



Universidade de Brasília

Instituto de Psicologia

Departamento de Psicologia Social e do Trabalho

Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações

EVIDÊNCIAS DE VALIDADE E PROPRIEDADES PSICOMÉTRICAS DO SON-R 6-40

Talita de Araújo Alves

Brasília, DF

Fevereiro de 2016



Universidade de Brasília

Instituto de Psicologia

Departamento de Psicologia Social e do Trabalho

Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações

EVIDÊNCIAS DE VALIDADE E PROPRIEDADES PSICOMÉTRICAS DO SON-R 6-40

Talita de Araújo Alves

Orientador: Jacob Arie Laros

Dissertação apresentada ao Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília, como requisito parcial à obtenção do título de Mestre em Psicologia.

Brasília, DF

Fevereiro de 2016

Universidade de Brasília
Instituto de Psicologia
Departamento de Psicologia Social e do Trabalho
Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações

Banca Examinadora:

Prof. PhD. Jacob Arie Laros (*Orientador*)
Universidade de Brasília - UnB

Prof^a. Dr^a. Cristiane Faiad (*Membro*)
Universidade de Brasília - UnB

Prof. Dr. Josemberg Moura de Andrade (*Membro*)
Universidade Federal da Paraíba - UFPB

Prof^a. Dr^a. Raquel Carvalho Hoersting (*Suplente*)
Universidade de Brasília - UnB

À minha família,

Às pessoas com deficiência,

Aos que são e serão beneficiados com os testes SON.

“A coisa mais indispensável a um homem é reconhecer o uso que deve fazer do seu próprio conhecimento.”

Platão.

AGRADECIMENTOS

À minha família, por todo suporte e exemplo. Mãe e pai, obrigada por serem reais exemplos de força, companheirismo, humildade e benevolência. Weverton e Carina, obrigada pelo cuidado e torcida. Miguel, meu pequeno e lindo sobrinho, obrigada por tornar nossos dias mais alegres. Tom, obrigada pelos momentos de carinho!

Ao meu orientador e amigo, professor Jacob Arie Laros, por quem tenho profunda admiração acadêmica e pessoal. MUITÍSSIMO obrigada por tudo que aprendi nos últimos anos, por todos os conselhos e suporte acadêmico. Em especial, agradeço por sua postura paciente e solícita, além da companhia sempre muito agradável!

Ao Hugo, pelo carinho e compreensão nesta reta final. Obrigada pelo incentivo, cuidado e assistência.

Ao Fábio William, meu grande conselheiro e amigo. Eu só posso agradecê-lo pelas palavras e pela força. Você fez e faz toda a diferença em minha caminhada (...).

Aos membros da banca, prof. Josemberg, profa. Cristiane e profa. Raquel. Obrigada por tão prontamente aceitarem o convite e pelas sugestões no trabalho.

Às pessoas que foram importantes desde o começo: Felipe Valentini, Natália e Gabriel. Obrigada por tudo que aprendi com vocês.

À profa. Cláudia Fukuda, pelo incentivo fundamental para o ingresso no mestrado. Sob sua orientação, me tornei psicóloga, sou muito grata e feliz por isso.

A todos os meus amigos do PSTO, por todos os momentos de troca e aprendizado. Em especial, agradeço aos amigos: Gina, Hannah, Renata, Jesselyn e Francisco. MUITÍSSIMO obrigada pelo suporte nos momentos difíceis e por todos os encontros de amizade. Vocês tornaram essa jornada muito mais especial!

Aos companheiros do Laboratório META, Gina, Renata, Luiz, Alexandre e Camila. Obrigada por todo auxílio e postura aberta para as trocas acadêmicas.

À Thais Vaz e Bianca Rogoski. MUITÍSSIMO obrigada pela forma organizada e eficaz com que vocês realizaram a coleta do segundo estudo. Um obrigada, também especial, para a Renata, que foi importantíssima na primeira coleta.

Aos psicólogos, Lívia, Cecília e Enildo. Muito obrigada pela abertura e pela confiança. Eu tenho grande admiração e respeito pelo trabalho de vocês!

Aos meus amigos, Allan, Layane, Noemi, Geraldo, Darlene e sua família. Obrigada pela compreensão nos momentos de ausência, pela admiração e pela maneira singular como cada um de vocês me ajudou nos últimos anos.

A todos os professores do PSTO, em especial aos professores, Laros, Juliana, Fábio, Pasquali, Kátia e Cristiane. Obrigada pelos ensinamentos e suporte!

Aos funcionários, Línea, Thiago, Lucas, Railson e Jheniffer. Obrigada por sempre atenderem minhas demandas de maneira tão eficiente e prestativa.

Aos meus alunos de TEP, pelo semestre de aprendizado mútuo.

À CAPES e ao Fundo-SON, pelo apoio financeiro para as pesquisas.

Sumário

LISTA DE TABELAS.....	9
RESUMO GERAL.....	10
GENERAL ABSTRACT	11
APRESENTAÇÃO.....	12
MANUSCRITO 1.....	15
MANUSCRITO 2.....	40
CONSIDERAÇÕES GERAIS	73
ANEXO 1	76
ANEXO 2	78
ANEXO 3	80

Lista de Tabelas

Manuscrito 1. Adequação do SON-R 6-40 para pessoas com deficiência intelectual

Tabela 1. Distribuição dos escores no SON-R 6-40 ($n = 50$)	29
Tabela 2. Relação entre o QI-SON e a avaliação do aplicador	30
Tabela 3. Cargas fatoriais (CF) e a comunalidade média (h^2) depois da Análise Fatorial Exploratória do SON-R 6-40 ($n = 50$)	32
Tabela 4. Índices de fidedignidade do SON-R 6-40	33

Manuscrito 2. Estudos da validade convergente do SON-R 6-40 com a EMMC, TONI-3 e MPC

Tabela 1. Amplitude e média das cargas fatoriais depois da rotação <i>Promin</i> e valores de comunalidade (h^2) para a solução fatorial do SON-R 6-40 ($n = 150$)	56
Tabela 2. Amplitude e média das cargas fatoriais no fator único e valores de comunalidade (h^2) para a solução fatorial encontrada nos testes: EMMC, MPC e TONI -3	58
Tabela 3. Cargas fatoriais de acordo com a solução <i>Schmid-Leiman</i> nos fatores de primeira ordem e segunda ordem para o Fator Espacial e o Fator de Raciocínio do SON-R 6-40 e os testes MPC, EMMC e TONI-3	59
Tabela 4. Média, desvio-padrão, erro padrão da média e intervalo de confiança de 95% das médias obtidas para os escores normatizados no SON-R 6-40, MPC, EMMC e TONI-3	60
Tabela 5. Propriedades psicométricas dos subtestes do SON-R 6-40 e dos escores totais do SON-R 6-40, EMMC, MPC e TONI-3	62
Tabela 6. Coeficientes de correlação de Pearson e coeficientes de determinação para os escores totais normatizados	64
Tabela 7. Correlações corrigidas e brutas entre os dois fatores do SON-R 6-40 e os testes MPC, EMMC e TONI-3	65

RESUMO GERAL

Os testes de inteligência são ferramentas muito empregadas em diversos contextos de atuação do psicólogo. Assim, estes instrumentos necessitam de aprimoramento, revisão e atualização, a fim de garantir a qualidade e a ética dos serviços prestados em psicologia. Considerando o impacto que os resultados alcançados em testes psicológicos podem ter sobre a vida dos sujeitos, são exigidos altos padrões de qualidade psicométrica dos instrumentos. O objetivo primordial dos dois manuscritos da presente dissertação foi fornecer informações acerca da adequação do teste não-verbal de inteligência SON-R 6-40 para grupos especiais e da validade dos seus escores. Para obter evidências da adequação do SON-R 6-40 para grupos especiais, o teste foi aplicado em um grupo de pessoas com o diagnóstico de deficiência intelectual ($n = 50$). Durante a aplicação, os examinadores realizaram uma avaliação da motivação, concentração, cooperação e compreensão das instruções dos participantes. As evidências de validade convergente foram obtidas a partir da aplicação do SON-R 6-40 com três outros testes não-verbais de inteligência para um grupo de estudantes do ensino básico ($n = 150$). Adicionalmente, foram analisadas as diferenças nos escores dos quatro testes para os diversos grupos. Os resultados do primeiro estudo indicam a adequação do SON-R para a avaliação das habilidades cognitivas de pessoas com deficiência intelectual. O segundo estudo apresentou coeficientes de fidedignidade altos para os escores do SON-R 6-40, e coeficientes satisfatórios de validade convergente com os outros instrumentos administrados. Em ambos os estudos foi encontrada uma solução bifatorial para o SON-R 6-40. Em geral, os resultados dos dois estudos são condizentes com pesquisas realizadas anteriormente com o instrumento. Ressalta-se que o SON-R 6-40 é um instrumento promissor para a avaliação das habilidades cognitivas no Brasil.

Palavras-chave: avaliação psicológica; teste psicológico; medidas de inteligência.

GENERAL ABSTRACT

Intelligence tests are frequently used tools in a wide variety of contexts in the daily practice of many psychologists. In order to guarantee the quality and ethical integrity of the services rendered in psychology these instruments need constant refinement, revision and updating. Considering the impact that the results achieved on psychological tests may have for the personal situation of subjects, high standards regarding their psychometric quality should be required. The main objective of the two manuscripts that are part of this dissertation was to supply information concerning the adequacy of the SON-R 6-40, a non-verbal test of intelligence, for special groups and on the score validity of the test. To obtain evidence of the adequacy of the SON-R 6-40 for special groups, the test was administered to a group of people with a diagnosis of intellectual disability ($n = 50$). During application, examiners conducted an evaluation of the motivation, concentration, cooperation and comprehension of instructions of the participants. Evidence of the convergent validity of the SON-R 6-40 was acquired by applying the SON-R 6-40 together with three other nonverbal intelligence tests to a group of basic education students ($n = 150$). The factor structure of the instruments, the psychometric properties of the test scores, the reliability coefficients and the convergent validity were analyzed. Additionally, differences between various groups of test scores on the four instruments were analyzed. The results of the first study indicate the adequacy of the SON-R 6-40 to assess the cognitive abilities of subjects with intellectual disability. The second study showed high reliability coefficients of the scores on the SON-R 6-40, and satisfactory convergent validity coefficients with the other administered tests of intelligence. In both studies a bi-factor solution of the SON-R 6-40 was found. In general, the results of the two studies are in accordance with previously realized studies with the test. It should be highlighted that the SON-R 6-40 is a promising instrument for the assessment of cognitive abilities in Brazil.

Keywords: psychological assessment; psychological test; intelligence test.

Apresentação

No Brasil, os psicólogos têm à disposição um número restrito de testes padronizados que avaliam a inteligência. Neste contexto, se destacam tanto instrumentos com tarefas verbais quanto a crescente disponibilização de testes não-verbais para a avaliação da inteligência (CFP, 2016). Os testes não-verbais de inteligência têm grandes vantagens se comparados com testes que fazem apelo à linguagem, tais como os testes de inteligência geral de Wechsler, como o WISC-IV para crianças (Wechsler, 2013) e o WAIS-III (Wechsler, 2011) para adultos. Os testes não-verbais de inteligência são mais adequados para pessoas com problemas ou necessidades especiais na linguagem, fala ou comunicação, com deficiência auditiva, autismo, com transtornos de desenvolvimento social e para pessoas analfabetas ou semi-analfabetas. Para estes grupos, um desempenho inferior em testes verbais de inteligência pode simplesmente representar um conhecimento verbal pobre, em vez de uma habilidade limitada de raciocínio intelectual (Tellegen & Laros, 2005).

Em primeiro lugar se destaca que mesmo nos casos em que a interpretação do desempenho no teste se restringe à escala de execução (a parte não-verbal), o problema levantado acima não está de todo superado, uma vez que as instruções destes testes são dadas verbalmente e podem ser menos inteligíveis para as pessoas com dificuldades relacionadas à linguagem (Tellegen, 1993). Em segundo lugar, os testes não-verbais são menos influenciados por escolarização, aculturação e desenvolvimento da linguagem em comparação com testes de inteligência geral que se baseiam fortemente em habilidades verbais (Tellegen & Laros, 1993). Além disso, para os testes não-verbais não há a necessidade de tradução dos itens para a língua portuguesa.

O SON-R 6-40 (*Snijders-Oomen Non-Verbal Intelligence Test – Revised*) é um instrumento não-verbal de inteligência para pessoas entre 6 e 40 anos (Tellegen & Laros, 2014). Atualmente o instrumento é utilizado em diversos países da Europa (Holanda,

Alemanha, Áustria, Suíça), na China e no Brasil. A pesquisa de normatização e validação do SON-R 6-40 encontra-se em andamento pelo Laboratório de Métodos e Técnicas de Avaliação - META da Universidade de Brasília, UnB. A pesquisa do SON-R 6-40 está sendo financiada pelo Fundo-SON, gerenciado pela *Rijksuniversiteit* Groningen, Holanda.

A presente dissertação objetiva reunir evidências da adequação e da validade dos escores do SON-R 6-40. Este trabalho está organizado em dois manuscritos. Os manuscritos são relatos de pesquisa referentes a dois diferentes estudos conduzidos com o instrumento SON-R 6-40.

O primeiro manuscrito refere-se à análise do SON-R 6-40 em um grupo de pessoas com o diagnóstico de deficiência intelectual. Neste estudo, objetivou-se obter indícios da adequação do instrumento para um grupo especial, contribuindo para o campo específico da avaliação das habilidades cognitivas nesta população. Para este grupo se destaca um comprometimento significativo das funções cognitivas e há carência de estudos que avaliem a adequação dos testes de inteligência.

O segundo manuscrito apresenta dados relativos às evidências de validade convergente do SON-R 6-40 com vários instrumentos utilizados na prática profissional no país. Os instrumentos escolhidos apresentam parecer favorável para uso profissional, sendo eles: a Escala de Maturidade Mental Colúmbia (EMMC), as Matrizes Progressivas Coloridas de Raven (MPC) e o teste TONI-3, forma A (*Test of Nonverbal Intelligence, Form A, Third Edition*). Além disso, foram avaliadas as propriedades psicométricas dos itens e dos escores nos instrumentos utilizados, e a estrutura fatorial dos testes. Assim, visou-se contribuir para o conhecimento acumulado sobre os testes já utilizados e para o uso e a interpretação dos escores do SON-R 6-40 no campo da avaliação psicológica.

Referências

- Conselho Federal de Psicologia (2016). *Sistema de Avaliação de Testes Psicológicos*. Retrieved from <http://satepsi.cfp.org.br/>.
- Tellegen, P. J. (1993). *A nonverbal alternative to the Wechsler scales: The Snijders –Oomen Nonverbal intelligence tests*. Paper presented at the First Annual South Padre Island Conference on Cognitive Assessment of Children and Youth in School and Clinical Settings. A Compendium of Proceedings (pp. 35-49).
- Tellegen, P. J. & Laros, J. A. (1993). The construction and validation of a nonverbal test of intelligence: The revision of the Snijders-Oomen tests. *European Journal of Psychological Assessment, 4*, 147-157.
- Tellegen, P. & Laros, J. A. (2005). Fair assessment of children from cultural minorities: A description of the SON-R non-verbal intelligence tests. In D. Kopcanova (Eds.), *Quality education for children from socially disadvantaged settings* (pp. 50-73). Bratislava, Slovakia: Commission of UNESCO.
- Tellegen, P. J., & Laros, J. A. (2014). *SON-R 6-40. Non-verbal intelligence test: Research report*. Göttingen, Germany: Hogrefe Verlag.
- 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.

