

Universidade de Brasília

Instituto de Psicologia

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

A RELAÇÃO ENTRE INTELIGÊNCIA FLUIDA, DESEMPENHO ACADÊMICO E
APRENDIZAGEM: UMA ABORDAGEM MULTINÍVEL

Doutorando: Felipe Valentini

Orientador: Jacob Arie Laros

Brasília - DF
2013

Felipe Valentini

A relação entre inteligência fluida, desempenho acadêmico e aprendizagem:
uma abordagem multinível

Tese elaborada sob orientação do Prof. PhD. Jacob Arie Laros, 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 Doutor em Psicologia.

Brasília

2013

Universidade de Brasília

Instituto de Psicologia

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

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

Prof. Dr. Ricardo Primi (*Membro*)
Universidade São Francisco - USF

Prof^a. Dr^a. Cláudia Cristina Fukuda (*Membro*)
Universidade Católica de Brasília - UCB

Prof^a. Dr^a. Juliana Barreiros Porto (*Membro*)
Universidade de Brasília - UnB

Prof^a. PhD. Isolda Araújo Günther (*Membro*)
Universidade de Brasília - UnB

Prof^a. Dr^a. Girlene Ribeiro de Jesus (*Membro Suplente*)
Universidade de Brasília - UnB

Brasília, 06 de setembro de 2013.

Errar é humano, perdoar é divino; mas incluir o erro na modelagem é estatística.

Leslie Kish

Agradecimentos

*Que viva la ciencia,
Que viva la poesia!*

*El agua esta en el barro,
El barro en el ladrillo,
El ladrillo está en la pared
Y en la pared tu fotografia.*

*Es cierto que no hay arte sin emoción,
Y que no hay precisión sin artesanía.*

*Hay cines,
Hay trenes,
Hay cacerolas,
Hay fórmulas hasta para describir la espiral de una caracola,*

***Hay tantas cosas
Yo sólo preciso dos:
Mi guitarra y vos
Mi guitarra y vos.***

(Autor: Jorge Drexler)

Bom, como o meu violão está jogado num canto há alguns anos, eu posso dizer que só preciso de uma coisa: ‘vos’. Um muitíssimo obrigado à Joana seria pouco para agradecer. Ainda que eu não seja exatamente um poeta, quero te agradecer o apoio durante estes anos tanto na minha vida pessoal quanto profissional. Já são 10 anos juntos... Eu, 10 kg a mais, algumas rugas nos olhos e alguns fios de cabelos a menos. Mudei algumas vezes de óculos, estilo de roupa, tamanho do cabelo e até de cidade. Mas todas as vezes que vejo essas mudanças nas fotografias, você está em todas elas... Não sei exatamente quanto tempo temos ainda pela frente, 50 ou 60 anos, 70 talvez. Sei apenas que quero vivê-los, todos, ao seu lado. Muito obrigado pela sua paciência, carinho e compreensão.

Gostaria de agradecer aos meus familiares. Pai e mãe, o apoio de vocês durante todos estes anos foi fundamental para a concretização desta tese. Ao Bruno, meu irmão, um agradecimento especial em função dos *downloads* (e churrasco)! Também agradeço a nona que sempre foi um grande exemplo de dedicação e coragem. Aos meus tios Toni e Mari muito obrigado pelo carinho e apoio (e carreteiro, também). Ao grande primo (e mais do que isso, um grande irmão) Diogo um muito obrigado por todos os anos de companheirismo e exemplo! Lembro-me das várias vezes que te protegi na escola (mesmo sabendo das tuas travessuras). Bom, faria tudo de novo se fosse preciso! Estamos juntos nessa! Espero que tenhas te esquecido das vezes que brigamos quando éramos crianças...

Também agradeço aos amigos de Porto Alegre, João Basso, Aline, Dani, Mônica, Cecília, Francis, Gustavo, Simone, Rafael Pereira e Eduardo Andrade. Ainda que os nossos encontros tenham diminuído ao longo dos anos, guardo vocês na memória. Aos amigos que deixo em Brasília e Natal, Renato, Fabiana Queiroga, Fabiana Damásio, Girlene, André Moniz, Fábio Cristo, Lilian, Renata Manuely, Gabriel Almeida, Aleksandra, Rodrigo Ferreira, Clara Cantal, Dudu e Hania, também, muito obrigado. Ao Daniel Kimpara e à Camila Akemi um abraço especial em função das sugestões e ‘trocas de ideias’ sobre a pesquisa.

Agradeço aos amigos e colegas Josemberg, João Alchieri, Adriana de Oliveira, Maycoln Teodoro, Marcos Balbinotti, Sílvio Vasconcelos, Valdiney Gouveia, Vicente, Isabel Vasconcelos e Raquel Diniz: vocês são exemplos de dedicação acadêmica. Quando crescer quero ser igual a vocês! Ao Fábio Iglesias, um agradecimento muito especial. Obrigado, Fábio, pelas ideias sobre pesquisa, vida acadêmica, carreira e, principalmente, amizade. Ao Amer e à Ana Paula devo agradecer-lhes a acolhida na UFPR.

Agradeço a dedicação dos meus (ex) alunos Wlad, Laizza, Ronny, Arthur, Talita, Mary, Márcia e Natália e da colega Renata M. Uma parte significativa desta tese pertence a vocês!

Sou grato a todos os professores do PSTO, principalmente ao Hartmut, Isolda, Kátia, Ronaldo, Juliana, Fábio e Jairo. Aprendi muito com vocês! Levarei comigo o exemplo de seriedade e dedicação à pesquisa. Também agradeço aos membros da banca Ricardo, Isolda, Cláudia, Juliana e Gírlene: obrigado pelas sugestões e ‘trocas de ideias’ nos corredores da universidade e nos congressos.

Bom, deixei para agradecer ao Prof. Laros por último, pois a minha gratidão a ele é muito especial. Muito obrigado, professor, pela sua paciência com os meus erros e teimosias... Já estudei e trabalhei com muitos professores. Todos eles foram muito importantes, mas, aproximadamente, 95% do que eu sei, aprendi com o senhor. Agradeço as horas compartilhadas no laboratório, todos os ensinamentos sobre psicometria, inteligência, avaliação (etc...), o exemplo de seriedade na condução de uma pesquisa e conselhos sobre carreira, trabalho e vida! Muito obrigado, mesmo! Como afirmei antes, não sou um poeta, contudo, quero que saibas que o senhor realmente fez muita diferença na minha vida. Gostaria de continuar trabalhando contigo em parcerias de pesquisa, manuscritos, construção de testes, congressos, bancas...

