

Evidências de validade convergente e discriminante dos escores do SON-R 6-40

Renata Manuely Feitosa de Lima¹

Universidade de Brasília, DF

Jacob Arie Laros

Universidade de Brasília, DF

Resumo: Um adequado processo de validação é essencial para possibilitar uma interpretação correta dos escores de um teste. Assim, o objetivo deste estudo foi obter evidências de validade convergente e discriminante dos escores do teste SON-R 6-40, um teste não verbal de inteligência de origem holandesa, para o qual estão sendo elaboradas as normas brasileiras. O teste foi administrado juntamente com o WISC-IV em uma amostra de 120 crianças de 10 a 14 anos. A correlação, corrigida para atenuação, entre escores totais do SON-R 6-40 e do WISC-IV foi de 0,73. Como esperado, a correlação mais alta foi obtida entre SON-R 6-40 e Índice de Organização Perceptual do WISC-IV ($r = 0,84$) e a correlação mais baixa foi entre SON-R 6-40 e Índice de Velocidade de Processamento ($r = 0,32$). Os resultados obtidos indicam evidências satisfatórias de validade convergente e discriminante dos escores do SON-R 6-40 para a faixa etária investigada.

Palavras-chave: testes de inteligência; validade convergente; validade discriminante; SON-R 6-40; WISC-IV.

EVIDENCE OF CONVERGENT AND DISCRIMINANT VALIDITY OF THE SON-R 6-40 SCORES

Abstract: An adequate validation process is fundamental to ensure sound interpretations of test results. Thus, the purpose of this study was to acquire evidence of the convergent and discriminant validity of the test scores on the SON-R 6-40, a non-verbal test of intelligence of Dutch origin, for which Brazilian rulings are being prepared. The test was performed together with the WISC-IV to a sample of 120 children from 10 up to 14 years old. The correlation, corrected for attenuation, between the total scores on the SON-R 6-40 and the WISC-IV was .73. As expected, a higher correlation ($r = .84$) was obtained between the SON-R 6-40 and the Perceptual Organization Scale of the WISC-IV, and the lowest correlation was between SON-R 6-40 and the Processing Speed Scale ($r = 0.32$). The results indicate satisfactory evidence of the convergent and discriminant validity of the SON-R 6-40 scores for the age group investigated in the research sample.

Keywords: intelligence tests; convergent validity; discriminant validity; SON-R 6-40; WISC-IV.

EVIDENCIAS DE VALIDEZ CONVERGENTE Y DISCRIMINANTE DE LOS PUNTAJES DEL TEST SON-R 6-40

Resumen: Un proceso de validación apropiado es esencial para conseguir una correcta interpretación de los resultados de un test. Así, el objetivo de este estudio fue obtener

¹ **Endereço para correspondência:** Renata Manuely Feitosa de Lima, 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, CEP: 70910-900, Telefone, (61) 3107-6902. E-mail: renatamanuely@gmail.com.

ner evidencias de validez convergente y discriminante de los puntajes del test SON-R 6-40, un test no verbal de inteligencia de origen holandés, para el cual están siendo elaboradas las normas brasileñas. El test fue administrado junto con el WISC-IV en una muestra de 120 niños de 10 a 14 años. La correlación, corregida para la atenuación, entre los puntajes totales del SON-R 6-40 y del WISC-IV, fue de 0,73. Como esperado, la correlación más alta fue obtenida entre el SON-R 6-40 y el Índice de Organización Perceptual del WISC-IV ($r = 0,84$), y la correlación más baja fue entre SON-R 6-40 y el Índice de Velocidad de Procesamiento ($r = 0,32$). Los resultados obtenidos indican una adecuada validez convergente y discriminante de los puntajes del test SON-R 6-40 para el grupo etario investigado.

Palabras clave: testes de inteligencia; validez convergente; validez discriminante; SON-R 6-40; WISC-IV.

Introdução

As análises sobre validade ocupam espaço central no contexto do desenvolvimento de testes psicológicos, pois são necessários estudos que fundamentem teoricamente e empiricamente a interpretação dos escores. De acordo com a visão atual descrita nos *Standards* (Aera, APA, & NCME, 2014), os estudos sobre as evidências de validade dos escores dos testes devem apresentar dados sobre o conteúdo do teste, evidências com base no processo de resposta, na estrutura interna, na relação com outras variáveis e nas consequências da testagem. Este artigo apresenta dados referentes a um estudo com base na relação com outras variáveis, mais especificamente, evidências de validade convergente e discriminante do teste SON-R 6-40 com o WISC-IV. Ressalta-se que o teste SON-R 6-40 é uma medida que está sendo adaptada para o Brasil e foram publicados apenas dois estudos como esse instrumento no contexto brasileiro, sendo importantes e necessários outros estudos com o SON-R 6-40 no país. A primeira parte do artigo apresenta o teste SON-R 6-40, e a segunda, detalha os procedimentos quantitativos de coleta e análise dos dados.