MANUSCRITO 1

Adequação do SON-R 6-40 para pessoas com deficiência intelectual

Título em inglês

Adequacy of the SON-R 6-40 for people with intellectual disabilities

Título abreviado

O SON-R 6-40 e deficiência intelectual

RESUMO

O diagnóstico da deficiência intelectual é baseado na presença de restrições consideráveis no comportamento adaptativo e no funcionamento intelectual, com o início do transtorno antes dos dezoito anos. Os testes de inteligência são essenciais para a avaliação das habilidades cognitivas deste grupo. Neste contexto, o presente estudo teve por objetivo avaliar a adequação do teste não-verbal de inteligência SON-R 6-40 para pessoas com deficiência intelectual. O SON-R 6-40 consiste de quatro subtestes: Categorias e Situações (subtestes de raciocínio), Mosaicos e Padrões (subtestes de execução). O teste foi administrado em 50 pessoas com o diagnóstico de deficiência intelectual e com idades entre 15 e 39 anos. A média de QI para este grupo foi de 73 ($DP = 2,1$). A análise fatorial exploratória indicou dois fatores, o primeiro constituído dos subtestes de raciocínio e o segundo dos subtestes de execução. A média dos escores dos homens no fator de execução foi significativamente maior do que a média dos escores para as mulheres ($p < 0,01$; d de Cohen = 0,58). Os resultados obtidos dão suporte aos estudos prévios do SON-R 6-40 na Holanda com este grupo, e sugerem a adequação do teste para pessoas com deficiência intelectual.

Palavras-chave: deficiência intelectual; teste não-verbal de inteligência; SON-R 6-40

ABSTRACT

The diagnosis of intellectual disability is based on the presence of considerable limitations in adaptive behavior and intellectual functioning while the onset of the disability should occur before the age of 18. For the assessment of the cognitive abilities of this group intelligence tests are essential. In this context, the present study aimed to evaluate the adequacy of the SON-R 6-40, a non-verbal test of intelligence, for people with intellectual disability. The SON-R 6-40 consists of four subtests: Categories and Situations (reasoning subtests) and Mosaics and Patterns (performance subtests). The test was administered to 50 people with intellectual disability with ages between 15 and 39 years. For this group a mean IQ score of 73 ($SD = 2.1$) was observed. Exploratory factor analysis indicated two factors, the first being composed of the performance subtests, and the second of the reasoning subtests. The mean score of men on the performance factor was significantly higher than the mean score of women ($p < .01$; Cohen's $d = .58$). The obtained results give support to previous studies of the SON-R 6-40 in The Netherlands with this group, and suggest the adequacy of this test for people with intellectual disability.

Keywords: intellectual disability; non-verbal test of intelligence; SON-R 6-40

RESUMEN

El diagnóstico de deficiencia intelectual se basa en la presencia de considerables dificultades en el comportamiento adaptativo y en el funcionamiento intelectual, con el inicio del trastorno ocurrir antes de los dieciocho años. Los testes de inteligencia son esenciales para evaluar las habilidades cognitivas en este grupo. En ese contexto, el presente estudio tuvo por objetivo evaluar la adecuación del test no-verbal de inteligencia SON-R 6-40 en personas con deficiencia intelectual. El SON-R 6-40 consiste en cuatro sub-testes: Categorías y Situaciones (sub-testes de raciocinio) y Mosaicos y Padrones (sub-testes de ejecución). El test fue administrado a 50 personas con diagnóstico de deficiencia intelectual con edades entre 15 y 39 años. En este grupo fue observado un puntaje promedio de IQ de 73 ($DE = 2,1$). El análisis factorial exploratorio indicó dos factores, relacionados a los sub-testes de raciocinio y ejecución del SON-R 6-40. El promedio de los puntajes de los hombres para el factor de ejecución fue significativamente mayor que el promedio de los puntajes de las mujeres ($p < 0,01$; d de Cohen = 0,58). Los resultados obtenidos dan soporte a los estudios previos del SON-R 6-40 con ese grupo en Holanda e indican la adecuación del test en personas con deficiencia intelectual.

Palabras clave: deficiencia intelectual, teste no-verbal de inteligência, SON-R 6-40.

A deficiência intelectual é caracterizada por limitações no funcionamento cognitivo e no comportamento adaptativo dos sujeitos (Schalock et al., 2010). Considera-se que o surgimento dos déficits intelectuais e adaptativos ocorre até os 18 anos de idade (AAIDD, 2015). De acordo com o DSM-V a deficiência intelectual também pode ser denominada de Transtorno do Desenvolvimento Intelectual e pode ser classificada de acordo com o nível de gravidade apresentada: leve, moderada, grave e profunda (Associação Americana de Psiquiatria, 2014). Este tipo de deficiência engloba diferentes condições que variam desde distúrbios metabólicos e genéticos até modificações funcionais em decorrência de traumas no sistema nervoso no nascimento ou traumas cerebrais que podem ocorrer mais tardiamente no período de desenvolvimento (Harris, 2006).

O comportamento adaptativo se refere a um conjunto de habilidades de ordem prática, social e conceitual que fazem parte do processo de aprendizado do sujeito e são desempenhadas no cotidiano (AAIDD, 2015). A alfabetização, a linguagem e os conceitos numéricos são exemplos de habilidades conceituais. Dentre as habilidades sociais destacam-se as habilidades interpessoais, as responsabilidades sociais, a autoestima, a culpabilidade, a resolução de problemas sociais e o seguimento das normas e regras. As habilidades práticas são atividades do dia a dia como o auto cuidado, habilidades ocupacionais, de cuidados com a saúde, mobilidade, planejamento, segurança, manipulação de dinheiro e comunicação (APA, 2015).

O funcionamento cognitivo da pessoa com deficiência intelectual apresenta déficits no raciocínio, planejamento, aprendizado, resolução de problemas, pensamento abstrato e juízo. Alguns componentes são críticos, tais como a compreensão verbal, a memória de trabalho, o raciocínio perceptivo, o raciocínio quantitativo e a eficiência cognitiva. De acordo a Associação Americana de Psiquiatria (2014), os indivíduos com deficiência intelectual apresentam escores cerca de dois ou mais desvios-padrão abaixo da média da população. Em

testes de inteligência normatizados com média 100 e desvio-padrão 15, os indivíduos com deficiência intelectual apresentam aproximadamente escores entre 65 e 75. De acordo com a Pesquisa Nacional de Saúde realizada pelo IBGE (2013) no Brasil estima-se que 6,2% da população total possui algum tipo de deficiência, dentre elas: a deficiência física, intelectual, visual e auditiva. Do total de pessoas que possuem deficiência intelectual, mais da metade ou 54,8% apresentam um grau intenso ou muito intenso de limitação. Ressalta-se ainda que apenas 30,4 % das pessoas com deficiência intelectual frequentam algum tipo de serviço de reabilitação. Neste contexto, considera-se a importância da identificação da pessoa com deficiência intelectual enquanto passo inicial do processo que culmina na prestação de serviços e suporte adequados para esta realidade.

O diagnóstico da deficiência intelectual é realizado pela observância dos seguintes critérios: o grau de comprometimento cognitivo, o comportamento adaptativo e o início do transtorno no ciclo de vida do sujeito (Associação Americana de Psiquiatria, 2014). A utilização de instrumentos psicológicos faz parte do processo de avaliação, sobretudo ao considerar o uso de testes de inteligência padronizados para avaliar as habilidades cognitivas. Para os testes utilizados, destaca-se a importância da apresentação de evidências de validade, propriedades psicométricas satisfatórias e normas que estejam condizentes com o contexto sociocultural e com o idioma nativo dos sujeitos (Conselho Federal de Psicologia, 2003; Noronha, Primi & Alchieri, 2004; Ortiz & Ochoa, 2005).

Os escores nos testes podem ser influenciados por diversos fatores, dentre eles o efeito Flynn, em que há uma superestimação dos escores quando são utilizadas normas desatualizadas (Flynn, 2000; Flynn, 2009). Além disso, Anastasi e Urbina (2000) já chamavam a atenção para o fato de que na testagem de populações especiais são necessários alguns cuidados específicos, considerando as possíveis limitações apresentadas pelos indivíduos, desde déficits motores até dificuldades relacionadas à comunicação e a

linguagem. A utilização de testes de inteligência que exigem habilidades verbais e/ou escritas pode limitar a avaliação cognitiva, através da subestimação dos escores dos sujeitos (Tellegen, 1993; Tellegen & Laros, 1993). Outras limitações no processo de avaliação advêm de testagens breves ou uso de instrumentos que apresentam apenas um subteste, impossibilitando a comparação entre escores em diferentes escalas.

Em um estudo de comparação entre a utilização de instrumentos verbais e não-verbais, Jansen (1991) apontou uma grande diferença entre os escores obtidos a partir de quatro subtestes verbais do WISC-R e os escores obtidos no teste não-verbal de inteligência SON-R. A pesquisa foi realizada com um grupo de crianças que apresentavam prejuízos específicos na linguagem. A diferença encontrada entre o QI verbal no WISC-R ($M = 83,1; DP = 14$) e o QI SON-R ($M = 97,5; DP = 14$) foi significativa ($p < 0,01$).

Em casos como a deficiência intelectual ou nos diagnósticos de dificuldades de aprendizado os sujeitos apresentam prejuízos no domínio acadêmico, pois o mesmo está relacionado a habilidades de leitura, escrita e raciocínio quantitativo. Neste sentido, os testes de inteligência tradicionalmente utilizados não são considerados adequados para a avaliação do funcionamento cognitivo dessa população. O WISC é o teste mais popular para a avaliação das habilidades cognitivas de crianças em diversos países e possibilita a medida de diferentes aspectos do desempenho dos sujeitos (Evers, et al., 2012). Em estudo realizado com o WISC-IV, Giofrè e Cornoldi (2015) apontam que o QI obtido através do WISC-IV pode não ser uma medida adequada do funcionamento cognitivo em pessoas com dificuldades específicas de aprendizagem. Entretanto, a literatura também apresenta estudos sugerindo que os índices obtidos para o WISC são estimativas válidas para a avaliação de grupos clínicos (Wechsler, 1991; Wechsler, 2013).

No campo da avaliação neuropsicológica os testes não-verbais de inteligência são alternativas favoráveis para a avaliação dos transtornos do desenvolvimento (Macedo, Mecca,

Valentini, Laros, Lima & Schwartzman, 2013). Em diversos transtornos do desenvolvimento são identificadas alterações motoras e de linguagem, de modo que a avaliação pode ser prejudicada pela utilização de testes verbais (Mecca, Orsati & Macedo, 2014). A não dependência da linguagem nos testes não-verbais permite, ainda, a testagem de indivíduos com dificuldade de comunicação, estrangeiros, pessoas que não apresentam total domínio da língua do país de residência, pessoas com deficiência auditiva e pessoas de difícil testagem (Laros, Jesus & Karino, 2013).

Dentre os instrumentos de inteligência não-verbal é possível destacar os testes SON (Snijders, Tellegen & Laros, 1989; Tellegen & Laros, 1993; Tellegen, Winkel, Wijnberg – Williams, Laros, 1998). As iniciais do nome SON fazem referência à primeira autora dos testes, Dra. Nan Snijders-Oomen. Os testes SON são de origem holandesa cuja primeira versão foi lançada no ano de 1943, voltada para a avaliação do funcionamento cognitivo de crianças surdas com idades entre 4 e 14 anos (Snijders-Oomen, 1943). Desde então, foram lançadas diferentes versões dos testes SON e foram realizadas pesquisas de normatização e validação em diversos países, tais como: Holanda, Alemanha, França, China, República Tcheca e República Eslováquia. Também foram realizados estudos de validação na Austrália, Estados Unidos e Reino Unido.

No Brasil a versão SON-R 2½-7[a] (Laros, Tellegen, Jesus & Karino, 2015) apresenta normas para crianças entre 2 anos e meio até 7 anos de idade e encontra-se disponível para comercialização e uso profissional. O SON-R 6-40 é a quinta versão dos testes SON, sendo destinada para crianças e adultos com idades entre 6 e 40 anos (Tellegen, Laros & Petermann, 2007). A pesquisa de normatização do SON-R 6-40 foi realizada simultaneamente para a Alemanha e Holanda ($n = 2.040$). No Brasil, a normatização do SON-R 6-40 encontra-se em andamento.

O SON-R 6-40 caracteriza-se como um teste não-verbal de inteligência pois sua aplicação pode ser realizada de maneira verbal ou não-verbal. O instrumento é composto por quatro subtestes, a saber: Analogias, Mosaicos, Categorias e Padrões, sendo aplicados nesta ordem. Cada subteste está organizado em duas ou três séries paralelas de itens. Os subtestes Analogias e Categorias são subtestes de múltipla escolha e estão relacionados ao raciocínio abstrato e visual. Já os subteste Mosaicos e Padrões são escalas de execução, onde se pretende avaliar o raciocínio espacial (Tellegen & Laros, 2014).

O SON-R 6-40 possui um procedimento adaptativo de testagem. Por meio deste procedimento as tarefas apresentadas durante a realização da testagem são adequadas para o nível de desenvolvimento cognitivo do sujeito (Tellegen & Laros, 2014). Durante a aplicação do teste é fornecido *feedback* ao sujeito após cada resposta. Este procedimento possibilita que o sujeito possa refletir sobre o próprio desempenho, buscando engajamento na atividade para alcançar respostas corretas (Tellegen & Laros, 2014).

Considerando o contexto de adequação normativa do SON-R 6-40 para o Brasil e, ainda, a necessidade de testes psicológicos que apresentem estudos adicionais que incorporem evidências de sua validade e uso para diferentes grupos, este estudo objetiva obter evidências da adequação do instrumento não-verbal de inteligência SON-R 6-40 em uma amostra de pessoas com deficiência intelectual. Pretende-se ainda, verificar possíveis associações entre variáveis comportamentais e o desempenho da amostra no SON-R 6-40, obter evidências da validade de construto do instrumento e averiguar a confiabilidade dos escores obtidos.

Método

Participantes

Participaram do estudo 50 pessoas, sendo 27 sujeitos (54%) do sexo masculino. A idade dos participantes variou entre 15 e 39 anos ($M = 26,2$; $DP = 7,92$). A amostra foi composta por pessoas provenientes de 23 bairros do Distrito Federal, vinculadas a uma organização não governamental que presta diversos serviços para adolescentes e adultos com deficiência intelectual ou múltipla. Todos os respondentes apresentaram junto à instituição o diagnóstico de deficiência intelectual de acordo com a Classificação Internacional de Doenças, CID-10 (OMS, 1997). Do total de participantes, os seguintes diagnósticos foram identificados: Retardo Mental (68%), Síndrome de Down (18%), Sequelas de Traumatismo Craniano (6%), Hidrocefalia (4%), Transtorno Global do Desenvolvimento (2%) e Epilepsia (2%).

Instrumentos

Para a avaliação das habilidades cognitivas dos participantes foi utilizada a bateria SON-R 6-40 (Tellegen & Laros, 2014). O SON-R 6-40 é a versão revisada do SON-R 5½ - 17, sendo a quinta versão dos testes SON. O instrumento é de aplicação individual e pode ser administrado de forma não-verbal, permitindo a avaliação de um espectro relativamente amplo das funções cognitivas independentemente das habilidades de linguagem. Os quatro subtestes do SON-R 6-40 (Analogias, Categorias, Mosaicos e Padrões) são descritos em seguida.

Analogias é um subteste de múltipla escolha, sendo composto por três séries com doze itens cada. Cada item apresenta como exemplo uma figura geométrica que sofre uma ou mais alterações, de forma que uma nova figura geométrica é formada. A tarefa do respondente é aplicar o mesmo princípio de mudança do exemplo para a figura seguinte. Assim, o sujeito

escolhe dentre as opções disponíveis aquela que apresenta a solução correta para a transformação.

Mosaicos é o segundo subteste da bateria que consiste em duas séries com treze itens cada. O sujeito deve copiar em uma moldura os diferentes padrões de mosaicos apresentados pelo examinador usando quadrados brancos e vermelhos. O tempo máximo para a resolução da tarefa é de dois minutos por item.