Sou grato também a todos os participantes da pesquisa que disponibilizaram seu tempo para colaborar com o estudo. Estendo esse agradecimento aos aplicadores e profissionais do CESPE-UnB (especialmente a Elisete, Sônia, prof. MV, Tati, Letícia, Luciana e Cássio) que me auxiliaram a coletar e organizar os bancos de dados.

Agradeço o financiamento da pesquisa realizado pela CAPES. Também agradeço ao Centro de Seleção e de Promoção de Eventos (CESPE-UnB) e à Secretaria de Educação pela parceria realizada para a coleta de dados.

Sumário

Lista de Figuras	IX
Lista de Tabelas	X
Lista de Abreviações	XI
Resumo	XII
Abstract	XIII
Apresentação	14
Manuscrito 1. Inteligência e desempenho acadêmico: Revisão de literatura	16
Manuscrito 2. Evidências de validade de construto do Teste de Raciocínio Abstrato e Espacial – TRAE	43
Manuscrito 3. Evidências de validade do Teste de Raciocínio Abstrato e Espacial com base na abordagem multitraço-multimétodo	70
Manuscrito 4. Análise multinível das relações entre inteligência fluida, desempenho acadêmico e aprendizagem	82
Considerações finais	118
Anexos.....	125

Lista de Figuras*Manuscrito 2. Evidências de validade de construto do Teste de Raciocínio Abstrato e Espacial – TRAE*

- Figura 1. Curva de informação do TRAE 55
- Figura 2. Análise fatorial confirmatória do TRAE 59

Manuscrito 3. Evidências de validade do Teste de Raciocínio Abstrato e Espacial com base na abordagem multitraço-multimétodo

- Figura 1. Abordagem multitraço-multimétodo 76

Manuscrito 4. Análise multinível das relações entre inteligência fluida, desempenho acadêmico e aprendizagem

- Figura 1. Coeficientes *pattern* padronizados do modelo explicativo do desempenho do aluno por meio da inteligência 94
- Figura 2. Coeficientes *pattern* padronizados do modelo randômico de equações estruturais multinível 102

Lista de Tabelas

Manuscrito 2. Evidências de validade de construto do Teste de Raciocínio Abstrato e Espacial – TRAE

Tabela 1.	Índices psicométricos da Teoria Clássica dos Testes da primeira versão do TRAE	51
Tabela 2.	Índices de dificuldade, discriminação, acerto ao acaso e informação da segunda versão do TRAE.....	53
Tabela 3.	Cargas fatoriais dos itens do TRAE para os modelos não-lineares, FIFA e bi-factor	56
Tabela 4.	Indicadores de ajuste dos modelos das análises de equações estruturais	58

Manuscrito 3. Evidências de validade do Teste de Raciocínio Abstrato e Espacial com base na abordagem multitraço-multimétodo

Tabela 1.	Estatísticas descritivas e fidedignidade do TRAE e BPR-5	73
Tabela 2.	Correlações entre os escores dos testes TRAE e BPR-5 e correlações corrigidas para a ausência de fidedignidade perfeita	75
Tabela 3.	Comparações entre os modelos multitraço-multimétodo	77

Manuscrito 4. Análise multinível das relações entre inteligência fluida, desempenho acadêmico e aprendizagem

Tabela 1.	Parâmetros do modelo explicativo do desempenho do aluno por meio da inteligência fluida	93
Tabela 2.	Parâmetros não-padronizados do modelo 1 (modelo vazio)	95
Tabela 3.	Parâmetros não-padronizados do modelo 2 e do modelo 3	97
Tabela 4.	Parâmetros não padronizados dos modelos randômicos	100
Tabela 5.	Parâmetros não-padronizados do modelo multinível de medidas repetidas 1 (modelo vazio)	103
Tabela 6.	Parâmetros não-padronizados dos modelos multinível de medidas repetidas 2)	104
Tabela 7.	Parâmetros não-padronizados dos modelos multinível de medidas repetidas 3 (inclusão dos efeitos randômicos)	105
Tabela Anexo	Estatísticas descritivas	117

Lista de Abreviações

2PL	Modelo de TRI de dois parâmetros logístico
3PL	Modelo de TRI de três parâmetros logístico
<i>a</i>	Parâmetro de discriminação da TRI
AF	Análise Fatorial
<i>b</i>	Parâmetro de dificuldade da TRI
BPR-5	Bateria de Provas de Raciocínio
<i>c</i>	Parâmetro de acerto ao acaso da TRI
CDS	Velocidade de Tomada de Decisão
CHC	Modelo de inteligência Cattell-Horn-Carroll
CN	Ciências Naturais
FIFA	<i>Full Information Factor Analysis</i>
<i>g</i>	Fator Geral de Spearman
<i>Ga</i>	Inteligência Auditiva
<i>Gc</i>	Inteligência Cristalizada
<i>Gf</i>	Inteligência Fluida
<i>Gf-Gc</i>	Inteligência Fluida e Cristalizada
<i>Glr</i>	Memória de Longo Prazo
<i>Gq</i>	Conhecimento quantitativo
<i>Gs</i>	Velocidade do Processamento cognitivo
<i>Gsm</i>	Memória de Curto Prazo
<i>Gv</i>	Inteligência Visual
LP	Língua Portuguesa
MT	Matemática
MTMM	Abordagem Multitraço-Multimétodo
NSE	Nível socioeconômico
NSEA	Nível socioeconômico agregado para o nível da escola
RA	Raciocínio Abstrato
RE	Raciocínio Espacial
<i>s</i>	Fator específico de Spearman
TCT	Teoria Clássica dos Testes
TRAE	Teste de Raciocínio Abstrato e Espacial
TRI	Teoria de Resposta ao Item

