

## Evidências de Validade Convergente e pela Comparação com Construtos Relacionados para o Teste de Inteligência

Evidence of Convergent Validity and of Validity by Comparison of Related Constructs for the Intelligence Test

Fabián Javier Marín Rueda<sup>1</sup>  
Nelimar Ribeiro de Castro<sup>1</sup>

### Resumo

Este estudo tem por objetivo a busca de evidências de validade convergente e de validade baseada na comparação com construtos relacionados para o Teste de Inteligência (TI), respectivamente pela correlação deste último com o Teste de Não-Verbal de Inteligência (R1) e o Teste de Atenção Concentrada (AC). Participaram do estudo 484 candidatos à Carteira Nacional de Habilitação, com idade entre 18 e 67 anos ( $M=26,56$ ,  $DP=8,81$ ), sendo 63,8% homens. Os resultados indicaram coeficientes de correlação moderados e positivos (acima de 0,509) entre o TI e o R1, o que, de acordo com a literatura, pode ser interpretado como evidência de validade convergente, bem como correlações variando entre 0,20 e 0,50 entre o TI e o AC, indicando evidência de validade pela comparação com testes que avaliam construtos relacionados.

**Palavras-chave:** Validade dos testes; inteligência; psicomетria.

### Abstract

This study aims to establish evidence of convergent validity and of validity by comparison of related constructs for the Intelligence Test (TI) by examining the correlation between the TI and the Test of Nonverbal Intelligence (R1), and the correlation between the TI and the Concentrated Attention Test (AC). Participants consisted of 484 candidates for National Drivers License, of both sexes (63.8% men), with ages ranging from 18 to 67 years old ( $m=26.56$ ,  $SD=8.81$ ). The results showed moderate and positive correlation coefficients (above 0.509) between the TI and the R1, which, according to the literature, can be interpreted as evidence of convergent validity, as well as correlation coefficients ranging from 0.20 to 0.50 between the TI and the AC, indicating evidence of validity by comparison with tests that assess related constructs.

**Keywords:** Validity of tests; intelligence; psychometrics.

<sup>1</sup>Universidade São Francisco

No início do século XX, interessado no estudo dos processos intelectuais, Spearman (1904) avaliou e relacionou o resultado de diversos testes de inteligência e observou que habilidades distintas, como as numéricas, espaciais e verbais, estavam altamente correlacionadas. A partir dessa informação, ele postulou a existência de um fator geral subjacente a toda habilidade específica, o que constituiu sua primeira teoria explicativa da inteligência (Spearman, 1904). Posteriormente, verificou que as habilidades específicas poderiam se agrupar em conjuntos mais ou menos homogêneos, mas diferenciados entre si (Spearman, 1927).

Assim, em sua versão final, a teoria de Spearman sobre a inteligência baseia-se em um fator geral, ou fator *g*, comum a todas as habilidades intelectuais; o fator específico, ou fator *e*, que corresponderia a cada uma delas em particular; e o fator de grupo, comum a um grupo de habilidades afins (Spearman, 1904, 1927).

Apesar da controvérsia contemporânea relacionada à teoria proposta por Spearman, e que ainda persiste, em torno da existência ou não de um fator *g* de inteligência, o fator *g* tem demonstrado

bom poder explicativo para as evidências empíricas a respeito dos mecanismos intelectuais, além de alta correlação com o desempenho em diversas atividades, tais como acadêmicas, profissionais e condução de veículos automotores (Alves, 1998; Alves, Colosio & Ruivo, 1992; Ceci & Willians, 1997; Silva, 2008). Por esse motivo, e também por serem instrumentos de fácil aplicação e correção, os testes de inteligência baseados nessa teoria têm sido amplamente utilizados em diversos contextos e populações (Alves, 2002; Angelini, Alves, Custódio, Duarte & Duarte, 1999; Raven, 2000). Com isso em vista, ainda hoje são desenvolvidos instrumentos de medida que se propõem a avaliar o fator *g* de Spearman, dentre os quais se encontra o Teste de Inteligência (TI) (Rueda & Castro, 2011), recentemente desenvolvido no Brasil.

Os estudos psicométricos iniciais do TI (Rueda & Castro, 2011) indicaram, entre outras propriedades, equivalência de desempenho entre os sexos, ou seja, não foram encontradas diferenças significativas no desempenho médio de homens e mulheres. Além disso, foi observada correlação negativa e significativa ( $r=-0,43$ ;

$p < 0,001$ ) entre o desempenho no TI e a idade, indicando decréscimo na habilidade intelectual em decorrência do envelhecimento. Em complementação a este estudo, determinou-se, por meio da análise de variância, a diferenciação de quatro faixas etárias [ $F(3,557)=41,02$ ,  $p < 0,001$ ], sendo elas 18 e 19 anos; de 20 a 25 anos; de 26 a 33; e, por fim, 34 anos ou mais. Verificou-se melhor desempenho nos grupos mais jovens, o que representou, para o TI, uma evidência de validade baseada na relação com outras variáveis, segundo a nomenclatura dos Padrões de Testagem da APA (American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME], 1999), ou de critério, de acordo com terminologias mais tradicionais (Pasquali, 2001).

Valeressaltarqueabusca de evidências de validade em seus variados tipos tem sido apontada como de extrema importância durante o desenvolvimento de testes psicológicos, principalmente quando utilizados para fins diagnósticos (AERA, APA & NCME, 1999). Deste modo, outros estudos psicométricos têm sido realizados para o TI. O presente artigo apresenta duas novas evidências de validade do tipo convergente e pela associação a construtos relacionados. Esses dois tipos caracterizam-se pela correlação entre o teste para o qual se deseja novas evidências de validade, neste caso o TI, e outros que já possuam um conjunto consistente de evidências.

