

 <p>ESCOLA DE CIÊNCIAS DA SAÚDE E DA VIDA</p>	<p><b>PSICO</b></p> <p>Psico, Porto Alegre, v. 53, n. 1, p. 1-12, jan.-dez. 2022 e-ISSN: 1980-8623   ISSN-L: 0103-5371</p>
<p><a href="http://dx.doi.org/10.15448/1980-8623.2022.1.36638">http://dx.doi.org/10.15448/1980-8623.2022.1.36638</a></p>	

SEÇÃO: ARTIGO

## Uso da modelagem multitraço-multimétodo no teste de raciocínio abstrato e espacial (TRAE)

*Multitrait-multimethod model applied on the reasoning and spatial intelligence test (TRAE)*

*Utilización del modelo multirrasgo-multimétodo en la prueba de razonamiento abstracto y espacial (TRAE)*

### Felipe Valentini<sup>1</sup>

[orcid.org/0000-0002-0198-0958](https://orcid.org/0000-0002-0198-0958)  
[valentini.felipe@gmail.com](mailto:valentini.felipe@gmail.com)

### Leonardo de Barros

#### Mose<sup>1</sup>

[orcid.org/0000-0002-5328-7442](https://orcid.org/0000-0002-5328-7442)  
[leonardo.mose@hotmail.com](mailto:leonardo.mose@hotmail.com)

### João Paulo Araújo

#### Lessa<sup>1</sup>

[orcid.org/0000-0003-0751-7662](https://orcid.org/0000-0003-0751-7662)  
[lessa.jpa@gmail.com](mailto:lessa.jpa@gmail.com)

### Jacob Arie Laros<sup>2</sup>

[orcid.org/0000-0001-8763-5401](https://orcid.org/0000-0001-8763-5401)  
[jalaros@gmail.com](mailto:jalaros@gmail.com)

### Ricardo Primi<sup>1</sup>

[orcid.org/0000-0003-4227-6745](https://orcid.org/0000-0003-4227-6745)  
[rprimi@mac.com](mailto:rprimi@mac.com)

**Recebido em:** 12 jan. 2021.

**Aprovado em:** 1 jul. 2021.

**Publicado em:** 7 mar. 2022.



Artigo está licenciado sob forma de uma licença  
[Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).

**Resumo:** O objetivo deste estudo foi obter evidências da fidedignidade e da validade convergente e discriminante do Teste de Raciocínio Abstrato e Espacial (TRAE). O TRAE e a BPR-5 (subtestes RA, RE, RV e RN) foram aplicados em 149 estudantes do ensino médio (52,3% do sexo masculino;  $M_{idade} = 16,98$ ,  $DP = 0,87$ ). Os escores da escala geral do TRAE apresentaram coeficiente de fidedignidade adequado (0,76), entretanto os quatro subtestes foram menos precisos. Utilizou-se a abordagem multitraço-multimétodo, por modelagem por equações estruturais, e verificou-se que a inclusão dos fatores raciocínio abstrato (RA) e espacial (RE) melhorou a adequação do modelo. Tais resultados oferecem suporte à validade convergente do TRAE. Porém, um modelo de correlação perfeita entre RA e RE também foi plausível, o que indica baixa validade discriminante. Portanto, os resultados respaldam a precisão e a validade convergente do TRAE, principalmente, para a escala geral. Todavia, sugere-se que os subtestes sejam interpretados com cautela.

**Palavras-chave:** raciocínio, inteligência, validade do teste, precisão do teste

**Abstract:** In the present study we obtained evidence of the reliability and the convergent and discriminant validity of the Abstract and Spatial Reasoning Test (TRAE, in Portuguese). The TRAE and the BPR-5 (subtests AR, SR, VR, and NR) were both administered to 149 high school students (52.3% male;  $M_{age} = 16,98$ ,  $SD = .87$ ). The total scores of the TRAE showed adequate reliability (.76), despite relatively low reliability of its subtests. The multitrait-multimethod approach in a structural equation modeling context showed that the inclusion of the abstract reasoning (AR) and spatial reasoning (SR) factors improved the model fit. These results support the convergent validity of the TRAE. However, an alternative model with perfect correlations between the AR and SR factors seemed plausible as well, indicating a lack of discriminant validity. These findings support the reliability and the convergent validity of the general scale of the TRAE. Meanwhile, caution is needed with the interpretation of the subtests.

**Keywords:** reasoning, intelligence, test validity, reliability

**Resumen:** En el presente estudio se obtuvo evidencia de la fiabilidad y de la validez convergente y discriminante de la Prueba de Razonamiento Abstracto y Espacial (TRAE, en portugués). El TRAE y la BPR-5 (subpruebas de RA, RE, RV y RN) fueron administrados a 149 estudiantes de secundaria (52,3% niños;  $M_{edad} = 16,98$ ,  $DE = 0,87$ ). Las puntuaciones totales de TRAE mostraron una fiabilidad adecuada (0,76), sin embargo, los escores de las subpruebas fueron menos fiables. El enfoque multi-rasgo/multi-método en ecuaciones estructurales mostró que la inclusión del razonamiento abstracto (AR) y espacial (SR) mejoraron el ajuste del modelo. Estos resultados apoyan la validez convergente del TRAE. Sin embargo, un modelo alternativo con correlaciones perfectas entre los factores

<sup>1</sup> Universidade São Francisco (USF), Campinas, SP, Brasil.

<sup>2</sup> Universidade de Brasília (UnB), Brasília, DF, Brasil.

AR y SR también fue plausible, lo que indica una falta de validez discriminante. Estos hallazgos respaldan la fiabilidad y la validez convergente de la escala general del TRAE. Mientras tanto, se necesita precaución con la interpretación de las subpruebas.