Resumo

A inteligência fluida refere-se à capacidade de solucionar problemas novos, para os quais a pessoa não tem conhecimento prévio. Esse construto psicológico está relacionado à capacidade de aprendizagem e prediz outras variáveis importantes, como desempenho acadêmico e desempenho no trabalho. Contudo, pouco se conhece sobre a influência de variáveis do nível contextual sobre a relação entre a inteligência e os diversos tipos de desempenho. Nesse sentido, esta tese investigou, principalmente, a moderação de variáveis da escola na relação entre a inteligência fluida, o desempenho acadêmico e a aprendizagem. Para tanto, foram conduzidos quatro estudos. O *primeiro estudo* teve como objetivo revisar criticamente a literatura sobre o tema. No que diz respeito às variáveis do estudante, a revisão apontou, principalmente, para as características de responsabilidade e de autodisciplina na predição do desempenho acadêmico. No que se refere às variáveis contextuais, os recursos da escola e o nível socioeconômico agregado também foram apontados como importantes para a predição do desempenho acadêmico. Entretanto, a revisão indicou a carência de investigações sobre a influência das variáveis contextuais na moderação da relação entre a inteligência, o desempenho acadêmico e a aprendizagem. Para que os modelos teóricos pudessem ser testados empiricamente, buscou-se, primeiramente, nos estudos dois e três, construir um instrumento para a avaliação da inteligência fluida. No *segundo estudo* buscou-se construir e obter evidências de validade de construto do Teste de Raciocínio Abstrato e Espacial (TRAE). A amostra principal foi composta de 1.069 estudantes do ensino fundamental com idade entre 11 e 17 anos (essa amostra foi selecionada a partir da amostra do quarto estudo). Por meio da Teoria de Resposta ao Item, estimou-se os parâmetros de discriminação (a) e dificuldade (b), cujas médias foram próximas de 1,00 e 0,70, respectivamente. As análises fatoriais sugeriram a adequação de uma estrutura hierárquica que permite estimação de dois fatores de primeira ordem associados ao raciocínio abstrato e ao raciocínio espacial, bem como um fator geral de segunda ordem associado à inteligência fluida. Por meio do *terceiro estudo* buscou-se evidências de validade convergente entre o TRAE e a Bateria de Provas de Raciocínio (BPR-5). Esses instrumentos foram aplicados em uma amostra de 149 estudantes do ensino médio. Obteve-se evidências de validade convergente do teste TRAE por meio da abordagem multitraço-multimétodo e por meio das correlações entre os instrumentos. Contudo, os resultados indicaram restrições quanto à validade discriminante entre raciocínio abstrato e espacial. Finalmente, o objetivo do *quarto estudo* foi construir e testar modelos multinível do desempenho acadêmico e da aprendizagem, considerando as variáveis explicativas inteligência fluida (Gf), infraestrutura na escola e nível socioeconômico agregado para a escola (NSEA). Participaram do estudo 1.295 estudantes do ensino fundamental. O modelo simples indicou relação forte entre Gf e desempenho acadêmico. Os modelos multinível apontaram para relação moderada entre a Gf e o desempenho acadêmico, bem como relação moderada entre a Gf e a aprendizagem. A infraestrutura na escola influenciou ambas as relações; o NSEA, apenas a relação entre a Gf e o desempenho acadêmico. Nesse sentido, uma infraestrutura melhor e um NSEA mais alto diminuíram o efeito de Gf sobre o desempenho e a aprendizagem. Esses resultados sugerem que a infraestrutura e o nível socioeconômico da escola reduzem as diferenças do desempenho e da aprendizagem entre os estudantes com baixos e altos escores de inteligência fluida.

Palavras-chave: inteligência fluida; desempenho acadêmico; aprendizagem; infraestrutura escolar; nível socioeconômico.

Abstract

Fluid intelligence refers to the ability to solve new problems without previous knowledge. This psychological construct is related to the ability to learn and predicts many other important variables such as academic and work achievement. However, little is known about the influence of contextual variables on the relationship between intelligence and the several types of achievement. In this sense, this thesis aims to assess the moderation effect of school variables on the relationship among fluid intelligence, academic achievement and learning. For that purpose, four studies were performed. The *first study* aimed to review the literature about this topic. With regard to the student variables the review showed the significance of responsibility and self-discipline in the prediction of academic achievement. With respect to the contextual variables, the school resources and the socioeconomic status of the school also predicted the school achievement. However, the literature review indicated a deficiency of research on the influence of contextual variables that moderate the relationship between intelligence, academic achievement and learning. Before to test the theoretical models, the studies two and three aimed to develop a test to assess the fluid intelligence. The *second study* intended to develop the Abstract and Spatial Reasoning Test, and obtain evidence of the construct validity of this test. The main sample consisted of 1,069 students varying in age from 11 to 17 years (this sample is part of the sample of the fourth study). The mean discrimination (*a*) and difficulty (*b*) parameter, based on IRT, showed an approximate value of 1.00 and 0.70, respectively. Factor analysis indicated that the structure of the instrument allows the estimation of two first-order factors, related to abstract and spatial reasoning, as well as a second-order factor related to fluid intelligence. The *third study* aimed to obtain evidence of convergent validity between the Abstract and Spatial Reasoning Test and the Battery of Reasoning Tests [Bateria de Provas de Raciocínio, BPR-5, in Portuguese]. These instruments were administered to a sample of 149 high school students. Based on the multitrait-multimethod approach and the correlations between the instruments, the results showed evidences of convergent validity of the Abstract and Spatial Reasoning Test. Nevertheless, the results also indicated a lack of discriminant validity between abstract and spatial reasoning. Finally, the *fourth study* aimed to develop and to test multilevel models explaining academic achievement and learning considering fluid intelligence (Gf) of the students, infrastructure and socioeconomic status of the school (SESS) as predictor variables. The study examined 1,295 elementary school students. The basic model showed a strong relationship between Gf and academic achievement. The multilevel models indicated a moderate relationship between Gf and academic achievement, and also between Gf and learning. Infrastructure influenced both relationships, while SESS influenced only the relationship between Gf and achievement, in the sense that a better infrastructure and a higher SESS had a diminishing effect on the relationships. This effect suggests that school infrastructure and SESS decreases existing differences in academic achievement and learning between the students with low and high intelligence scores.

Keywords: fluid intelligence; academic achievement; learning; school infrastructure; socioeconomic status.

Apresentação

Estima-se que o Brasil gaste, anualmente, mais de R\$ 500 milhões para a realização de avaliações educacionais, considerando apenas a rede pública de educação. Entretanto, ainda é necessário aprofundar o conhecimento científico sobre quais são, especificamente, os aspectos pessoais e contextuais que podem influenciar o desempenho do estudante na escola.