Espera-se, contudo, níveis de relação distintos nos estudos de validade convergente e pela associação com testes que avaliam construtos relacionados. No caso da validade convergente, o teste deve ser correlacionado a outros instrumentos que avaliem o mesmo construto. Já levando em conta o TI, são necessários outros testes de inteligência, preferencialmente que adotem a proposta do fator  $g$  de Spearman. Por consequência, é aguardado que esses dois testes apresentem alta correlação entre si, contudo, considerando que índices de correlação superiores a 0,60 são raros, devido às diferenças encontradas entre instrumentos e amostras, de acordo com Nunes e Primi (2010) valores acima de 0,509 são aceitos como indicativo de que tais instrumentos mensuram o mesmo construto, o que caracteriza a nomenclatura de evidência de validade convergente.

Por sua vez, na evidência de validade pela relação com testes que avaliam construtos relacionados associa-se o teste a outros que avaliem construtos semelhantes mas que não são convergentes nem

discriminantes ao construto medido pelo teste em processo de validação. Eles devem avaliar processos psicológicos com dimensões relacionadas, mas distintas, apresentando, por consequência, índices de correlações baixas e moderados, ou seja, com valores entre 0,20 e 0,50 (Nunes & Primi, 2010). Todavia, ao se avaliar os coeficientes de correlação encontrados, deve-se ter em mente que são afetados por condições externas ao teste, como, por exemplo, a idade, o sexo e o nível educacional (Anastasi & Urbina, 2000).

Na literatura podem ser encontrados alguns estudos de validade convergente para testes de inteligência. Jesus (2009), por exemplo, investigou as qualidades psicométricas do Teste Não-verbal de Inteligência SON-R 2 ½ -7[a], correlacionando a ele quatro outros testes de inteligência utilizados em populações infantis, *Wechsler Preschool and Primary Scale of Intelligence-Third Edition* (WPPSI-III), Escala Wechsler de Inteligência para Crianças – Terceira Edição (WISC-III), Escala de Maturidade Mental Columbia (EMMC), e Matrizes Progressivas Coloridas de Raven. Embora se trate de um único trabalho, foram usadas três amostras diferentes. Assim, na pesquisa com a WPPSI-III participaram 49 crianças com média de 4,7 anos, na WISC-III foram 50 participantes com 6 e 7 anos e, por fim, foi utilizado um grupo com 120 voluntários entre 5 e 7 anos na EMMC e na Matrizes Progressivas Coloridas de Raven. Em todos os casos houve predominância de meninas, e os resultados indicaram correlações significativas entre a SON-R 2 ½ -7[a] e a WPPSI-III, a WISC-III, a EMMC e as Matrizes Progressivas Coloridas de Raven, respectivamente iguais a 0,75, 0,67, 0,68 e 0,56.

Também estudando o SON-R 2 ½ -7[a], Laros, Reis e Tellegen (2010) associaram-no às Matrizes Progressivas Coloridas de Raven e à EMMC. Deste estudo participaram 120 crianças entre 5 e 7 anos, sendo 68 meninos. No caso, a correlação entre o SON-R 2 ½ -7[a] e as Matrizes Progressivas Coloridas de Raven foi igual a 0,77, e com a EMMC chegou a 0,62. Em ambos os estudos concluiu-se que houve evidências de validade convergente para o SON-R 2 ½ -7[a] por meio da associação com outros testes de inteligência (Jesus, 2009; Laros et al., 2010).

Considerando amostras adultas, Santos, Noronha e Sisto (2005) associaram o Teste Não-verbal de Inteligência R1 — forma B e o Teste G36, ambos construídos segundo o modelo de Spearman. A pesquisa contou com a participação de 78 pessoas entre 15 e 64 anos que cursavam o ensino médio

supletivo e obteve correlação muito alta e positiva entre os instrumentos ( $r=0,80$ ;  $p<0,001$ ), sugerindo alta comunalidade entre os testes. Por sua vez, Noronha, Santos e Sisto (2005) também associaram o Teste Não-verbal de Inteligência R1 — forma B ao Teste dos Relógios, ambos se propondo a avaliar o fator  $g$ , segundo proposição de Spearman. Desse estudo participaram 68 pessoas entre 16 e 65 anos, sendo 51,5% do sexo masculino. Os autores realizaram correlações entre a medida geral do Teste dos Relógios e a pontuação total e os fatores 1 e 2 do R1, obtendo, respectivamente, os índices 0,64, 0,58 e 0,62, todos significativos a 1%. Desse modo, concluíram que o nível de comunalidade entre os testes sugeriu que eles avaliam o mesmo construto.

Em um estudo semelhante, Sisto, Ferreira e Matos (2006) correlacionaram o Teste Não-Verbal de Inteligência — R1 ao Teste Conciso de Raciocínio (TCR) em uma amostra de 65 candidatas à Carteira Nacional de Habilitação, entre 18 e 48 anos. Foram encontradas correlações altas e positivas para a amostra total ( $r=0,60$ ;  $p<0,001$ ) e masculina ( $r=0,65$ ;  $p<0,001$ ) e moderada e positiva para as mulheres ( $r=0,53$ ;  $p<0,001$ ). Quando avaliada a associação entre os testes por faixa etária observou-se que para o grupo entre 18 e 19 anos não foram registradas mudanças importantes, pois o índice de correlação manteve-se alto e positivo ( $r=0,63$ ;  $p<0,001$ ), diferentemente do que ocorreu com o grupo entre 20 e 31 anos ( $r=0,54$ ;  $p<0,001$ ) e, principalmente, entre 32 e 55 anos ( $r=0,40$ ;  $p<0,001$ ), que apresentaram redução. Sobre o último resultado apresentado, os autores concluíram que ambos os testes avaliavam, em sua maior parte, o mesmo construto, ou seja, a inteligência segundo os pressupostos teóricos de Spearman.

Por fim, Rabelo (2008) associou as Matrizes Progressivas de Raven — Escala Avançada e o Teste de Raciocínio Inferencial (RIIn), os quais foram aplicados em 291 estudantes de ensino superior, com 17 a 63 anos de idade. A magnitude de correlação encontrada foi alta e positiva para a população total ( $r=0,61$ ;  $p<0,001$ ), assim como para os homens ( $r=0,63$ ;  $p<0,001$ ) e as mulheres ( $r=0,60$ ;  $p<0,001$ ), separadamente. Esses valores mantiveram-se estáveis quando controlado o efeito das variáveis sexo ( $r=0,61$ ;  $p<0,001$ ) e idade ( $r=0,61$ ;  $p<0,001$ ).