Categorias é um subteste de múltipla escolha, constituído por 36 itens dividido em três série de 12 itens. São apresentadas três figuras que possuem algo em comum e o sujeito deve descobrir o conceito subjacente às imagens e escolher duas figuras da página seguinte que pertençam à mesma categoria.

Padrões é o último subteste da bateria, formado por duas séries com treze itens. Os itens consistem de padrões de desenhos formados por linhas, com a parte do meio omitida. A tarefa do sujeito é completar com o lápis a parte do desenho que falta. Neste subteste são disponibilizados dois minutos para a realização dos itens iniciais e três minutos para os itens finais.

Os subtestes do SON-R 6-40 podem ser divididos em dois tipos: testes de raciocínio abstrato (Analogias e Categorias) e testes de raciocínio espacial (Mosaicos e Padrões). Cada subteste é introduzido com dois ou mais itens de exemplo e iniciado a partir dos primeiros itens da primeira série. Os itens do teste estão dispostos em ordem crescente de dificuldade e há interrupção da série quando o sujeito comete dois erros no total, não necessariamente consecutivos.

Durante a administração do SON-R 6-40, os aplicadores avaliaram os participantes acerca da motivação, cooperação, concentração e compreensão das instruções. Para tanto, os aplicadores pontuaram estes aspectos como: 1 (Muito baixa), 2 (Baixa), 3 (Bom) e 4 (Muito bom). Também se obteve acesso as fichas individuais dos participantes para o levantamento

de informações adicionais, tais como: data de nascimento, local de residência, renda familiar e diagnóstico encaminhado à instituição.

Procedimento

A coleta de dados foi realizada entre o segundo semestre de 2014 e o primeiro semestre de 2015 e contou com a participação de dois aplicadores treinados. Todas as aplicações ocorreram na instituição da qual o grupo faz parte. Após o contato com a instituição foram enviados aos pais e responsáveis uma Carta de Apresentação e o Termo de Consentimento para Participação na Pesquisa. A Carta de Apresentação continha dados que contextualizam a pesquisa e apresentava os principais objetivos, bem como o contato do pesquisador. O Termo de Consentimento para Participação na Pesquisa apresentava as condições de coletas de dados e os riscos envolvidos na participação, para que os alunos pudessem participar da pesquisa foi necessário que os pais e/ou responsáveis permitissem a sua participação.

A seleção do grupo foi feita a partir do critério da idade. A amostra dos sujeitos foi uma amostra de conveniência. As aplicações do SON-R 6-40 foram realizadas em sessões únicas e individuais, com tempo médio de aplicação de 30 minutos. Para a testagem os participantes foram retirados das salas de aula e oficinas durante o período de permanência na instituição.

Análise dos dados

Inicialmente foi realizada a transformação dos escores brutos do SON-R 6-40 em escores normatizados. O escore total do teste (QI-SON) é expresso em uma escala padronizada com média 100 e desvio-padrão 15, enquanto os escores normatizados nos subtestes são expressos em uma escala com média 10 e desvio-padrão 3. As transformações dos escores foram realizadas usando o CD versão 5.4 com a correção informatizada fornecida com a bateria SON-R 6-40 (Tellegen, 2013).

As normas do SON-R 6-40 foram baseadas na amostra Holandesa junto com a amostra Alemã ($n = 1.933$). Foram utilizadas estas normas uma vez que as normas brasileiras ainda não estavam disponíveis. A normatização brasileira do SON-R 2½-7 [a] apresentou uma diferença entre a população de normatização da Holanda e do Brasil de 16,7 na escala de QI (100,15). Com base deste achado foi realizada uma correção nos escores normatizados do SON-R 6-40, aumentando 16,7 pontos para cada escore normatizado. A correção dos escores foi realizada a fim de analisar a adequação das normas disponíveis do SON-R 6-40 para a indicação do nível de funcionamento cognitivo do grupo investigado. Ressalta-se que esta correção tem por base achados empíricos e representa uma estimativa para a diferença entre os países, entretanto, os dados serão representativos quando as normas para o Brasil forem disponibilizadas.

As análises exploratórias dos dados foram realizadas a fim de verificar a integridade da base de dados. Dessa forma, foram feitas análises de frequência, normalidade univariada, dados *missing* e *outliers*. Todas as análises foram realizadas com o *software* SPSS, versão 20. O pressuposto da normalidade da distribuição dos escores no SON-R 6-40 foi verificado de acordo com os índices de assimetria e curtose (Miles & Shevlin, 2001). O critério para avaliar a assimetria e curtose foi adaptado do Miles e Shevlin: os autores defendem que um índice de assimetria e curtose menor de que 2,00 geralmente não traz problemas à análise dos dados.

A análise da fidedignidade dos escores do teste foi feita a partir do coeficiente *Lambda* 2 de Guttman. De acordo com (Tellegen & Laros, 2004; Ten Berge & Zegers, 1978; Sijtsma, 2009; Sijtsma, 2012) a escolha deste coeficiente é mais indicada para amostras pequenas ou instrumentos que contêm poucos itens. Para as análises de correlação dos escores normatizados no SON-R 6-40 com as variáveis comportamentais avaliadas pelos examinadores durante a testagem foi utilizado o coeficiente de correlação bivariada r de Pearson (Thompson, 2006).

Foi conduzida uma análise fatorial exploratória (AFE) a fim de verificar a dimensionalidade do SON-R 6-40 e ainda, obter evidências de validade de construto do instrumento para a amostra estudada (Laros, 2005). Considerou-se o índice KMO (*Kaiser Meyer Olkin*) para verificar se a matriz dos dados é passível de fatoraçoão, de acordo com essa medida os valores ideais de KMO giram entre 0,80 e 0,90, de modo que valores abaixo de 0,50 são considerados inaceitáveis (Pasquali, 2009). A significância do teste de esfericidade de Bartlett também foi verificada, esta medida é bastante sensível ao tamanho da amostra, entretanto pode ser considerada quando são observados menos de 5 casos por variável (Tabachnick & Fidell, 2007). O método utilizado foi a análise fatorial dos eixos principais (*Principal Axis Factoring, PAF*) e para a rotação dos eixos foi utilizado o método *Promax*. Na solução fatorial foram desconsiderados os itens com cargas fatoriais abaixo de 0,32.

Resultados e Discussão

Para a apresentação dos resultados, inicialmente são exibidas informações acerca da distribuição dos escores dos participantes a fim de caracterizar a amostra. A média do QI-SON para todo o grupo foi de 73,18 ($DP = 2,08$), com escores variando entre 72 e 80. Para os subtestes do SON-R 6-40 os escores normatizados variaram entre 1 e 5. Do número total de participantes, 25 sujeitos (50%) atingiram o escore normatizado mais baixo de 72 e apenas 1 sujeito obteve o escore mais alto de 80. A tabela 1 apresenta os valores de média, desvio padrão e amplitude para os escores brutos e normatizados obtidos para os quatro subtestes e para o escore total do SON-R 6-40.

Tabela 1
Distribuição dos escores no SON-R 6-40 (n = 50)

Distribuição dos escores brutos			
	Média	(DP)	Amplitude
Analogias	6,02	4,4	0 – 18
Mosaicos	3,26	3,18	0 – 11
Categorias	5,14	4,44	0 – 15
Padrões	2,82	3,05	0 – 11
Soma dos escores	17,24	15,07	2 – 48
Distribuição dos escores padronizados			
	Média	(DP)	Amplitude
Analogias	4,81	0,93	1 – 5
Mosaicos	4,51	0,48	1 – 3
Categorias	4,89	1,03	1 – 5
Padrões	4,59	0,49	1 – 3
QI-SON	73,18	2,08	72 – 80

Notas. DP = Desvio-padrão.

Os subtestes Mosaicos e Padrões apresentam os escores padronizados mais baixos, variando de um a três. Nesses dois subtestes são encontradas distribuições de escores padronizados com maior nível de assimetria. Uma porcentagem considerável de sujeitos obteve escores padronizados iguais a 1, Mosaicos (86%) e Padrões (76%). Em todos os subtestes, do número total da amostra 66% obtiveram escores de um e dois.

Foi investigada a relação entre o escore total normatizado no SON-R 6-40 (QI-SON) e as avaliações de motivação, concentração, cooperação e compreensão das instruções. Para tanto, foi utilizado o coeficiente de correlação bivariada r de Pearson, considerando que as variáveis estão no nível intervalar e atendem aos pressupostos de distribuição. Para as correlações foram calculados os intervalos de confiança de 95%, adicionando informação ao

valor p . Os limites superiores e inferiores dos intervalos de confiança não devem incluir o zero, pois este valor indicaria a possibilidade de não haver relação entre as variáveis (Field, 2009). A Tabela 2 apresenta a distribuição dos escores dos sujeitos para a avaliação dos aplicadores e respectivas correlações com o QI-SON.

Tabela 2
Relação entre o QI-SON e a avaliação do aplicador

Avaliação do Administrador	Motivação		Concentração		Cooperação		Compreensão *	
Muito baixa	-	-	6%	3	-	-	26%	13
Baixa	4%	2	8%	4	2%	1	14%	7
Bom	32%	16	36%	18	26%	13	34%	17
Muito bom	50%	25	36%	18	58%	29	12%	6
<i>r</i>	0,21		0,39***		0,33**		0,52***	

Notas. * Compreensão das instruções; **Correlações significativas = $p < 0,05$; *** Correlações significativas = $p < 0,01$.

A Tabela 2 apresenta que a maior parte dos sujeitos obteve a pontuação referente a “muito bom” para cooperação (58%) e para a motivação (50%). A avaliação dos aplicadores referentes à motivação e a cooperação pode ser um indicador da predisposição e aceitação dos participantes frente ao material do teste e a situação de testagem. Verifica-se ainda que em nenhum dos casos a motivação e cooperação dos sujeitos durante a aplicação do SON-R 6-40 foram pontuadas como “muito baixas”.

As análises de correlação apontaram uma relação de efeito pequeno e não significativa entre a motivação e o QI-SON, $r = 0,21$, IC 95% [- 0,05, 0,44], $p = 0,17$. As correlações entre cooperação e concentração com o QI-SON foram significativas e representam um efeito médio, para a cooperação $r = 0,33$, IC 95% [0,03, 0,57], $p = 0,03$ e para a concentração $r = 0,39$, IC 95% [0,10, 0,62], $p = 0,01$. Das quatro observações feitas pelo aplicador, a compreensão das instruções demonstrou ser aquela que tem maior associação com o

desempenho no SON-R 6-40, $r = 0,52$, IC 95% [0,26, 0,71], $p < 0,01$. A correlação alta indica que a compreensão das instruções está positivamente relacionada ao desempenho no SON-R 6-40.

Os resultados encontrados são semelhantes ao estudo de Tellegen & Laros (2014) com crianças holandesas que apresentavam dificuldades severas no aprendizado ($n = 57$). A associação entre o QI-SON e a motivação foi a mais baixa, $r = 0,20$, seguida por efeitos médios para a cooperação e concentração $r = 0,34$ e $r = 0,30$. A correlação foi alta para a compreensão das instruções, $r = 0,62$.

O mesmo padrão de associação entre a avaliação do aplicador e o desempenho no SON-R 6-40 foi encontrado no estudo com a amostra de normatização alemã ($n = 938$). Neste estudo foi encontrada uma correlação significativa entre os escores de QI e a compreensão das instruções $r = 0,29$, $p < 0,01$ (Tellegen & Laros, 2014). A amostra de normatização brasileira do SON-R 2½-7[a] ($n = 1.200$) também apresentou maior correlação entre o QI-SON e compreensão das instruções, $r = 0,44$, $p < 0,01$ (Laros, Tellegen, Jesus & Karino, 2015).

Em seguida foi conduzida a Análise Fatorial Exploratória (AFE) a fim de verificar a estrutura fatorial do SON-R 6-40. A partir da análise dos eixos principais (*Principal Axis Factoring, PAF*) com a rotação *Promax*, verificou-se uma solução fatorial de dois fatores. Os resultados da AFE podem ser observados na Tabela 3.

Tabela 3
Cargas fatoriais (CF) e a comunalidade média (h^2) depois da análise fatorial exploratória do SON-R 6-40
($n = 50$)

Subteste	CF Fator 1	Média	CF Fator 2	Média	h^2
Analogias	-	-	0,33 - 0,66	0,46	0,25
Mosaicos	0,57 - 0,79	0,67	-	-	0,38
Categorias	-	-	0,47 - 0,78	0,59	0,50
Padrões	0,34 - 0,87	0,66	-	-	0,45

Notas. CF = carga fatorial; h^2 = Comunalidade.

O critério para a retenção do número de fatores teve por base pesquisas realizadas anteriormente com o instrumento (Tellegen & Laros, 2014; Laros, Tellegen, Jesus & Karino, 2015). Os resultados da análise fatorial exploratória indicaram a adequação da solução fatorial com 2 fatores, sendo capaz de explicar 40,80% da variância comum. Esse valor pode ser considerado uma medida de ajuste do modelo aos dados (Baglin, 2014). As cargas fatoriais para os quatro subtestes do SON-R 6-40 variaram entre 0,33 e 0,87. A média das cargas fatoriais é mais alta para os subtestes Mosaicos (0,67) e Padrões (0,66). O primeiro fator apresentado pela solução fatorial apresenta itens dos subtestes Mosaicos e Padrões e o segundo fator apresenta itens agrupados dos subtestes Categorias e Analogias. A correlação entre os dois fatores encontrados foi de 0,33, resultado que confirma a escolha pela rotação oblíqua dos fatores.

O primeiro fator, composto por itens dos subtestes Mosaicos e Padrões, está relacionado ao raciocínio espacial e abstrato, pois as atividades destes subtestes envolvem a análise visual e manipulação do material do teste para a resolução da tarefa. Enquanto o segundo fator, com itens dos subtestes Analogias e Categorias está associado ao raciocínio abstrato, pois os sujeitos devem aplicar a lógica observada para as figuras ou imagens seguintes.

Na pesquisa brasileira de normatização do SON-R 2½-7[a] também foi encontrada uma solução fatorial com dois fatores. No SON-R 2½-7[a] os escores nos subtestes Mosaicos e Padrões formam uma escala de execução SON-EE, enquanto os subtestes Analogias e Categorias formam uma escala de raciocínio SON-ER (Laros, Tellegen, Jesus & Karino, 2015).

De acordo com Hogan (2006), a fidedignidade dos escores também deve ser considerada a fim de avaliar a validade de um teste, de modo que a fidedignidade pode ser considerada como um pré-requisito para a validade. Assim, também foi verificada a fidedignidade dos escores para os quatro subtestes do SON-R 6-40, para o QI total e para os escores fatoriais obtidos a partir da solução fatorial. A fidedignidade dos escores obtidos foi calculada por meio do índice Lambda 2 de Guttman. A Tabela 4 apresenta os resultados da fidedignidade para os escores obtidos no SON-R 6-40.

Tabela 4

Índices de fidedignidade do SON-R 6-40

Parâmetros Psicométricos	SON-R 6-40: Escala geral e subtestes				
	Analogias	Mosaicos	Categorias	Padrões	Escala Geral
Número de itens	36	26	36	26	124
Consistência interna (λ_2)	0,88	0,91	0,91	0,88	0,95
r média entre itens	0,25	0,37	0,30	0,35	0,18
r média item-total	0,46	0,57	0,51	0,55	0,41

Notas. λ_2 = Lambda 2 de Guttman; r = correlações.