O teste SON-R 6-40 é um teste não verbal de inteligência de origem holandesa, para o qual estão sendo elaboradas normas brasileiras. Além do estudo de normatização, outros estudos de validação desse instrumento estão em andamento. No nome da bateria SON-R 6-40, as três primeiras letras fazem menção ao nome da primeira autora (Nan Snijders-Oomen), a letra R indica que ele é um teste revisado e os números fazem referência à faixa etária do público-alvo do teste.

A primeira versão dos testes SON surgiu em 1943, quando a primeira autora desenvolveu um teste independente das habilidades verbais para avaliar o potencial de aprendizagem de crianças surdas com idades entre 4 e 14 anos (Snijders-Oomen, 1943). Outras versões dos testes SON surgiram com o passar dos anos e, hoje, há uma versão já adaptada para o contexto brasileiro, o teste SON-R 2½-7[a], destinado para a avaliação de crianças entre 2½ a 7 anos de idade (Laros, Jesus, & Karino, 2013).

O foco de mensuração do teste SON-R 6-40 é a inteligência fluida, que pode ser definida como a capacidade para resolver problemas para os quais a pessoa tem pouco conhecimento prévio (Carroll, 2005). A inteligência fluida é capaz de permear muitos tipos de atividades mentais, e as operações que representam essa capacidade são

aquelas que exigem a formação e o reconhecimento de conceitos, a identificação de relações complexas e a realização de inferências (Schelini, 2006).

O teste SON-R 6-40 é composto por quatro subtestes que avaliam um espectro das habilidades cognitivas sem a utilização da linguagem falada ou escrita e, de forma específica, avaliam raciocínio abstrato e concreto, habilidade espacial e percepção visual. As tarefas propostas nos subtestes podem ser explicadas por uma dimensão geral, de acordo com estudos preliminares sobre a estrutura interna do teste (Laros, Almeida, Valentini, & Lima, 2015). Os itens do teste foram desenvolvidos segundo uma teoria de dificuldade (Snijders, Tellegen, & Laros, 1989) para conferir uma ampla variedade de dificuldade dos itens e para possibilitar uma aplicação adaptativa. A ampla variedade de dificuldade dos itens permite que os itens fiquem posicionados em diferentes pontos da escala (itens muito fáceis, medianos e difíceis, por exemplo), possibilitando, assim, a mensuração do traço latente de forma mais precisa.

Como vantagens do teste SON-R 6-40, podemos citar a diversificação das tarefas, a inclusão de um procedimento adaptativo de aplicação que limita a apresentação de itens que são difíceis ou fáceis demais para a pessoa, repercutindo tanto no tempo necessário para responder os subtestes quanto no engajamento do respondente para resolver as atividades, o fornecimento de *feedback* após a resolução de cada item e a possibilidade de avaliar a inteligência geral sem envolver o uso da linguagem falada ou escrita. Essa última característica do teste faz com que ele seja muito adequado para a realidade brasileira, em que ainda existe uma quantidade considerável de pessoas com dificuldades com a linguagem. De acordo com o relatório *Educação Brasileira: Indicadores e Desafios*, produzido pelo Fórum Nacional de Educação (2013), 20,4% da população brasileira com 15 anos ou mais são analfabetos funcionais, e esse índice de *deficit* educacional aumenta nas regiões Norte e Nordeste, 25,3% e 30,8%, respectivamente.

Pessoas com baixa escolaridade, com deficiência auditiva, com dificuldades de expressão ou com algum transtorno no desenvolvimento podem ter seu resultado subestimado quando são avaliadas com testes tradicionais de inteligência que geralmente fazem forte apelo para as habilidades específicas da linguagem, dificultando o entendimento e cumprimento da atividade proposta pelo avaliador (Karino, Laros, & Jesus, 2011). Estudos apontam que indivíduos com transtornos no desenvolvimento tendem a apresentar maiores escores em testes não verbais (Duarte, Covre, Braga, & Macedo, 2011; Flanagan, Alfonso, Mascolo, & Sotelo-Dynega, 2012). Além disso, o estudo de normatização do teste SON-R 6-40 envolve respondentes de mais de 40 municípios das cinco regiões brasileiras, introduzindo métodos diferenciados na composição do grupo amostral e no cálculo das normas.

Sobre os estudos realizados no Brasil com o teste SON-R 6-40, o primeiro estudo foi realizado com a forma A da Bateria de Provas de Raciocínio (BPR-5) em uma amostra composta de 68 adolescentes, com idade variando entre 11,7 e 16,1 anos ($M = 14,2$; $DP = 1,17$) (Laros *et al.*, 2015). As correlações mais altas foram encontradas entre o SON-R 6-40

e os subtestes de Raciocínio Abstrato e Raciocínio Espacial da BPR-5 ($r = 0,68$ para os dois subtestes). O segundo estudo foi com o SON-R 2½-7[a], uma versão dos testes SON já normatizada para o contexto brasileiro (Laros *et al.*, 2015). Os dois testes foram aplicados em 44 crianças com idade variando entre 5,9 e 7,8 anos ($M = 6,9$; $DP = 0,47$) e o SON-R 6-40 mostrou uma correlação corrigida de 0,75 com o SON-R 2½-7[a].