Destaca-se a influência da inteligência fluida (*Gf*) na predição do desempenho acadêmico, cuja literatura científica apresenta diversos estudos que indicam relações moderadas e fortes entre esses dois construtos. A *Gf* está mais relacionada à solução de problemas novos, do que ao conhecimento já adquirido. Portanto, a relação entre a *Gf* e o desempenho acadêmico não é tão óbvia quanto possa parecer. No que se refere às variáveis de contexto, o nível socioeconômico é um dos aspectos mais estudados na relação com o desempenho acadêmico. Todavia, a literatura apresenta contradições: por um lado, algumas pesquisas, que desconsideraram a variável inteligência, encontraram relações fortes entre nível socioeconômico (agregado para o nível de análise da escola) e o desempenho acadêmico; por outro lado, pesquisas também indicaram efeito muito pequeno do nível socioeconômico do estudante sobre o seu desempenho acadêmico, após o controle da variável inteligência. Portanto, a relação entre a inteligência, o nível socioeconômico e o desempenho acadêmico parece complexa e a literatura científica carece de aprofundamento sobre o tema.

A presente tese insere-se nesse contexto e tem como objetivo principal avaliar o impacto de variáveis pessoais (principalmente inteligência fluida) e de variáveis contextuais (principalmente infraestrutura na escola e nível socioeconômico) sobre o desempenho acadêmico e a aprendizagem. Para responder ao objetivo principal, foram conduzidos estudos, que, por sua vez, foram organizados em quatro manuscritos.

O primeiro manuscrito refere-se à revisão crítica de literatura sobre a relação entre a inteligência, o desempenho acadêmico e a aprendizagem, destacando as principais variáveis

que podem influenciar essas relações. Ressalta-se, ainda, que o primeiro manuscrito fora apresentado no exame de qualificação e, nesta tese, é reapresentado após as revisões sugeridas pela banca examinadora.

Ainda que bons instrumentos para a avaliação da inteligência fluida estejam disponíveis aos pesquisadores brasileiros, a maior parte dos instrumentos é de aplicação individual ou necessita de um tempo de aplicação superior a uma hora e meia. Entretanto, nem todos os participantes têm disponibilidade para responder a uma pesquisa com tamanha duração. Ademais, as instituições nas quais os participantes estão vinculados tendem a negar o pedido de uma pesquisa longa. Para evitar esse problema, o segundo e o terceiro manuscrito apresentam os resultados da construção do Teste de Raciocínio Abstrato e Espacial (TRAPE), bem como as evidências de validade e precisão desse instrumento. Tal instrumento visa avaliar a inteligência fluida por meio de dois tipos de raciocínio, de maneira acessível, principalmente, às pessoas que dispõem pouco tempo para responder a uma pesquisa.

Após a construção do instrumento, o quarto manuscrito buscou construir e testar modelos teóricos de compreensão do desempenho acadêmico. Para tanto, foram conduzidos dois estudos: o primeiro buscou a construção de modelos multinível para a compreensão da relação entre a inteligência, a infraestrutura escolar e o desempenho acadêmico; o segundo investigou a relação multinível entre a inteligência, a infraestrutura na escola e a aprendizagem. A aprendizagem foi mensurada, neste estudo, por meio dos escores nas provas de língua portuguesa e de matemática nas avaliações educacionais nos anos de 2011 e 2012. Tais avaliações foram realizadas pela Secretaria de Educação de uma grande cidade brasileira.

Os quatro manuscritos serão apresentados a seguir. Após a descrição dos mesmos, na seção ‘Considerações Finais’, serão discutidas as principais conclusões dos estudos.

Manuscrito 1

Inteligência e Desempenho Acadêmico: Revisão de Literatura

Intelligence and Academic Achievement: A Literature Review

Sugestão de Título Abreviado: Inteligência e Desempenho Acadêmico

Resumo

A inteligência é um aspecto importante para a compreensão do desempenho acadêmico dos estudantes. Nesse contexto, buscou-se, por meio da presente pesquisa, revisar a literatura sobre a relação entre a inteligência e o desempenho acadêmico, destacando as variáveis que podem influenciar essa relação. Tais variáveis têm sido investigadas tanto em nível individual, quanto contextual (professores, escolas etc). No nível de análise individual (do aluno), existem evidências que as características de personalidade associadas à responsabilidade e à autodisciplina estão relacionadas, de maneira positiva, ao desempenho acadêmico. Outros aspectos como a criatividade, a motivação, a autopercepção de desempenho, a esperança e os pensamentos positivos também podem impactar positivamente o desempenho. No que se refere ao nível contextual, os recursos da escola exercem um papel importante. No entanto, o efeito do nível socioeconômico e da etnia, sobre a relação entre a inteligência e o desempenho carece de mais estudos.

Palavras-Chave: inteligência; habilidades cognitivas; desempenho acadêmico.

Abstract

Intelligence is an important variable in the search for comprehension of academic achievement. The present research aimed to review the literature about the relation between intelligence and academic achievement, highlighting variables that may influence this relation. These variables have been studied at an individual level and at a context level (teachers, school, etc.). Considering the individual (student) level of analysis, evidence exists that personality characteristics of responsibility and self-discipline are positively associated with academic achievement. Other factors like creativity, motivation, self-perceived performance, hope and positive thoughts also seem to influence academic achievement. At the context level, the school resources play an important role. To reach conclusions on the effect of other contextual variables, such as socioeconomic status and ethnical background more conclusive studies are needed.

Keywords: intelligence; cognitive abilities; academic achievement.

As relações entre as habilidades cognitivas, definidas operacionalmente como inteligência, e o desempenho acadêmico foram avaliadas por diversos estudos. Na maioria deles foi encontrada associação positiva com tamanho de efeito forte (Floyd, Evans, & McGrew, 2003; Hattie, 2009; Primi, Ferrão, & Almeida, 2010; Rohde & Thompson, 2007). Embora essas relações estejam razoavelmente bem estabelecidas na literatura científica, ainda é necessário aprofundar o conhecimento especializado sobre quais as variáveis que influenciam a associação entre a inteligência e o desempenho de estudantes. Por exemplo, um professor, com alta proficiência no conteúdo que leciona, pode contribuir para o aumento no desempenho acadêmico de um estudante com habilidades cognitivas abaixo da média? Tais discussões poderiam ampliar a compreensão teórica sobre o fenômeno, bem como subsidiar políticas públicas eficientes.

