

# Evidências de Validade Convergente do SON-R 2½-7[a] com o WPPSI-III e WISC-III

## *Evidences of Convergent Validity of SON-R 2½-7[a] with WPPSI-III and WISC-III*

Camila Akemi Karino\*, Jacob Arie Laros & Girlene Ribeiro de Jesus  
*Universidade de Brasília, Brasília, Brasil*

### **Resumo**

Um processo de validação adequado é fundamental para assegurar uma interpretação correta dos resultados de um teste. Assim, este estudo investigou a validade convergente do teste não-verbal de inteligência SON-R 2½-7[a] com o WPPSI-III e o WISC-III. O teste é a versão abreviada do SON-R 2½-7, de origem holandesa, normatizado e validado em vários países da Europa. No total, 109 crianças de três estados brasileiros participaram do estudo. A correlação entre o SON-R 2½-7[a] e o WPPSI-III foi igual a 0,75 e com o WISC-III foi de 0,69. Esses coeficientes de validade são muito semelhantes aos encontrados em estudos feitos com o SON-R 2½-7 em outros países. Conclui-se que os resultados indicam uma validade convergente satisfatória do teste SON-R 2½-7[a].

*Palavras-chave:* Testes de Inteligência; Validação de Instrumentos Psicológicos; Habilidades Cognitivas.

### **Abstract**

An adequate validation process is fundamental to ensure sound interpretations of test results. Thus, this study aimed to investigate the convergent validity of the SON-R 2½-7[a] with the WPPSI-III and the WISC-III. The test is an abridged edition of the SON-R 2½-7, a Dutch non-verbal intelligence test, with normatization and validation studies done in several European countries. A total of 109 children from three Brazilian states participated in the study. The correlations of the SON-R 2½-7[a] with the WPPSI-III and with the WISC-III were .75 and .69, respectively. The validation coefficients encountered are very similar to the ones found in studies with the SON-R 2½-7 in other countries. The results imply a satisfactory convergent validity of the SON-R 2½-7[a].

*Keywords:* intelligence tests; validation of psychological instruments; cognitive abilities.

Os testes psicológicos desempenham hoje papel proeminente nos diversos contextos de atuação dos psicólogos: uso clínico, contextos educacionais, seleção de pessoal, pesquisas, etc. Hogan (2006), por exemplo, afirma que psicólogos que trabalham com saúde mental gastam de 15 a 18% de seu tempo em atividades de avaliação e psicólogos escolares gastam em média metade de seu tempo profissional em avaliação. Com o aumento da demanda de avaliação psicológica, tem crescido também a necessidade de um aprimoramento de instrumentos e métodos.

Para que os testes sejam úteis e eficientes, eles devem passar por estudos que comprovem sua qualidade psicométrica (Noronha & Vendramini, 2003). No entan-

to, verifica-se no cenário brasileiro a necessidade de melhoria da qualidade da maioria dos testes no que tange a sua construção, normatização e validação.

Em uma revisão dos testes psicológicos comercializados no Brasil, Noronha, Primi e Alchieri (2004) verificaram que uma parte significativa dos instrumentos carece de fundamentação teórica e não atende aos requisitos mínimos definidos nas perspectivas nacionais e internacionais. Assim, alguns testes psicológicos têm sido publicados sem os critérios mínimos exigidos.

Especificamente no contexto de avaliação de habilidades cognitivas de crianças pré-escolares e escolares, vários estudos têm apontado a necessidade no Brasil de testes psicológicos válidos e com normas nacionais (Alves, 2002; Muñiz, Prieto, Almeida, & Bartram, 1999; Noronha & Alchieri, 2002; Noronha & Vendramini, 2003; Oakland, Wechsler, Bensuan, & Stafford, 1994). Nesse cenário, o presente estudo pretende verificar a validade convergente do teste não-verbal de inteligência SON-R 2½-7[a] com os testes WPPSI-III (*Wechsler Preschool and Primary Scale of Intelligence – Third*

\* Endereço para correspondência: Universidade de Brasília, Instituto de Psicologia, Instituto Central de Ciências Sul, Campus Universitário Darcy Ribeiro, prédio Minhocão, Brasília, DF, Brasil, CEP 70 910-900. E-mails: [camilaakarino@gmail.com](mailto:camilaakarino@gmail.com), [jalaros@gmail.com](mailto:jalaros@gmail.com) e [girlene.ribeiro@gmail.com](mailto:girlene.ribeiro@gmail.com)

Edition) e WISC-III (Escala de Inteligência Wechsler para Crianças).

O teste SON-R 2½-7[a] para crianças na faixa etária entre 2 anos e meio e 7 anos, normatizado para a população brasileira (Jesus, 2009; Laros, Tellegen, Jesus, & Karino, in press), é a versão reduzida do teste não-verbal de inteligência SON-R 2½-7, de origem holandesa. Este é um teste de inteligência geral para crianças novas, que avalia um espectro largo de habilidades sem envolver o uso da linguagem.

Os testes SON (Snijders-Oomen Não-verbal) devem seu nome à criadora do primeiro teste, Dr<sup>a</sup>. Nan Snijders-Oomen. Ela iniciou o desenvolvimento do primeiro teste SON na Holanda no início dos anos 40 para avaliação de crianças surdas (Snijders-Oomen, 1943). Na construção do teste, a inteligência foi operacionalizada em termos da habilidade para aprender, visando mensurar a denominada inteligência fluida (J. L. Horn, 1994; Primi, 2002). Buscava-se permitir que funções como abstração, simbolismo e entendimento de situações comportamentais pudessem ser mais acessíveis por meio de uma avaliação não-verbal de inteligência.