Levando em conta a validade pela relação com construtos relacionados para testes de inteligência, usualmente utilizam-se testes que avaliam outros processos cognitivos, em especial a atenção pela

consistente documentação a respeito da associação entre esses dois construtos. Sobre isso, Rosselló e Roca (1994) afirmaram que a inteligência geral ou fator  $g$ , estaria relacionada a diversos aspectos da atenção, tais como o tempo de reação em tarefas de detecção de sinal e na interferência observada no teste de Stroop. Além disso, os autores salientaram que a associação entre os processos atencionais e intelectuais seria maior em grupos com grande variação no desempenho em inteligência e em crianças. Essas afirmações têm sido corroboradas por evidências empíricas que indicam níveis moderados de correlação entre testes de atenção e inteligência, independente da amostra e do tipo de atenção avaliada (Arruda, 2008; Castro & Rueda, 2009; Cecílio-Fernandes & Rueda, 2007; Rueda & Castro, 2010; Rueda, Lamounier, Sisto, Bartholomeu & Noronha, 2006; Silva, 2008; Sisto, 2006; Sisto, Noronha, Lamounier, Bartholomeu & Rueda, 2006; Sisto, Rueda, Noronha & Bartholomeu, 2007).

Especificamente no contexto do trânsito, podem ser encontrados alguns estudos associando o fator  $g$  a diversos tipos de atenção. Sisto et al. (2006), por exemplo, relacionaram o TCR ao teste de Atenção Sustentada (AS) em um grupo de 432 candidatas à CNH, com idade entre 18 e 73 anos. Eles obtiveram correlações moderadas e positivas com a medida de concentração ( $r=0,43$ ;  $p<0,001$ ) e velocidade com qualidade do AS ( $r=0,47$ ;  $p<0,001$ ). Resultados semelhantes foram encontrados por Rueda et al. (2006) ao associarem esses mesmos testes. Eles obtiveram correlações moderadas e positivas para a medida de Velocidade com Qualidade ( $r=0,40$ ;  $p<0,001$ ) e para a de Concentração ( $r=0,42$ ;  $p<0,001$ ) do AS.

Por sua vez, Sisto (2006) correlacionou o TCR ao Teste de Atenção Concentrada (AC) em uma amostra de 118 candidatas à CNH e obteve correlação moderada e positiva ( $r=0,45$ ;  $p<0,001$ ) entre eles. Já Cecílio-Fernandes e Rueda (2007) associaram o TCR ao Teste de Atenção Concentrada (TEACO-FF) ao avaliarem 78 universitários entre 18 e 43 anos. Nesse caso, as correlações foram baixas e positivas ( $r=0,27$ ;  $p=0,018$ ) mesmo quando controlado o efeito da idade ( $r=0,25$ ;  $p=0,049$ ). Ao se verificar o desempenho por sexo, as mulheres obtiveram melhores resultados.

Também com candidatas à CNH, nesse caso 178 pessoas entre 18 e 73 anos, Sisto et al. (2007) correlacionaram o TCR ao Teste de Atenção Dividida (AD). Foram encontradas correlações moderadas e positivas entre o TCR e a medida de concentração ( $r=0,46$ ;  $p<0,001$ ), e de velocidade com qualidade ( $r=0,38$ ;

$p < 0,001$ ) do AD. Ao se realizar correlações separadamente por sexo, as mulheres obtiveram magnitudes maiores que os homens, sendo, para elas, alta e positiva para a medida de concentração ( $r = 0,63$ ;  $p < 0,001$ ) e moderada e positiva para velocidade com qualidade ( $r = 0,45$ ;  $p < 0,001$ ), enquanto para os homens foram baixas e positivas ( $r = 0,39$ ;  $p < 0,001$  e  $r = 0,33$ ;  $p < 0,001$ ). Contudo, os resultados para a amostra total não apresentaram diferenças ao se controlar o efeito do sexo.

Castro e Rueda (2009) também associaram o TCR ao desempenho em tarefas de atenção, no caso, o Teste de Atenção Alternada. Eles submeteram 48 candidatos à CNH, com idade entre 18 e 46 anos, a ambos os testes e, posteriormente, correlacionaram seus escores. Os resultados apontaram correlação moderada e positiva ( $r = 0,44$ ;  $p = 0,004$ ) quando se avaliou a população geral, e baixa e positiva para os participantes de até 22 anos ( $r = 0,26$ ;  $p = 0,379$ ) e entre 23 e 32 anos ( $r = 0,26$ ;  $p = 0,278$ ), mas alta e positiva para aqueles a partir de 33 anos ( $r = 0,67$ ;  $p = 0,050$ ). Ao se controlar o efeito da idade, não foram observadas mudanças importantes nas magnitudes encontradas anteriormente. Essas correlações foram consideradas indicativas de evidências de validade, apesar dos altos valores  $p$  em alguns casos, pois, como ressaltaram os autores, o valor  $p$  é altamente influenciado pelo número de pessoas, o que justificou os valores  $p$  encontrados em seu estudo.

Por sua vez, Silva (2008) relacionou o desempenho de motoristas entre 18 e 41 anos no Teste Não-verbal de Inteligência – R1 e o AC e o Teste de Atenção Difusa (TADIM). A avaliação ocorreu durante o processo de obtenção da CNH e de renovação da licença, em 2002 e 2007, respectivamente. Com o TADIM foram observadas correlações moderadas e positivas ( $r = 0,49$ ;  $p = 0,058$  e  $r = 0,44$ ;  $p = 0,080$ ), em 2002 e 2007, e com o AC, correlação baixa e positiva ( $r = 0,29$ ;  $p = 0,185$ ) em 2002, e moderada e positiva em 2007 ( $r = 0,39$ ;  $p = 0,065$ ), porém, em nenhum desses casos, foi significativa.