No cenário internacional, há diferentes estudos com o teste SON-R 6-40 que indicam evidências de sua validade convergente (Tellegen & Laros, 2014). Com relação aos estudos com as escalas Wechsler, foi realizada uma pesquisa na Alemanha e na Holanda com 118 indivíduos, com idade entre 6 e 16 anos. A correlação encontrada entre o QI total do SON-R 6-40 e o QI total do WISC-IV foi de 0,80. Todos os subtestes do SON-R 6-40 apresentaram uma correlação mais forte com o índice do WISC voltado para avaliação de raciocínio fluido (Índice de Organização Perceptual – IOP) e, dos quatro subtestes do SON-R 6-40, o subteste Analogias apresentou correlação mais forte com o QI total do WISC-IV ($r = 0,75$). Os resultados encontrados indicam que há evidências de validade convergente do SON-R 6-40, apesar do fato que as duas baterias são compostas por subtestes diferentes.

A versão WISC-IV foi publicada em 2013 no Brasil e tem como objetivo avaliar a capacidade intelectual de pessoas com idade entre 6 anos e 16 anos e 11 meses. Nessa versão, além do QI total, há quatro índices que podem ser mensurados. O Quadro 1 descreve os subtestes que compõem cada índice e o tipo de raciocínio predominante demandado nos subtestes que compõem os respectivos índices, de acordo com a Teoria CHC (Schneider & McGrew, 2012).

Quadro 1. Índices que compõem o teste WISC-IV.

WISC-IV		
Índices	Subtestes principais	Raciocínio predominante
1. Índice de Compreensão Verbal (ICV)	Semelhanças, Vocabulário e Compreensão	Conhecimento-compreensão
2. Índice de Organização Perceptual (IOP)	Cubos, conceitos figurativos e raciocínio matricial	Raciocínio fluido
3. Índice de Memória Operacional (IMO)	Dígitos, sequência de números e letras	Memória de curto prazo
4. Índice de Velocidade de Processamento (IVP)	Código e procurar símbolos	Velocidade de processamento

Fonte: Elaborado pelos autores.

O teste WISC-IV foi escolhido para ser aplicado juntamente com o teste SON-R 6-40 por ser largamente utilizado para avaliação da inteligência de crianças e adultos em diferentes contextos (Figueiredo, Mattos, Pasquali, & Freire, 2008; Fiorello *et al.*, 2007)

e, principalmente, por proporcionar tanto uma medida de evidências de validade convergente quanto discriminante. Neste estudo, espera-se obter correlações altas e positivas entre o QI total do SON-R 6-40 e QI total do WISC-IV. Além disso, espera-se obter correlações altas entre o QI total do SON-R 6-40 e o Índice de Organização Perceptual (IOP), pois os dois índices buscam avaliar raciocínio fluido e correlações mais baixas entre o QI do SON-R 6-40 e os demais índices do WISC-IV. Os tipos de raciocínios que são empregados para solucionar as tarefas propostas nos subtestes que compõem o Índice de Memória Operacional (IMO), o Índice de Velocidade de Processamento (IVP) e o Índice de Compreensão Verbal (ICV) não são reconhecidos como raciocínios representativos do raciocínio fluido, por isso correlações mais baixas são esperadas entre esses índices e o teste SON-R 6-40.

Método

Participantes

Responderam aos testes 120 crianças (61 meninas), residentes de diferentes regiões administrativas do Distrito Federal, matriculadas em duas escolas públicas ($N = 80$) e em uma escola particular ($N = 40$). A idade das crianças variou de 10 a 14 anos ($M = 11,6$; $DP = 0,96$), e a idade entre o intervalo citado foi o critério de seleção adotado para a triagem dos respondentes.

Instrumentos

Foram utilizados dois instrumentos na pesquisa: o SON-R 6-40 e o WISC-IV.

SON-R 6-40 (Tellegen & Laros, 2014) é um teste de aplicação individual, composto por quatro subtestes: Analogias (36 itens), Mosaicos (26 itens), Categorias (36 itens) e Padrões (26 itens). Os subtestes são aplicados nessa ordem, com Analogias e Categorias avaliando raciocínio abstrato e os subtestes Mosaicos e Padrões avaliando o raciocínio espacial. Nesta pesquisa, todos os subtestes foram aplicados.