Na Tabela 4 são indicados valores altos para a fidedignidade dos escores no SON-R 6-40 (índices variaram de 0,88 a 0,95). Após encontrar uma solução fatorial composta por dois fatores: Fator 1 (Mosaicos e Padrões) e Fator 2 (Analogias e Categorias), buscou-se verificar se há diferença significativa no desempenho nos dois fatores para o sexo feminino e para o sexo masculino. Para tanto, foi realizado um teste t da diferença entre os escores fatoriais no

Fator 1 e no Fator 2 para homens e mulheres. Os resultados indicaram uma diferença significativa entre homens ($M = 46,06$; $DP = 33,9$) e mulheres ($M = 22,8$; $DP = 20,71$) no Fator 1, $t(48) = 2,92$, $p < 0,01$. Entretanto, para o Fator 2 a diferença entre homens ($M = 35,27$; $DP = 26,29$) e mulheres ($M = 43,68$; $DP = 25,06$) não é significativa, $t(48) = 1,15$, $p = 0,25$. A diferença encontrada para homens e mulheres encontra-se associada aos subtestes de execução, Mosaicos e Padrões que compõem o Fator 1 e avaliam aspectos do raciocínio espacial.

As diferenças entre sexos relacionadas às habilidades cognitivas são mínimas para escalas gerais do funcionamento mental. Entretanto, pode haver diferenças para escalas de habilidades cognitivas mais específicas. O raciocínio espacial se destaca por apresentar maiores escores para os homens do que para as mulheres (Halpern, 1992; Hogan, 2006). Na pesquisa de normatização do SON-R 6-40 para a Alemanha e Holanda essa diferença também foi verificada para o subteste Mosaicos ($p < 0,05$). Em pesquisa com crianças brasileiras, Laros, Tellegen, Jesus & Karino (2015) encontram diferenças significativas para o subteste Categorias e para a Escala de Raciocínio (ER). Neste último caso, as meninas apresentaram escores significativamente mais altos que os meninos, para a Escala de Raciocínio (SON-ER), $t = 1,98$, $p < 0,05$ e para Categorias, $t = 2,82$, $p < 0,01$.

Considerações finais

O principal objetivo deste estudo foi obter evidências da adequação do teste não-verbal de inteligência SON-R 6-40 para pessoas com deficiência intelectual. Para avaliar os resultados do presente estudo os princípios de construção e desenvolvimento do SON-R 6-40 foram considerados, buscando garantir a confiabilidade da avaliação e o bem-estar daqueles que são submetidos à testagem. Em segundo lugar destacou-se a acurácia da medida. Ao considerar que a avaliação por meio da testagem culmina em resultados numéricos

padronizados é necessário garantir que a medida seja a mais precisa possível. E finalmente, o princípio de que os resultados do teste precisam ser contextualizados. Isso quer dizer que no processo de diagnóstico outras fontes de informação precisam ser levadas em consideração.

Os resultados do estudo oferecem indícios da adequação da utilização do SON-R 6-40 para a amostra estudada. É possível destacar que os escores obtidos no teste refletem o comprometimento cognitivo abaixo da média populacional, característico da deficiência intelectual. Esta informação é um critério fundamental para a realização do diagnóstico. Os resultados também apontam para valores adequados e satisfatórios da precisão dos escores obtidos no teste. As avaliações de motivação e cooperação para a realização da testagem foram positivas neste estudo, podendo ser interpretadas como indícios da atratividade da tarefa e do material do SON-R 6-40.

É importante considerar que o SON-R 6-40 apresenta uma série de características que garantem que a testagem ocorra de maneira mais suave, dentre elas: o aspecto não-verbal da testagem, o procedimento adaptativo, o fornecimento de feedback e a atratividade do material. No caso de amostras especiais, como na deficiência intelectual, estas características garantem a administração de itens cuja dificuldade corresponde ao nível de desenvolvimento cognitivo dos sujeitos e também permite que durante a aplicação o examinador trabalhe com o intuito de dar *feedbacks* que encorajem e motivem o examinando durante o processo.

O SON-R 6-40 permite uma avaliação relativamente ampla do funcionamento cognitivo. Os resultados do presente estudo indicaram que os quatro subtestes do SON-R 6-40 avaliam aspectos do raciocínio viso-espacial, concreto e abstrato. No campo da avaliação psicológica, sobretudo na avaliação neuropsicológica, este instrumento representa uma importante ferramenta para a avaliação de diferentes distúrbios do desenvolvimento, tais como os transtornos do espectro do autismo, os transtornos de aprendizagem e a deficiência intelectual, considerando ainda, as limitações de linguagem próprias para estes grupos.

É importante notar algumas limitações deste estudo. Os participantes foram selecionados com base na disponibilidade, e não de maneira aleatória. Além disso, os participantes foram encaminhados com diagnósticos realizados fora da instituição, fato que limita as considerações sobre os critérios e procedimentos adotados para o diagnóstico. O número de participantes e a faixa etária dos participantes neste estudo são restritos. Por último, se destaca que a avaliação realizada pelos aplicadores acerca da motivação, cooperação, concentração e compreensão das instruções, foi subjetiva, e não obedeceu a um padrão de definição operacional para estes aspectos.

Para estudos futuros com amostras de pessoas com deficiência intelectual destaca-se a importância de trabalhar com um número maior de participantes a fim de aumentar a variabilidade dos resultados. Além disso, é possível trabalhar com diferentes faixas etárias para compreender o funcionamento cognitivo de pessoas com deficiência intelectual em diferentes fases do ciclo de vida. O uso de instrumentos disponíveis para uso profissional também é uma alternativa viável para estudos de evidências de validade convergente e comparação entre o desempenho de pessoas com deficiência intelectual em testes verbais e não-verbais de inteligência.

Por fim, destaca-se a relevância do SON-R 6-40 para o processo de avaliação psicológica, sendo capaz de fazer estimativas válidas acerca das habilidades cognitivas em grupos especiais. Entretanto, os escores no instrumento não podem ser utilizados como os únicos critérios para fins de classificação e diagnóstico.

Referências

- American Association on Intellectual and Developmental Disabilities (2015). *Definition of intellectual disability*. Retrieved from <http://aaidd.org/intellectual-disability/definition#.Vhv1FPiViko>.
- American Psychological Association (2015). *APA dictionary of psychology*. Washington, DC: American Psychological Association.
- Anastasi, A., & Urbina, S. (2000). *Testagem Psicológica*. 7ª Edição. Porto Alegre, RS: Artmed.
- Associação Americana de Psiquiatria (2014). *Manual diagnóstico e estatístico de transtornos mentais*. Porto Alegre, RS: Artmed.
- Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research & Evaluation*, 19(5), 1-14. Retrieved from <http://pareonline.net/getvn.asp?v=19&n=5>
- Conselho Federal de Psicologia (2003). *Resolução CFP nº 002/2003*. Retrieved from <http://satepsi.cfp.org.br/legislacao.cfm>.
- Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., et al. (2012). Testing practices in the 21st century: Developments and European psychologists' opinions. *European Psychologist*, 17(4), 300–319.
- Field, A. (2009). *Discovering statistics using SPSS*. London: Sage Publications.
- Flynn, J. R. (2000). IQ trends over time: Intelligence, race and meritocracy. In K. Arrow, S. Bowles & S. Durlauf (Eds.), *Meritocracy and economic inequality*. Princeton, NJ: Princeton University Press.
- Flynn, J. R. (2009). *O que é inteligência? Além do efeito Flynn*. Porto Alegre, RS: Artmed.
- Giofrè, D., & Cornoldi, C. (2015). The structure of intelligence in children with specific learning disabilities is different as compared to typically development children. *Intelligence*, 52, 36-43.
- Halpern, D. F. (1992). *Sex differences in cognitive abilities*, Hillsdale, NJ: LEA.
- Harris, J. C. (2006). *Intellectual disability: Understanding its development, causes, classification, evaluation, and treatment*. New York: Oxford University Press.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. RJ, LTC – Livros Técnicos e Científicos Editora S.A.
- Instituto Brasileiro de Geografia e Estatística (2013). *Pesquisa Nacional de Saúde 2013: Ciclos de vida*. Retrieved from <http://biblioteca.ibge.gov.br/visualizacao/livros/liv94522.pdf>
- Jansen, B. (1991). *Een exploratief onderzoek naar de werkbaarheid van een diagnose: Hoe specifiek is de diagnose "Specifieke Taalontwikkelingsstoornis"?* [An explorative study into the usefulness of a diagnosis]. Groningen: Internal report, Academic Hospital.

- Laros, J. A. (2012). O uso da análise fatorial: algumas diretrizes para pesquisadores. In L. Pasquali (Ed.), *Análise fatorial para pesquisadores* (pp. 141-160). Brasília: LabPAM / UnB.
- Laros, J. A., Jesus, G. R., & Karino, A. K. (2013). Validação brasileira do teste não-verbal de inteligência SON-R 2½-7[a]. *Avaliação Psicológica*, 12 (2), 233-242.
- Laros, J. A., Tellegen, P. J., Jesus, G.R., & Karino, C. A. (2015). *SON-R 2½-7[a]. Manual. Teste não-verbal de inteligência*. São Paulo, SP: Hogrefe CETEPP.
- Macedo, E. C., Mecca, T. P., Valentini, F., Laros, J. A., Lima, R. M.F., & Schwartzman, J. S. (2013). Utilizando o teste não verbal de inteligência SON-R 2½-7[a] para avaliar crianças com Transtornos do Espectro do Autismo. *Revista Educação Especial*, 47, 603-618.
- Mecca, T. P., Orsati, F. T., & Macedo, E. C. (2014). Inteligência e transtornos do desenvolvimento In A. G. Seabra, J. A. Laros, E. C. Macedo & N. Abreu (Eds.) *Inteligência e funções executivas: Avanços e desafios para a avaliação neuropsicológica* (pp. 95-112). São Paulo, SP: Memnon Edições Científicas Ltda.
- Miles, J., & Shevlin, M. (2001). *Applying regression and correlation. A guide for students and researchers*. London: Sage publications.
- Moore, C., O'Keef, S. L., Lawhon, D., & Tellegen, P. (1998). Concurrent validity of the Snijders-Oomen nonverbal intelligence test 2½-7 - Revised with the Wechsler Preschool and Primary Scale of Intelligence - Revised. *Psychological Reports*, 82, 619-625.
- 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.
- Organização Mundial da Saúde (1997). *Classificação internacional de doenças e problemas relacionados à saúde*. São Paulo: Universidade de São Paulo.
- Ortiz, S. O., & Ochoa, S. H. (2005). Advances in cognitive assessment of culturally and linguistically diverse individuals: A nondiscriminatory interpretive approach. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (2nd Edition, pp. 234-250). New York: The Guilford Press.
- Pasquali, L. (2012). *Análise fatorial para pesquisadores*. Brasília: LabPAM/UnB.
- Schalock et al. (2010). *Intellectual disability: Definition, classification, and systems of supports*. Washington: American Association on Intellectual and Developmental Disabilities.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74, 107-120.
- Sijtsma, K. (2012). Future of psychometrics: Ask what psychometrics can do for psychology. *Psychometrika*, 77, 4-20.

- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen [The examination of intelligence with deaf-mute children]*. Nijmegen: Berkhout.
- Snijders, J. T., Tellegen, P., & Laros, J. A. (1989). *Snijders–Oomen non-verbal intelligence test: SON-R 5½-17. Manual and research report*. Groningen: Wolters–Noordhoff.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5^a ed.). Boston: Pearson Education.
- Tellegen, P. J. (1993). *A nonverbal alternative to the Wechsler scales: The Snijders –Oomen Nonverbal intelligence tests*. Paper presented at the First Annual South Padre Island Conference on Cognitive Assessment of Children and Youth in School and Clinical Settings. A Compendium of Proceedings (pp. 35-49).
- Tellegen, P. J. & Laros, J. A. (1993). The Snijders-Oomen Nonverbal intelligence tests: general intelligence test or tests for learning potential? In J. H. M. Hamers, K. Sijtsma & A. J. J. M. Ruijsenaars (Eds.), *Learning potential assessment, theoretical, methodological and practical issues* (pp. 267–283). Amsterdam: Swets & Zeitlinger.
- Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. J., & Laros, J. A. (1998). *Snijders-Oomen nonverbal intelligence test, SON-R 2½-7: Manual & research report*. Lisse: Swets & Zeitlinger.
- Tellegen, P. J., & Laros, J. A. (2004). Cultural bias in the SON-R test: Comparative study of Brazilian and Dutch children. *Psicologia: Teoria e Pesquisa*, 20, 103-111.
- Tellegen, P. J., Laros, J. A., & Petermann, F. (2007). *SON-R 2½-7, non-verbaler intelligenztest: Testmanual mit deutscher normierung und validierung*. Göttingen, Bern, Wien, Paris, Oxford: Hogrefe.
- Tellegen, P. J. (2013). *SON-R 5.4 (Computer program)*. Göttingen: Hogrefe.
- Tellegen, P. J., & Laros, J. A. (2014). *SON-R 6-40. Non-verbal intelligence test: Research report*. Göttingen, Germany: Hogrefe Verlag.
- Ten Berge, J. M. F., & Zegers, F. E. (1978). A serie of lower bounds to the reliability. *Psychometrika*, 43, 575-579.
- Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach*. New York: The Guilford Press.
- Wechsler, D. (1991). *Wechsler intelligence scale for children – 3^a ed. (WISC-III): Manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2013). *WISC-IV - Escala Wechsler de inteligência para crianças. Manual técnico*. São Paulo: Casa do Psicólogo.

MANUSCRITO 2

Estudos da validade convergente do SON-R 6-40
com a EMMC, TONI-3 e MPC

Título em inglês

Studies on the convergent validity of the SON-R 6-40
with EMMC, TONI-3 and MPC

Sugestão de título abreviado:

Validade convergente do SON-R 6-40

RESUMO

A obtenção de evidências de validade dos escores de testes psicológicos é fundamental para fins de diagnóstico e propósitos de pesquisa. Assim, o presente estudo teve por objetivo reunir evidências da validade convergente do teste não-verbal de inteligência SON-R 6-40 com três outros testes de inteligência: Matrizes Progressivas Coloridas de Raven, a Escala de Maturidade Mental Colúmbia e o TONI-3. O estudo também buscou avaliar a estrutura fatorial dos instrumentos, a precisão dos seus escores e as diferenças entre os grupos avaliados nos escores totais. Participaram do estudo 150 crianças com idades entre 6 e 9 anos ($M = 7,7$; $DP = 1,0$) provenientes da rede pública e privada das escolas do Distrito Federal. Análises fatoriais exploratórias apontaram uma estrutura bifatorial para o SON-R 6-40 e unifatorial para os demais instrumentos. Os dois fatores do SON-R 6-40 foram interpretados como Fator Espacial e Fator de Raciocínio. O coeficiente de fidedignidade do escore total do SON-R foi de 0,96 e os coeficientes dos escores totais dos outros testes variaram entre 0,84 e 0,92. Foram encontrados para o SON-R 6-40 coeficientes de validade corrigidos para atenuação entre 0,61 e 0,72. Considera-se que o SON-R 6-40 apresenta evidências adequadas de validade convergente com os instrumentos utilizados.

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

ABSTRACT

To obtain evidence of the validity of the scores of psychological tests is fundamental for diagnostic and research purposes. Therefore, the present study had as objective to acquire evidence of the convergent validity of the SON-R 6-40, a non-verbal test of intelligence with three other tests of intelligence: Raven's Colored Progressive Matrices, Columbia Scale of Mental Maturity and TONI-3. The study also sought to evaluate the factor structure of the instruments, their score reliability and the differences between existent groups with respect to their total scores. 150 children participated with ages varying between 6 and 9 years ($M = 7.7$; $SD = 1.0$) attending public and private schools in the Federal District of Brazil. Exploratory factor analyses indicated a bi-factor structure for the SON-R 6-40 and a unifactor structure for the other instruments. The two factors of SON-R 6-40 were interpreted as a Spatial and a Reasoning Factor. The reliability coefficient of the total score of the SON-R 6-40 was .96, the coefficients of the total scores of the other three instruments varied between .84 and .92. For the SON-R 6-40 validity coefficients - corrected for attenuation - were found between .61 and .72. The results indicate satisfactory evidence of the convergent validity of the SON-R 6-40 with the other three intelligence tests.