O Teste Não-verbal de Inteligência – R1 também foi associado ao Teste de Atenção Dividida (TEADI) em uma amostra de 116 candidatos à CNH, entre 18 e 52 anos. Nesse estudo, Rueda e Castro (2010) verificaram a relação entre os dois testes por faixa etária, com e sem o efeito da idade. No caso, os resultados indicaram correlações moderadas e positivas entre o R1 e o TEADI para o grupo até 25 anos ( $r = 0,46$ ;  $p < 0,001$ ), a partir de 36 anos ( $r = 0,43$ ;  $p = 0,050$ ) e para a amostra total ( $r = 0,46$ ;  $p < 0,001$ ), e baixa e positiva para os participantes entre 26 e 35

anos ( $r = 0,38$ ;  $p = 0,011$ ). Ao retirar o efeito da idade na correlação não ocorreram modificações importantes, exceto para a amostra total, cuja magnitude passou de moderada para baixa ( $r = 0,38$ ;  $p < 0,001$ ).

Por fim, Arruda (2008) associou o Teste Não-verbal de Inteligência – R1 aos de AC, AD e AS. Participaram do estudo 150 candidatos à CHN, entre 18 e 74 anos. O R1 – Forma A obteve correlação de magnitudes moderadas e positivas com a medida de concentração ( $r = 0,58$ ;  $p < 0,001$ ) e velocidade com qualidade do AD ( $r = 0,56$ ;  $p < 0,001$ ), o mesmo ocorrendo com as medidas de concentração ( $r = 0,54$ ;  $p < 0,001$ ) e velocidade com qualidade do AS ( $r = 0,53$ ;  $p < 0,001$ ) e com o AC ( $r = 0,52$ ;  $p < 0,001$ ).

Em suma, considerando a necessidade de avaliar as qualidades psicométricas de novos testes psicológicos, com destaque à validade (AERA, APA & NCME, 1999), este trabalho teve como objetivo apresentar dois estudos de validade para o TI (Rueda & Castro, 2011). O primeiro, que buscou evidências de validade convergente, foi realizado por meio da correlação com o Teste Não-verbal de Inteligência – R1, enquanto o outro investigou evidências de validade pela relação com construtos relacionados, sendo utilizado o Teste de AC.

## Método

### *Participantes*

Fizeram parte do estudo 484 indivíduos que estavam passando pelo processo de obtenção, mudança, renovação ou adição de categoria da CNH nos estados da Bahia (63,6%) e de São Paulo (36,4%). A idade variou de 18 a 67 anos, com média de 26,56 anos e desvio-padrão de 8,81, sendo 63,8% homens. Em relação à escolaridade, os participantes tinham desde o ensino fundamental incompleto até o ensino superior completo.

### *Instrumentos*

Teste de Inteligência (Rueda & Castro, 2012)

O TI avalia a capacidade de educação de relação, definida por Spearman (1904, 1927) como um dos elementos do fator  $g$ , ou inteligência geral. Destina-se a pessoas com idades entre 18 e 67 anos, de ambos os sexos, e nível de escolaridade entre ensino fundamental incompleto e pós-graduação. A aplicação pode ser



realizada de forma coletiva ou individual, sendo observadas as condições adequadas de testagem.

O instrumento é composto por 30 desenhos com um pedaço faltando, sendo que o respondente deve escolher, dentre seis alternativas constantes na parte inferior do caderno de respostas, a que melhor os completa. Todos os itens são compostos por figuras abstratas. O tempo de aplicação é de 20 minutos.

No que se refere aos estudos psicométricos, o TI apresenta um relacionado à estrutura interna por meio do funcionamento diferencial do item em relação ao sexo e ao estado de residência das pessoas. Quanto às evidências de validade baseadas na relação com variáveis externas, um estudo relacionando o desempenho no TI com a idade das pessoas mostrou correlação negativa e estatisticamente significativa ( $r=-0,43$ ,  $p<0,001$ ). Com base no resultado da correlação, os autores formaram agrupamentos de idade, realizando posteriormente vários estudos para verificar quais faixas etárias melhor se diferenciavam. Dessa forma, e tendo como base o resultado da ANOVA, a melhor configuração das faixas etárias, e que diferenciou as pontuações obtidas, foi 18 e 19 anos, dos 20 aos 25 anos, dos 26 aos 33 e, por fim, a partir de 34 [ $F(3,557)=41,02$ ,  $p<0,001$ ]. Assim, além de obter evidência de validade para o TI, os autores sugerem que as normas do teste sejam apresentadas para cada uma dessas faixas etárias. E também foi verificada a relação entre a pontuação do teste e a variável escolaridade, observando correlação negativa e estatisticamente significativa ( $r=-0,35$ ,  $p<0,001$ ). Nesse sentido, ao aumento da escolaridade correspondeu um aumento também na pontuação no TI.

No que se refere aos estudos de precisão, foram obtidos por meio do coeficiente alfa de Cronbach e pelo método das metades de Spearman-Brown e Guttman, tanto para cada faixa etária quanto para a amostra total. Os coeficientes variaram de 0,87 a 0,94 no que diz respeito às faixas etárias, e de 0,90 e 0,93 considerando a amostra total.

#### Teste Não Verbal de Inteligência – R1 (Alves, 2002)

O Teste Não Verbal de Inteligência – R1 foi criado em 1963 por Rynaldo de Oliveira como instrumento para avaliar a inteligência em motoristas. Atualmente encontra-se disponível para utilização uma versão ampliada e revisada por Alves (2002). Aprovado para uso pelo Conselho Federal de Psicologia, tem como base os pressupostos teóricos do fator  $g$  de Spearman e

compreende 40 problemas a serem resolvidos no tempo limite de 30 minutos. Os itens são compostos por vários tipos de estímulos, dentre eles figuras concretas, um objeto, por exemplo, e abstratas.