WISC-IV – Escala Wechsler de Inteligência para Crianças – quarta edição – (Wechsler, 2013) é um teste de aplicação individual, composto por 15 subtestes, que são os subtestes identificados como principais (10) e suplementares (5). Para obter os cinco escores compostos do WISC-IV (QI Total e quatro índices), foram aplicados nesta pesquisa os 10 subtestes principais, a saber: Cubos, Semelhanças, Dígitos, Conceitos Figurativos, Código, Vocabulário, Sequência de Números e Letras, Raciocínio Matricial, Compreensão e Procurar Símbolos.

Procedimentos

O projeto de pesquisa foi inicialmente submetido ao comitê de ética e, após a sua aprovação, um grupo de três psicólogas foi treinado para auxiliar na aplicação dos dados. Após as unidades escolares concordarem em participar da pesquisa, foram enviadas cartas aos pais explicando os objetivos da pesquisa, os procedimentos e os ter-

mos de consentimento. Aqueles que concordavam com a participação do seu filho ou tutelado deveriam devolver o termo assinado.

No final da pesquisa, os pais receberam um relatório descrevendo o desempenho do filho na pesquisa, e a escola recebeu um relatório geral descrevendo como foi o desempenho das crianças e sugestões de atividades que podem ser desenvolvidas. No caso do relatório das escolas, as recomendações éticas foram observadas, e os dados foram tratados de forma conjunta, garantindo o sigilo e a não divulgação de dados individuais.

Foram observados os procedimentos de aplicação e avaliação dos subtestes descritos no manual de cada teste. Todas as aplicações foram individuais e os testes foram aplicados nas escolas, durante o horário das aulas. A aplicação dos testes ocorreu da seguinte forma para evitar vieses de aplicação: 60 crianças responderam ao primeiro o teste SON-R 6-40 e depois o teste WISC-IV; a segunda metade da amostra ($N = 60$) respondeu ao primeiro o teste WISC-IV e depois o teste SON-R 6-40. Os testes foram aplicados em duas sessões e o intervalo de tempo entre a aplicação dos testes foi de três a quatro semanas.

Análise dos dados

Inicialmente, foram realizadas as transformações dos escores brutos em escores normatizados levando em consideração as informações presentes nos manuais de cada teste. O coeficiente Lambda 2 de Guttman foi utilizado para estimar a fidedignidade dos escores dos testes por ser considerado o índice mais adequado para tal análise (Sijtsma, 2009). Testes de normalidade foram realizados e os valores da *skewness* foram observados para avaliar e assegurar os pressupostos de normalidade dos dados (Miles & Shevlin, 2001). Utilizou-se a correlação bivariada de Pearson para a realização das análises de validade convergente e discriminante dos escores dos testes.

Resultados

O primeiro passo foi verificar as estatísticas descritivas para os dois testes utilizados. A Tabela 1 apresenta os valores da média, desvio-padrão, erro padrão da média, o intervalo de confiança de 95% das médias e a amplitude dos escores normatizados de cada subteste e do QI do SON-R 6-40 e de cada subteste do WISC-IV, dos quatro índices e do QI Total do WISC-IV.

É necessário utilizar escores normatizados em vez de escores brutos, uma vez que nos escores brutos ainda existe variância compartilhada com a variável idade, o que resultará em uma superestimação da correlação entre dois testes. É importante destacar que, para transformar os escores brutos do SON-R 6-40 em escores normatizados, as normas da Holanda/Alemanha ($N = 1.933$) foram utilizadas, uma vez que ainda não existem normas para o Brasil. Os escores normatizados dos subtestes do SON-R 6-40 estão numa escala com $M = 10$ e $DP = 3$. Para o escore total (o QI-SON), a escala é $M = 100$ e $DP = 15$. Os escores normatizados dos subtestes do WISC-IV usam a escala com $M = 10$ e $DP = 3$. Para os índices e escore total do WISC-IV, utiliza-se a escala $M = 100$ e $DP = 15$.

Tabela 1. Estatísticas descritivas dos escores normatizados do SON-R 6-40 e WISC-IV.

