpha, really. Em B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 1-24). Sage.
- Thompson, B. (2005). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209-220. <https://doi.org/10.1037/a0023353>
- Tran, U. S., & Formann, A. K. (2009). Performance of parallel analysis in retrieving unidimensionality in the presence of binary data. *Educational and Psychological Measurement*, 69(1), 50-61. <https://doi.org/10.1177/0013164408318761>
- Urbina, S. (2014). *Essentials of Psychological Testing: Second Edition*. Wiley.
- Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, 51(4), 473-481. <https://doi.org/10.1037/0022-0167.51.4.473>
- Weng, L. J., & Cheng, C. P. (2005). Parallel analysis with unidimensional binary data. *Educational and Psychological Measurement*, 65(5), 697-716. <https://doi.org/10.1177/0013164404273941>

recebido em abril de 2019
aprovado em dezembro de 2020

Sobre os autores

Talita de Araújo Alves é psicóloga e possui mestrado em psicologia pelo Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília, UnB. É docente do curso de graduação em psicologia do Centro Universitário Unieuro.

Jacob Arie Laros é professor titular aposentado do Departamento de Psicologia Social e do Trabalho da Universidade de Brasília (UnB). Atualmente é pesquisador colaborador sênior do Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações da UnB.

Como citar este artigo

Alves, T., & Laros, J. (2021). SON-R 6-40: dimensionalidade e validade convergente com a CMMS, TONI-3 e COM. *Avaliação Psicológica*, 20(3), 289-298. <http://dx.doi.org/10.15689/ap.2021.2003.18095.03>