O estudo da validade de critério do R1 foi realizado por meio da correlação com o Teste de Matrizes Progressivas de Raven – Escala Geral. Ambos foram aplicados em um grupo de 50 sujeitos, com idade de 18 a 42 anos, sendo que metade deles respondeu inicialmente o R1 e a outra metade o Raven, com intervalo de duas horas entre as aplicações. O coeficiente de correlação de Pearson obtido foi de 0,76 ( $p<0,001$ ), considerado satisfatório.

No que se refere aos índices de precisão, foram obtidos por meio de dois métodos: o teste-reteste e o das metades, por meio da fórmula de Spearman-Brown. Para a amostra total, o coeficiente de correlação entre o teste e o reteste foi de 0,68 ( $p<0,001$ ). O resultado do índice de precisão obtido pelo método das metades na correlação entre números pares e ímpares foi 0,83, que, quando corrigido pela fórmula de Spearman-Brown, corresponde a 0,91. Assim, observa-se que os coeficientes de precisão verificados nessas duas pesquisas foram satisfatórios e podem ser considerados altos.

Para a correção do teste é atribuído um ponto para cada resposta correta e zero para as erradas. Posteriormente eles devem ser somados e o resultado final é transformado em percentil. O manual do R1 apresenta tabelas de percentis adequadas para diversas regiões do Brasil e uma geral abrangendo as localidades.

#### Teste de Atenção Concentrada (Cabraia, 2004)

O Teste de AC de Cabraia foi publicado pela primeira vez em 1967 e conta hoje com edição revisada e ampliada. Tem como objetivo avaliar a atenção concentrada, ou seja, a capacidade de selecionar uma fonte de informação entre todas as disponíveis e conseguir dirigir sua atenção para este estímulo.

Nos estudos de validade relatados no manual encontram-se correlações com o Teste de Atenção Concentrada para Motoristas, formas A e B (TACOM-A e TACOM-B). Para isso utilizou-se uma amostra composta por 439 candidatos à CNH. O coeficiente de correlação de Pearson obtido entre os resultados dos testes TACOM-A e TACOM-B com o AC foram de 0,63 e 0,66, respectivamente, ambos estatisticamente significantes ao nível de 0,01.

A precisão do teste foi obtida por meio do método de teste-reteste, sendo o coeficiente de correlação observado de 0,73 ( $p < 0,001$ ).

A correção do AC é feita obtendo o total de acertos, de erros e omissões, e subtraindo a somatória de erros e omissões dos acertos. Dessa forma, chega-se à pontuação total, que posteriormente é transformada em percentil. Destaca-se que o tempo limite para responder o teste é cinco minutos.

### Procedimentos

Os instrumentos utilizados foram aplicados de forma individual em cada sujeito que procurou a avaliação para a CNH, após o consentimento deles para a participação na pesquisa. O tempo médio foi de aproximadamente 60 minutos, em uma única sessão para todas as aplicações. Os casos tiveram início pelo teste AC, por ser um instrumento mais rápido, com duração de cinco minutos, e que avalia a atenção. Dessa forma, optou-se por apresentá-lo no começo, para então evitar o efeito da fadiga. No caso da aplicação do TI e do R1, houve alternância na aplicação, ou seja, metade da amostra respondeu primeiro o TI e depois o R1, e a outra metade dos participantes começou pelo

R1 seguido pelo TI. A aplicação de cada instrumento levou em conta a orientação específica e padronizada de seu respectivo manual.

### Resultados

Na análise dos resultados, primeiramente são apresentadas as estatísticas descritivas dos testes utilizados, e, posteriormente, encontra-se a correlação entre o TI com o R1 e o AC, tanto para as pontuações totais de cada teste quanto para as faixas etárias apresentadas no manual do TI. Os dados relativos à estatística descritiva encontram-se na Tabela 1.

Verifica-se que as pontuações variaram entre 2 e 30 no caso do TI, entre 5 e 40 no R1, e entre 23 e 147 no AC. A média no TI foi 15,00 (DP=6,69), sendo exatamente o ponto médio de corte do instrumento, uma vez que o mesmo possui 30 exercícios para serem resolvidos. Já a média de pontos no R1 foi 28,80 (DP=5,41) e no AC de 85,89 (DP=22,49). No caso específico do TI, houve concentração entre 8 e 18 pontos, representando 50,4% das respostas totais. Os indivíduos que apresentaram menos de 8 pontos totalizaram 18,4%, sendo que nenhum sujeito fez 0 pontos. No que se refere à pontuação máxima possível (30 pontos), apenas uma pessoa a alcançou.

No R1, as pontuações ficaram concentradas entre 23 e 37 (85,4% do total de sujeitos); apenas duas pessoas obtiveram a pontuação máxima possível. Por fim, no caso do AC, foi observado que apenas um participante alcançou a pontuação máxima possível, ou seja, 147.

Para verificar a relação entre os instrumentos da pesquisa, foi realizada correlação de Pearson adotando como nível de significância 0,05. Levou-se em conta a amostra total estudada, assim como também cada uma das faixas etárias descritas no manual do TI.

**Tabela 1.** Média, desvio padrão, pontuação mínima e máxima para o Teste de Inteligência, Teste Não-verbal de Inteligência – R1 e Teste de Atenção Concentrada