Este estudo insere-se nesse contexto e tem como objetivo revisar criticamente a literatura sobre as relações entre a inteligência e o desempenho acadêmico, principalmente nas disciplinas relacionadas à matemática e à linguagem. Buscou-se analisar os diversos aspectos apontados na literatura que podem influenciar, mediando ou moderando, as relações citadas.

Este texto está organizado em três seções. Na primeira são apresentadas, brevemente, as principais definições e modelos teóricos da inteligência. Na segunda discutem-se os estudos empíricos e de meta-análise das relações entre a inteligência e o desempenho acadêmico. Na terceira seção são apresentadas as conclusões da revisão de literatura.

Inteligência: Definição e Principais Modelos Teóricos

Em que pese à discordância de alguns profissionais, o estudo da inteligência é um dos empreendimentos mais bem sucedidos da psicologia moderna. Nas décadas de pesquisa sobre o tema, foi possível evidenciar os aspectos genéticos e ambientais que embasam as habilidades cognitivas dos indivíduos. A literatura também ressalta a importância da inteligência na compreensão do desenvolvimento psicológico (Berg, 2000; Chen & Siegler, 2000). Além disso, por meio desse construto, é possível prever, com razoável eficiência, várias formas de desempenho (Neisser et al., 1996).

No que se refere à definição de inteligência, os principais conceitos envolvem, de algum modo, as proposições de Galton ou as de Spearman (Almeida, 1994). Galton entendia a inteligência como força mental, e o seu conceito pode ser considerado um precursor dos modelos neuropsicológicos. O conceito de Spearman, por sua vez, embasa grande parte dos modelos fatoriais. Nessa perspectiva, a inteligência é definida, basicamente, como a capacidade para aprender. Conceitos semelhantes são adotados por diversos instrumentos psicológicos, tais como o RAVEN, o SON e o WISC. Para a bateria SON, por exemplo, o construto em questão é definido como a habilidade para aprender e a chance de sucesso na escola (Laros, Reis, & Tellegen, 2010; Snijders, Tellegen, & Laros, 1989). Em suma, a própria definição da inteligência parece estar associada aos processos de aprendizagem formais ou informais. Grande parte dos testes e estudos sobre a relação entre a inteligência e o desempenho acadêmico utilizou algum modelo fatorial de inteligência. Portanto, o presente estudo focará a revisão teórica dos modelos fatoriais.

Modelos de Spearman e Thurstone

As proposições de Spearman e Thurstone são consideradas as primeiras teorias fatoriais da inteligência. Ambos se interessaram em estudar a taxonomia das habilidades cognitivas (Hogan, 2006). Em outras palavras, ambos queriam responder à mesma pergunta: quantas e quais dimensões são necessárias para descrever, de forma eficiente e parcimoniosa, as habilidades cognitivas? Contudo, suas conclusões são bastante distintas.

Spearman (1904) realizou estudos a partir de testes de funções sensoriais já existentes, buscando avaliar as correlações entre os testes. Por meio das análises dessas pesquisas, Spearman reuniu evidências de relações significativas entre todos os testes avaliados. Tais correlações, na visão do pesquisador, poderiam ser explicadas por uma dimensão latente subjacente aos testes específicos. Em outras palavras, Spearman concluiu que, embora os testes avaliassem habilidades cognitivas distintas, suas associações eram tão elevadas a ponto de sustentar a existência de uma ‘habilidade geral’, denominada fator *g*. Spearman também nomeou os seus achados de teoria dos dois fatores. Nessa perspectiva, um dos fatores é composto pela habilidade geral (fator *g*) e o outro, pelas habilidades específicas (*s*). Apesar da proposta de Spearman, a maioria dos pesquisadores refere-se a ela como a teoria do fator *g* ou teoria de um fator (Brody, 2000; Hogan, 2006).

Thurstone (1938) discordava amplamente de Spearman. Segundo ele, uma dimensão não explicaria as especificidades das habilidades cognitivas. Thurstone desenvolveu um método de análise fatorial múltipla e, por meio desse método, concluiu que as correlações encontradas entre os testes não eram suficientemente altas para sustentar a existência de uma dimensão geral da inteligência. Assim, ele desenvolveu uma teoria de sete fatores: (1) compreensão verbal, que envolve o conhecimento das

definições das palavras; (2) fluência verbal associada à rapidez na reprodução das palavras; (3) facilidade numérica, referente à capacidade de solução de problemas matemáticos; (4) raciocínio espacial, que envolve a habilidade para lidar mentalmente com figuras e objetos tridimensionais; (5) velocidade perceptual associada à velocidade no reconhecimento dos estímulos; (6) indução, referente à capacidade de descobrir os princípios de uma situação e aplicá-los aos novos problemas; (7) memória, que envolve a capacidade de recordar uma lista de estímulos (Brody, 2000; Davidson & Downing, 2000; Thurstone, 1938).

A teoria de Spearman ganhou força na primeira metade do século passado e tornou-se uma das teorias mais aceitas do período. Esse modelo também embasa, atualmente, instrumentos e pesquisas. Em suma, após décadas de estudos, não é possível reunir evidências que refutem a teoria de Spearman. A teoria de Thurstone galgou ‘menos fama’ entre os acadêmicos. Todavia, algumas das suas dimensões (tais como raciocínio verbal e memória) estão representadas, de alguma maneira, em teorias modernas da inteligência (Brody, 2000).

Modelo de Inteligência Fluida e Cristalizada (*Gf-Gc*)

As teorias de Spearman e Thurstone eram antagônicas e geraram inúmeras discussões sobre as dimensões teóricas da inteligência. Modelos posteriores ajudaram a conciliar as duas teorias. Um desses modelos foi teorizado por Cattell (1943, 1963), que organizou as habilidades cognitivas em duas dimensões. A primeira delas, inteligência fluida (*Gf*), envolve as habilidades de raciocínio e a capacidade de solução de problemas novos. Nesse sentido, a inteligência fluida refere-se à capacidade de raciocinar indutiva e dedutivamente, formar e testar hipóteses para problemas novos e identificar relações e conceitos subjacentes às situações. A *Gf* é a dimensão que está mais associada ao fator *g*

de Spearman (Carroll, 2005; McGrew, 2005). A segunda dimensão, denominada inteligência cristalizada (*Gc*) está associada à aquisição e à solidificação dos conhecimentos formais e informais, aprendidos nas escolas ou por meio da transmissão cultural (Almeida, 1994; Cattell, 1943, 1963; Horn, 1994).