Em 1998, a versão completa do teste – SON-R 2½-7 – foi normatizada e validada na Holanda com base em uma amostra de 1.124 crianças (Tellegen, Winkel, Wijnberg-William, & Laros, 1998). A letra R foi inserida para indicar que se trata de uma versão revisada e os números referem-se a faixa de idade das crianças a qual o teste contempla. Também foram realizadas normatizações do teste na Alemanha (Tellegen, Laros, & Petermann, 2007) e na República Tcheca (Tellegen, Laros, & Heider, 2008). A normatização para a população brasileira do teste SON-R 2½-7 ocorreu em 2008 com uma amostra de 1.200 crianças de todas as regiões do país (Laros et al., in press). Na ocasião foi normatizada uma versão reduzida do teste, o SON-R 2½-7[a] (a letra “a” refere-se a palavra inglesa *abbreviated*). Essa versão é composta por quatro subtestes: Categorias, Situações, Mosaicos e Padrões.

Outra característica do teste não-verbal de inteligência SON-R 2½-7[a] é que ele pode ser aplicado em crianças com problemas auditivos e de linguagem. Os testes tradicionais de inteligência geralmente fazem apelo para habilidades de linguagem específicas, tanto nos conteúdos quanto nas instruções do teste. Com isso, crianças com problemas auditivos e de linguagem podem ter seu resultado enviesado. Para esses grupos, um baixo desempenho no teste poderia refletir um conhecimento verbal pobre, em vez de raciocínio ou habilidade para aprendizagem pobre (Tellegen, 1993; Tellegen & Laros, 2005).

Os testes SON-R foram citados pela especialista em avaliação neuropsicológica, M. D. Lezak, em sua obra intitulada “*Neuropsychological Assessment*”, como ferramentas muito úteis porque podem ser utilizados na exploração de várias deficiências neuropsicológicas (Lezak, Howieson, & Loring, 2004). O teste também

recebeu avaliação máxima da Comissão Nacional de Avaliação de Testes dos Países Baixos (Evers, Van Vliet-Mulder, & De Groot, 2000). Dentre os aspectos avaliados estavam: base teórica para construção do teste, manual, material do teste, normas, validade e fidedignidade. Diante das boas avaliações e dos diferenciais que o SON-R 2½-7[a] apresenta, estudos com o teste no Brasil são fundamentais para a disponibilização de mais um instrumento de avaliação psicológica e neuropsicológica infantil.

#### *Estudos de Validação do SON-R 2½-7 com as Escalas Wechsler*

Já foram realizados estudos de validação do SON-R 2½-7 em países como Holanda, Austrália, Estados Unidos, Reino Unido e Alemanha (Tellegen et al., 2007; Tellegen et al., 1998). Nestes estudos foram analisadas as correlações entre o SON-R 2½-7 e outros testes de inteligência, entre eles, o WPPSI-III e WISC-III.

A opção por realizar um estudo de validade convergente do SON-R 2½-7[a] com o WPPSI-III e o WISC-III deve-se à ampla utilização dessas baterias de testes no Brasil. Na área de avaliação da inteligência, as escalas Wechsler são internacionalmente reconhecidas quanto ao valor de suas características psicométricas, comprovadas pela ampla produção científica relacionada à escala (Figueiredo, 2001; Figueiredo, Mattos, Pasquali, & Freire, 2008).

O início do desenvolvimento do teste Wechsler de inteligência ocorreu por volta da década de 30 e surgiu com muitas inovações, tais como a provisão de uma medida geral de capacidade e escalas verbais e de execução separadas (Wechsler, 1991). No Brasil, o uso das Escalas Wechsler começou pela tradução em 1964 do teste WISC (*Wechsler Intelligence Scale for Children*) de 1949. Não consistia ainda de uma padronização para o contexto brasileiro, mas a carência de testes adequados levou a uma enorme utilização do WISC por diversos profissionais (Figueiredo, 2001).

A normatização brasileira do WISC somente veio a ocorrer em 2001, quando a versão WISC-III foi adaptada e validada para uma amostra do Rio Grande do Sul por Figueiredo (2001). Já o WPPSI (*Wechsler Preschool and Primary Scale of Intelligence*), destinado a crianças um pouco mais novas, ainda não foi validado e normatizado para o Brasil. Apesar da inexistência de normas brasileiras, o WPPSI-III foi utilizado porque há vários estudos em outros países entre o SON-R 2½-7 e o WPPSI-III devido à maior correspondência entre as idades das crianças. Esse fator foi determinante para a escolha do WPPSI-III nesse estudo.

Na Austrália, um estudo com 59 crianças acerca da validade convergente do SON-R 2½-7 com a versão revisada do WPPSI (WPPSI-R) apontou correlação alta entre os testes. A correlação entre o QI geral do SON-R 2½-7 com a escala verbal, de execução e completa fo-

ram, respectivamente: 0,54, 0,74 e 0,75 (Jenkinson, Roberts, Dennehy, & Tellegen, 1996; Tellegen, 1997).

Na Holanda, foram administrados em 53 crianças com atrasos no desenvolvimento cognitivo o SON-R 2½-7 e o WPPSI-R. A correlação com o QI total do WPPSI-R foi de 0,60 e as correlações com a escala verbal e de execução foram: 0,49 e 0,59 (Tellegen et al., 1998).

Ainda na Holanda, o SON-R 2½-7 foi aplicado com o WISC-R (versão revisada do WISC) em 20 crianças especiais. A correlação entre os escores totais dos dois testes foi 0,62. A correlação do QI total do SON-R 2½-7 com a escala verbal do WISC-R foi 0,47 e com a escala de execução, 0,76.