**Palabras clave:** razonamiento, inteligencia, validez, fiabilidad

O estudo da inteligência é um dos empreendimentos bem-sucedidos da psicologia moderna (Wasserman, 2018). Por meio das pesquisas sobre o tema foi possível reunir evidências sobre o funcionamento de diversos aspectos genéticos e ambientais que embasam a inteligência (Lima & Laros, 2017; Primi, 2002; Schneider & McGrew, 2018). Neste sentido, a construção de instrumentos para avaliação cognitiva com evidências de validade e precisão dos escores é de suma importância para a prática e para a ciência psicológica, visto que a inteligência tem um papel relevante na compreensão das diferenças individuais.

A avaliação da inteligência é compreendida como de fundamental relevância, pois é a partir dela que consequências significativas para a vida de um sujeito podem acontecer – como recomendação para programas de intervenção, licença para dirigir, dentre outros. Com isso em mente, é necessário utilizar-se de um modelo que consiga apresentar evidências de validade para assegurar as inferências realizadas (Schneider & McGrew, 2018; Sternberg, 2011). Assim, o modelo Cattell-Horn-Carroll (CHC) tem sido o mais comumente utilizado para embasar instrumentos que visam avaliar o construto da inteligência (Keith & Reynolds, 2010; Laros et al., 2013).

O modelo CHC (Cattell-Horn-Carroll), teorizado por McGrew e Flanagan (cf. Schneider & McGrew, 2018), estabelece uma estrutura multidimensional para a inteligência. Com base em teorias anteriores, os autores propõem uma organização hierárquica dos fatores de inteligência em três níveis: o terceiro nível é formado por uma habilidade geral, cuja dimensão equivalente ao fator  $g$  de Spearman, sendo ele um indicio das operações cognitivas comum a todos os sujeitos (Primi, 2002); o segundo nível envolve 17 dimensões, nas quais estão inclusas a inteligência fluída e a inteligência cristalizada ( $Gf$ - $Gc$ ); e o primeiro nível envolve mais de 80 habilidades específicas.

Entre as dimensões do segundo estrato, a inteligência fluída é a que apresenta maior relação com o fator geral (Gignac, 2014; Primi et al., 2019; Schneider & McGrew, 2018). O conceito de inteligência fluída foi elaborado por Cattell – cuja concepção é anterior ao modelo CHC (Primi, 2002) – e diz respeito às habilidades de adaptação às situações novas e às soluções de problemas que requerem autonomia intelectual por meio da capacidade de raciocínio. Dentre as habilidades específicas relacionadas a este estrato, destacam-se, na literatura científica, os raciocínios abstrato e espacial (Laros et al., 2013; Primi, 2002). O raciocínio abstrato caracteriza-se pela capacidade do indivíduo resolver situações e problemas para os quais possui pouco conhecimento prévio. O raciocínio espacial, por sua vez, está relacionado à capacidade de visualização, manipulação mental de imagens e de orientação no espaço tridimensional (Primi et al., 2019).

A inteligência fluída tem se mostrado preditora de diversas variáveis de desfecho, como comportamento e realização criativa (Beatty et al., 2014), habilidades básicas de escrita (Cormier et al., 2016) e desempenho em linguagem (Li & Shi, 2019). Na metanálise de Peng et al. (2019), a inteligência fluída se demonstrou moderadamente associada com leitura ( $r = 0,38$ ) e matemática ( $r = 0,41$ ). A relação entre inteligência fluída e leitura, quando mediada pelo raciocínio visuoespacial ( $Gv$ ) ou raciocínio não matricial (o qual inclui a categoria de raciocínio abstrato) demonstraram ser significativas ( $r = 0,33$  e  $0,35$ ; respectivamente). Estas duas habilidades também se relacionaram significativamente com o desempenho em matemática ( $r = 0,38$  e  $0,43$ ; respectivamente).

No que se refere à mensuração, alguns instrumentos no Brasil são disponíveis para realizar a avaliação da inteligência, mais especificamente a  $Gf$ . Dentre os instrumentos que estão favoráveis para uso na avaliação da inteligência e do raciocínio (Sistema de Avaliação de Testes Psicológicos [SATEPSI], 2019) é possível citar o Beta-III (Rabelo et al., 2015), TCR (Sisto, 2012), BPR-5 (Almeida & Primi, 2015), APM – RAVEN (Nunes & Nunes, 2015) e WISC-IV (Rueda et al., 2016)

Em relação ao Beta-III e TCR, esses são testes curtos e de fácil aplicação, mas conseguem avaliar apenas um aspecto do raciocínio, qual seja, o raciocínio indutivo. O RAVEN, mesmo tendo uma duração média, avalia apenas o fator geral de inteligência (Nunes & Nunes, 2015). Por sua vez, a BPR-5 e o WISC-IV avaliam os tipos de raciocínio de maneira detalhada (Almeida & Primi, 2015; Primi & Almeida, 2000; Rueda et al., 2016); contudo, são baterias de provas longas, que além de tomar tempo do examinando, podendo desmotivá-lo na sua avaliação, também têm um custo maior para aplicação. Desta forma, um instrumento desenvolvido para sanar a lacuna entre tempo de aplicação e avaliação de mais de um aspecto de raciocínio foi o Teste de Raciocínio Abstrato e Espacial (TRAE; Valentini et al., 2020). No entanto, esse instrumento ainda carece de evidências de validade convergente e discriminante, bem como as estimativas de precisão dos escores.