A teoria de Cattell ficou amplamente conhecida pelas dimensões *Gf-Gc*. No entanto, após a publicação do primeiro artigo, em 1943, a teoria sofreu algumas ampliações. Juntamente com o seu aluno de doutorado, John Horn, Cattell buscou evidências para a existência de outras sete dimensões: (1) aquisição e recuperação da memória de curto prazo - *Gsm*; (2) memória de longo prazo - *Glr*; (3) inteligência visual - *Gv*; (4) inteligência auditiva - *Ga*; (5) velocidade do processamento cognitivo - *Gs*; (6) velocidade de tomada de decisão - *CDS*; (7) e conhecimento quantitativo - *Gq* (Horn, 1994; Horn & Blankson, 2005).

Cattell ainda propôs uma organização da teoria em fatores de segunda e de terceira ordem (Hakstian & Cattell, 1978). Sendo assim, a teoria de Cattell e Horn pode ser considerada como um dos primeiros modelos de organização hierárquica da inteligência (Carroll, 2005). Os autores ressaltaram, no entanto, que a organização em fatores de segunda e terceira ordem carecia de estudos que oferecessem maior suporte empírico às conclusões (Hakstian & Cattell, 1978).

Teoria dos Três Estratos e Modelo CHC

A teoria dos três estratos foi proposta por Carroll (1993, 2005) a partir de uma meta-análise de mais de 450 estudos que continham análises fatoriais de instrumentos para avaliação da inteligência. Carroll propôs uma estrutura dividida em três estratos hierarquizados. O estrato III é composto por uma grande habilidade geral que embasa as demais capacidades intelectuais. As características da inteligência descritas no estrato

III são praticamente equivalentes ao fator *g* de Spearman. O estrato II agrega oito dimensões gerais: inteligência fluida, inteligência cristalizada, memória, percepção visual, percepção auditiva, habilidade de apreensão, velocidade cognitiva e velocidade de processamento. Percebe-se que as dimensões do estrato II se assemelham amplamente ao modelo de Cattell e Horn. Em outras palavras, a proposta do Carroll engloba boa parte da teoria do Cattell. O estrato I é composto por habilidades mais específicas, tais como raciocínio quantitativo, compreensão de leitura, memória visual e tempo de reação.

Ressalta-se que os três estratos organizam fatores de primeira, de segunda e de terceira ordem. Ou seja, todas as habilidades específicas do primeiro estrato estão relacionadas a alguma habilidade do segundo estrato. Além disso, as oito habilidades do estrato II podem ser explicadas, em grande parte, pelo fator geral do estrato III (Carroll, 1993, 2005).

Conforme indicado pelo próprio Carroll (2005), o modelo era provisório. Segundo o pesquisador, seria importante definir melhor as habilidades específicas do estrato I, bem como encontrar evidências da sua existência. Ademais, seria necessário aprofundar as discussões sobre as relações entre o estrato III e o fator *g* de Spearman.

Nesse contexto, McGrew e Flanagan (1998) propuseram uma reorganização e ampliação da teoria dos estratos por meio do modelo Cattell-Horn-Carroll (CHC). Na realidade, o modelo CHC, também organizado em uma estrutura de três níveis, é considerado uma integração da teoria dos três estratos e da teoria *Gf-Gc*. O nível mais amplo (fator de terceira ordem) é composto pela capacidade geral. No segundo nível, as habilidades são organizadas em 10 dimensões gerais: inteligência fluida, inteligência cristalizada, conhecimento quantitativo, leitura escrita, memória de curto prazo, processamento visual, processamento auditivo, capacidade da memória de longo prazo,

velocidade do processamento e rapidez na decisão. Finalmente, o nível mais baixo é composto por aproximadamente 70 fatores bastante específicos, tais como raciocínio sequencial, raciocínio piagetiano, sensibilidade gramatical, proficiência em línguas estrangeiras e sensibilidade cinestésica. Os três níveis da estrutura do modelo CHC seguem uma ordem de especialização, variando do mais geral -nível III- até as dimensões específicas -nível I- (McGrew, 2005; Primi, 2003).

Perecebe-se que grande parte das 10 dimensões de segunda ordem do modelo CHC é semelhante aos fatores do estrato II da teoria de Carroll. Adicionalmente, os 70 fatores do nível I do CHC parecem mais bem definidos, se comparados aos de Carroll. Entretanto, ainda não foram encontradas evidências científicas suficientes para sustentar a existência de todos os fatores do modelo CHC.

Pesquisas sobre a Relação entre a Inteligência e o Desempenho Acadêmico

O construto inteligência é estudado em relação ao desempenho acadêmico desde as primeiras investigações sobre as habilidades cognitivas. Tais pesquisas buscaram avaliar quais outras variáveis, além da inteligência, poderiam explicar o desempenho acadêmico. A presente revisão de literatura tem como objetivo recuperar e classificar os artigos que abarcaram a relação entre inteligência e desempenho acadêmico. Tais estudos foram recuperados dos bancos de dados do PsycInfo, Bireme, BVS-Psi, Pepsic e Scielo, cujos artigos foram publicados até o ano de 2012. Utilizou-se os seguintes critérios de inclusão: (1) o artigo deveria apresentar dados de uma pesquisa empírica ou de meta-análise; (2) os testes de inteligência utilizados deveriam ser embasados em algum modelo fatorial da inteligência; (3) o desempenho acadêmico deveria ser mensurado por, ao menos, um teste de linguagem ou de matemática.

A revisão dos artigos permitiu classificá-los em categorias relacionadas ao conteúdo investigado e o nível de análise das variáveis: (1) estudos gerais sobre a relação entre inteligência e desempenho acadêmico; (2) personalidade; (3) criatividade; (4) variáveis da psicologia positiva (pensamento positivo e motivação, entre outras); (5) variáveis contextuais (etnia e nível socioeconômico, entre outras). Primeiramente, serão apresentados os estudos gerais sobre a relação inteligência e o desempenho acadêmico. Posteriormente, serão discutidos os estudos que apresentaram uma ou mais variáveis com o potencial de moderar ou mediar essas relações.