No Oeste da Virgínia (Estados Unidos da América), também foi realizado um estudo de validação entre o SON-R 2½-7 e o WPPSI-R. Os testes foram aplicados em 75 crianças normais que tinham entre 3 e 6 anos de idade. A correlação entre o QI do SON-R 2½-7 com o QI total do WPPSI-R e com as escalas (verbal e de execução) foi nesta mesma ordem de 0,59, 0,43 e 0,60 (J. ten Horn, 1996).

Todos os estudos acima apresentados corroboram com a hipótese de uma adequada validade dos testes SON-R, incentivando, assim, pesquisas no contexto brasileiro. Adicionalmente, vale ressaltar as inovações trazidas por esse teste e a possibilidade de o mesmo contribuir para o preenchimento de uma lacuna existente no Brasil no que se refere à avaliação cognitiva de crianças pré-escolares e escolares. Diante do exposto, o presente estudo tem como objetivo verificar a validade convergente do teste não-verbal de inteligência SON-R 2½-7[a] no Brasil com os testes Wechsler: o WPPSI-III e o WISC-III. Tem-se como hipóteses que: (a) haverá correlações altas e positivas entre o SON-R 2½-7[a] e os testes Wechsler; (b) es-

pera-se especificamente correlações mais altas entre o SON-R 2½-7[a] e as escalas de execução em comparação as escalas verbais dos testes Wechsler e (c) os subtestes do SON-R 2½-7[a] que mais se assemelham em relação ao tipo de tarefa aos subtestes das escalas Wechsler terão correlações maiores.

## Método

### Participantes

A amostra foi composta por 49 crianças que responderam aos testes SON-R 2½-7[a] e WPPSI-III (*Wechsler Preschool and Primary Scale of Intelligence – Third Edition*) e por 59 crianças que responderam aos testes SON-R 2½-7[a] e WISC-III (Escala de Inteligência Wechsler para Crianças). A amostra foi por conveniência e o critério de seleção das crianças foi ter entre 3 e 7 anos de idade. O nível socioeconômico das crianças era diverso.

As crianças que responderam ao WPPSI-III ( $N = 49$ ) possuíam entre 3 e 6 anos de idade e 51% eram do sexo feminino. A maioria (77,5%) era de alunos de uma escola pública de Brasília (DF) e destes, 75% cursava a pré-escola. Os demais 22,5% eram de uma creche de Brasília.

Já as crianças que responderam ao WISC-III ( $N = 59$ ) tinham 6 e 7 anos de idade e eram residentes de Brasília-DF ( $N=8$ ), Porto Alegre-RS ( $N=2$ ), Curitiba-PR ( $N=18$ ), Dom Feliciano-RS ( $N=8$ ), São José dos Pinhais-PR ( $N=15$ ) e São Gabriel-RS ( $N=8$ ). A maior parte das crianças estava na 2ª série do ensino fundamental (75%) e eram do sexo feminino (52,5%). Havia crianças de escolas públicas e particulares. Apenas uma criança não frequentava a escola. Na Tabela 1, são apresentadas as características da amostra.

Tabela 1  
*Características da Amostra Pesquisada*

Amostra	Sexo		Idade (anos)					Série Escolar		Localidade				
	F	M	3	4	5	6	7	NE	PE	1 <sup>a</sup>	2 <sup>a</sup>	DF	RS	PR
SON-R 2½-7[a] & WPPSI-III	25	24	11	13	10	12	3	11	36	2	0	49	0	0
SON-R 2½-7[a] & WISC-III	31	28	0	0	0	25	34	1	11	43	2	8	18	33
Total	56	52	11	13	10	37	37	12	47	45	2	57	18	33

Nota. F = Feminino; M = Masculino; NE = Não Escolar; PE = Pré-Escolar; Faltou informação quanto a série escolar de duas crianças que responderam ao WISC-III.

As cidades onde ocorreram as aplicações foram selecionadas por possuírem estudantes de psicologia e psicólogos vinculados ao grupo de pesquisa da Universidade de Brasília com experiência na aplicação do WISC-III e do WIPPSI-III. Todos participaram de um treinamento em Brasília para aprender a aplicar o SON-R 2½-7[a].

### Instrumentos

Foram utilizados três instrumentos na pesquisa:

*SON-R 2½-7[a]* (Laros et al., in press). Consiste na versão reduzida e normatizada para a população brasileira do teste não-verbal de inteligência Snijders-Oomen para a faixa etária de crianças de 2½ a 7 anos. O processo de normatização brasileira ocorreu em 2008 por Laros et al.

(in press) em uma amostra de 1.200 crianças estratificada por região, situação educacional, sexo e variáveis demográficas. Antes do início da normatização foram feitos estudos pilotos para adaptação dos itens a realidade brasileira. Também já foi desenvolvido estudo de validação do SON-R 2½-7[a] com outros testes de inteligência no Brasil (Laros, Reis, & Tellegen, 2010). O teste é de aplicação individual e possui procedimento adaptativo de aplicação. As instruções podem ser dadas tanto de maneira verbal quanto não-verbal. Sua versão reduzida é composta por quatro subtestes: Categorias (15 itens) e Situações (14 itens) que avaliam raciocínio abstrato e concreto respectivamente, e Mosaicos (15 itens) e Padrões (6 itens) que avaliam raciocínio espacial. Essas habilidades são tradicionalmente mensuradas em testes de inteligência, porém, nos testes SON, a avaliação é realizada através de tarefas que não exigem qualquer tipo de explicação verbal ou nomeação de figuras. Assim, em Categorias, desenhos devem ser organizados em dois grupos de acordo com a categoria a que pertencem. Em Situações, partes faltantes de uma figura ou situação devem ser completadas. No subteste Mosaicos, seqüências mosaicas são copiadas numa moldura. Por fim, em Padrões, formas devem ser copiadas utilizando lápis. Exemplos dos itens dos quatro subtestes podem ser encontrados no *website* dos testes SON ([www.testresearch.nl](http://www.testresearch.nl)).