### Evidências de validade convergente e multitraço-multimétodo (MTMM)

O TRAE ainda é um instrumento em processo de angariar evidências de validade e estimativas de precisão dos escores. Deste modo, ressalta-se a necessidade de obter evidências de validade convergente com outros instrumentos que avaliam as mesmas dimensões da inteligência (i.e., raciocínio espacial e raciocínio abstrato). Todavia, o uso da correlação de Pearson para este fim pode levar a interpretações equivocadas, tendo em vista que os coeficientes são atenuados pelo erro de medida (Fan, 2003). Uma das maneiras de corrigir a correlação pelo efeito de atenuação é por meio da equação (1), na qual  $r_{xy}$  = correlação bruta sem correção;  $r_{xx}$  = fidedignidade do teste X; e  $r_{yy}$  = fidedignidade do teste Y.

$$r \text{ corrigida} = r_{xy} \div \sqrt{r_{xx} \cdot r_{yy}} \quad (1)$$

Além disso, quando são utilizadas escalas no mesmo formato, a associação entre duas variáveis pode ser inflada pela variância compartilhada de método (Lance et al., 2010). Uma das estratégias

para amenizar este efeito se dá por meio de um estudo multitraço-multimétodo (MTMM). Nesta abordagem, as relações entre os traços latentes são mensuradas por meio de diferentes fontes empíricas, a fim de verificar, separadamente, qual a variância específica do construto e dos tipos de medidas utilizados na avaliação (Kline, 2016). Este modelo surge com Campbell e Fiske (1959), cujas propostas consistem em avaliar a convergência e a divergência de um teste por meio de uma matriz MTMM, na qual os coeficientes de validade representam a variância comum entre o método e o traço avaliado. Quando a variável latente apresenta uma correlação de maior magnitude em diferentes medidas, há uma evidência de validade convergente do instrumento, ao passo que ela deve exibir menor associação com outros traços independentes para indicar validade discriminante.

Apesar das vantagens que a proposta de Campbell e Fiske (1959) trouxeram para o estudo das medidas, o uso da matriz de MTMM foi criticado por não especificar qual a parcela de variância que concerne aos métodos e aos traços inseridos no delineamento do estudo (Widaman, 1985). Não obstante, o nível de associação entre as variáveis pode ser alterado pelos erros de medida, visto que as correlações da matriz são calculadas apenas por meio dos escores observados (Eid, 2010; Fan, 2003). Para superar estes obstáculos, a abordagem MTMM pode ser utilizada no contexto de modelagem de equações estruturais (SEM). Nesta estratégia de análise, é possível distinguir as variáveis latentes e o método de mensuração dos atributos psicológicos, ponderando o erro de medida. Além disso, a SEM permite testar o nível em que fatores externos explicam a variância dos componentes do modelo, seja de método ou conteúdo (Eid et al., 2006; Koch et al., 2018).

Para analisar a validade convergente e discriminante na abordagem MTMM em equações estruturais, é proposta uma metodologia na qual quatro modelos diferentes são comparados. O primeiro modelo, menos restritivo, possui traços e fatores de método correlacionados livremente. Esse é comparado com um segundo modelo sem a presença de traços latentes, em que há livre cor-

relação entre os fatores de método. Se a diferença entre os ajustes (principalmente o  $\chi^2$ ) de ambos os modelos for significativa, os dados empíricos suportam a convergência entre os construtos.

Em seguida, o primeiro modelo é comparado com o terceiro, no qual a correlação entre os traços é perfeita (fixada em 1) e os fatores de método são correlacionados livremente. Caso os dois modelos se ajustem aos dados na mesma medida, a divergência entre os traços é pouco plausível. Por fim, o primeiro modelo é contraposto com o quarto modelo, no qual os fatores de métodos são ortogonais e a correlação entre os traços é livre. No caso de diferença significativa entre o  $\chi^2$  dos dois, pode haver influência de viés de método na estrutura de covariância. Como resultado, a validade discriminante entre os métodos inseridos no modelo possuirá pouca sustentação empírica (Byrne, 2012; Widaman, 1985). Assim, a aplicação do MTMM no contexto de SEM consegue separar (*disentangle*) a variância do método e do traço, ponderando o erro de medida. Consequentemente, as relações entre os construtos são estimadas de maneira controlada, removendo o efeito da imprecisão e do método.

Apesar de ser uma abordagem ainda pouco utilizada para obter evidências de validade em instrumentos que avaliam habilidades cognitivas, o uso da metodologia MTMM contribuiu relevantemente em algumas pesquisas embasadas no modelo CHC. No estudo de Beaujean et al. (2010) foi investigada a validade de construto de um teste de inteligência fluida e cristalizada e de outros instrumentos que avaliam memória de curto prazo (*Gsm*), os quais divergiram na forma de aplicação (individual ou coletiva). O uso de análises MTMM culminou em um modelo com três fatores de traço correlacionados e nenhum de método, sustentando a hipótese de que este não influenciou os resultados. Por sua vez, Jacobs e Roodenburg (2014) investigaram validade convergente e discriminante de uma escala de autorrelato, outra de autoestima e um teste de desempenho (designados para mensurar *Gf*, *Gc* e *Gv*). Um modelo com os três traços correlacio-

nados e dois fatores de método (autoestima e desempenho) se ajustou aos dados, sendo que os parâmetros dos itens evidenciaram a validade convergente entre os domínios *Gc* e *Gf*.

## O presente estudo

Para ampliar a compreensão teórica da inteligência, bem como subsidiar a prática psicológica, é necessário construir instrumentos adequados à população brasileira. Nesse contexto, nosso objetivo é avaliar os indicadores de fidedignidade do Teste de Raciocínio Abstrato e Espacial (TRA-E), bem como as evidências de validade convergente e discriminante desse instrumento por meio do método MTMM. Para tanto, avaliou-se a relação entre os escores do TRA-E e da BPR-5. Justificamos a utilização da BPR-5 pela semelhança das tarefas e forma de aplicação. Portanto, a parte "multimétodo" do presente estudo foi composta por instrumentos distintos, mas com tarefas semelhantes. Ainda assim, a parte "multimétodo" foi configurada apenas para abarcar a validade convergente dos escores do TRA-E, controlando o efeito de instrumentos distintos. Categorizamos este modelo como multitraço-multimétodos correlacionados, baseado na taxonomia de Widaman (1985).