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

RESUMEN

Para el uso de testes psicológicos en procesos diagnósticos, es fundamental la obtención de evidencias de validez de sus puntajes. Así, el presente estudio tuvo por objetivo reunir evidencias de validez convergente del test no-verbal de inteligencia SON-R 6-40 con otros tres testes de inteligencia: Matrices Progresivas Coloridas de Raven, Escala de Madurez Mental Columbia y el TONI-3. El estudio también buscó evaluar la estructura factorial de los cuatro instrumentos, la precisión de sus puntajes y las diferencias entre los grupos evaluados en los puntajes totales. Participaron en el estudio 150 niños con edades entre 6 y 9 años ($M = 7,7$; $DE = 1,0$) provenientes de la red pública y privada del Distrito Federal. Los análisis factoriales exploratorios indicaron una estructura bifactorial para el SON-R 6-40 y unifactorial para los demás instrumentos. Los dos factores del SON-R 6-40 fueron interpretados como Factor Espacial y Factor Raciocinio. El coeficiente de confiabilidad del puntaje total del SON-R fue de 0,96: los coeficientes de los puntajes totales de los otros testes variaron entre 0,84 y 0,92. Fueron encontrados para el SON-R 6-40 coeficientes de validez corregidos para atenuación entre 0,61 y 0,72 para el SON-R 6-40. Se considera que el SON-R 6-40 presenta evidencias adecuadas de validez convergente con los instrumentos utilizados.

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

Os testes psicológicos necessitam ser avaliados do ponto de vista científico e técnico acerca da construção, da qualidade psicométrica e aspectos relacionados aos comportamentos e construtos psicológicos que buscam avaliar (Urbina, 2007). O processo de avaliação dos testes psicológicos é anterior a sua utilização. Este procedimento visa atender requisitos mínimos para a confiabilidade dos resultados atingidos por meio da testagem.

No Brasil, o Conselho Federal de Psicologia - CFP é o órgão reconhecido por definir e regulamentar o uso, a elaboração e a comercialização dos testes psicológicos, por meio da resolução CFP nº 002/2003 (CFP, 2003), e alterada pela resolução CFP nº 005/2012. Para a elaboração da resolução foram consultados documentos de instituições internacionais reconhecidas por desenvolver diretrizes sobre construção, aspectos técnicos e uso dos testes psicológicos. A resolução em questão destaca a necessidade de aprimorar e revisar periodicamente os instrumentos de trabalho dos psicólogos a fim de garantir a qualidade técnica e ética dos serviços prestados à população.

Dentre os requisitos mínimos exigidos para o uso dos testes psicológicos, a resolução ressalta a apresentação de fundamentação teórica do instrumento e do construto mensurado, evidências empíricas da validade e precisão dos escores do teste, informações sobre as propriedades psicométricas dos itens, apresentação dos procedimentos de aplicação, do sistema de correção e a interpretação dos escores com base nas normas ou no referencial teórico adotado.

Para os requisitos destacados, a validade pode ser compreendida como o aspecto mais importante a ser considerado durante a construção e avaliação dos testes psicológicos (*American Educational Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 2014*). É a respeito da validade que se considera o uso dos escores obtidos por meio da testagem (Pacico & Hutz, 2015). Dessa forma, as inferências feitas a partir dos resultados estão relacionadas às

evidências de validade reunidas que fornecem subsídios à interpretação e uso dos escores no contexto da avaliação psicológica.

Trabalhos clássicos no campo da validade dos instrumentos, como o de Cronbach e Meehl (1955) discutem que o processo de validação deve dar maior ênfase à validade de construto, envolvendo critérios empíricos e teóricos no desenvolvimento de teorias sobre os construtos psicológicos. Para além das evidências empíricas e teóricas, Messick (1989) ressalta que a validade dos instrumentos está relacionada às consequências do uso dos escores dos testes. Perspectivas contemporâneas destacam a relação de causa e efeito entre o construto psicológico e os testes (Borsboom, Mellenbergh & Heerden, 2004). Entretanto, a ideia comum a todas as propostas é de que o processo de validação de um teste objetiva assegurar que a medida é uma boa representação do construto psicológico que busca avaliar.

Os principais critérios adotados para avaliar a validade dos testes psicológicos são propostos pelo atual *Standards for Educational and Psychological Testing*. Neste documento a validade de um teste é concebida como um conjunto de evidências que fornecem base para a interpretação dos escores obtidos nos testes psicológicos (*American Educational Research Association [AERA], APA & National Council on Measurement & Education [NCME], 2014*). De acordo com o mesmo documento, o esboço contemporâneo sobre a validade destaca os seguintes tipos: (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.

Ao considerar a relevância em demonstrar aspectos da validade do teste não-verbal de inteligência SON-R 6-40 para o uso e interpretação dos seus escores, este trabalho objetivou reunir evidências de validade convergente desse instrumento com três outros testes de inteligência: Matrizes Progressivas Coloridas de Raven (MPC), a Escala de Maturidade Mental Colúmbia (EMMC) e o teste de inteligência não-verbal TONI-3, forma A (*Test of*

Nonverbal Intelligence, Form A, Third Edition). Além disso, objetivou-se avaliar a estrutura fatorial dos instrumentos utilizados, verificar a fidedignidade dos escores obtidos e analisar as diferenças nos escores totais para os grupos investigados.

Os testes MPC, EMMC e TONI-3 apresentam parecer favorável para utilização no país, de acordo com o Sistema de Avaliação de Testes Psicológicos (SATEPSI). A pesquisa de normatização do SON-R 6-40 para o contexto brasileiro está em andamento (Laros, Almeida, Valentini & Lima, 2015). Foi concluído um estudo sobre a validade convergente do SON-R 6-40 com a Bateria de Provas de Raciocínio, forma A (BPR-5) e com o SON-R 2½-7[a] (Laros, Almeida, Valentini & Lima, 2015). Ainda não foram realizados estudos relacionados à validade convergente do SON-R 6-40 com os instrumentos utilizados no presente estudo.

No Brasil, a versão SON-R 2½-7[a], voltada para crianças com idades entre 2 anos e 6 meses até 7 anos e 11 meses, encontra-se disponível para uso profissional (Laros, Tellegen, Jesus & Karino, 2015). Durante a pesquisa de normatização e validação do SON-R 2½-7[a] para o Brasil foram analisadas evidências da validade convergente do instrumento com as Matrizes Progressivas Coloridas de Raven (MPC) e com as Escalas de Maturidade Mental Colúmbia (EMMC) (Laros, Reis & Tellegen, 2010), além dos testes WISC-III e WPPSI-III (Karino, Laros & Jesus, 2011).

Em seguida serão brevemente relatadas informações adicionais acerca da origem, aspectos teóricos e alguns estudos sobre os quatro instrumentos utilizados no presente estudo.

Os testes SON

A bateria de avaliação da inteligência SON tem origem holandesa, datada de 1943. Desde então, diversas versões foram desenvolvidas com a finalidade de abranger diferentes faixas etárias e público alvo. O nome SON (Snijders-Oomen Não-verbal) dos testes é uma homenagem à criadora do primeiro teste, a doutora N. Snijders-Oomen (Laros, Tellegen, Jesus & Karino, 2015). Dado que a inteligência foi definida por Snijders-Oomen em termos

do potencial para a aprendizagem, a autora buscou construir um instrumento em que a influência da escolaridade das crianças e das vivências familiares nos seus resultados fosse mínima. Dessa forma, o primeiro teste SON buscou medir a inteligência fluida em vez da inteligência cristalizada. De acordo com a teoria CHC (Cattell-Horn-Carroll) a inteligência fluida é aquela que depende menos do conhecimento prévio e que mais se aproxima do fator geral de inteligência (Flanagan & Harrison, 2012). A inteligência cristalizada se refere à aquisição e a solidificação de conhecimentos formais e informais (Davidson & Downing, 2000).

Uma das principais características dos testes SON é que sua administração pode ser realizada de maneira verbal ou não-verbal, tornando os testes especialmente adequados para pessoas com deficiência auditiva, comprometimento da fala, imigrantes, pessoas com limitações no desenvolvimento cognitivo e pessoas analfabetas (Tellegen & Laros, 2014). A primeira versão do teste SON foi concebida para a avaliação de crianças e adolescentes entre 4 e 14 anos com deficiência auditiva e limitações na comunicação (Snijders-Oomen, 1943). O SON-R 6-40 é a última versão dos testes SON e foi desenvolvida para pessoas na faixa etária entre 6 e 40 anos de idade. A pesquisa de normatização e validação do teste na Holanda e Alemanha foi concluída em 2011 (Tellegen & Laros, 2014).

Matrizes Progressivas de Raven

As Matrizes Progressivas de Raven são utilizadas em diversos países há mais de 50 anos. Os testes Raven apresentam três escalas: *Standard*, *Colorida* e *Avançada*. As Matrizes Progressivas Coloridas de Raven - MPC (também chamada de Escala Especial) foi publicada na Inglaterra em 1947 e revisada em 1956. Esta versão é voltada para avaliar habilidades cognitivas de crianças na faixa etária de 5 aos 11 anos. O teste MPC também pode ser

utilizado para pessoas com deficiência mental e idosos (Angelini, Alves, Custódio, Duarte & Duarte, 1999).

No Brasil, o Raven pode ser destacado como um dos instrumentos mais ensinados nos cursos de psicologia (Alves, Alchieri & Marques, 2001). No campo teórico, as Matrizes Progressivas buscam avaliar o fator geral de inteligência, ou fator *g*, como foi conceituado por Spearman. Segundo os autores do manual (Angelini, Alves, Custódio, Duarte & Duarte, 1999) o fator geral de inteligência apresenta dois componentes: a capacidade edutiva e de reprodução. A educação diz respeito ao processo de extrair novas informações e significados de uma situação não estruturada. Por outro lado, a reprodução está no domínio da memória e diz respeito à capacidade de reproduzir materiais, com base cultural de conhecimentos.

A normatização brasileira da MPC foi realizada em duas fases. A primeira fase foi em 1987 e contou com a participação de 1.417 sujeitos. Posteriormente, foram realizadas novas coletas com o intuito de complementar as normas. No total, a amostra foi de 1.547 sujeitos para a população geral e para as escolas públicas. Todas as aplicações ocorreram no Estado de São Paulo. Os critérios para a seleção da amostra foram o tipo de escola do aluno (pública ou privada), o sexo e a idade. As normas desenvolvidas com base na primeira fase foram publicadas em 1988 em uma versão abreviada do manual. Atualmente o manual utilizado no Brasil apresenta os resultados das duas fases da normatização e foi publicado no ano de 1999.

Escala de Maturidade Mental Colúmbia

A Escala de Maturidade Mental Colúmbia - EMMC teve sua primeira publicação no ano de 1954. O desenvolvimento da EMMC está relacionado à necessidade de obter estimativas da capacidade mental que fossem adequadas para crianças com deficiência, tais como a paralisia cerebral e outros comprometimentos das funções motoras e/ou verbais (Burgemeister, Blum & Lorge, 2001). A EMMC oferece uma estimativa da capacidade de raciocínio geral para fins de tomada de decisão em diversos contextos.

A normatização brasileira da EMMC foi realizada com a versão de 1972 (terceira versão). Os dados para a normatização foram coletados de 1992 até 1993. Até então, o teste era utilizado no país com as normas desenvolvidas para crianças americanas. A pesquisa de normatização no Brasil contou com a participação de 1.535 crianças com idades entre 3 anos e 6 meses até 9 anos e 11 meses (Alves & Duarte, 2001). As aplicações ocorreram na cidade de São Paulo a fim de estabelecer normas regionais. O critério de seleção da amostra foi pelo tipo de escola: estadual, municipal e particular. O Manual da EMMC foi publicado no Brasil em 2001.

TONI- 3

A primeira versão do teste TONI foi desenvolvida em 1982 (Brown, Sherbenou & Johnsen, 1982). O teste TONI foi construído para avaliar pessoas cujo desempenho em um teste de inteligência poderia ser afetado por limitações linguísticas ou motoras. A versão de 1982 foi destinada para pessoas na faixa etária dos 5 até 85 anos e 11 meses. A segunda versão do teste, TONI-2 é datada de 1990 e o TONI-3 foi publicado em 1997, apresentando duas formas paralelas, forma A e forma B.

O TONI-3 objetiva avaliar a inteligência geral. O foco do teste está em avaliar o raciocínio abstrato e a resolução de problemas. Na construção do instrumento não foi adotada uma teoria particular sobre a inteligência, entretanto segundo o manual (Brown, Sherbenou & Johnsen, 2006) o teste se adapta a 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.

No Brasil o teste TONI-3, forma A foi publicado no ano de 2006 (Santos, Noronha & Sisto, 2006). O manual do instrumento não traz a descrição detalhada do plano amostral e aspectos sócio demográficos da amostra de normatização brasileira. As normas fornecidas no

manual são relativas a uma amostra de 382 crianças com idades entre 6 e 10 anos, testadas em escolas de cidades próximas à São Paulo.

Os aspectos técnicos e os estudos de validade e fidedignidade para os testes, MPC, SON-R 6-40, EMMC e TONI-3 são descritos na sessão acerca dos instrumentos utilizados no presente estudo. Outras informações são encontradas em seus manuais (Angelini, Alves, Custódio, Duarte & Duarte, 1999; Alves & Duarte, 2001; Santos, Noronha & Sisto, 2006; Tellegen & Laros, 2014).

Método

Participantes

O presente estudo analisou dados coletados em uma amostra de 150 estudantes (54,7 % do sexo feminino). A idade dos participantes variou entre seis e nove anos ($M = 7,7$; $DP = 1,0$). Os estudantes cursavam do 1º ao 5º ano do Ensino Fundamental. Os dados foram coletados em quatro escolas do Distrito Federal, sendo três escolas da rede pública de ensino e uma escola da rede privada. Salienta-se que 74,7 % dos estudantes analisados são provenientes da rede pública de ensino.

Instrumentos

Cada participante respondeu a cinco diferentes instrumentos de avaliação. Quatro instrumentos foram destinados para a avaliação das habilidades cognitivas e um instrumento para avaliar aspectos contextuais das crianças. A seguir são caracterizados os instrumentos utilizados.

1. As Matrizes Progressivas Coloridas de Raven - MPC (Angelini, Alves, Custódio, Duarte & Duarte, 1999) é um instrumento destinado para crianças de 5 até 11 anos de idade, para pessoas com deficiência intelectual e para idosos. A escala é dividida em três séries: A, Ab e B, com 12 itens cada. A tarefa do teste depende da habilidade de completar padrões

contínuos de figuras. Os itens são apresentados para o sujeito em um caderno de estímulos e a criança precisa apontar a resposta correta entre as alternativas apresentadas. O escore total das Matrizes Progressivas Coloridas - MPC na amostra de normatização brasileira apresentou coeficientes de fidedignidade usando o método das duas metades variando de 0,52 a 0,93 nas faixas etárias pesquisadas. As evidências de validade do instrumento MPC foram obtidas por meio do critério de diferenciação de idade (Angelini, Alves, Custódio, Duarte & Duarte, 1999), por meio da análise fatorial exploratória (Pasquali, Wechsler & Bensusan, 2002) e com a aplicação do teste não-verbal de inteligência SON-R 2½-7[a] (Laros, Reis & Tellegen, 2010). Em uma amostra de crianças nos Estados Unidos, Martin e Wiechers (1954) encontraram correlações maiores entre o teste MPC e os subtestes de execução do WISC. A correlação encontrada entre o MPC e o subteste de informação do WISC foi de 0,47, e com o subteste Desenho de Blocos foi de 0,74.