SON-R 6-40							
Subtestes	M	DP	EP	IC 95%	Mín.	Máx.	
Analogias	10,01	(6,68)	0,18	9,66 – 10,36	4	13	
Mosaicos	9,98	(6,65)	0,20	9,59 – 10,37	2	12	
Categorias	10,56	(7,23)	0,22	10,13 – 10,99	2	14	
Padrões	10,18	(6,85)	0,18	9,83 – 10,53	2	15	
QI-SON	98,30	(81,65)	0,78	96,77 – 99,82	65	111	
WISC-IV							
Subtestes	M	DP	EP	IC 95%	Mín.	Máx.	
Cubos	9,71	2,63	0,24	9,24 – 10,18	3	16	
Semelhanças	9,38	2,80	0,26	8,88 – 9,88	3	17	
Dígitos	9,79	2,99	0,27	9,26 – 10,32	2	18	
Conceitos Figurativos	9,27	2,71	0,25	8,79 – 9,75	1	14	
Código B	9,90	2,55	0,23	9,45 – 10,35	4	19	
Vocabulário	8,46	3,10	0,28	7,91 – 9,01	1	16	
Seq. de Números e Letras	8,55	2,90	0,26	8,04 – 9,06	3	19	
Raciocínio Matricial	9,37	2,64	0,24	8,90 – 9,84	4	17	
Compreensão	8,41	3,07	0,28	7,87 – 8,95	1	15	
Procurar Símbolos B	11,17	2,17	0,20	10,79 – 11,55	4	18	
Índices	M	DP	EP	IC 95%	Mín.	Máx.	
Índice de Compreensão Verbal	92,63	15,09	1,37	89,95 – 95,31	55	130	
Índice de Organização Perceptual	96,60	12,46	1,13	94,39 – 98,81	69	126	
Índice de Memória Operacional	95,18	14,65	1,33	92,58 – 97,78	65	138	
Índice de Velocidade de Processamento	103,11	11,85	1,08	98,98 – 107,24	64	147	
QI Total do WISC-IV	95,38	13,15	1,20	93,03 – 97,73	61	131	

Notas. M = média; DP = desvio padrão; EP = erro padrão da média; IC 95% = intervalo de confiança de 95% da média. As médias dos escores normatizados do SON-R 6-40 entre parênteses são os escores observados em base das normas da Holanda/Alemanha. Ao lado encontram-se as médias corrigidas.

Fonte: Elaborado pelos autores.

Os resultados da Tabela 1 mostram que as médias dos escores normatizados dos subtestes e do escore total do SON-R 6-40 (apresentadas entre parênteses) são bem inferiores aos valores das médias ($M = 10$) na amostra normativa da Holanda/Alemanha. Esse resultado foi esperado, uma vez que na normatização do SON-R 2½-7[a] no Brasil a diferença observada entre as crianças brasileiras e holandesas/alemãs na Escala Geral foi 16,7 pontos o que equivale a 1,11 desvio-padrão (Laros, Jesus, & Karino, 2013). Os escores normatizados corrigidos foram obtidos por um aumento de 1,11 desvio padrão nos valores das médias brutas. No caso das médias dos subtestes, isso resultou em um aumento de 3,33 pontos ($1,11 \times 3$) e no caso do escore total, em um aumento de 16,65 pontos ($1,11 \times 15$).

A Tabela 1 mostra também que a variância dos escores normatizados dos subtestes do SON-R 6-40 é bem inferior à variância nesses escores na amostra normativa da Holanda/Alemanha. A mesma observação aplica-se a variância do escore total do SON-R 6-40.

Verificando os valores das médias dos escores normatizados dos subtestes do WISC-IV, podemos observar que, com exceção do subteste Procurar símbolos B, os valores são inferiores aos valores da amostra normativa do WISC-IV. Em geral, a variância dos escores normatizados dos subtestes também são inferiores aos valores da amostra normativa do WISC-IV.

Com relação aos escores compostos do WISC-IV, a Tabela 1 mostra que três dos quatro índices (ICV, IOP e IMO) e o QI Total do WISC-IV possuem valores abaixo do valor da média (100) da amostra normativa do WISC-IV. Os valores do desvio padrão de três índices (IOP, IMO e IVP) e do QI Total também são menores do que o desvio padrão (15) na amostra normativa do WISC-IV. Essas observações indicam que, em geral, na amostra do estudo atual, existe menos variabilidade nas habilidades cognitivas avaliadas em comparação com a amostra normativa do WISC-IV e que, em comparação com a amostra normativa, as médias no QI Total e nos três dos quatro índices do WISC-IV são mais baixas.

A Tabela 2 mostra os valores dos coeficientes de fidedignidade e a correlação média entre os itens de todos os subtestes e da escala geral. Ressalta-se que alguns itens não foram incluídos no momento do cálculo da fidedignidade por não apresentarem variância, o que pode dificultar a estimação da fidedignidade. Além disso, a fidedignidade dos subtestes Códigos e Procurar Símbolos B não foi estimada pelo método da consistência interna porque não é correto medir a fidedignidade de testes de velocidade com esse método, mas com o método teste-reteste (Karino, Laros, & Jesus, 2011). Na pesquisa com o WAIS-III, por exemplo, a fidedignidade dos subtestes Códigos e Procurar Símbolos foi estimada a partir do método teste-reteste, que é a correlação entre os resultados da primeira e da segunda aplicação.

A inspeção da Tabela 2 revela que tanto os coeficientes de fidedignidade dos escores nos subtestes do SON-R 6-40 quanto os índices de fidedignidade dos subtestes do WISC-IV são satisfatórios. Além disso, os escores na escala total do SON-R 6-40 e WISC-IV foram de 0,92 e 0,95, respectivamente.