A relação geral entre a inteligência e o desempenho acadêmico foi estudada por Rohde e Thompson (2007). Os autores investigaram o desempenho dos estudantes nas provas do sistema de avaliação educacional SAT e do Teste de Desempenho Amplo [*Wide Range Achievement Test - WRAT III*]. No estudo foram encontradas correlações moderadas e fortes ($0,30 \leq r \leq 0,71$) entre os testes de desempenho e as avaliações das habilidades cognitivas realizadas por meio do teste RAVEN. Adicionalmente, quando controlados os efeitos da memória de trabalho, da velocidade de processamento e da habilidade espacial, as habilidades cognitivas gerais continuaram acrescentando explicação para a variabilidade do desempenho acadêmico. Em outras palavras, ao remover os efeitos de outras variáveis cognitivas, os autores ainda encontraram relações significativas entre a inteligência geral e o desempenho acadêmico.

Um estudo de meta-análise foi conduzido por McGrew e Wendling (2010) sobre a relação entre o modelo Cattell-Horn-Carroll (CHC) de inteligência e o desempenho acadêmico. Embasados em 19 estudos anteriores, os pesquisadores concluíram que os fatores *g* (geral) e *Gf* (inteligência fluida) relacionaram-se positivamente ao desempenho acadêmico em todos os estudos ($p < 0,05$).

Relações importantes entre a inteligência e o desempenho acadêmico também foram encontradas por Primi, Ferrão e Almeida (2010). Os autores utilizaram um teste para a avaliação da inteligência fluida e associaram-no a quatro avaliações longitudinais de matemática. Em concordância com a maior parte da literatura científica, foram encontradas relações positivas entre o desempenho em matemática e as dimensões da inteligência geral, raciocínio abstrato, numérico, verbal e espacial ($r \geq 0,27$, para a dimensão geral). Ademais, o maior crescimento na proficiência de matemática foi observado nos estudantes com os maiores escores de inteligência (maior *slope*). Em outras palavras, além do desempenho mais elevado, os estudantes com altos escores de inteligência parecem adquirir mais conhecimento em matemática. Resultados semelhantes foram encontrados por Geary (2011). Esse autor conduziu um estudo longitudinal para observar o desempenho acadêmico de crianças do ensino fundamental. Os resultados indicaram que a inteligência e a velocidade de processamento previam a evolução da proficiência em matemática após cinco anos da realização da primeira avaliação. Ressalta-se ainda que a velocidade do processamento também é considerada uma dimensão específica da inteligência, segundo o modelo CHC.

Uma das variáveis mais investigadas em relação à inteligência e ao desempenho acadêmico é a personalidade. Fhurham e Chamorro-Premuzic (2004), por exemplo, verificaram que a personalidade explica uma parte importante da variância do modelo da relação entre a inteligência e o desempenho acadêmico.

Nesse contexto, Fabio e Busoni (2007) desenvolveram um modelo, a partir dos dados coletados com o teste RAVEN e o Questionário *Big Five* (BFQ). Além da inteligência, os resultados indicaram que o fator de personalidade Realização (ou Conscienciosidade) é um preditor dos exames finais de estudantes do ensino secundário italiano. No modelo final, a inteligência explicou 17% da variância do desempenho

acadêmico, e os traços de personalidade acrescentaram mais 11% de explicação. Ou seja, características como perseverança, senso de dever e autodisciplina (associadas ao fator Realização) foram relacionadas positivamente ao desempenho acadêmico.

Estudo semelhante ao de Fabio e Busoni foi conduzido por Rosander, Bäcktröm e Stenberg (2011). Esses últimos autores utilizaram a modelagem por equações estruturais para analisar o desempenho de crianças suecas. As conclusões indicaram que, além da inteligência e do fator de personalidade Realização, os fatores Extroversão e Neuroticismo associaram-se significativamente às provas de linguagem, ciências sociais, matemática, artes e esportes. O fator Neuroticismo teve um impacto positivo (coeficiente padronizado = 0,14), e a Extroversão, negativo (coeficiente padronizado = - 0,14). Portanto, as características de assertividade, festividade e gregarismo (fator Extroversão), bem como a instabilidade emocional (fator Neuroticismo) também estão associadas ao desempenho acadêmico. Os autores argumentaram que, provavelmente, ambos os fatores estejam influenciando a motivação para o estudo e, indiretamente, o desempenho acadêmico.

As direções das relações encontradas por Rosander foram semelhantes às apresentadas por Furnham e Monsen (2009), que investigaram a relação entre a inteligência, a personalidade e as notas de inglês (linguagem), literatura, matemática, ciências e outras disciplinas eletivas de adolescentes britânicos. Além da relação positiva entre inteligência e desempenho acadêmico ($Beta = 0,34$), Furnham e Monsen também encontraram associação negativa entre o fator de personalidade Extroversão e a nota geral das provas escolares ($Beta = - 0,18$). Todavia, quando considerado apenas o desempenho nas disciplinas eletivas (Francês, Geografia e Alemão), a personalidade Extrovertida não apresentou relação estatisticamente significativa com o desempenho acadêmico.

Ainda no que se refere ao contexto da personalidade, Chamorro-Premuzic e Arteche (2008) apresentaram um modelo no qual a inteligência foi estudada como variável mediadora da relação entre a personalidade e o desempenho acadêmico. Para tanto, uma amostra de aproximadamente 450 estudantes universitários britânicos respondeu ao teste RAVEN e ao inventário de personalidade NEO-PI-R. A inteligência fluida apresentou um efeito de mediação parcial entre o fator de personalidade Realização e o desempenho acadêmico. Tais resultados foram interpretados pelos autores como um ‘efeito de compensação’. Em outras palavras, os estudantes que são mais eficientes e com boa capacidade de raciocínio abstrato tendem a ser menos organizados e autodisciplinados. A redução da autodisciplina poderia, por sua vez, diminuir o desempenho acadêmico.

Entre os aspectos individuais, a criatividade também foi estudada em relação à inteligência e ao desempenho acadêmico. A criatividade pode ser definida como a fluidez e a flexibilidade na produção de ideias originais. Nesse contexto, Rindermann e Neubauer (2004) testaram um modelo teórico, por meio de equações estruturais, cuja variável dependente era o desempenho acadêmico em linguagem, matemática, física e ciências humanas. Apesar do grande efeito da inteligência (coeficiente padronizado = 0,53), a criatividade também influenciou o desempenho acadêmico dos estudantes (coeficiente padronizado = 0,19).