2. A Escala de Maturidade Mental Colúmbia – EMMC (Burgemeister, Blum & Lorge, 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. Os itens da EMMC estão dispostos em uma série de oito escalas ou níveis sobrepostos. Os níveis são apresentados para as crianças de acordo com a idade. Dependendo do nível, a criança precisa responder entre 55 e 66 itens. Os itens da EMMC são imagens impressas em cartões, constituídos de três a cinco desenhos. A tarefa é apontar a figura que é diferente ou que não está relacionada às demais figuras. A aplicação da EMMC tem duração média de 15 a 20 minutos. Estudos de fidedignidade dos escores totais na EMMC indicaram coeficientes que variam de 0,82 a 0,93 usando o método das duas metades. O manual do instrumento apresenta que, durante a pesquisa de normatização para o Brasil, foram obtidas evidências de validade convergente da EMMC aplicando o teste junto com as Matrizes Progressivas Coloridas de Raven - MPC ($r = 0,60$).

3. O TONI-3, forma A é um teste de aplicação individual e é destinado para crianças na faixa etária de 6 até 10 anos. A tarefa do teste é a resolução de problemas usando figuras abstratas. Trate-se de um teste não-verbal de inteligência, pois não requer leitura, escrita, fala ou a audição do testando. Por meio de gestos, o examinador apresenta as imagens que estão no caderno de estímulos. O sujeito deve apontar ou indicar a alternativa escolhida, não existe tempo limite para responder os itens. O teste consiste de 45 itens, organizados de acordo com o nível de dificuldade. A aplicação do TONI-3 é interrompida depois de três respostas incorretas em cinco itens consecutivos. Na pesquisa de normatização brasileira foram encontrados coeficientes de fidedignidade (Alfa de Cronbach) variando de 0,80 até 0,83 (Santos, Noronha & Sisto 2006). As evidências de validade do teste foram obtidas por meio do aumento dos escores brutos conforme a idade e através da correlação com os seguintes testes: Teste Cloze ($r = 0,46$), Desenho da Figura Humana- Escala Sisto ($r = 0,49$), Escala de Reconhecimento de Palavras ($r = 0,31$) e a Escala de Avaliação da Escrita – Forma A ($r = 0,46$).

4. O SON-R 6-40 é um teste não-verbal de inteligência destinado para pessoas na faixa etária de 6 até 40 anos de idade (Tellegen & Laros, 2014). O teste apresenta quatro subtestes: Analogias, Mosaicos, Categorias e Padrões, sendo administrados nesta ordem. No total, o SON-R 6-40 possui 124 itens. Os subtestes Analogias e Categorias são constituídos de três séries paralelas com 12 itens cada, totalizando 36 itens em cada subteste. Os subtestes Mosaicos e Padrões são compostos por duas séries paralelas com 13 itens cada, totalizando 26 itens em cada um dos subtestes. Os itens estão organizados em ordem crescente de dificuldade, dessa forma são apresentados os itens cuja dificuldade corresponde ao nível de habilidade dos sujeitos. Os subtestes Analogias e Categorias são subtestes de múltipla escolha que avaliam raciocínio abstrato. Mosaicos e Padrões são subtestes de execução e avaliam habilidades viso-espaciais. Os escores nos subtestes do SON-R 6-40 apresentam, em média,

um coeficiente de fidedignidade (Alpha de Cronbach) de 0,87 e para o escore total normatizado (QI-SON) foi encontrado um valor médio de 0,95. O instrumento foi normatizado inicialmente para Alemanha e Holanda ($n = 1.200$). Os estudos de evidências de validade do SON-R 6-40 nesses países incluem correlações 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* ($r = 0,79$). Também foram conduzidas análises para verificar a estrutura fatorial do instrumento e estudos adicionais de validade com grupos especiais (Tellegen & Laros, 2014).

5. O questionário contextual é constituído por questões sobre o contexto escolar do aluno, tais como: tipo de escola, série em curso e o número de vezes que o aluno repetiu o ano escolar.

Procedimento

A coleta de dados foi realizada durante o primeiro e o segundo semestre letivo do ano de 2015. Foram treinados dois aplicadores para o processo de coleta e tabulação de dados. Inicialmente, as escolas foram contatadas a fim de realizar a apresentação da pesquisa e a entrega da documentação. Cada escola recebeu uma carta de apresentação da pesquisa, uma cópia da carta do comitê de ética com a aprovação do projeto, as cartas de apresentação da pesquisa para os pais e responsáveis e o Termo de Consentimento Livre e Esclarecido (TCLE) para a participação na pesquisa. As escolas que concordaram em participar da pesquisa entregaram para os alunos a carta de apresentação da pesquisa para os pais e o TCLE. O critério de seleção dos estudantes foi a idade. Portanto, as turmas escolhidas foram àquelas cuja faixa etária estava de acordo com os limites estabelecidos pelo delineamento de pesquisa. A faixa de idade escolhida (6 aos 9 anos) está associada as normas brasileiras disponíveis para os testes que foram aplicados: MPC, EMMC e o TONI-3.

As aplicações ocorreram durante o período das aulas em ambientes reservados pelas escolas. Cada estudante foi testado em duas sessões individuais com duração aproximada de 30 minutos, cada uma. Em cada escola foi adotado um procedimento de alternância na ordem de aplicação dos testes, a fim de evitar vieses no processo de testagem. Dessa forma, a metade dos estudantes em cada escola iniciou respondendo o SON-R 6-40 e depois os demais testes: EMMC, MPC e TONI-3. Para a outra metade de alunos o processo inverso foi realizado, a fim de evitar vieses de aprendizado e o cansaço dos participantes. Ao final do procedimento de coleta foram confeccionados e entregues aos responsáveis relatórios individuais de participação na pesquisa.

Análise dos dados

As análises exploratórias e descritivas foram viabilizadas pelo pacote estatístico SPSS, versão 20.0. Inicialmente foram conduzidas análises exploratórias e descritivas a fim de verificar a integridade da base de dados, a presença de dados ausentes, *outliers* e normalidade. Os escores brutos dos quatro testes (SON-R 6-40, TONI-3, EMMC e MPC) foram transformados em escores normatizados. O TONI-3 e o teste MPC fornecem nos seus manuais apenas escores totais normatizados por meio do escore percentil. A utilização deste tipo de escore não é aconselhada devido ao problema da desigualdade das unidades (Hogan, 2006). Assim, para o MPC e TONI-3 o escore percentil foi transformado em QI desvio. Os escores totais normatizados no SON-R 6-40 (QI-SON) foram corrigidos, uma vez que as normas holandesas /alemãs foram utilizadas e na pesquisa de normatização brasileira do SON-R 2½-7[a] foi encontrada uma diferença de 16,7 pontos na escala de QI entre as amostras o que equivale a 1,1 desvio-padrão.

Para a análise fatorial exploratória foi utilizado o programa computacional FACTOR, versão 10.3.01 (Baglin, 2014; Lorenzo-Seva & Ferrando, 2013). Para a análise fatorial foram construídas parcelas de itens para todos os testes. As parcelas possuíam entre 3 e 4 itens

consecutivos. A justificativa para juntar itens individuais em parcelas de itens é para evitar o surgimento de fatores artificiais, relacionados à dificuldade dos itens dicotômicos (Little, Cunningham, Shahar & Widaman, 2002; Holt, 2004; Rocha & Chelladurai, 2012). Vários estudos mostraram que na análise fatorial de itens dicotômicos podem surgir fatores que não são relacionados com o conteúdo, mas apenas com o grau de dificuldade (De Ayala, 2009; Tran & Formann, 2009; Weng & Cheng, 2005). Nas análises fatoriais não foram incluídos os 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), EMMC (17 parcelas de itens), MPC (9 parcelas de itens) e TONI-3 (5 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 10.03.01. A Análise Paralela de Horn é considerada uns dos procedimentos mais adequados para determinar o número de fatores a serem retidos (Courtney, 2013; Damásio, 2012; Laros, 2012; Timmerman & Lorenzo-Seva, 2011). A acurácia do procedimento AP de Horn fica ainda melhor quando o percentil 95% dos eigenvalues aleatórios é considerado em vez da média desses valores (Crawford & cols., 2010). 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. O MRFA é fortemente recomendado com base em estudos de simulação (Timmerman & Lorenzo-Seva, 2011). Utilizou-se a rotação *Promin* (Lorenzo-Seva, 1999). A precisão dos escores dos quatro testes (fidedignidade) foi avaliada por meio do cálculo do Lambda 2 de Guttman (Tellegen & Laros, 2004; Ten Berge & Zegers, 1978; Sijtsma, 2012).

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 normalizados foi corrigida por meio da fórmula de correção dupla para atenuação (Osborne, 2003; Thompson, 2003). A fórmula da correção é: $r_{xy'} = r_{xy} \div \sqrt{r_{xx} \cdot r_{yy}}$, onde, $r_{xy'}$ é a correlação após a correção, r_{xy} é a correlação bruta entre os testes e r_{xx} e r_{yy} são os coeficientes de fidedignidade dos testes. De acordo com Osborne (2003) e Thompson (2003), a correção para atenuação é necessária para a obtenção de estimativas mais próximas da relação verdadeira entre as variáveis na população, uma vez que essa relação pode ser subestimada quando parte da variância observada é erro. Também foi calculado o coeficiente de determinação para a relação entre os testes, que é a correlação elevada ao quadrado. Este valor representa a variância compartilhada entre os dois testes.

Resultados e Discussão

Inicialmente foi conduzida uma análise fatorial exploratória a fim de verificar a estrutura fatorial dos quatro instrumentos: SON-R 6-40, EMMC, MPC e TONI-3. As análises foram realizadas para cada um dos instrumentos usando parcelas de itens. 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

Amplitude e média das cargas fatoriais depois da rotação Promin e valores de comunalidade antes da rotação (h^2) para a solução fatorial do SON-R 6-40 ($n = 150$).

Subteste	CF Fator 1	Média	CF Fator 2	Média	h^2
Analogias	0,05 – 0,17	0,12	0,40 - 0,63	0,52	0,39
Mosaicos	0,57 - 0,73	0,66	0,07 – 0,15	0,10	0,55
Categorias	(-0,16) - 0,09	-0,04	0,50 - 0,73	0,65	0,39
Padrões	0,75 - 0,86	0,81	(- 0,06) - (-0,01)	-0,03	0,62

Notas. CF = carga fatorial; Percentagem explicada da variância comum= 47%; Método de extração de fatores= *Minimum Rank Factor Analysis* (MRFA).

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. Estes resultados indicam uma solução de dois fatores para o SON-R 6-40, explicando 47% da variância comum.

Os subtestes Mosaicos e Padrões apresentam cargas acima de 0,50, o que pode ser considerado um valor alto (Reise, Waller & Comrey, 2000). Assim, o primeiro fator pode ser interpretado como um fator espacial. Os subtestes Analogias e Categorias apresentam cargas fatoriais altas no segundo fator. Este fator pode ser interpretado como um fator de raciocínio.

A solução bifatorial 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 bifatorial para o instrumento (Laros, Tellegen, Jesus & Karino, 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 viso-espaciais 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, Valentini e Lima (2015) encontram 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). Os autores observaram no manual do SON-R 6-40 (Tellegen & Laros, 2014) que a estrutura fatorial bifatorial dos testes SON tende a desaparecer com o avanço da idade dos sujeitos. Esta observação está em acordo com os resultados do presente estudo, uma vez que os participantes da pesquisa pertencem a faixa etária mais jovem que

pode ser testada com o SON-R 6-40. Os resultados da análise fatorial exploratória para os testes EMMC, MPC e TONI-3 são resumidos na Tabela 2.

Tabela 2

Amplitude e média das cargas fatoriais no fator único e valores de comunalidade (h^2) para a solução fatorial encontrada nos testes: EMMC, MPC e TONI-3.

Teste	Cargas Fatoriais	Média	h^2
EMMC	0,23 – 0,61	0,48	0,59
MPC	0,53 – 0,74	0,60	0,53
TONI-3	0,56 – 0,89	0,74	0,71

Notas. Método de extração de fatores = *Minimum Rank Factor Analysis* (MRFA).

O método da Análise Paralela de Horn indicou uma estrutura unifatorial nos testes EMMC, MPC e TONI-3. Para os três instrumentos, os índices de KMO variaram de 0,73 até 0,87. Em média, as cargas fatoriais para os instrumentos apresentam valores satisfatórios. Entretanto, os dados da Tabela 2 permitem verificar que a Escala de Maturidade Mental Colúmbia (EMMC) apresenta alguns valores de cargas fatoriais insatisfatórios ($< 0,30$). A média das cargas fatoriais do EMMC ficou consideravelmente abaixo das médias dos outros dois testes. O teste que apresentou a maior proporção de variância explicada foi o TONI -3 (56%), seguido pelos testes MPC (37%) e EMMC (24%).

A análise fatorial de todos os instrumentos juntos resultou em uma solução com um fator geral de segunda ordem e dois fatores de primeira ordem. Foi utilizada a solução *Schmid-Leiman*, proposta por Thompson (2005). Nesta solução os fatores de segunda ordem são expressos em termos das variáveis mensuradas, removendo a variância dos fatores de primeira ordem nos fatores de segunda ordem. A solução *Schmid-Leiman* para o SON-R 6-40 encontra-se na Tabela 3.

Tabela 3

Cargas fatoriais de acordo com a solução Schmid-Leiman nos fatores de primeira e segunda ordem para o Fator Espacial e o Fator de Raciocínio do SON-R 6-40 e os testes MPC, EMMC e TONI-3.

Escala	Segunda ordem	Primeira ordem	
	Fator Geral	Fator 1	Fator 2
SON-FE	0,64	0,30	0,06
SON-FR	0,56	-0,09	0,24
MPC	0,49	0,27	0,03
EMMC	0,38	0,23	-0,03
TONI-3	0,57	0,17	0,10
Média das cargas	0,53	0,21	0,09

Notas. SON-FE = Fator Espacial do SON-R 6-40; SON-FR = Fator de Raciocínio do SON-R 6-40.

A partir dos resultados da Tabela 3 observa-se que nos fatores de primeira ordem as cargas fatoriais são inferiores quando comparadas com as cargas no fator de segunda ordem. A diferença entre os valores das cargas fatoriais na Tabela 3 e os valores das cargas fatoriais na solução encontrada anteriormente se deve ao fato de que a solução *Schmid-Leiman* não apresenta variância compartilhada entre os fatores de primeira ordem e de segunda ordem.

O Fator Espacial do SON-R 6-40 apresenta a maior carga fatorial (0,64) no fator de segunda ordem e o EMMC apresenta a menor carga fatorial (Fator Geral). Assim, esses resultados indicam que dos testes aplicados neste estudo, o EMMC está medindo um aspecto mais específico do construto inteligência e o SON-R 6-40 está medindo aspectos mais gerais desse construto. Ao comparar as cargas fatoriais dos demais instrumentos nos fatores de primeira ordem verifica-se que as cargas fatoriais mais altas se encontram no Fator 1. Este resultado fornece indícios de que na amostra estudada os testes MPC, EMMC e TONI-3 medem o mesmo construto que o Fator Espacial do SON-R 6-40.

Os resultados da análise fatorial de segunda ordem usando o método *Schmid-Leiman* sugerem que o SON-R 6-40 está medindo um espectro maior do construto de inteligência em comparação com os outros três testes de inteligência. Este achado está em concordância com

as pressuposições teóricas do modelo CHC, onde há um fator geral de inteligência que está relacionado com fatores amplos e específicos das habilidades cognitivas (Flanagan & Harrison, 2012).

A Tabela 4 apresenta algumas estatísticas descritivas para os escores normatizados nos quatro testes. 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, erro padrão da média e intervalo de confiança de 95% das médias obtidas para os escores normatizados no SON-R 6-40, MPC, EMMC e TONI-3.