*WPPSI-III (Wechsler Preschool and Primary Scale of Intelligence – Third Edition; Wechsler, 2002a)*. Como ainda não foi normatizado para o Brasil, utilizou-se uma versão traduzida do WPPSI-III. O teste é composto por 14 subtestes: 7 Verbais, 5 de Execução e 2 subtestes de Velocidade de Processamento. O teste é subdividido por faixa etária: para crianças com idades de 2 anos e 6 meses a 3 anos e 11 meses e para crianças com idade de 4 anos e 0 mês a 7 anos e 3 meses. Há três tipos de subtestes para as duas faixas etárias: Essenciais, Suplementares e Opcionais. Foram aplicados os subtestes necessários para a obtenção do QI Verbal, de Execução e Geral. As crianças com menos de 4 anos responderam os seguintes subtestes: Cubos, Informação, Armar objeto e Vocabulário Receptivo. Já as crianças com idades entre 4 e 7 anos responderam a: Cubos, Informação, Raciocínio Matricial, Conceitos Pictóricos, Raciocínio Verbal, Vocabulário e Código.

*WISC-III (Escala de Inteligência Wechsler para Crianças; Wechsler, 2002b)*. Esse teste foi desenvolvido para uso com crianças entre 6 anos e 0 mês e 16 anos e 11 meses. O WISC-III foi traduzido e adaptado para o Brasil por Figueiredo (2001) a partir de uma pesquisa realizada com 801 crianças de escolas do estado do Rio Grande do Sul. O teste é composto por 13 subtestes, sendo um conjunto de subtestes verbais e outro de subtestes de execução. Foram aplicados os subtestes necessários para o cálculo do QI Verbal, QI de Execução e QI Geral. Dessa forma, as crianças responderam aos seguintes subtestes: Completar Figuras, Informação, Código, Semelhanças, Arranjo de Figuras, Aritmética, Cubos, Vocabulário, Armar Objetos e Compreensão.

### Procedimento

A aplicação ocorreu entre o segundo semestre de 2007 e o final do ano de 2008. A aplicação foi realizada em escolas, creches e na casa das crianças. As escolas e creches foram selecionadas por conveniência e as crianças foram selecionadas de acordo com o critério de idade dos testes. Foram enviados aos pais Termos de Consentimento e uma carta de explicação sobre a pesquisa. Aqueles que concordavam com a participação de seu filho devolviam o Termo de Consentimento assinado e receberam após a pesquisa um relatório com os resultados da criança.

Para a aplicação dos testes, as crianças eram retiradas de sala no horário de aula, sendo proibido que o procedimento ocorresse durante o intervalo, pois poderia acarretar em uma falta de cooperação por parte da criança. Nos casos das aplicações realizadas na casa da criança, marcava-se um horário conveniente para o aplicador, os pais e a criança. Todas as aplicações foram individuais.

Em um dia era aplicado o SON-R 2½-7[a], com duração de aproximadamente 30 minutos. Em outro dia aplicava-se ou o WPPSI-III ou o WISC-III, a aplicação desses testes duravam aproximadamente 45 minutos. Em apenas 8 casos foi necessário aplicar ambos os testes em um único dia. Entre os motivos para essa mudança no procedimento estão: o aplicador e o avaliado não residirem na mesma cidade, fim do ano letivo na escola e dificuldades de encontrar a criança.

Em média o intervalo entre a aplicação do SON-R 2½-7[a] e a aplicação do WPPSI-III foi de 3 semanas. Entre as crianças que responderam ao WISC-III, o intervalo entre as duas aplicações foi em média de 6 semanas.

### Análise de Dados

A correção dos testes e a conversão dos escores brutos em escores normatizados se basearam nas normas e informações presentes em seus manuais (Laros et al., in press; Wechsler, 2002a, 2002b). Os escores normatizados dos testes de inteligência são expressos em uma escala com média 100 e desvio padrão 15 e o dos subtestes, em uma escala com média 10 e desvio padrão 3. É importante ressaltar que o teste WPPSI-III foi corrigido com base nas normas dos Estados Unidos, uma vez que até o período de elaboração deste artigo, o teste ainda não havia sido normatizado para o Brasil.

Inicialmente, foram realizadas análises exploratórias e descritivas para avaliar a consistência dos dados e características dos testes e da amostra. A análise de fidedignidade foi realizada com base no coeficiente *Lambda 2* de Guttman. Esse coeficiente foi escolhido porque estudos apontam que o *Lambda 2* de Guttman é uma estimativa melhor da fidedignidade, principalmente quando os instrumentos contêm poucos itens ou quando a amostra é pequena (Tellegen & Laros, 2004; Ten Berge & Zegers, 1978).

A fidedignidade do subteste Código do WISC-III e do WPPSI-III, que afere velocidade de processamento, foi analisada por meio do método das duas metades e coeficiente Spearman Brown. Este procedimento foi adotado

porque não é correto medir a fidedignidade de testes de velocidade utilizando o método da consistência interna (Nunnally & Bernstein, 1994). Não foram analisados os subtestes que apenas as crianças de 3 anos respondiam, pois o número restrito de crianças dessa faixa etária ( $N=10$ ) inviabilizaria uma análise adequada da fidedignidade.