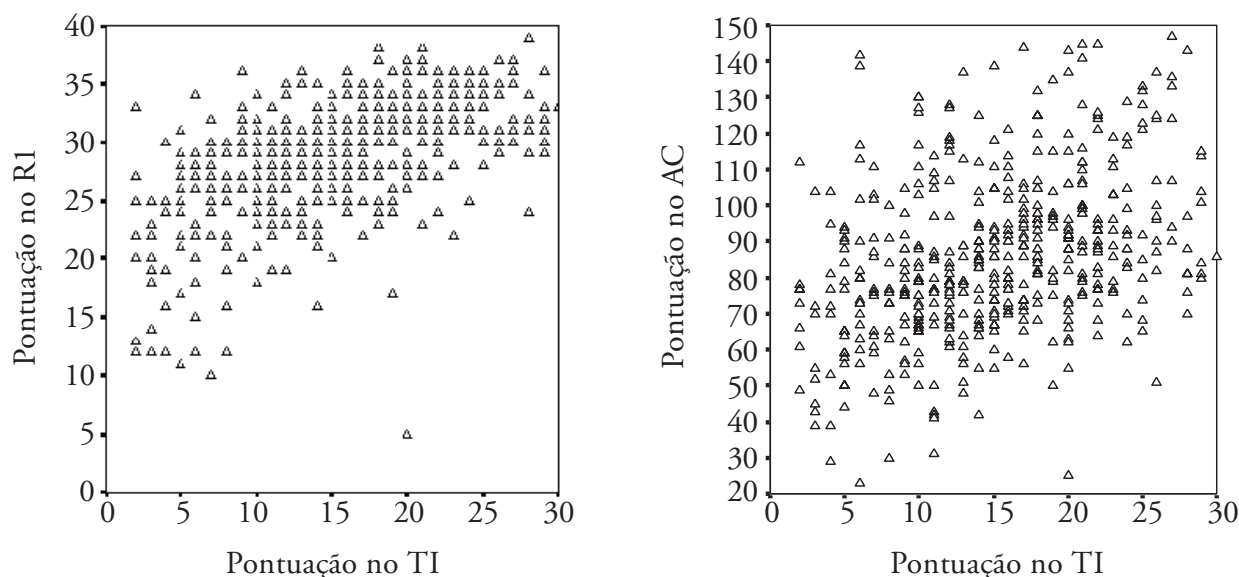
	TI	R1	AC
Média	15,00	28,80	85,89
Desvio padrão	6,69	5,41	22,49
Mínimo	2	5	23
Máximo	30	50	147

TI: Teste de Inteligência; R1: Teste Não-verbal de Inteligência – R1; AC: Teste de Atenção Concentrada

**Tabela 2.** Coeficientes de correlação de Pearson e níveis de significância por faixa etária e para a amostra total entre o Teste de Inteligência com o Teste Não-verbal de Inteligência – R1 e o Teste de Atenção Concentrada

	Faixa etária	R1			AC	
		n	r	p	r	p
TI (anos)	18 e 19	123	0,44	<0,001	0,18	0,047
	20 a 25	134	0,53	<0,001	0,50	<0,001
	26 a 33	126	0,46	<0,001	0,33	<0,001
	34 ou mais	101	0,58	<0,001	0,31	0,003
	Geral	484	0,56	<0,001	0,38	<0,001

TI: Teste de Inteligência; R1: Teste Não-verbal de Inteligência – R1; AC: Teste de Atenção Concentrada; r: coeficiente de correlação de Pearson; p: nível de significância



R1: Teste Não-verbal de Inteligência – R1; TI: Teste de Inteligência; AC: Teste de Atenção Concentrada

**Figura 1.** Gráfico de dispersão da relação entre as pontuações do Teste de Inteligência com o Teste Não-verbal de Inteligência – R1 e do Teste de Inteligência com o Teste de Atenção Concentrada

A variável sexo não foi levada em conta, pois não se observaram diferenças no desempenho de homens e mulheres nesse teste. Os resultados da análise podem ser observados na Tabela 2.

Todos os coeficientes de correlação apresentados na Tabela 2 foram positivos e estatisticamente significativos, indicando que conforme aumentou a pontuação no TI houve incremento também na pontuação no R1 e no AC. No que se refere à intensidade da relação dos coeficientes, entre o TI e o R1 observou-se que todas podem ser classificadas como moderadas, uma vez que se encontram no intervalo de 0,40 a 0,69 (Dancey & Reidy, 2006). Ao transformar os coeficientes em porcentagens, verifica-se que o TI e o R1 têm variância partilhada de 31,3%, o que indica que os instrumentos avaliam boa parte do mesmo construto, porém há uma parcela de especificidades que própria de cada instrumento.

Em relação ao resultado do TI com o AC, observa-se que, com exceção da faixa de 20 a 25 anos, a intensidade das correlações foi fraca, ficando no intervalo de 0,18 a 0,39. Especificamente nessa faixa, a intensidade foi moderada ( $r=0,50$ ) (Dancey & Reidy, 2006). No que se refere à comunalidade de ambos os instrumentos, foi de aproximadamente 15%, indicando que embora possuam aspectos em comum, a maior parte de cada teste avalia construtos diferentes. Para ilustrar a correlação obtida entre os instrumentos foi realizado um gráfico de dispersão

considerando a amostra total da pesquisa, que pode ser observado na Figura 1.

Nela, fica evidenciado que a correlação entre o TI e o R1 foi, de fato, positiva, mostrando que ao aumento da pontuação em um deles correspondeu incremento no outro. Dessa forma, ambos os instrumentos têm em comum parte dos mecanismos implicados na resolução dos itens, ainda que parte da variância indique a presença de outro mecanismo necessário para resolver os itens dos testes. Quanto à relação entre o TI e o AC, observa-se que, embora haja tendência que indique que ao aumentar a pontuação no TI aumentou a pontuação também no AC, a dispersão dos resultados dos testes é maior do que na correlação entre o TI e o R1. Dessa forma, a Figura 1 mostra claramente o resultado da correlação de Pearson observado na Tabela 2.

### Discussão

Considerando o objetivo deste trabalho de realizar dois estudos de validade (convergente e pela relação com construtos relacionados) para o TI (Rueda & Castro, 2012), foram conduzidas as análises estatísticas realizadas. A procura por esses tipos de evidências de validade é considerada de extrema importância na construção de instrumentos de medida que serão utilizados principalmente para fins de diagnóstico (AERA, APA & NCME, 1999).