Subtestes SON-R 6-40	Média	DP	EP	IC 95%
Analogias	10,90 (7,57)	2,82	0,23	10,45 – 11,35
Mosaicos	9,58 (6,25)	2,58	0,21	9,17 – 9,99
Categorias	11,34 (8,01)	3,14	0,26	10,83 – 11,85
Padrões	10,04 (6,71)	2,57	0,21	9,63 – 10,45
Testes	Média	DP	EP	IC 95%
SON-R 6-40	99,3 (82,6)	12,12	0,99	97,7 - 101,5
MPC	106,5	13,11	1,07	104,4 - 108,6
EMMC	105,1	11,56	0,94	103,3 – 106,9
TONI-3	87,3	15,83	1,29	84,8 - 89,8

Notas. DP = desvio-padrão; EP = erro padrão da média; IC 95% = intervalo de confiança de 95% da média. As médias em parênteses são os valores observados utilizando as normas holandesas/alemãs. Ao lado, estão os valores de média corrigidos.

Os valores da Tabela 4 mostram que para os testes MPC e EMMC são observados escores normatizados mais altos (106,5 e 105,1 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.

Para explicar os escores normatizados mais altos nos testes MPC e EMMC comparados com os escores normatizados do SON-R 6-40 e do TONI-3 é possível destacar que os dois

primeiros instrumentos têm normas anteriores as do TONI-3 e do SON-R 6-40. Neste contexto, é necessário considerar a influência do efeito Flynn (Flynn, 2000) em que se reconhece um aumento nos escores brutos em testes de inteligência com o passar dos anos. Uma consequência prática do efeito Flynn é que testes com normas desatualizadas produzem escores normatizados mais altos. Assim, o efeito Flynn pode oferecer uma explicação para o fato dos testes MPC e EMMC mostrarem escores normatizados mais elevados do que o SON-R 6-40 e o TONI-3. As normas do teste MPC foram estabelecidas há 26 anos e as normas do EMMC há 21 anos. Corrigindo para o efeito Flynn, que representa um aumento médio de 0,3 pontos na escala de QI, por ano, a média corrigida dos escores normatizados no teste MPC seria de 98,67 (106,47 – 7,8) e no EMMC seria de 98,83 (105,13 – 6,3). Depois da correção para o efeito Flynn, os testes MPC e EMMC mostrariam médias nos escores normatizados semelhantes às médias dos escores normatizados do SON-R 6-40. O fato de que as normas dos testes MPC e EMMC foram estabelecidas há respectivamente 26 e 21 anos atrás significa que estes instrumentos carecem de atualização das normas para a população brasileira, e que no processo de tomada de decisões por psicólogos o efeito Flynn precisa ser levado em consideração.

O teste que apresentou o menor valor para a média, e maior desvio-padrão e erro-padrão da média foi o TONI-3. Na pesquisa de normatização brasileira do TONI-3 foi encontrada uma média para os escores brutos totais de 16,57 e desvio-padrão de 6,32 (Santos, Noronha & Sisto, 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). Além disso, o TONI-3 foi o teste que perdeu mais itens durante o agrupamento em parcelas de itens, através do critério de dificuldade, restando apenas 5 parcelas de itens dos 45 itens originais. O manual brasileiro do teste não traz uma caracterização detalhada da amostra,

como por exemplo, se as escolas da amostra eram públicas ou privadas, o que impossibilita a melhor compreensão das diferenças de médias encontradas no presente estudo.

Outro resultado que pode ser destacado é a diferença encontrada para os escores normatizados dos subtestes do SON-R 6-40. Os valores em parênteses são inferiores as médias da amostra de normatização da Holanda/Alemanha ($M = 10$). Os valores foram corrigidos com base na diferença de desempenho das crianças brasileiras e holandesas encontrada na pesquisa de normatização do SON-R 2½-7[a] no Brasil (Laros, Tellegen, Jesus & Karino, 2015). A diferença de 16,7 pontos para a Escala Geral equivale a 1,11 desvio-padrão. Assim, para a correção dos escores normatizados foram acrescentados 3,33 pontos nas médias dos subtestes.

Foram avaliados os índices de fidedignidade dos subtestes do SON-R 6-40 e dos escores totais normatizados dos demais subtestes. De acordo com Hogan (2006) os índices de fidedignidade por meio da consistência interna dos itens podem ser considerados como um pré-requisito 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 do SON-R 6-40, EMMC, MPC e TONI-3.

Propriedades psicométricas	Subtestes do SON-R 6-40			
	Analogias	Mosaicos	Categorias	Padrões
Consistência interna (λ_2)	0,88	0,88	0,89	0,92
Correlação item-total (r_{it})	0,40	0,40	0,39	0,52
Correlação entre os itens	0,18	0,20	0,18	0,30
Proporção de acertos	0,47	0,43	0,41	0,41
Número de itens	36	26	36	26

Parâmetros psicométricos	Escores totais normalizados			
	SON-R 6-40	EMMC	TONI-3	MPC
Consistência interna (λ_2)	0,96	0,86	0,92	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	0,11
Proporção de acertos	0,43	0,58	0,35	0,60
Número de itens	124	92	45	36

Nota. λ_2 = Lambda 2 de Guttman.

Os coeficientes de fidedignidade (Lambda 2 de Guttman) para os escores dos subtestes no SON-R 6-40 variaram de 0,88 a 0,92. O subteste com o maior índice de fidedignidade foi Padrões (0,92). Dos escores totais, o SON-R 6-40 apresenta o maior índice de fidedignidade (0,96), seguido do TONI-3 (0,92). O menor índice é verificado para o teste MPC (0,84). Em relação à proporção de acertos, o teste TONI-3 apresenta 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 MPC (60%) e EMMC (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 normalizado do SON-R 6-40 e os escores normalizados 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 normalizados.

	Coeficientes de correlação (r_s) e coeficientes de determinação (r_s^2)			
	SON-R 6-40	MPC	EMMC	TONI-3
SON-R 6-40	----	0,52 (0,41)	0,45 (0,37)	0,37 (0,32)
MPC	0,72 (0,64)	----	0,61 (0,44)	0,35 (0,27)
EMMC	0,67 (0,61)	0,78 (0,66)	----	0,32 (0,26)
TONI-3	0,61 (0,57)	0,59 (0,52)	0,57 (0,51)	----

Notas. No triângulo inferior são encontrados 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 verificar que todas as correlações corrigidas do SON-R 6-40 com os outros testes são altas, acima de 0,60. O coeficiente de validade mais alto pôde ser observado entre o SON-R 6-40 e o teste MPC ($r = 0,72$), indicando que os dois testes compartilham 52% da variância. A correlação entre os dois testes (MPC e SON-R 6-40) é semelhante aos resultados encontrados por Laros, Reis e Tellegen (2010), onde foi encontrada uma correlação corrigida alta ($r = 0,77$) entre o SON-R 2½-7[a] e o teste MPC. O mesmo estudo encontrou uma correlação corrigida semelhante aos resultados do presente estudo para a correlação entre a EMMC ($r = 0,62$) e o SON-R 2½-7[a].

Também foram analisadas as correlações entre os testes MPC, EMMC e TONI-3 com os dois Fatores do SON-R 6-40, o Fator Espacial (SON-FE) e o Fator de Raciocínio (SON-FR). Os resultados encontram-se na Tabela 7.

Tabela 7

Correlações corrigidas e brutas entre os dois fatores do SON-R 6-40 e os testes MPC, EMMC e TONI-3.

Testes	SON-FE	IC 95%	SON-FR	IC 95%
MPC	0,72 (0,64)	0,63 – 0,79	0,58 (0,51)	0,46 – 0,68
EMMC	0,63 (0,57)	0,52 – 0,72	0,59 (0,54)	0,47 – 0,68
TONI-3	0,56 (0,52)	0,44 – 0,66	0,50 (0,46)	0,37 – 0,61

Notas. SON-FE = Fator Espacial do SON-R 6-40; SON-FR = Fator de Raciocínio do SON-R 6-40; IC 95% = Intervalo de confiança de 95% das correlações entre os escores normatizados.

De acordo com resultados na Tabela 7, verifica-se que a correlação mais alta encontrada foi para o Fator Espacial do SON-R 6-40 e o teste MPC ($r = 0,72$). Este achado é semelhante aos resultados obtidos por Laros, Reis e Tellegen (2010), onde foi encontrada uma correlação corrigida para atenuação maior entre a Escala de Execução do SON-R 2½-7[a] e o teste MPC ($r = 0,68$). Os testes EMMC e o TONI-3 também apresentam maior correlação com o Fator Espacial do SON-R 6-40 (SON-FE) do que com o Fator de Raciocínio (SON-FR). Este achado está em concordância com o que é afirmado nos manuais dos testes MPC, EMMC e TONI-3, de que as tarefas exigem habilidades viso-espaciais (Angelini, Alves, Custódio, Duarte & Duarte, 1999; Burgemeister, Blum & Lorge, 2001; Brown, Sherbenou & Johnsen, 2006).

Em último lugar, buscou-se analisar as diferenças dos escores normatizados entre os grupos que podem ser diferenciados no estudo, tendo em vista a composição da amostra (alunos provenientes da rede pública e privada de ensino). Os resultados indicaram uma diferença significativa entre os alunos da escola privada e das escolas públicas nos escores normatizados dos testes EMMC, SON-R 6-40 e TONI-3 O que significa que para estes instrumentos, os alunos da rede privada de ensino obtiveram médias significativamente superiores quando comparados com os alunos da rede pública. Entretanto, para o teste MPC, a diferença entre escolas públicas e privadas não foi significativa, $t(148) = -1,51$, $p = 0,13$, IC 95% (-8,62; - 1,15). A média dos escores normatizados no teste MPC apresentou uma

diferença significativa entre os alunos que já repetiram o ano escolar e os alunos que nunca repetiram o ano escolar, $t(135) = -2,05$, $p = 0,04$, IC 95% (-10,10 - 0,18). Ressalta-se, ainda, que em nenhum dos quatro instrumentos utilizados neste estudo foram encontradas diferenças significativas nos escores normatizados para o grupo de meninos e meninas.

As análises realizadas para o tipo de escola são pertinentes, pois na literatura são identificados estudos que apontam para a importância de se considerar variáveis de contexto dos alunos (Soares, 2005; Laros, Marciano & Andrade, 2012). Além disso, diversos manuais de testes de inteligência trazem normas distintas para o tipo de escola frequentada, tais como o teste MPC e a EMMC.

Considerações finais

O objetivo principal deste trabalho foi reunir evidências da validade convergente do SON-R 6-40 com os testes MPC, EMMC e TONI-3. Os três últimos instrumentos são testes não-verbais de inteligência e são de grande relevância para a prática profissional do psicólogo no Brasil. O principal resultado esperado foi um alto grau de associação entre os escores obtidos no SON-R 6-40 com os demais testes utilizados. Além disso, foram avaliadas a precisão e a estrutura fatorial dos instrumentos. Dessa forma, buscou-se investigar se os instrumentos estavam mensurando um ou mais fatores de inteligência, nesta amostra. Segundo os manuais, os quatro testes estão medindo um único fator de inteligência. É interessante destacar que a o SON-R 2½-7[a], para crianças entre 2,5 e 7 anos de idade está medindo dois fatores. Neste teste, além do QI geral, são calculados os escores normatizados para a Escala de Execução e a Escala de Raciocínio.

Os resultados encontrados fornecem indícios satisfatórios da validade convergente do SON-R 6-40 com os instrumentos pesquisados. O coeficiente de validade mais alto foi encontrado para o SON-R 6-40 e o teste MPC de Raven (0,72). Em relação à estrutura

fatorial, o SON-R 6-40 apresentou uma estrutura bifatorial, com um fator espacial e um fator de raciocínio. A correlação mais alta para os fatores do SON-R 6-40 foi encontrada entre o Fator Espacial e o teste MPC (0,72). O SON-R 6-40 demonstrou ser o teste que avalia um aspecto mais amplo das habilidades cognitivas, se comparado aos demais instrumentos. Os testes MPC, EMMC e TONI-3 são mais específicos e apresentam maiores correlações com o Fator Espacial do SON-R 6-40. Além disso, foram verificados índices adequados de precisão dos escores obtidos nos instrumentos. Os resultados obtidos são semelhantes aos achados de pesquisas realizadas com o SON-R 6-40 no Brasil e em outros países.

Verificou-se que há carência de informações relevantes nos manuais utilizados. A insuficiência de dados e a falta de atualização das normas são fatores que dificultam a realização de inferências mais adequadas a partir dos escores nos testes psicológicos. Nesse sentido, Hogan (2006) discute que a validade de um teste e as normas são questões diferentes. Mesmo que um teste apresente evidências da sua validade, pode ser que as normas não tenham sido construídas de maneira adequada. É importante que haja um refinamento nos manuais dos testes psicológicos, considerando aspectos teóricos, metodológicos e práticos.

Outra limitação deste estudo está relacionada à faixa etária estudada. O SON-R 6-40 foi desenvolvido para pessoas na faixa etária dos 6 aos 40 anos, entretanto, na presente pesquisa foi avaliada uma amostra com faixa etária restrita. Além disso, a amostra foi de conveniência, com maior número de estudantes de escolas públicas. Para estudos futuros com o SON-R 6-40 se destacam pesquisas com amostras especiais, sobretudo pessoas com deficiência auditiva. O SON-R 6-40 é especialmente adequado para esta população, uma vez que sua aplicação pode ser realizada de maneira verbal ou não-verbal. Também se destaca a importância de avaliar a relação entre os escores normatizados obtidos no SON-R 6-40 e variáveis de critério. Por último, se ressalta a avaliação complementar do instrumento por meio dos modelos da Teoria de Resposta ao Item (TRI), e análises confirmatórias, a fim de se

obter mais informações acerca das propriedades psicométricas dos escores no SON-R 6-40 em amostras representativas da população brasileira. É digno de nota que as evidências de validade aqui apresentadas, não encerram de qualquer maneira, as possibilidades de pesquisas futuras que visem contribuir para o uso e a interpretação adequados dos escores do SON-R 6-40.

Referências

- 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*. São Paulo: CETEPP - Centro Editor de Testes e Pesquisas em Psicologia.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Alves, I. C. B., Alchieri, J. C., & Marques, K. (2001). Panorama geral do ensino das Técnicas de Exame Psicológico no Brasil. In *I Congresso de Psicologia Clínica – Programas e Resumos* (pp. 10-11). São Paulo: Universidade Presbiteriana Mackenzie.
- Alves, I. C. B., & Duarte, J. L. M. (2001). Padronização brasileira da Escala de Maturidade Mental Colúmbia. In 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). São Paulo: Casa do Psicólogo.
- Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research & Evaluation, 19*(5), 1-14. Retrieved from <http://pareonline.net/getvn.asp?v=19&n=5>
- Borsboom, D., Mellenbergh, G. J., & Heerden, V. J. (2004). The concept of validity. *Psychological Review, 111*(4), 1061-1071.
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1982). *Test of nonverbal intelligence*. Austin, TX: PRO-ED.
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2006). *TONI-3 (forma A). Teste de inteligência não verbal. Manual do examinador*. São Paulo: Vetor Editora.
- Burgemeister, B., Blum, L. H., & Lorge, I. (2001). *Escala de Maturidade Mental Colúmbia – CMMS. Manual para aplicação e interpretação* (1ª edição brasileira). São Paulo: Casa do Psicólogo.
- Conselho Federal de Psicologia (2003). *Resolução 002/2003*. Retrieved from <http://site.cfp.org.br/resolucoes/resolucao-n-2-2003/>.
- Courtney, M. G. R. (2013). Determining the number of factors to retain in EFA: Using the SPSS R-menu to make more judicious estimations. *Practical Assessment, Research & Evaluation, 18*(8), 1-14. Retrieved from <http://pareonline.net/getvn.asp?v=18&n=8>
- Crawford, A. V., Green, S. B., Levy, R., Lo, W., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement, 70*(6), 885-901.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281- 302.

- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213 -228.
- Davidson, J. E., & Downing, C. L. (2000). Contemporary models of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 34-52). New York: Cambridge University Press.
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: Guilford.
- Flanagan, D. P., & Harrison, P. L. (2012). *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed.). New York: Guilford Press.
- 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.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora S.A.
- Holt, J. K. (2004). *Item parceling in structural equation modeling for optimum solutions*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus: Ohio, USA.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179- 185.
- 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.
- Laros, J. A. (2012). O uso da análise fatorial: algumas diretrizes para pesquisadores. In L. Pasquali (Ed.), *Análise fatorial para pesquisadores* (pp. 141-160). Brasília: LabPAM.
- 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.
- Laros, J. A., Marciano, J. L. P., & Andrade, J. M. (2012). Fatores associados ao desempenho escolar em português: um estudo multinível por regiões. *Ensaio: Avaliação e Políticas Públicas em Educação*, 20, 623-646.
- 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.
- Laros, J. A., Tellegen, P. J., Jesus, G. R., & Karino, C. A. (2015). *SON-R 2½-7[a]. Manual - Teste não-verbal de inteligência*. São Paulo, SP: Hogrefe-CETEPP.
- 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.

- Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research*, 34, 347-356.
- 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.
- Martin, A. W., & Wiechers, J. E. (1954). Raven's Progressive Matrices and the Wechsler Intelligence Scale for Children. *Journal of Consulting Psychology*, 18(2), 143-144.
- Messick, S. J. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: 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, 19-30.
- Osborne, J. W. (2003). Effect sizes and the disattenuation of correlation and regression coefficients: Lessons from Educational Psychology. *Practical Assessment, Research & Evaluation*, 8(11), 1-5. Retrieved from <http://pareonline.net/getvn.asp?v=8 &n=11>
- Pacico, J. C., & Hutz, C. S. (2015). Validade. In C. S. Hutz, D. R. Bandeira, & C. M. Trentini (Eds.), *Psicometria* (pp. 71 - 84). Porto Alegre, RS: Artmed Editora Ltda.
- Pasquali, L., Wechsler, S., & Bensusan, E. (2002). Matrizes progressivas do Raven infantil: Um estudo de validação para o Brasil. *Avaliação Psicológica*, 2, 95-110.
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12(3), 287-297.
- 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.
- Santos, A. A. A., Noronha, A. P. P., & Sisto, F. F. (2006). Estudos psicométricos no Brasil. In L. Brown, R. J. Sherbenou & S. K. Johnsen (Eds.), *TONI-3 Forma A – Teste de Inteligência Não-Verbal* (pp. 67-120). São Paulo: Vetor Editora.
- Sijtsma, K. (2012). Future of psychometrics: Ask what psychometrics can do for Psychology. *Psychometrika*, 77(1), 4-20.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen [The examination of intelligence with deaf-mute children]*. Nijmegen: Berkhout.
- Soares, J. F. (2005). O efeito da escola no desempenho cognitivo de seus alunos. In A. M. Souza (Ed.), *Dimensões da avaliação educacional* (pp. 174-204). Petrópolis, RJ: Vozes.
- 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. (2014). *SON-R 6-40. Non-verbal intelligence test: Research report*. Göttingen, Germany: 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*, 309-315.
- Ten Berge, J. M. F., & Zegers, F. E. (1978). A serie of lower bounds to the reliability. *Psychometrika*, *43*(4), 575-579.
- Thompson, B. (2003). Understanding reliability and coefficient alpha, really. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 1-24). Thousand Oaks, CA: Sage.
- Thompson, B. (2005). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: 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.
- 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.
- Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.
- Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, *51*(4), 473-481.
- Weng, L. J., & Cheng, C. P. (2005). Parallel analysis with unidimensional binary data. *Educational and Psychological Measurement*, *65*, 697-716.

CONSIDERAÇÕES GERAIS

No Brasil, diversos instrumentos encontram-se disponíveis para o processo de avaliação psicológica. A crescente disponibilização desses instrumentos implica na necessidade de maior conhecimento, por parte dos profissionais, acerca dos seus conteúdos, a fim de avaliar e identificar aqueles instrumentos que melhor atendem a demandas específicas. No campo das habilidades cognitivas, o avanço das pesquisas está diretamente relacionado à adequação dos instrumentos de mensuração para os novos modelos teóricos. Além disso, as medidas de inteligência são influenciadas pelo efeito Flynn, onde há um aumento dos escores brutos de QI com o passar dos anos. Dessa maneira, as normas para testes de inteligência requerem revisão periódica.

Neste contexto, se destaca que a qualidade do processo de avaliação psicológica está relacionada à acurácia e eficácia dos instrumentos de medida. A qualidade dos instrumentos, por sua vez, é analisada por meio das diferentes pesquisas que são realizadas no intuito de se obter evidências de validade e qualidade psicométrica dos escores dos testes para o uso apropriado dos seus resultados. Nesse sentido, se destaca o desenvolvimento do campo da psicometria para a construção e disponibilização de instrumentos que possam garantir a qualidade dos serviços prestados em psicologia.

O objetivo essencial dos manuscritos apresentados nessa dissertação foi obter evidências empíricas da validade dos escores do SON-R 6-40. Os resultados dos dois estudos indicaram uma estrutura bi-fatorial para o instrumento. Este achado é consistente com os resultados encontrados para a versão SON-R 2½-7[a] no Brasil, e reforça o fato de que o SON-R 6-40 tende a apresentar uma estrutura bi-fatorial para pessoas que estão nas faixas etárias iniciais do instrumento. A estrutura de dois fatores do SON-R 6-40 pode estimular mais pesquisas sobre a dimensionalidade do teste e a relação da dimensionalidade com a

variável idade. Além disso, a distinção entre o Fator Espacial e o Fator de Raciocínio é importante ao se considerar as diferentes facetas que constituem a inteligência.

Ao verificar a associação entre os escores normatizados obtidos no SON-R 6-40 e os escores normatizados nos testes MPC, EMMC e TONI-3, observou-se que o SON-R 6-40 é o instrumento que avalia um aspecto mais amplo das habilidades cognitivas, enquanto os demais instrumentos estão voltados para um fator mais específico, as habilidades visoespaciais. Assim, o segundo estudo não só fornece informações acerca do SON-R 6-40, mas também apresenta evidências de validade dos demais instrumentos.

Os resultados atingidos nos manuscritos evidenciam a qualidade psicométrica dos escores do SON-R 6-40. Os coeficientes de fidedignidade dos escores totais no instrumento foram satisfatórios nos dois estudos, sendo 0,95 e 0,96, respectivamente. Os achados são coerentes com as pesquisas realizadas anteriormente com o instrumento quanto à estrutura fatorial, a precisão dos escores e as evidências de validade convergente do teste.

No primeiro manuscrito, chama a atenção o fato de que, apesar do número pequeno de participantes e da baixa variabilidade nos dados, foi possível realizar a análise fatorial exploratória e foram obtidos coeficientes altos de fidedignidade (0,95). Além disso, destaca-se que foi encontrada uma diferença significativa entre o grupo de homens e mulheres para os dois fatores encontrados. Os resultados deste estudo são satisfatórios e fornecem indícios da adequação do uso dos escores do SON-R 6-40 para o processo de diagnóstico da deficiência intelectual.

Os dois estudos corroboram para o processo de validação do SON-R 6-40 no contexto brasileiro, pois reúne importantes evidências de sua validade. Do ponto de vista metodológico, buscou-se empregar técnicas de análise de dados que fornecem informações adicionais sobre a precisão e a estrutura fatorial dos instrumentos utilizados. Aqui que se ressalta que nos manuais brasileiros dos testes MPC, EMMC e TONI-3 não são identificados

estudos sobre a estrutura fatorial dos instrumentos por meio da análise fatorial exploratória. Também não são identificados nos manuais destes instrumentos, estudos adicionais sobre o uso dos escores com amostras de pessoas com deficiência intelectual no Brasil.

A finalidade última dos estudos sobre a validade e precisão dos escores do SON-R 6-40 é disponibilizar para os profissionais de psicologia um instrumento não-verbal de avaliação das habilidades cognitivas, garantindo os princípios de qualidade e rigor psicométrico dos instrumentos de medida. Destaca-se, ainda, que o processo de normatização do SON-R 6-40 para o Brasil objetiva fornecer normas atualizadas e representativas da população brasileira.

Dentre as vantagens de se utilizar o SON-R 6-40 se destacam a abordagem amigável do teste e a testagem adaptativa. O procedimento de testagem possibilita que o examinador estabeleça interação com o testando por meio das instruções nos itens de exemplo e *feedback*. Dessa forma, a aplicação do instrumento oferece a oportunidade adicional de observação e avaliação de aspectos comportamentais durante a testagem. A faixa etária do teste é ampla e permite a avaliação das habilidades cognitivas desde crianças até adultos. A correção do teste pode ser feita tanto com base nas tabelas de normas, quanto pelo *software* de correção.

O SON-R 6-40 é um teste não-verbal de inteligência. Por consequência, a avaliação pode ser realizada de maneira independente da linguagem, assim, o instrumento pode ser aplicado para pessoas com deficiência auditiva, analfabetos, imigrantes sem domínio pleno da língua anfitriã e pessoas com problemas de aprendizado. Para estes grupos, os escores são frequentemente subestimados por instrumentos que exigem habilidades verbais e de conhecimento prévio.

Diante do exposto, é possível destacar a relevância acadêmica e prática deste trabalho, bem como o potencial do teste não-verbal de inteligência SON-R 6-40 para o campo da testagem psicológica no Brasil.

Anexo 1

Carta de apresentação da pesquisa SON-R 6-40 para escolas



Carta de apresentação da pesquisa SON-R 6-40 para escolas

Prezados (as),

Estamos realizando uma pesquisa cujo objetivo é obter evidências de validade do SON-R 6-40 para o Brasil. O SON-R 6-40 é um instrumento não-verbal de avaliação das habilidades cognitivas uma vez que não depende das habilidades de escrita e fala do examinando. O instrumento é voltado para pessoas com idades entre 6 e 40 anos. A pesquisa está vinculada ao Instituto de Psicologia da Universidade de Brasília (UnB), sendo coordenada pelo professor Ph.D. Jacob Arie Laros e pela mestrandia Talita de Araújo Alves.

O teste SON-R 6-40 é aplicado individualmente, com duração média de 1h. Esse teste avalia dois fatores e é composto por quatro subtestes, a saber: testes de raciocínio (Categorias e Analogias) e testes de raciocínio espacial (Mosaicos e Padrões). O projeto foi aprovado pelo Comitê de Ética em Pesquisa do Instituto de Ciências Humanas da UnB (CEP/IH). Os pais dos participantes deverão assinar um termo caso não permitam a participação dos seus filhos na pesquisa. Além disso, a criança/adolescente selecionada também será questionada se deseja participar da pesquisa. Os pais ou a criança/adolescente poderão desistir da participação em qualquer momento da pesquisa sem nenhuma penalidade.

Em contrapartida, a equipe de pesquisa se responsabiliza em oferecer: (a) para a escola: parceria com a UnB - Instituto de Psicologia, bem como contribuir com uma pesquisa desenvolvida por uma Universidade pública do país; (b) para os pais e crianças/adolescentes: um relatório de desempenho no SON- R 6-40.

Dúvidas podem ser enviadas para:

- e-mail: jalaros@gmail.com; talitaalves.psi@gmail.com;

- telefone: (61) 3107-6902.

Atenciosamente,

Jacob Arie Laros
Universidade de Brasília
Laboratório META/UnB – Métodos e Técnicas de Avaliação

Anexo 2

Termo de consentimento para participação na pesquisa



Termo de consentimento para participação na pesquisa

Prezados pais e/ou responsáveis,

Seu filho/tutelado _____ foi convidado para participar da pesquisa de validade do SON-R 6-40. Esta pesquisa é coordenada pelo professor Ph.D. Jacob Arie Laros e pela mestrandia Talita de A. Alves. Esta pesquisa tem como objetivo a construção de um instrumento para a avaliação do desenvolvimento cognitivo de brasileiros na faixa entre 6 e 40 anos. Destaca-se a relevância do estudo para aumentar a compreensão do desenvolvimento cognitivo da população brasileira nesta faixa etária.

Caso permita que seu filho/tutelado participe da pesquisa, ele poderá ser sorteado para responder uma bateria de avaliação das habilidades cognitivas SON-R 6-40 e a três instrumentos complementares. O horário e local das sessões de avaliação serão combinados com o aplicador responsável. Eles são psicólogos ou estudantes de psicologia devidamente treinados.

Os riscos envolvidos com a participação são mínimos, visto que eles responderão, basicamente, a problemas de raciocínio. Ao final da pesquisa ele receberá um relatório final individual com as conclusões das avaliações realizadas, contendo informações úteis sobre suas habilidades cognitivas. Além disso, a participação poderá contribuir significativamente com as discussões científicas sobre o tema.

Todas as informações obtidas serão sigilosas e os nomes dos participantes não serão divulgados em nenhum momento, sob hipótese alguma. A participação é voluntária, o que significa que você (ou o seu filho/tutelado) poderá desistir a qualquer momento, sem que isso traga prejuízo ou penalidade.

Qualquer dúvida que você tiver a respeito desta pesquisa, poderá perguntar ao Dr. Jacob Arie Laros ou a mestrandia Talita de Araújo Alves, pelos e-mails jalaros@gmail.com e talitaalves.psi@gmail.com; ou telefone: (61) 31076903; ou pelo seguinte endereço: Campus Darcy Ribeiro, ICC, sala A1 061/4, Brasília DF. Dúvidas a respeito da ética dessa pesquisa poderão ser encaminhadas ao CEP do Instituto de Ciências Humanas da UnB, localizado no Campus Universitário Darcy Ribeiro, ICC, sala B1 683.

Considerando que a pesquisa oferece o mínimo de risco às crianças e aos adolescentes, e que a escola/instituição do seu filho autorizou a realização da pesquisa, caso você **NÃO PERMITA** que o seu filho participe da pesquisa, você deve assinalar o quadro abaixo, assinar este termo e devolvê-lo no prazo de uma semana. Caso permita que seu filho participe, comunique-o da sua decisão e devolva o termo à escola.

NÃO PERMITO a participação do meu filho _____ na pesquisa

Local, Data e Assinatura

Pesquisadores responsáveis:

Jacob Arie Laros

Talita de Araújo Alves

Anexo 3

Questionário Contextual

Questionário Contextual

Quem respondeu o questionário contextual: () O próprio sujeito () Pai () Mãe
() Outro responsável: _____

Escolaridade (do sujeito): _____

Escolaridade do pai ou responsável: _____

Escolaridade da mãe ou responsável: _____

Renda familiar mensal (soma dos rendimentos mensais de todas as pessoas da casa): _____

Quantas pessoas com mais de 16 anos vivem na casa: _____

Cor da pele: () Branco () Pardo/Mulato () Negro () Amarelo () Indígena

Quantidade em casa de:

Item	Quantidade	Item	Quantidade
televisão	_____	computador	_____
rádio	_____	automóvel próprio	_____
aparelho de DVD	_____	banheiro	_____
máquina de lavar roupa	_____	empregada doméstica	_____
geladeira	_____		

Apenas para os pais de crianças e adolescentes:

Série da criança ou adolescente: _____

Tipo de escola: () Privada - Confessional () Privada - Laica
() Pública - Federal () Pública - Municipal ou Estadual

O aluno já repetiu de ano? () Nunca () Uma vez () Duas vezes ou mais

Estado Civil dos pais: () Solteiros () Casados/união estável () Separados () Viúvo(a)