### Resultados e Discussão

Na Tabela 2 é apresentado a média, o desvio-padrão, o erro padrão da média e o intervalo de confiança de 95% das médias das crianças que responderam ao SON-R 2½-7[a] e ao WISC-III. Nota-se que as médias no teste SON-R 2½-7[a] são menores do que as obtidas no WISC-III. Contudo, a análise dos intervalos de confiança evidencia que há intersecção entre os intervalos das diversas escalas, o que indicaria não existir diferenças significativas entre as médias.

Deve-se considerar também a influência do Efeito Flynn, que aponta a tendência de uma elevação nos esco-

res brutos em testes de inteligência (Flynn, 2000). De acordo com Flynn (2009), a média populacional aumenta cerca de 0,3 pontos de QI a cada ano. Por isso, faz-se necessários estudos de normatização periódicos dos testes de inteligência, pois a cada normatização o escore médio é colocado novamente em 100, de modo que o aumento devido ao efeito Flynn desaparece de vista. Em suma, de acordo com o efeito Flynn, se passados 10 anos desde a última vez em que o teste foi normatizado, as pessoas que agora obtêm um escore de 100 na nova versão, provavelmente obteriam um escore cerca de 103 na versão antiga (Neisser, 1998). Como o WISC-III foi normatizado entre 7 a 8 anos antes do SON-R 2½-7[a], uma diferença em torno de 2,1 a 2,4 pontos entre as médias obtidas nos dois testes era esperada.

Os valores encontrados de desvio-padrão estão dentro dos valores esperados, considerando que a escala normatizada tem como base média e desvio-padrão iguais a 100 e 15, respectivamente. Com relação ao erro padrão da média, observa-se erros mais altos nas escalas do WISC-III em comparação ao SON-R 2½-7[a].

Tabela 2

*Média, Desvio-padrão, Erro Padrão da Média e Intervalo de Confiança de 95% das Médias das Crianças no SON-R 2½-7[a] e no WISC-III [N=59]*

	M	DP	EP	IC (95%)	
				Inferior	Superior
SON-R QI Geral	99,34	12,46	1,62	96,09	102,59
SON-R ER	98,49	13,64	1,78	95,24	101,75
SON-R EE	100,46	12,48	1,63	96,90	104,01
WISC-III QI Geral	103,71	13,77	1,79	100,12	107,30
WISC-III QI Verbal	102,81	14,75	1,92	98,97	106,66
WISC-III QI de Execução	102,66	16,84	2,19	98,27	107,05

*Nota.* DP = desvio-padrão; EP = erro-padrão da média; IC95% = intervalo de confiança de 95% da média; SON-R ER = Escala de Raciocínio do SON-R; SON-R EE = Escala de Execução do SON-R.

Tabela 3

*Média, Desvio-padrão, Erro Padrão da Média e Intervalo de Confiança de 95% das Médias das Crianças no SON-R 2½-7[a] e no WPPSI-III [N=49]*

	M	DP	EP	IC (95%)	
				Inferior	Superior
SON-R QI Geral	106,98	15,56	2,22	102,51	111,45
SON-R ER	106,65	14,60	2,09	102,46	110,85
SON-R EE	105,86	16,67	2,38	101,07	110,65
WPPSI-III QI Geral	93,18	14,67	2,10	88,97	97,40
WPPSI-III QI Verbal	90,27	12,28	1,75	86,74	93,79
WPPSI-III QI de Execução	95,39	15,09	2,16	91,05	99,72

*Nota.* DP = desvio-padrão; EP = erro-padrão da média; IC (95%) = Intervalo de Confiança de 95% da média; SON-R ER = Escala de Raciocínio do SON-R; SON-R EE = Escala de Execução do SON-R.

Na Tabela 3, são apresentadas as mesmas análises da Tabela 2 para as crianças que responderam o SON-R 2½-7[a] e o WPPSI-III. Diferentemente da análise anterior, as médias das crianças no SON-R 2½-7[a] foram superiores às do WPPSI-III e o intervalo de confiança entre as diversas escalas não se intersectam. Esse resultado era esperado, pois o WPPSI-III ainda não foi normatizado no Brasil e foram utilizadas as normas dos EUA.

Antes da normatização do SON-R 2½-7[a] no Brasil, foi realizado um estudo piloto com crianças do Distrito Federal e os escores foram analisados de acordo com as normas holandesas (Laros et al., in press). Observou-se que os QIs das crianças brasileiras eram cerca de 15 pontos abaixo da média holandesa, padrão semelhante ao encontrado nas

crianças que responderam ao WPPSI-III neste estudo. Ou seja, as crianças brasileiras, em média, quando comparadas a crianças estadunidenses e holandesas apresentam um desvio-padrão de desempenho inferior.

A Tabela 4 apresenta os valores dos coeficientes de fidedignidade de todos os subtestes e os respectivos números de itens. Ao considerar os valores do coeficiente *Lambda* 2 de Guttman torna-se evidente que a maioria dos subtestes apresenta fidedignidade satisfatória, entre 0,70 e 0,90. Há apenas um subteste com fidedignidade muito elevada (informação – WPPSI-III) e quatro subtestes com fidedignidade baixa (menor que 0,70). Os valores de fidedignidade mais baixos foram observados para os subtestes do WISC-III.