Como hipóteses, espera-se: (h1) Correlações fortes entre as dimensões gerais do TRA-E e da BPR-5, indicando validade convergente; (h2) Os modelos MTMM com um fator latente se ajustarão melhor aos dados do que o modelo MTMM apenas com o fator de método, indicando validade convergente; (h3) Correlações significativas, no entanto, mais fracas entre diferentes fatores específicos (por exemplo, RA x RV) se comparadas às correlações dos mesmos construtos de diferentes testes (por exemplo, RA-TRA-E x RA-BPR-5), indicando validade discriminante; (h4) O modelo MTMM com uma correlação forçada em 1 entre os fatores RA e RE não se ajustará bem aos dados, indicando validade discriminante.

## Método

### Participantes

Participaram deste estudo 149 estudantes (52,3% do sexo masculino) do ensino médio de duas escolas do Distrito Federal, sendo uma delas da rede pública. Adotou-se o procedimento de amostragem não probabilística. Os participantes tinham idades entre 14 e 19 anos ( $M = 16,98$ ;  $DP = 0,87$ ). Os estudantes cursavam o segundo e o terceiro ano do ensino médio. Destes, 19,9% relataram ter repetido de ano escolar alguma vez. No que se refere à família, 44% dos estudantes relataram que a mãe possuía ensino médio completo e, em 21% dos casos a mãe possuía ensino superior. A maior parte da família (44%) era composta por três pessoas.

### Instrumentos

O primeiro instrumento, o TRAE, busca avaliar as habilidades cognitivas relacionadas aos raciocínios abstrato (RA) e espacial (RE), por meio de dois subtestes. Cada subteste contém 15 questões de múltipla escolha e o tempo limite de aplicação é de 30 minutos (sendo 12 minutos para RA e 18 minutos para RE). Há evidências da estrutura interna e precisão dos escores do TRAE, por meio de Teoria de Resposta ao Item (Valentini et al., 2020). Os indicadores de precisão para este estudo são discutidos nos resultados.

O segundo instrumento, a BPR-5 (Almeida & Primi, 2015), é utilizada para avaliar a habilidade cognitiva geral e as habilidades cognitivas em cinco áreas específicas: raciocínio abstrato (RA), raciocínio verbal (RV), raciocínio espacial (RE), raciocínio numérico (RN) e raciocínio mecânico (RM). Cada subteste contém entre 20 e 25 questões, sendo que o tempo de aplicação total é de aproximadamente 1 hora e 30 minutos. Na presente pesquisa, o subteste de raciocínio mecânico, por ser opcional, não foi utilizado. A BPR-5 pode ser administrada a estudantes entre o sétimo ano do ensino fundamental (Forma A) e o terceiro ano do ensino médio (Forma B). No presente estudo, utilizou-se a Forma B, em

lápiz e papel. Os índices de precisão reportados no estudo original são  $\alpha = 0,83$  para raciocínio abstrato,  $\alpha = 0,80$  para raciocínio verbal,  $\alpha = 0,84$  para raciocínio espacial e  $\alpha = 0,88$  para raciocínio numérico (Almeida & Primi, 2015).

### Procedimentos e análise de dados

Os instrumentos foram aplicados em sessões coletivas, realizadas nas salas de aula. Os estudantes foram voluntários e assinaram o Termo de Assentimento Livre e Esclarecido (TALE), e as escolas permitiram a coleta de dados. Além disso, foi enviado aos pais um termo de informação, conforme orientação do Comitê de Ética. O projeto foi aprovado pelo Comitê de Ética de Ciências Humanas da Universidade de Brasília (parecer 03-06/2012). Para a aplicação, os estudantes responderam, primeiramente, ao Teste de Raciocínio Abstrato e Espacial (TRAE) e, em seguida, à Bateria de Provas de Raciocínio 5 (BPR-5).

Os dados de precisão foram analisados por meio do indicador Lambda 2 de Guttman. As evidências da validade convergente e discriminante foram verificadas por meio de correlações de Pearson e modelagem multitraço-multimétodo (MTMM) no contexto da SEM, descrito na introdução. Os dados foram analisados por meio do *software* MPLus 8.0, com estimador *Maximum Likelihood* (ML) uma vez que foram utilizados os escores brutos dos subtestes, ou seja, a soma do número de acertos. Os ajustes dos modelos foram avaliados por meio dos seguintes critérios:  $CFI \geq 0,95$ ;  $RMSEA \leq 0,06$  (Hu & Bentler, 1999). Para avaliar diferença na comparação de modelos, utilizou-se o DCFI  $\geq 0,01$ , bem como o DAIC, sendo preferido o modelo com menor AIC (Byrne, 2012; Widaman, 1985).

## Resultados

No intuito de responder ao objetivo específico referente à consistência interna do instrumento, calculou-se o coeficiente Lambda 2 de Guttman ( $\lambda_2$ ) para os subtestes e para a escala geral. Esses valores são indicados na Tabela 1 juntamente com as estatísticas descritivas do TRAE e da BPR-5.

**Tabela 1** – Estatísticas descritivas e precisão do TRAE e BPR-5

		Mínimo	Máximo	Média	DP	Precisão *
	RA	4	15	11,91	2,31	0,67
TRAE	RE	3	15	8,91	2,63	0,67
	Geral	8	30	20,83	4,20	0,76
	RV	7	23	15,97	3,09	0,61
BPR-5 (amostra deste estudo)	RA	7	23	16,47	3,10	0,64
	RE	3	20	12,16	4,09	0,81
	RN	2	20	10,89	3,70	0,81
	Geral	25	82	55,49	10,62	0,88
	RV	0	23	15,66	3,81	0,82
BPR-5 (manual)	RA	0	24	15,84	4,25	0,87
	RE	0	20	10,88	4,28	0,84
	RN	0	19	10,28	4,21	0,91
	Geral	0	80	52,67	13,12	0,95

*Nota.* \* Para a amostra do presente estudo, a precisão do TRAE e da BPR-5 foi estimada pelo coeficiente lambda 2 de Guttman ( $\lambda_2$ ); os valores de precisão do manual da BPR-5 referem-se ao método de consistência interna para os estudantes do segundo ano do ensino médio.