Nesse sentido, as correlações encontradas entre o TI e o Teste Não-verbal de Inteligência – R1 (Alves, 2002) apresentaram incongruências entre os resultados obtidos para a amostra total e as faixas etárias. De fato, para a amostra total e para os participantes entre 20 e 25 anos, e a partir de 34 anos, os índices de correlação estiveram acima de 0,509, valor apontado por Nunes e Primi (2010) como indicativo de evidência de validade convergente. Contudo, para os grupos entre 18 e 19 e 26 e 33 anos, as correlações, ainda que com magnitudes moderadas, foram inferiores ao critério adotado.

Apesar da aparente contradição, os resultados são semelhantes aos relatados em literatura similar, pois, embora sejam esperados coeficientes de correlação altos, os estudos correlacionando TI em populações brasileiras têm encontrado valores bastante distintos entre si, variando de magnitudes moderadas a altas, sendo que os relatados nesses manuscritos apresentaram índices entre 0,53 e 0,80 (Laros et al., 2010; Jesus, 2009; Rabelo, 2008; Santos et al., 2005; Sisto, et al., 2006).

Somado a isso, esses índices sofreram alterações quando mensurados em amostras distintas. Como exemplo pode-se citar a relação encontrada entre o SON-R 2 ½ -7[a] e os testes EMMC e Matrizes Progressivas Coloridas de Raven nos trabalhos de Jesus (2009) e Laros et al. (2010). No primeiro, os valores foram, respectivamente, 0,68 e 0,56, e no segundo, 0,62 e 0,77. Além disso, semelhante ao que ocorreu entre o TI e o R1, Sisto et al. (2006) obtiveram correlações entre o R1 e o TCR menores entre subgrupos formados por idade e mesmo por sexo, quando comparadas com a encontrada para a amostra total da pesquisa. Nas correlações aferidas, tais achados destacam a interferência das características amostrais, em especial, no caso de inteligência, da idade e principalmente da variável sexo (Anastasi & Urbina, 2000).

Além disso, as diferenças observadas entre os diversos coeficientes de correlação podem ser atribuídas, principalmente, a dois motivos: o erro da medida, próprio a cada instrumento; e diferenças na operacionalização do construto, ou seja, os itens dos testes. Nesse sentido, os itens do TI parecem ser mais homogêneos que os do R1, segundo descrições constantes em seus manuais (Alves, 2002; Rueda & Castro, 2012). Assim, enquanto o R1 apresenta itens compostos por vários tipos de estímulos, como figuras concretas e abstratas, o TI conta apenas com figuras abstratas.

Apesar da aparente incongruência, os resultados obtidos entre o TI e R1 foram ao encontro dos relatados em literatura similar. Desse modo, ainda que as correlações entre o TI e o R1 sugiram, principalmente ao se considerar a totalidade da amostra, evidências de validade convergente, devem ser avaliadas com cuidado, principalmente porque a relação entre esses testes parece sofrer influência considerável da variável idade, talvez pela homogeneidade dos grupos com menores índices, o que deve ser melhor investigado posteriormente.

Por sua vez, no que se refere às análises que compararam o TI com o AC, permitiram atribuir evidência de validade pela relação com construtos relacionados para o TI. Esse tipo de evidência se caracteriza por utilizar testes como variáveis externas que avaliam construtos teórica ou logicamente relacionados ao instrumento, mas que não são construtos convergentes nem discriminantes. A correlação esperada nesses casos encontra-se entre 0,20 e 0,50 (Nunes & Primi, 2010). No caso desta pesquisa, o coeficiente obtido entre os instrumentos foi 0,38.

Este resultado esteve em consonância com os estudos de Castro e Rueda (2009), Rueda e Castro (2010), Silva (2008), Sisto (2006), Sisto et al. (2006) e Sisto et al. (2007), que encontraram correlações inferiores a 0,50, mas superiores a 0,20. Uma diferença deste estudo com as pesquisas citadas também se refere à intensidade dos coeficientes de correlação, uma vez que no presente artigo a intensidade foi classificada, segundo Dancy e Reidy (2006), como fraca, enquanto nos estudos mencionados pode ser classificada como moderada. De qualquer forma, evidência de validade pela relação com construtos relacionados também foi verificada para o TI (Rueda & Castro, 2012).

Ao se observar as correlações das diferentes faixas etárias entre o TI e AC nota-se grande variação nas magnitudes encontradas. Sabe-se que a idade interfere de modo significativo no desempenho em tarefas intelectuais e atencionais, sendo que as faixas etárias utilizadas nesse estudo diferenciaram-se em seu desempenho no TI. Além disso, segundo Rosselló e Roca (1994), a relação entre esses dois construtos tende a ser mais elevada quando existe grande variação na inteligência. Assim, pode-se inferir que o grupo entre 20 e 25 anos apresentou maior variação em seu desempenho no TI, enquanto o grupo entre 18 e 19 anos seria o mais homogêneo. Essa hipótese, contudo, deve ser melhor investigada posteriormente.



Sobre os objetivos propostos terem sido alcançados, sugere-se que novos estudos sejam realizados avaliando amostras com características diferentes, assim como também com grupos critério (categorias profissionais e níveis de escolaridade, dentre outros), a fim de conferir novas evidências de validade para o TI (Rueda & Castro, 2012), uma vez que estudos que realcem as qualidades psicométricas dos instrumentos utilizados devem ser constantemente desenvolvidos, não só para a melhoria do teste, como também para contribuir com o desenvolvimento da área de avaliação psicológica.