Tabela 4

Coeficientes de Fidedignidade e Número de Itens por Subteste do SON-R 2½-7[a], WISC-III e WPPSI-III

Teste	Subteste	nº Itens	Fidedignidade
SON-R 2½-7[a][N=108]	Mosaicos	15	0,89
	Categorias	15	0,87
	Situações	14	0,85
	Padrões	16	0,88
WISC-III [N=59]	Arranjo de Figuras	13	0,78
	Completar Figuras	27	0,83
	Informação	18	0,68
	Códigos	59	0,65
	Semelhanças	15	0,70
	Aritmética	16	0,85
	Cubos	11	0,81
	Vocabulário	19	0,71
	Armar Objetos	5	0,66
Compreensão	14	0,64	
WPPSI-III[N=49]	Blocos	18	0,89
	Informação	32	0,95
	Raciocínio Matricial	22	0,90
	Vocabulário	18	0,79
	Conceitos Pictóricos	27	0,89
	Raciocínio Verbal	27	0,90
	Códigos	59	0,74

*Nota.* A análise de fidedignidade foi realizada com base no coeficiente *Lambda* 2 de Guttman; Alguns itens foram excluídos dos subtestes devido à ausência de variância; Para o cálculo da fidedignidade do subteste Armar Objetos foram desconsiderados os dois primeiros itens do teste por falta de variabilidade, uma vez que todas as crianças acertaram esses itens.

Para a maioria dos subtestes, os valores encontrados estão muito próximos aos valores dos seus respectivos estudos de normatização (Laros et al., in press; Wechsler, 2002a, 2002b). Ressalta-se que no subteste Armar Objetos do WISC-III os dois primeiros itens não foram incluídos para ao cálculo da fidedignidade devido à falta de variabilidade, pois esses itens foram acertados por todas as crianças avaliadas.

A validade convergente do SON-R 2½-7[a] foi verificada por meio da correlação com o WISC-III e

WPPSI-III. Como apresentado na Tabela 5, nota-se que as correlações do SON-R 2½-7[a] com as escalas do WPPSI-III foram mais altas, variando de 0,48 a 0,78, do que as correlações com as escalas do WISC-III, cujas correlações variam entre 0,38 e 0,69. Esse resultado pode ser explicado pela decisão na última revisão do WPPSI de focalizar mais na avaliação da inteligência fluida, assim como o SON-R 2½-7[a]. No manual do WPPSI-III é citado que em decorrência de diversas teorias do funcionamento cognitivo enfatizar a medida da inteligência flui-

da, na revisão do WPPSI-III buscou-se implementar e aumentar o número de subtestes que avaliam este domínio da inteligência (Wechsler, 2002a).

As menores correlações foram entre as escalas verbais do WISC-III e WPPSI-III com a escala de execução do SON-R 2½-7[a], iguais a 0,38 e 0,48, respectivamente. De modo geral, as maiores correlações ocorreram entre as escalas de execução e completa dos testes Wechsler e o QI total do SON-R 2½-7[a]. As correlações do QI total do SON-R 2½-7[a] com as escalas de execução do WISC-

III e WPPSI-III foram respectivamente, 0,65 e 0,73 e com as escalas completas foram 0,69 e 0,75.

Esses resultados corroboram com o esperado. Considerando que o SON-R 2½-7[a] é um teste não-verbal de inteligência, eram esperadas correlações mais baixas com as escalas verbais do que em relação às escalas de execução. Todavia, correlações moderadas do teste SON-R 2½-7[a] com as escalas verbais do WISC-III e WPPSI-III são previstas porque é necessário considerar que tanto as escalas verbais quanto o SON-R 2½-7[a] buscam mensurar o mesmo construto geral: inteligência.

Tabela 5

*Correlações entre as Escalas do SON-R 2½-7[a] com as Escalas WISC-III e WPPSI-III*

Escalas do WISC-III	Escalas do SON-R 2½-7[a]		
	SON-R EE	SON-R ER	SON-R QI
WISC-III QI Total	0,57	0,61	0,69
WISC-III QI Verbal	0,38	0,51	0,52
WISC-III QI de Execução	0,53	0,59	0,65
Escalas do WPPSI-III	SON-R EE	SON-R ER	SON-R QI
WPPSI-III QI Total	0,59	0,78	0,75
WPPSI-III QI Verbal	0,48	0,73	0,66
WPPSI-III QI de Execução	0,61	0,70	0,73

*Nota.* SON-R EE = Escala de Execução do SON-R; SON-R ER = Escala de Raciocínio do SON-R; SON-R QI = QI total do SON-R 2½-7[a]. Correlações significativas ao nível de significância de 0,01.

Também foram analisadas as correlações entre os subtestes que compunham cada um dos testes (Tabelas 6 e 7). A análise das correlações aponta que as baixas correlações entre as escalas verbais dos testes Wechsler e a escala de execução do SON-R 2½-7[a] estão mais relacionadas com o subteste Mosaicos, uma vez que as correlações com o subteste Padrões são de moderadas a alta, sobretudo com o WPPSI-III.

Os subtestes Mosaicos e Padrões do SON-R 2½-7[a] apresentaram correlações elevadas com o subteste Cubos do WISC-III. Tais correlações eram esperadas, uma vez que em Cubos a criança precisa reproduzir padrões geométricos usando cubos de duas cores, tarefa muito semelhante às exigidas em Mosaicos e Padrões. Ademais, todos são subtestes que avaliam raciocínio espacial.

Categorias e Situações mostraram uma relação mais forte com Vocabulário e Arranjo de Figuras do WISC-

Tabela 6

*Correlações entre os Subtestes do SON-R 2½-7[a] com os Subtestes do WISC-III*

WISC-III	SON-R 2½-7[a]			
	Mosaicos	Categorias	Situações	Padrões
Completar Figuras	0,40**	0,44**	0,43**	0,44**
Informação	0,17	0,24	0,14	0,18
Código	0,19	0,34**	0,35**	0,23
Semelhanças	0,38**	0,37**	0,44**	0,25
Arranjo de Figuras	0,38**	0,36**	0,49 <sup>a</sup> **	0,38**
Aritmética	0,28*	0,42**	0,30*	0,56 <sup>a</sup> **
Cubos	0,50 <sup>a</sup> **	0,36**	0,26*	0,53 <sup>a</sup> **
Vocabulário	0,27*	0,51 <sup>a</sup> **	0,32*	0,40**
Armar Objetos	0,38**	0,21	0,28*	0,40**
Compreensão	0,31*	0,38**	0,25	0,21

<sup>a</sup> Correlações mais altas entre os subtestes do SON-R 2½-7[a] e do WISC-III. \* = correlações significativas a 0,05; \*\* = correlações significativas a 0,01.