Observa-se, na Tabela 1, que a precisão dos escores gerais é maior do que a precisão dos escores dos subtestes. Isso era esperado, uma vez que a estimação da consistência interna é sensível ao aumento do número de itens (o que, de fato, ocorre na escala geral). No que se refere às estatísticas descritivas, nota-se que as médias e os desvios-padrão dos escores da BPR-5 do manual são semelhantes às da amostra da presente pesquisa.

Adicionalmente, avaliou-se, por meio do teste *t*, a diferença entre as médias deste estudo e as do estudo apresentado no manual da BPR-5. Verificou-se diferenças estatisticamente significativas ( $p < 0,05$ ) apenas para a escala geral

e para o subteste RE. Contudo, tais diferenças apresentaram tamanho de efeito pequeno: para a escala geral,  $d = 0,24$  e  $r = 0,12$ ; e para RE,  $d = 0,31$  e  $r = 0,15$ .

Para responder ao objetivo referente à validade convergente e discriminante, buscou-se avaliar as correlações de Pearson entre os escores gerais do TRAE e da BPR-5, bem como entre os subtestes. Ademais, esperava-se que a fidedignidade não perfeita de ambos os testes pudesse influenciar a relação entre o TRAE e a BPR-5. Neste sentido, buscou-se atenuar as correlações para a ausência de fidedignidade perfeita.

**Tabela 2** – Correlações entre os escores dos testes TRAE e BPR-5 e correlações corrigidas para a ausência de fidedignidade perfeita

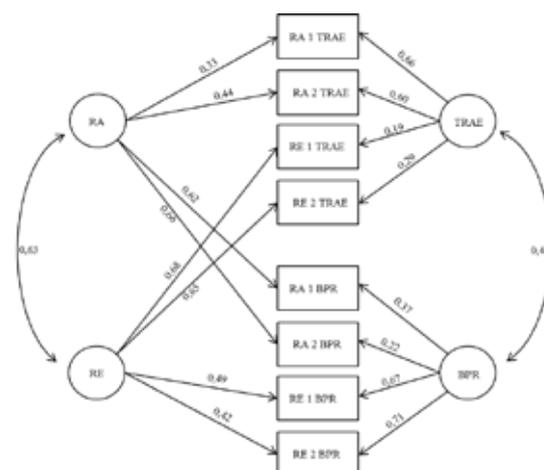
	TRAE-Espacial	TRAE-Abstrato	TRAE-Geral
BPR-Verbal	0,30 (0,47)	0,27 (0,42)	0,34 (0,50)
BPR-Abstrato	0,41 (0,63)	0,41 (0,63)	0,48 (0,69)
BPR-Espacial	0,50 (0,68)	0,43 (0,58)	0,55 (0,70)
BPR-Numérico	0,34 (0,46)	0,37 (0,50)	0,42 (0,54)
BPR-Geral	0,52 (0,68)	0,49 (0,64)	0,60 (0,73)

*Nota.* n = 149; na tabela, são apresentados os coeficientes das correlações de Pearson; entre parênteses são apresentadas as correlações atenuadas para a ausência de fidedignidade perfeita.

Observa-se, na Tabela 2, que a maior correlação ocorreu entre os escores gerais do TRAE e da BPR-5. As correlações atenuadas, apresentadas entre parênteses, indicam a estimativa da correlação caso ambos os testes produzissem escores livres de erro de medida. Nesse sentido, a correlação atenuada entre os escores gerais é forte ( $r$  corrigida = 0,73), o que aponta para a evidência de validade convergente do TRAE. As menores correlações foram entre os escores de raciocínio verbal e numérico da BPR-5 e os escores de raciocínio abstrato e espacial do TRAE. Esses resultados apontam para a evidência de validade discriminante entre os tipos de raciocínio do TRAE e o raciocínio verbal e o numérico da BPR-5. Destaca-se, ainda que, mesmo controlando a imprecisão dos escores, a correlação entre os escores gerais do TRAE e da BPR-5 não foi perfeita (apenas 53% da variância foi compartilhada, i.e.,  $0,73 \times 0,73 = 0,532$ ). Assim, torna-se importante controlar o efeito do método sobre as associações.

Para adicionar o controle do método (diferentes instrumentos) nas análises, os dados também foram analisados por meio da abordagem multitraço-multimétodo (MTMM). No intuito de atender aos pressupostos da análise, os subtestes do TRAE foram avaliados, primeiramente, no que se refere à unidimensionalidade por meio de uma análise fatorial *full information* (FIFA), utilizando o *software* TestFact versão 4.0. Excluíram-se três

itens de cada subteste (dentre os 15 itens por subteste), porque apresentaram cargas fatoriais inferiores a 0,30. Os itens remanescentes foram agrupados em quatro parcelas (*item parcels*): duas parcelas por subteste. As parcelas foram necessárias, pois um modelo MTMM de SEM apenas com dois traços não seria identificável. Finalmente, por meio do coeficiente de Mardia, assumiu-se a normalidade multivariada da distribuição das parcelas (Coeficiente de Mardia de normalidade multivariada = -0,05; Razão Crítica do coeficiente de Mardia = -0,02).



**Figura 1** – Abordagem multitraço-multimétodo para os fatores RA e RE do TRAE e BPR-5. As variâncias residuais foram omitidas da figura. TRAE = Teste de Raciocínio Abstrato e Espacial; BPR= Bateria de Provas de Raciocínio; RA = Raciocínio Abstrato; RE = Raciocínio Espacial; CFI = 1,00; RMSEA (IC 90%) = 0,03 (0,00 - 0,09).

No que se refere à abordagem MTMM, utilizou-se a adaptação do método para a modelagem por equações estruturais (SEM) proposta por Widaman (1985). Nesse sentido, os subtestes são modelados e explicados por fatores latentes, bem como por fatores relacionados aos diferentes métodos (ou diferentes testes), conforme a Figura 1.