## Referências

- Alves, I. C. B. (1998). Variáveis significativas na avaliação da inteligência. *Psicologia Escolar e Educacional*, 2, 109-114.
- Alves, I. C. B. (2002). *R-1 – Teste não verbal de inteligência de Rynaldo de Oliveira: manual*. São Paulo: Vetor Editora Psicopedagógica.
- Alves, I. C. B., Colosio, R., & Ruivo, R. (1992). O Teste R-I: um estudo das variáveis idade, sexo e escolaridade. *Anais da 44ª Reunião Anual da Sociedade Brasileira para o Progresso da Ciência*, 44, 904.
- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME] (1985). *Standards for Psychology and Educational Testing*. Washington, DC: American Psychology Association.
- American Education Research Association [AERA], American Psychology Association [APA] & National Council on Measurement in Education [NCME] (1999). *Standards for Psychology and Educational Testing*. Washington, DC: American Education Research Association.
- Anastasi, A., & Urbina, S. (2000) *Testagem psicológica*. Porto Alegre: Artes Médicas.
- 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*. São Paulo: Centro Editor de Testes e Pesquisa em Psicologia.
- Arruda, J. R. F. S. (2008). *Avaliação da atenção: estudos de validade no contexto do trânsito*. Dissertação de Mestrado. Programa de Pós-Graduação em Psicologia da Universidade São Francisco, Itatiba.
- Cambraia, S. V. (2004). *Teste de Atenção Concentrada*. São Paulo: Vetor Editora Psicopedagógica.
- Castro, N. R., & Rueda, F. J. M. (2009). Testes TEALT e TCR no contexto da avaliação psicológica para motoristas. *Transporte: Teoria e Aplicação*, 2, 27-41.
- Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist*, 52, 1051-1058.
- Cecílio-Fernandes, D., & Rueda, F. J. M. (2007). Evidência de validade concorrente para o Teste de Atenção Concentrada (TEACO-FF). *Revista de Psicologia da Vetor Editora*, 8, 167-174.
- Dancey, C. P., & Reidy, J. (2006). *Estatística sem matemática para psicologia*. Porto Alegre: Artes Médicas.
- Jesus, G. R. (2009). *Normatização e validação do Teste Não-verbal de Inteligência SON-R 2 ½ -7[a] para o Brasil*. Tese de Doutorado Programa de Pós-Graduação em Psicologia da Universidade de Brasília, Brasília.
- Laros, J. A., Reis, R. F., & Tellegen, P. J. (2010). Indicações da validade convergente do Teste Não-Verbal de Inteligência SON-R 2½-7[a]. *Avaliação Psicológica*, 9, 43-52.
- Noronha, A. P. P., Santos, A. A. A., & Sisto, F. F. (2005). Teste dos Relógios e R1 – Forma B: evidências de validade. *Psico*, 36, 243-250.
- Nunes, C. H. S. S., & Primi, R. (2010). Aspectos técnicos e conceituais da ficha de avaliação dos testes psicológicos. In Conselho Federal de Psicologia, *Avaliação psicológica: diretrizes na regulamentação da profissão*, (pp. 101-128). Brasília: Conselho Federal de Psicologia.
- Pasquali, L. (2001). Parâmetros psicométricos dos testes psicológicos. In L. Pasquali (Ed.). *Técnicas de Exame Psicológico – TEP: manual*, (pp. 111-136). São Paulo: Casa do Psicólogo/Conselho Federal de Psicologia.
- Rabelo, I. S. (2008). *Matrizes progressivas avançadas de Raven e teste de raciocínio inferencial: evidências de validade*. Dissertação de Mestrado. Programa de Pós-Graduação em Psicologia da Universidade São Francisco, Itatiba.
- Raven J. C. (2000). *Matrizes progressivas – escala geral*. Rio de Janeiro: CEPA.
- Rosselló, J., & Roca, E. M. (1994). El mecanismo atencional: estudio de las diferencias individuales. *Revista de Psicología General y Aplicada*, 47, 383-390.
- Rueda, F. J. M., & Castro, N. R. (2010). Atención dividida e inteligencia en el contexto del tránsito. *Ciencias Psicológicas* 4, 17-26.

- Rueda, F. J. M., & Castro, N. R. (2012). Teste de Inteligência (TI). Manual. Relatório técnico. Itatiba: Universidade São Francisco.
- Rueda, F. J. M., Lamounier, R., Sisto, F. F., Bartholomeu, D. & Noronha, A. P. P. (2006). Raciocínio inferencial e atenção sustentada: construtos convergentes? In C. Machado, L. S. Almeida, M. A. Guisande, M. Gonçalves & V. Ramalho (Orgs.). *XI Conferência Internacional. Avaliação psicológica: formas e contextos*. Braga: Psiquilibrios.
- Santos, A. A. A., Noronha, A. P. P., & Sisto, F. F. (2005). Teste de Inteligência R1-Forma B e G36: evidência de validade convergente. *Estudos de Psicologia, 10*, 191-197.
- Silva, F. H. V. C. (2008). *Medida da preditividade de instrumentos psicológicos nas avaliações psicológicas de condutores*. Dissertação de Mestrado. Programa de Pós-Graduação em Psicologia da Universidade Federal do Rio Grande do Norte, Natal.
- Sisto, F. F. (2006). *Teste Conciso de Raciocínio*. São Paulo: Vetor Editora Psicopedagógica.
- Sisto, F. F., Ferreira, A., & Matos, M. P. B. (2006). TCR e R1: duas medidas do fator g. *Revista de Psicologia da Vetor Editora, 7*, 69-77.
- Sisto, F. F., Noronha, A. P. P., Lamounier, R., Bartholomeu, D., & Rueda, F. J. M. (2006). *Testes de atenção dividida e sustentada: manual*. São Paulo: Vetor Editora Psicopedagógica.
- Sisto, F. F., Rueda, F. J. M., Noronha, A. P. P., & Bartholomeu, D. (2007). Atenção dividida e inteligência: evidências de validade. *Revista de Psicologia, XI*, 117-131.
- Spearman, C. (1904). "General Intelligence", objectively determined and measured. *The American Journal of Psychology, 15*, 201-292.
- Spearman, C. (1927). *Las habilidades del hombre: su naturaleza y medición*. Buenos Aires: Paidós.

#### **Endereço para correspondência:**

Fabián Javier Marín Rueda  
 Programa de Pós-Graduação Stricto Sensu em  
 Psicologia – Universidade São Francisco  
 Rua Alexandre Rodrigues Barbosa, 45  
 CEP 13251-900 – Itatiba/SP  
 E-mail: fabian.rueda@usf.edu.br

Recebido em 09/02/2012  
 Revisto em 24/09/2012  
 Aceito em 20/10/2012