Tabela 7  
Correlações entre os Subtestes do SON-R 2½-7[a] com os Subtestes do WPPSI-III

WPPSI-III	SON-R 2½-7[a]			
	Mosaicos	Categorias	Situações	Padrões
Cubos	0,56 <sup>a</sup> **	0,29*	0,54 <sup>a</sup> **	0,38*
Informação	0,38*	0,47**	0,48**	0,48 <sup>a</sup> **
Raciocínio Matricial	0,37*	0,48**	0,43**	0,39*
Vocabulário	0,21*	0,50**	0,46**	0,23*
Conceitos Pictóricos	0,22*	0,68 <sup>a</sup> **	0,30*	0,34*
Raciocínio Verbal	0,25*	0,46**	0,57 <sup>a</sup> **	0,38*
Códigos	0,31*	0,28*	0,39*	0,25*

<sup>a</sup>Correlações mais altas entre os subtestes do SON-R 2½-7[a] e do WPPSI-III. \* = correlações significativas a 0,05; \*\* = correlações significativas a 0,01.

III, respectivamente. É interessante a ocorrência de uma correlação alta entre Vocabulário e Categorias, possivelmente o fato de ambos procurarem investigar o desenvolvimento de conceitos tenha corroborado com o resultado. Vale apontar que a correlação alta entre Situações e Arranjo de Figuras deve-se, provavelmente, ao fato de que ambos buscam medir raciocínio concreto por meio de figuras que mostram uma situação ou uma estória com seqüência lógica.

Em referência aos subtestes do WPPSI-III, Cubos e Conceitos Pictóricos são os subtestes que apresentaram as maiores correlações com Mosaicos e Padrões, respectivamente. Observa-se também que entre os subtestes da escala de raciocínio do SON-R 2½-7[a], Categorias apresenta correlações maiores com os subtestes do WPPSI-III do que Situações. Novamente, os subtestes cujas tarefas são mais similares, possuem correlações mais elevadas. Indicativos de que os testes estão mensurando construtos muito semelhantes.

### Considerações Finais

As análises realizadas neste estudo apontam a boa qualidade psicométrica do teste SON-R 2½-7[a] e mostram evidências de adequada validade convergente entre o teste SON-R 2½-7[a] e os testes Wechsler. Todas as hipóteses inicialmente levantadas foram confirmadas.

Nos estudos de validação convergente do SON-R 2½-7 com o WPPSI-R realizados na Austrália, Holanda e Estados Unidos da América verifica-se resultados semelhantes. Nesses estudos a correlação do QI total do SON-R 2½-7 variou entre 0,43 e 0,54 com a escala verbal, entre 0,59 e 0,74 com a escala de execução e entre 0,59 e 0,75 com a escala completa do WPPSI-R (J. ten Horn, 1996, Jenkinson et al., 1996; Tellegen, 1997; Tellegen et al., 1998). Em comparação com os resultados deste estudo, os valores de correlação estão dentro dos intervalos encontrados em outros países, com exceção da correlação com a escala verbal que foi levemente mais alta.

Com relação aos estudos entre o SON-R 2½-7 e o WISC-R, um estudo na Holanda com crianças especiais verificou correlação de 0,62 entre o QI total dos dois testes. Resultado semelhante ao encontrado no Brasil. Não foram encontrados estudos de validação internacionais entre o SON-R 2½-7 e o WISC-III. Os outros países priorizaram mais o teste WPPSI do que o WISC, uma vez que o WPPSI atende crianças com idades mais próximas às do SON-R. No Brasil, no entanto, devido ao fato do WPPSI ainda não ser validado e normatizado para a população brasileira, optou-se por realizar o estudo com ambos os testes.

Além do uso de um teste ainda não validado para o Brasil, outra limitação deste estudo é o quantitativo de crianças avaliadas. Os custos com a pesquisa de validação acabaram por limitar o número de crianças avaliadas. Essas limitações sugerem a realização de outras pesquisas buscando mais evidências de qualidade do teste, o que fortaleceria os achados aqui encontrados.

Não obstante, pode-se concluir com os resultados encontrados no presente estudo que há evidências de validade convergente do teste SON-R 2½-7[a] com dois importantes testes de inteligência utilizados no mundo: WISC-III e WPPSI-III. Tais indicativos são fundamentais para uma adequada utilização e disseminação do teste, pois são somente a partir de evidências de validade que se podem realizar inferências confiáveis acerca do resultado de um teste.

### Referências

- Alves, I. C. B. (2002). Instrumentos disponíveis no Brasil para avaliação da inteligência. In R. Primi (Ed.), *Temas em avaliação psicológica* (pp. 80-102). Campinas, SP: Imprensa Digital do Brasil.
- Evers, A., Van Vliet-Mulder, J. C., & De Groot, J. (2000). *Documentation of tests and test research in The Netherlands*. Assen, Netherlands: Van Gorcum.
- Figueiredo, V. L. M. (2001). *Uma adaptação brasileira do teste de inteligência WISC-III*. Tese de Doutorado não-publicada, Programa de Pós-Graduação em Psicologia, Instituto de Psicologia, Universidade de Brasília, DF.