Adicionalmente, a MTMM versa a comparação de modelos aninhados, conforme relatado na introdução. Os resultados apresentados na Tabela 3

indicam que o modelo 1 (multitraço-multimétodo) se ajusta significativamente melhor aos dados do que o modelo 2 (apenas controlando o efeito de método). Ademais, a diferença de ajuste entre os modelos 1 (multitraço-multimétodo) e 4 (com efeito 0 de método) é pequena, conforme o esperado. Estes resultados apontam para a validade convergente dos escores, mesmo controlando o efeito do método.

**Tabela 3** – Comparações entre os modelos multitraço-multimétodo

Modelo	$\chi^2$ (gl)	CFI	RMSEA (IC 90%)	AIC
1: latente livre / método livre	15,469 (14)	0,996	0,027 (0,00 - 0,09)	59,469
2: sem latente / método livre	61,923 (23)	0,883	0,110 (0,08 - 0,14)	87,923
3: latente $r = 1$ / método livre	19,326 (15)	0,987	0,045 (0,00 - 0,10)	61,326
4: latente livre / método $r = 0$	19,887 (15)	0,985	0,048 (0,00 - 0,10)	61,887

Comparação dos Modelos	Diferença ( $\Delta$ )		
	$\chi^2$ (gl)	CFI	AIC
Teste da Validade Convergente			
Modelo 1 versus Modelo 2	46,45 (9)	0,113	28,454
Teste da Validade Discriminante			
Modelo 1 versus Modelo 3	3,857 (1)	0,009	1,857
Modelo 1 versus Modelo 4	4,418 (1)	0,011	2,418

*Nota.* Modelo 1 = livre estimação das correlações entre os fatores latentes e os fatores dos métodos; Modelo 2 = exclusão dos fatores latentes; Modelo 3 = correlação entre fatores latentes é fixada em 1; Modelo 4 = livre estimação das correlações entre fatores latentes e correlação entre os fatores dos métodos é fixada em 0.

No entanto as diferenças dos indicadores de ajuste entre os modelos 1 (multitraço-multimétodo) e 3 (monotraço-multimétodo) são pequenas, embora estatisticamente significativas ( $p < 0,05$ ). Portanto, para os subtestes de RA e RE não há diferença relevante em modelá-los como um ou dois traços, após controlar o efeito do método. Este resultado aponta para a baixa validade discriminante dos escores de RA e RE.

### Discussão e considerações finais

Este estudo buscou verificar as evidências de fidedignidade e de validade convergente e discriminante dos escores do Teste de Raciocínio

Abstrato e Espacial (TRAE). O TRAE apresentou consistência interna razoavelmente adequada para a escala geral. Todavia, os coeficientes de fidedignidade das escalas específicas de RA e RE foram abaixo do esperado. Indica-se o baixo número de itens como o principal responsável pelos fracos indicadores de fidedignidade das escalas específicas. Além disso, a variância dos escores foi pequena, principalmente para os subtestes, o que também reduz o índice de consistência interna.

No que se refere aos aspectos de validade, os resultados apontaram para evidências da convergência dos escores do TRAE e da BPR-5, prin-

principalmente para a escala geral, cuja correlação desatenuada entre os instrumentos foi de 0,73. Esse resultado confirma a primeira hipótese, e o tamanho do efeito é semelhante ao apresentado por alguns estudos anteriores, tais como o BETA-III e WAIS-III ( $r_{\text{corrigido}} = 0,77$ ) BETA-III e APM-RAVEN ( $r_{\text{corrigido}} = 0,75$ ) (Rabelo et al., 2015), SON-R 6-40 e BPR-5 ( $r_{\text{corrigido}} = 0,71$ ; Laros, Almeida, Valentini, & Lima, 2015). Destaca-se que, embora o efeito da correlação obtida no presente estudo seja forte, uma parte considerável da variância dos escores não foi compartilhada. A modelagem MTMM pode adicionar o controle do efeito de método na análise das correlações.

Os resultados da modelagem multitraço-multimétodo (MTMM) também oferecem suporte à validade convergente dos instrumentos, confirmando a segunda hipótese. O modelo apresentou significativa diminuição dos indicadores de ajuste frente à remoção dos fatores latentes de RA e RE no modelo 2. Em outras palavras, a estimação de fatores latentes relacionados aos fatores RA e RE é importante para a adequação do modelo, ainda que estes traços tenham sido avaliados por diferentes testes (TRAE e BPR-5). No entanto, por meio da Figura 1, percebe-se que algumas parcelas de itens são mais bem explicadas pelo fator relacionado ao método do que pelo fator latente. Tais resultados indicam que os escores de RA e de RE são influenciados pelas diferenças entre o TRAE e a BPR-5. Isso pode ser explicado parcialmente pela diferença da dificuldade dos instrumentos. O subteste RE da BPR-5, por exemplo, contém alguns itens de dupla rotação, ao passo que os cubos no TRAE rotacionam em apenas uma direção. Ademais, alguns itens de RA do TRAE foram pensados para apresentar fronteira compartilhada, o que ocorre quando as linhas de duas figuras se unem em uma única linha na figura transformada, o que também pode explicar as diferenças entre o TRAE e a BPR-5.

No que diz respeito à validade discriminante, por um lado, RA e RE apresentaram correlações, no máximo, moderadas com o raciocínio verbal e com o raciocínio numérico. Portanto, confirmando a terceira hipótese, RA e RE do TRAE associaram-se de maneira mais forte aos raciocínios da

BPR-5 associados à inteligência fluida (RA e RE da BPR-5) e de maneira mais fraca aos tipos de raciocínio associados à inteligência cristalizada – raciocínio verbal e numérico da BPR-5 (Almeida & Primi, 2015). Esses resultados oferecem suporte à validade discriminante. Por outro lado, as correlações BPR-Abstrato *versus* TRAE-Espacial e BPR-Espacial *versus* TRAE-abstrato foram acima do esperado ( $r > 0,30$ ). Tais resultados são indícios de que essas dimensões não apresentaram clara distinção entre si.