Tabela 2. Índices de fidedignidade do SON-R 6-40 e do WISC-IV.

Teste	Subteste	n° de itens	Lambda 2	r média entre itens
SON-R 6-40	Analogias	36	0,81	0,12
	Mosaicos	26	0,85	0,20
	Categorias	36	0,84	0,14
	Padrões	26	0,83	0,17
	Escala total	124	0,92	0,10
WISC-IV	Cubos	14	0,83	0,27
	Semelhanças	23	0,89	0,22
	Dígitos	32	0,84	0,17
	Conceitos figurativos	28	0,81	0,12
	Código B	119	----	0,26
	Vocabulário	36	0,88	0,17
	Seq. de números e letras	30	0,87	0,19
	Raciocínio matricial	35	0,88	0,18
	Compreensão	21	0,85	0,20
	Procurar símbolos B	60	----	0,17
	Índice de Compreensão Verbal	71	0,94	0,16
	Índice de Organização Perceptual	67	0,91	0,10
	Índice de Memória Operacional	43	0,90	0,14
	Índice de Velocidade de Processamento	96	----	0,17
	Escala Total	219	0,95	0,08

Fonte: Elaborado pelos autores.

Foram calculadas correlações de Pearson, a fim de buscar associações entre o QI do SON-R 6-40 com os índices e com o escore total do WISC-IV. A Tabela 3 apresenta as correlações brutas e as correlações depois da aplicação da fórmula de correção para atenuação e falta de variância. Osborne (2003) e Thompson (2003) argumentam que as relações reais entre variáveis podem ser subestimadas caso uma parte da variância seja variância de erro e quando a variância na amostra pesquisada é restrita. Segundo esses autores, pesquisadores precisam corrigir para atenuação no intuito de obter uma estimativa melhor da relação verdadeira entre as variáveis na população. A correlação entre o QI do SON-R 6-40 e o QI total do WISC-IV corrigida foi de 0,73. Como apresentado na Tabela 3, observa-se que a correlação corrigida entre o QI do SON-R 6-40 e o Índice de Organização Perceptual foi a correlação mais alta (0,84). Esse resultado corrobora com o esperado e pode ser explicado pela última revisão do WISC-IV

de incluir mais subtestes que avaliam a inteligência fluida. O manual desse teste relata que foram incorporados cinco novos subtestes, dois deles (Raciocínio Matricial e Conceitos Figurativos) servindo para medir raciocínio fluido e fazendo parte do IOP. Como esperado, foram encontradas correlações baixas entre o QI do SON-R 6-40 e os demais índices do WISC-IV, sendo menor com o Índice de Velocidade de Processamento ($r = 0,32$).

Tabela 3. Correlações entre o SON-R 6-40 com os Índices do WISC-IV

Índices do WISC-IV	QI SON-R 6-40	IC 95%
Índice de Compreensão Verbal	0,48 (0,45)	0,33 – 0,61
Índice de Organização Perceptual	0,84 (0,71)	0,79 – 0,89
Índice de Memória Operacional	0,44 (0,39)	0,28 – 0,57
Índice de Velocidade de Processamento	0,32 (0,26)	0,15 – 0,47
WISC-IV QI Total	0,73 (0,63)	0,63 – 0,80

Nota : A correlação bruta é apresentada entre parênteses. Os intervalos de confiança de 95% foram calculados para as correlações corrigidas.

Fonte: Elaborado pelos autores.

Por fim, a Tabela 4 apresenta os valores encontrados das correlações entre os subtestes que compõem cada um dos testes. Foram utilizados os escores normatizados para estimar a correlação entre os testes. Os valores indicam maior correlação entre os subtestes do SON-R 6-40 e os subtestes Cubos e Raciocínio Matricial do WISC-IV. Os subtestes Cubos e Raciocínio Matricial compõem o Índice de Organização Perceptual, índice que é destinado para avaliação de raciocínio fluido no WISC-IV.

Tabela 4. Correlações entre os subtestes do SON-R 6-40 com os subtestes do WISC-IV

WISC-IV	SON-R 6-40			
	Analogias	Mosaicos	Categorias	Padrões
Cubos	0,58 (0,44)	0,73 (0,57)	0,49 (0,37)	0,62 (0,47)
Semelhanças	0,35 (0,28)	0,38 (0,31)	0,44 (0,37)	0,19 (0,15)
Dígitos	0,35 (0,29)	0,33 (0,28)	0,27 (0,23)	0,24 (0,20)
Conceitos figurativos	0,51 (0,39)	0,50 (0,38)	0,50 (0,38)	0,37 (0,28)
Código B	0,14 (0,11)	0,24 (0,21)	0,25 (0,21)	0,24 (0,20)

(continua)

Tabela 4. Correlações entre os subtestes do SON-R 6-40 com os subtestes do WISC-IV (conclusão)