- Figueiredo, V. L. M. de, Mattos, V. L. D. de, Pasquali, L., & Freire, A. P. (2008). Propriedades psicométricas dos itens do teste WISC-III. *Psicologia em Estudo*, 13(3), 585-592.
- Flynn, J. R. (2000). IQ trends over time: Intelligence, race, and meritocracy. In K. Arrow, S. Bowles, & S. M. Durlauf (Eds.), *Meritocracy and economic inequality* (pp. 35-61). Princeton, NJ: Princeton University Press.
- Flynn, J. R. (2009). *O que é inteligência? Além do efeito Flynn*. Porto Alegre, RS: Artmed.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro, RJ: LTC.
- 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. ten (1996). *Amerikaanse validering van de Snijders-Oomen niet-verbale intelligentietest voor jonge kinderen, de SON-R 2½-7*. Groningen, Netherlands: Intern Verslag.
- Jenkinson, J., Roberts, S., Dennehy, S., & Tellegen, P. (1996). Validation of the Snijders-Oomen Nonverbal Intelligence Test – Revised 2½-7 for Australian Children with Disabilities. *Journal of Psychoeducational Assessment*, 14, 276-286.
- Jesus, G. R. de (2009). *Normalização e validação do teste não-verbal de inteligência SON-R 2½-7[a] para o Brasil*. Tese de Doutorado não-publicada, Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações, Instituto de Psicologia, Universidade de Brasília, DF.
- Laros, J. A., Reis, R. F., & Tellegen, P. J. (2010). Indicações da validade convergente do teste não-verbal de inteligência SON-R 2½-7[a]. *Avaliação Psicológica*, 9(1), 43-52.
- Laros, J. A., Tellegen, P. J., Jesus, G. R. de, & Karino, C. A. (in press). *SON-R 2½-7[a], Teste não-verbal de inteligência. Manual com normalização e validação brasileira*.
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4<sup>th</sup> ed.). New York: Oxford University Press.
- Muñiz, J., Prieto, G., Almeida, L., & Bartram, D. (1999). Test use in Spain, Portugal and Latin American Countries. *European Journal of Psychological Assessment*, 15, 151-157.
- Neisser, U. (1998). *The rising curve: Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.
- Noronha, A. P. P., & Alchieri, J. C. (2002). Reflexões sobre os instrumentos de avaliação psicológica. In R. Primi (Ed.), *Temas em avaliação psicológica* (pp. 7-16). Campinas, SP: Imprensa Digital do Brasil.
- Noronha, A. P. P., Primi, R., & Alchieri, J. C. (2004). Parâmetros psicométricos: Uma análise dos testes psicológicos comercializados no Brasil. *Psicologia: Ciência e Profissão*, 24, 88-99.
- Noronha, A. P. P., & Vendramini, C. M. M. (2003). Parâmetros psicométricos: Estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16(1), 177-182.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3<sup>rd</sup> ed.). New York: McGraw-Hill.
- Oakland, T., Wechsler, S., Bensuan, E., & Stafford, M. (1994). The construct of intelligence among Brazilian children: An exploratory study. *School Psychology International*, 15, 361-370.
- Primi, R. (2002). Inteligência fluida: Definição fatorial, cognitiva e neuropsicológica. *Padeia*, 12, 57-77.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen*. Nijmegen, Netherlands: Berkhout.
- Tellegen, P. (1997). An addition and correction to the Jenkinson et al. (1996) Australian SON-R 2½-7 validation study. *Journal of Psychoeducational Assessment*, 15, 67-68.
- Tellegen, P. J. (1993, November). *A nonverbal alternative to the Wechsler scales: The Snijders-Oomen Nonverbal Intelligence Tests*. Paper presented at the First Annual South Padre Island International Conference on Cognitive Assessment of Children and Youth in School and Clinical Settings, San Antonio, TX.
- Tellegen, P. J., & Laros, J. A. (2004). Cultural bias in the SON-R test: Comparative study of Brazilian and Dutch children. *Psicologia: Teoria e Pesquisa*, 20(2), 103-111.
- Tellegen, P. J., & Laros, J. A. (2005). Fair assessment of children from cultural minorities: A description of the SON-R nonverbal intelligence tests. In D. Kopcanová (Ed.), *Quality education for children from socially disadvantaged settings* (pp. 50-73). Bratislava, Slovakia: Commission for United Nations Educational, Scientific and Cultural Organization.
- Tellegen, P. J., Laros, J. A., & Heider, D. (2008). *SON-R 2½-7, Neverbalni inteligencni test: Prirucka*. [SON-R 2½-7, nonverbal test of intelligence: Manual]. Praha, Czech Republic: Testcentrum-Hogrefe.
- Tellegen, P. J., Laros, J. A., & Petermann, F. (2007). *SON-R 2½-7, Non-verbaler intelligenz-test: Testmanual mit deutscher normierung und validierung*. Göttingen, Germany: Hogrefe.
- Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. J., & Laros, J. A. (1998). *Snijders-Oomen Nonverbal Intelligence Test, SON-R 2½-7, Manual & Research Report*. Lisse, Netherlands: Swets & Zeitlinger.
- Ten Berge, J. M. F., & Zegers, F. E. (1978). A serie of lower bounds to the reliability. *Psychometrika*, 43(4), 575-579.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition Manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2002a). *WPPSI-III – Technical and interpretive manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (Ed.). (2002b). *WISC-III – Escala de Inteligência Wechsler para Crianças: Manual* (3. ed.). São Paulo, SP: Casa do Psicólogo.

Recebido: 03/05/2010  
1ª revisão: 20/07/2010  
Aceite final: 23/07/2010