Ademais, por meio da MTMM, ao estimar uma correlação perfeita entre RA e RE (correlação fixa em 1), esperava-se que o modelo 3 apresentasse significativa redução de ajuste em comparação ao modelo 1. No entanto, a semelhança dos valores de  $c^2$ , CFI e AIC indica que a colinearidade de RA e RE é bastante plausível. Em outras palavras, RA e RE podem apresentar problema de validade discriminante. Portanto, não foi possível confirmar a quarta hipótese. Embora não esperados, estes resultados são condizentes com a teoria. Schneider e McGrew (2018) apontam para a sobreposição de  $G_v$  (associado ao RE) e  $G_f$  associado ao RA. Laros et al. (2015) tampouco evidenciaram clara distinção entre estes dois tipos de raciocínio nos estudos do teste SON-R 6-40. Primi e Almeida (2000) também encontraram correlações moderadas entre RA e RE ( $r > 0,45$ ) nas pesquisas da BPR-5. Portanto, a distinção entre RA e RE não é clara na literatura científica (Laros et al., 2015; Primi & Almeida, 2000; Schneider & McGrew, 2018) indicando que as pessoas, provavelmente, resolvem as tarefas de ambos os subtestes utilizando processos mentais semelhantes, possivelmente, relacionados ao raciocínio hipotético-dedutivo.

Conforme o esperado, a diferença de ajuste entre os modelos 1 e 4 é pequena. Este resultado indica que a modelagem é capaz de discriminar os instrumentos. Tal evidência também é importante para sustentar o pressuposto da independência dos métodos, necessário à MTMM (Campbell & Fiske, 1959; Widaman, 1985).

Como limitação do estudo, ressalta-se que a BPR-5 e o TRAE apresentam estímulos e tarefas semelhantes. Portanto, não podem ser conside-

rados multimétodo no sentido literal do termo. Contudo, ao se avaliar a validade convergente de construtos psicológicos, parte das associações podem ser atribuídas ao método. Assim, as correlações, sem o controle do método, tenderiam à superestimação (Lance et al., 2010). Por outro lado, a imprecisão dos instrumentos tenderia a subestimar as associações (Fan, 2003). Assim, o MTMM no contexto de SEM foi utilizado na presente pesquisa exclusivamente para controlar o efeito dos instrumentos (ou método) e da imprecisão. Em suma, ainda que sejam métodos semelhantes, a utilização do MTMM permite avaliar a validade convergente, controlando o efeito de diferentes instrumentos. Estudos futuros podem ampliar a modelagem utilizando métodos e variáveis externas distintas. Por exemplo, seria importante avaliar se a relação entre RA, RE e desempenho acadêmico (variável externa) é alterada controlando o efeito do método por MTMM.

Outra limitação do estudo diz respeito à amostra: relativamente pequena e selecionada por conveniência. Esta limitação pode gerar instabilidade nos modelos apresentados e restringir a generalização. Sugere-se, portanto, a condução de estudos que utilizem amostras mais representativas, em especial, com delineamentos de pesquisa multiníveis com o uso do desempenho acadêmico dos participantes, como uma variável de critério. Ainda, estudos que busquem evidências de validade baseada no processo de resposta podem ser úteis para a interpretação dos escores.

De maneira geral, os resultados discutidos nesta pesquisa respaldam a precisão e a validade convergente do TRAE, principalmente, no que se refere à escala geral de raciocínio. No entanto, em função das limitações relacionadas à validade discriminante e à precisão das escalas de RA e RE, sugere-se que a interpretação dos resultados específicos dos subtestes seja feita com cautela.

## Referências

Almeida, L., & Primi, R. (2015). *BPR-5: Baterias de provas de raciocínio. Manual técnico* (2nd ed.). Casa do Psicólogo.

Beaty, R. E., Nusbaum, E. C., & Silvia, P. J. (2014). Does insight problem solving predict real-world creativity? *Psychology of Aesthetics, Creativity, and the Arts*, *8*(3), 287-292. <https://doi.org/10.1037/a0035727>

Beaujean, A. A., Firmin, M. W., Michonski, J. D., Berry, T., & Johnson, C. (2010). A multitrait-multimethod examination of the reynolds intellectual assessment scales in a college sample. *Assessment*, *17*(3), 347-360. <https://doi.org/10.1177/1073191109356865>

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications and programming*. Routledge.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105. <https://doi.org/10.1037/h0046016>

Cormier, D. C., Bulut, O., McGrew, K. S., & Frison, J. (2016). The role of Cattell-Horn-Carroll (CHC) cognitive abilities in predicting writing achievement during the school-age years. *Psychology in the Schools*, *53*(8), 787-803. <https://doi.org/10.1002/pits.21945>

Eid, M. (2010). Multitrait-multimethod-matrix. In N. Salkind (Ed.), *Encyclopedia of research design* (pp. 850-855). Sage.

Eid, M., Lischetzke, T., & Nussbeck, F. W. (2006). Structural Equation Models for Multitrait-Multimethod Data. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 283-299). American Psychological Association. <https://doi.org/10.1037/11383-019>

Fan, X. (2003). Two approaches for correcting correlation attenuation caused by measurement error: Implications for research practice. *Educational and Psychological Measurement*, *63*(6), 915-930. <https://doi.org/10.1177/0013164403251319>

Gignac, G. E. (2014). Fluid intelligence shares closer to 60% of its variance with working memory capacity and is a better indicator of general intelligence. *Intelligence*, *47*, 122-133. <https://doi.org/10.1016/j.intell.2014.09.004>

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. <https://doi.org/10.1080/10705519909540118>

Jacobs, K. E., & Roodenburg, J. (2014). The development and validation of the Self-Report Measure of Cognitive Abilities: A multitrait-multimethod study. *Intelligence*, *42*(1), 5-21. <https://doi.org/10.1016/j.intell.2013.09.004>

Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools*, *47*(7), 635-650. <https://doi.org/10.1002/pits.20496>

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). The Guilford Press.