WISC-IV	SON-R 6-40			
	Analogias	Mosaicos	Categorias	Padrões
Vocabulário	0,50 (0,42)	0,25 (0,22)	0,36 (0,31)	0,22 (0,19)
Seq. números e letras	0,33 (0,27)	0,16 (0,13)	0,23 (0,19)	0,08 (0,07)
Raciocínio matricial	0,77 (0,54)	0,60 (0,47)	0,59 (0,46)	0,41 (0,32)
Compreensão	0,35 (0,29)	0,20 (0,17)	0,24 (0,20)	0,18 (0,15)
Procurar símbolos B	0,28 (0,21)	0,30 (0,22)	0,37 (0,27)	0,33 (0,24)

Nota: As correlações brutas são apresentadas entre parênteses; as correlações corrigidas para atenuação são apresentadas ao lado.

Fonte: Elaborado pelos autores.

Discussão

O objetivo principal deste estudo foi estimar evidências de validade convergente e discriminante dos escores obtidos no teste SON-R 6-40 com o WISC-IV. Os resultados indicaram alta correlação com o Índice de Organização Perceptual (IOP) ($r = 0,84$), revelando evidências de validade convergente. A correlação alta entre o IOP e o QI Total do SON-R 6-40 se justifica porque as atividades propostas nos subtestes que compõem o IOP são semelhantes às tarefas dos subtestes do SON-R 6-40, por exemplo, identificar e reproduzir padrões e reconhecer o conceito subjacente de um grupo de figuras.

As correlações foram mais baixas com os demais índices do WISC-IV, porque as atividades propostas nos subtestes que compõem esses índices avaliam aspectos diferentes do raciocínio mensurado no teste SON-R 6-40: nesse teste, o foco é o raciocínio fluido; no WISC-IV, o IVP mensura velocidade de processamento, o IMO avalia memória de curto prazo, e o ICV conhecimento-compreensão. Quanto às tarefas que precisam ser realizadas nesses índices, no Índice de Velocidade de Processamento ($r = 0,32$) o examinando precisa indicar se um dos elementos do grupo de estímulo está presente no grupo de busca (subteste Procurar Símbolos). No caso do Índice de Memória Operacional ($r = 0,44$), o examinando precisa repetir uma ordem de números enunciados pelo aplicador (subteste Dígitos) ou repetir uma sequência de números e letras em ordem crescente e em ordem alfabética (subteste Sequência de Números e Letras). Por fim, no Índice de Compreensão Verbal, os subtestes demandam que o examinando responda perguntas tais como "Por que as pessoas escovam os dentes?" (subteste Compreensão) ou expressem uma definição clara das palavras que o aplicador lê em voz alta, por exemplo, "O que é o alfabeto?" (subteste Vocabulário). O teste SON-R 6-40 não apresenta nenhum subteste semelhante aos três últimos índices citados.

Considerando a correlação entre os subtestes das duas baterias, observou-se maior correlação entre o subtestes Mosaicos e Cubos ($r = 0,73$). Tal resultado era esperado devido à similaridade das tarefas desses subtestes: o examinando precisa reproduzir padrões que são apresentados utilizando peças ou cubos coloridos que lhe são oferecidos. Além de confirmar as hipóteses inicialmente levantadas, os resultados indicaram boa qualidade psicométrica dos itens do teste SON-R 6-40.

Vários estudos foram realizados na Holanda e na Alemanha com o SON-R 6-40 e outros testes de inteligência, tais como WISC-III, WISC-IV, WAIS-III, WNV (Tellegen & Laros, 2014). De forma geral, os resultados relatados aqui apresentam tendência semelhante aos resultados encontrados nessas pesquisas e principalmente com a pesquisa realizada na Alemanha com o SON-R 6-40 e as escalas Wechsler.

No que se refere aos estudos sobre validade, Messick (1989) influenciou profundamente o campo teórico da avaliação. Ele defendeu essa característica como o grau em que todas as evidências de validade obtidas apoiam a interpretação dos escores de um teste. Portanto, as diferentes fontes de validade devem ser avaliadas e acumuladas para conferir validade para o construto e apoiar as interpretações dos escores.

Na testagem psicológica, as características investigadas são deduzidas a partir da forma como as pessoas respondem aos instrumentos. Por esse motivo, percebe-se a importância e a necessidade das pesquisas sobre a qualidade dos testes utilizados em grupos ou aplicados individualmente. Procedimentos tais como a validade convergente e discriminante contribuem para a obtenção de evidências de validade baseada nas relações com outras variáveis, pois permitem observar tanto a correlação entre os escores obtidos de testes que investigam o mesmo construto ou construtos correlatos quanto a baixa correlação entre testes que examinam diferentes fatores (Hogan, 2006). Além disso, a utilização desse tipo de análise tende a contribuir para a construção de um conjunto de estudos que tratam das características psicométricas dos itens de novos testes.