Koch, T., Eid, M., & Lochner, K. (2018). Multitrait-Multimethod-Analysis: The Psychometric Foundation of CFA-MTMM Models. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development* (pp. 781-846). Wiley-Blackwell. <https://doi.org/10.1002/9781118489772.ch18>

Lance, C. E., Dawson, B., Birkelbach, D., & Hoffman, B. J. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods*, 13(3), 435-455. <https://doi.org/10.1177/1094428109352528>

Laros, J. A., Almeida, G. O. M., 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. <https://doi.org/10.9788/tp2015.4-10>

Laros, J. A., Jesús, G. R., & Karino, C. A. (2013). Validação brasileira do teste não-verbal de inteligência SON-R 2½-71a. *Avaliação Psicológica: Interamerican Journal of Psychological Assessment*, 12(2), 233-242.

Li, D., & Shi, J. (2019). Fluid intelligence, trait emotional intelligence and academic performance in children with different intellectual levels. *High Ability Studies*, 1-19. <https://doi.org/10.1080/13598139.2019.1694493>

Lima, R. M. F., & Laros, J. A. (2017). Evidências de validade convergente e discriminante dos escores do SON-R 6-40. *Psicologia - Teoria e Prática*, 19(1). <https://doi.org/10.5935/1980-6906/psicologia.v19n1p107-120>

McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR) - Gc-Gf cross battery assessment*. Allyn and Bacon.

Nunes, C. S. S., & Nunes, M. F. O. (2015). *Matrizes Progressivas Avançadas de Raven* (1st ed.). Casa do Psicólogo.

Peng, P., Wang, T., Wang, C. C., & Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/ mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin*, 145(2), 189-236. <https://doi.org/10.1037/bul0000182>

Primi, R. (2002). Inteligência fluída: definição fatorial, cognitiva e neuropsicológica. *Paidéia*, 12(23), 57-75. <https://doi.org/10.1590/S0103-863X2002000200005>

Primi, R., & Almeida, L. (2000). Estudo de validação da Bateria de Provas de Raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa*, 16(2), 165-173. <https://doi.org/10.1590/S0102-37722000000200009>

Primi, R., Correia, T. A., & Almeida, L. (2019). Bateria de Provas de Raciocínio (BPR-5). In C. S. Hutz, D. R. Bandeira, & C. M. Trentini (Eds.), *Avaliação da Personalidade e Inteligência* (1st ed., pp. 109-122). Artmed.

Rabelo, I. S., Pacanaro, S. V., Leme, I. F. A., Ambiel, R. A. M., & Alves, G. A. da S. (2015). *Teste não verbal de inteligência geral: BETA-III: Subtestes raciocínio matricial e códigos* (1st ed.). Casa do Psicólogo.

Rueda, F. J. M., Noronha, A. P. P., Sisto, F., Santos, A. A., & Castro, N. R. (2016). *Manual Técnico WISC-IV* (1st ed.). Casa do Psicólogo.

Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll Theory of Cognitive Abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: theories, tests, and issues* (4th ed., pp. 73-163). Guilford Press.

Sistema de Avaliação de Testes Psicológicos - SATEPSI. (2019). *Testes favoráveis*. Recuperado de <http://satepsi.cfp.org.br/testesFavoraveis.cfm>

Sisto, F. (2012). *Teste Conciso de Raciocínio* (1st ed.). Casa do Psicólogo.

Sternberg, R. J. (2011). The theory of successful intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge Handbook of Intelligence* (pp. 504-527). Cambridge University Press.

Valentini, F., Laros, J. A., & de Barros Mose, L. (2020). Validity Evidence of the Abstract and Spatial Reasoning Test. *Trends in Psychology*, 1-16. <https://doi.org/10.1007/s43076-020-00056-w>

Wasserman, J. D. (2018). A History of Intelligence Assessment: The Unfinished Tapestry. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: theories, tests, and issues* (4th ed., pp. 3-55). Guilford Press.

Widaman, K. F. (1985). Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data. *Applied Psychological Measurement*, 9(1), 1-26. <https://doi.org/10.1177/014662168500900101>

---

### Felipe Valentini

Doutor em Psicologia pela Universidade de Brasília (UnB), em Brasília, DF, Brasil. Mestre em Psicologia pela Universidade Federal do Rio Grande do Norte (UFRN), em Natal, RN, Brasil. Professor do Programa de Pós-Graduação em Psicologia da Universidade São Francisco (USF), em Campinas, SP, Brasil.

---

### Leonardo de Barros Mose

Mestre em Psicologia pela Universidade São Francisco (USF), em Campinas, SP, Brasil. Doutorando em Psicologia pela mesma universidade.

---

### João Paulo Araújo Lessa

Mestre em Psicologia pela Universidade São Francisco (USF), em Campinas, SP, Brasil. Doutorando em Psicologia pela mesma universidade.

---

### Jacob Arie Laros

Doutor em Psicologia pela Universidade Rijksuniversiteit Groningen, em Groninge, Holanda. Professor titular do Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília (UnB), em Brasília, DF, Brasil.

---

**Ricardo Primi**

Doutor em Psicologia Escolar e do Desenvolvimento Humano pela Universidade de São Paulo (USP), em São Paulo, SP, Brasil. Mestre em Psicologia pela Pontifícia Universidade Católica de Campinas (Unicamp), em Campinas, SP, Brasil. Professor do Programa de Pós-Graduação em Psicologia da Universidade São Francisco (USF), em Campinas, SP, Brasil.

---

**Endereço para correspondência****Felipe Valentini**

Universidade São Francisco

Programa de Pós-Graduação *stricto sensu* (mestrado e doutorado) em Psicologia

Rua Waldemar César da Silveira, 105, 2º andar

Cura D'Ars, 13045-510

Campinas, SP, Brasil

*Os textos deste artigo foram revisados pela Poá Comunicação e submetidos para validação do autor antes da publicação.*