Conclusão

A avaliação cognitiva é um passo fundamental tanto para diagnosticar quanto para planejar uma intervenção. Assim, considerando a pertinência e a contribuição dos testes de inteligência no diagnóstico precoce de deficiências intelectuais e acreditando que intervenções bem elaboradas podem cooperar para melhorar as chances de alguém atingir sua capacidade plena, espera-se que esta pesquisa contribua para o campo da avaliação psicológica no Brasil, mais especificamente na área do desenvolvimento de instrumentos que buscam mensurar a inteligência.

Como limitação da pesquisa, destaca-se que o estudo não apresentou uma amostra representativa do público-alvo do teste, pouca variabilidade da amostra e ausência de análises estatísticas mais sofisticadas. Apesar da pouca variabilidade da amostra, os resultados mostram-se consistentes com os estudos previamente apresentados, obtidos com outras amostras. Sugere-se que futuros estudos empreguem diferentes métodos

de amostragem, além de diferentes estratégias para avaliar o teste psicológico destacado aqui, tais como, avaliar evidências de validade de construto com Análise Fatorial Confirmatória (CFA), Análise Fatorial de Informação Plena (FIFA) e evidências de validade de critério já que a rede nomológica do construto investigado é bem conhecida.

Referências

- Aera, APA, & NCME. American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.). *Contemporary intellectual assessment: theories, tests and issues* (pp. 69-76). New York: The Guilford Press.
- Duarte, C. P., Covre, P., Braga, A. C., & Macedo, E. C. (2011). Visuospatial support for verbal short-term memory in individuals with Down syndrome. *Research in Developmental Disabilities, 32*(5), 1918-1923. doi: 10.1016/j.ridd.2011.03.024.
- Figueiredo, V. L. M., Mattos, V. L. D., Pasquali, L., & Freire, A. P. (2008). Propriedades psicométricas dos itens do teste WISC-III. *Psicologia em Estudo, 13*(3), 585-592. doi: 10.1590/S1413-73722008000300020.
- Fiorello, C. A., Hale, J. B., Holdnack, J. A., Kavanagh, J. A., Terrell, J., & Long, L. (2007). Interpreting intelligence test results for children with disabilities: Is global intelligence relevant? *Applied Neuropsychology, 14*(1), 2-12. doi: 10.1080/09084280701280338.
- Flanagan, D. P., Alfonso, V. C., Mascolo, J. T., & Sotelo-Dynega, M. (2012). Use of ability tests in the identification of specific learning disabilities within the context of an operational definition. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: theories, tests and issues* (pp. 643-669). New York: The Guilford Press.
- Fórum Nacional de Educação. (2013). *Educação brasileira: Indicadores e desafios*. Recuperado em 3 de Julho, 2017, de http://conae2014.mec.gov.br/images/pdf/educacao_brasileira_indicadores_e_desafios.pdf.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A.
- 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. doi: 10.1590/S0102-79722011000400001.
- Laros, J. A., Jesus, G. R., & Karino, C. A. (2013). Validação brasileira do teste não verbal de inteligência SON-R 2½-7[a]. *Avaliação Psicológica, 12*(2), 233-242.

- 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. doi: 10.9788/TP2015.4-10
- Messick, S. (1989). Validity. In R. L. Linn (Ed.). *Educational measurement* (3rd ed., pp. 13-103). New York: American Council on Education, MacMillan.
- Miles, J., & Shevlin, M. (2001). *Applying regression & correlation. A guide for students and researchers*. London: Sage Publications.
- 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 July 3rd, 2017, from: <http://pareonline>.
- Schelini, P. W. (2006). Teoria das inteligências fluida e cristalizada: início e evolução. *Estudos de Psicologia (Natal)*, 11(3), 323-332. doi: 10.1590/S1413-294X2006000300010.
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Ed.). *Contemporary intellectual assessment: theories, tests, and issues* (pp. 99-144). New York: Guilford Press.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107-120. doi: 10.1007/s11336-008-9101-0.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen* [The examination of intelligence with deaf-mute children]. Nijmegen: Berkhout.
- Snijders, J. Th., Tellegen, P. J., & Laros, J. A. (1989). *Snijders-Oomen Non-verbal intelligence test. SON-R 5½-17. Manual and research report*. Groningen: Wolters Noordhoff.
- Tellegen, P. J., & Laros, J. A. (2014). *SON-R 6-40. Non-verbal intelligence test: research report*. Göttingen, Germany: Hogrefe Verlag.
- Thompson, B. (2003). *Score reliability: contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage Publications.
- Wechsler, D. (2011). *WAIS-III – Escala de inteligência Wechsler para adultos. Manual técnico*. São Paulo: Casa do Psicólogo.
- Wechsler, D. (2013). *WISC-IV – Escala Wechsler de inteligência para crianças. Manual técnico*. São Paulo: Casa do Psicólogo.

Submissão: 11.12.2015

Aceitação: 10.4.2017