Resultados semelhantes ao estudo de Rindermann e Neubauer foram encontrados por Freund, Holling e Preckel (2007). Esses autores investigaram 1.500 estudantes alemães no que se refere às habilidades cognitivas, ao desempenho acadêmico e à criatividade, entre outros aspectos. Os autores encontraram um impacto significativo da inteligência sobre o desempenho acadêmico ($Beta = 0,47$, para ciências naturais; e $Beta = 0,16$, para as ciências sociais). Além disso, a criatividade também se

associou ao desempenho, principalmente na área de ciências sociais ($Beta = 0,23$). Entretanto, o impacto da criatividade sobre o desempenho em matemática e ciências naturais foi menor ($Betas = 0,10$).

Características associadas à psicologia positiva também foram estudadas na relação entre a inteligência e o desempenho acadêmico. Day, Hanson, Maltiby, Proctor e Wood (2010), por exemplo, estudaram a influência da variável esperança nessa relação. A esperança, nesse contexto, foi definida como uma característica da personalidade associada aos pensamentos positivos sobre a capacidade de alcançar os objetivos planejados. Aproximadamente, 150 estudantes universitários britânicos participaram da pesquisa. Por meio de uma regressão hierárquica, os autores evidenciaram que o fator esperança, além da personalidade e da inteligência, associou-se ao desempenho acadêmico. Para a inteligência e para a personalidade, os valores de $Beta$ (da análise de regressão linear) variaram de 0,03 a 0,34; e, para a esperança, os $Betas$ variaram de 0,16 e 0,30. Ou seja, a crença de que é possível criar planos de sucesso para alcançar um objetivo específico ou resolver um problema mostrou-se associada ao desempenho acadêmico superior.

Dentro do contexto da psicologia positiva, Leclerc, Larivée, Archambault e Janosz (2010) estudaram a relação entre a inteligência, as crenças de autocompetência e o desempenho acadêmico. Segundo os autores, a autocompetência se refere à autopercepção das próprias habilidades e das capacidades escolares. Os autores avaliaram aproximadamente 900 estudantes universitários canadenses francófonos. Por meio de correlações e de modelos de regressão, Leclerc et al. (2010) concluíram que, além da inteligência ($Beta = 0,07$), a autocompetência relacionou-se positivamente, e, de maneira forte, com o desempenho acadêmico ($Beta = 0,79$). Ademais, o fator de interação entre as variáveis inteligência e autocompetência apresentou predição

positiva sobre o desempenho nas escolas ($Beta = 0,08$). Em outras palavras, a relação entre a inteligência e o desempenho acadêmico é mais forte para os estudantes com altos escores de autocompetência.

Ainda no contexto da psicologia positiva, Leeson, Ciarrochi e Heaven (2008) estudaram, entre outros aspectos, a relação entre o pensamento positivo, a inteligência e o desempenho acadêmico. Os pensamentos positivos dizem respeito às características pessoais de autoestima, estilos atribucionais e esperança. Durante a pesquisa, os autores investigaram 784 estudantes universitários australianos. O modelo testado indicou que a inteligência e o pensamento positivo associaram-se ao desempenho acadêmico (coeficientes padronizados de 0,50 e 0,13, respectivamente). Especificamente, a esperança e os estilos atribucionais são preditores mais fortes do desempenho acadêmico se comparados à autoestima.

Além das variáveis associadas à psicologia positiva que contribuem para a compreensão do desempenho acadêmico, é importante destacar outras variáveis do nível do estudante, tais como a autopercepção da habilidade cognitiva, a motivação e o sexo. Chamorro-Premuzic, Harlaar, Greven e Plomin (2010), por exemplo, conduziram um estudo longitudinal com gêmeos britânicos para investigar a autopercepção das habilidades cognitivas. Além da inteligência, a autopercepção também apresentou relação positiva com o desempenho acadêmico (coeficientes padronizados entre 0,11 e 0,41). Além disso, o impacto do desempenho acadêmico na primeira avaliação (DA_1) sobre a autopercepção de desempenho na segunda avaliação (APD_2) foi semelhante ao impacto do efeito da autopercepção na primeira avaliação (APD_1) sobre o desempenho na segunda (DA_2) [$DA_1 \rightarrow APD_2 \approx APD_1 \rightarrow DA_2$, coeficientes padronizados de 0,13 e 0,11, respectivamente]. Ou seja, o desempenho e a autopercepção se influenciam mutuamente dentro de um sistema de retroalimentação. Diante disso, os autores

afirmam que tais relações podem ser explicadas pela autoeficácia e pelo *insight*. Ou seja, o alto desempenho poderia aumentar a autopercepção do desempenho (*insight*), que, por sua vez, poderia retroalimentar o sistema, aumentando o desempenho.

A variável motivação foi objeto de estudo de Steinmayr e Spinath (2009). Os autores pesquisaram a proficiência em matemática e linguagem de estudantes alemães. No modelo de regressão do desempenho acadêmico foram apresentados os valores de *Beta* para os preditores inteligência (0,08), esperança de sucesso (0,04), medo de falhar (-0,05) e competitividade (0,07). Portanto, uma parte do desempenho acadêmico pode ser explicada pelos aspectos relacionados à motivação acadêmica do estudante.

No que se refere à relação entre as variáveis sexo, gênero e desempenho acadêmico as pesquisas têm apontado para diferentes conclusões. Kuhn e Holling (2009) avaliaram o desempenho de aproximadamente 1.100 estudantes alemães. Os autores concluíram que as meninas obtiveram desempenho maior nas provas de linguagem (tamanho de efeito $d = 0,24$), e os meninos, nas provas de ciências ($d = 0,13$). Especificamente para as meninas, o desempenho em linguagem continua alto, após controlar o efeito da inteligência. Kaufman, Kaufman, Liu e Johnson (2009) também encontraram um efeito do gênero sobre o desempenho de adultos americanos. Nesse estudo, as mulheres obtiveram um desempenho maior nas provas de escrita, e os homens, nas provas de matemática. Ademais, não foram encontradas diferenças estatisticamente significativas entre homens e mulheres no que se refere à inteligência fluida e à cristalizada.

Ainda no que se refere ao gênero e ao sexo, a literatura é um pouco controversa. Por um lado, no estudo de Furham e Mosen (2009), as meninas obtiveram um desempenho superior nas provas de história e tecnologia, e os meninos, nas provas de literatura, matemática e ciências. Ademais, a combinação entre a inteligência, a

personalidade e o sexo explicou mais de 25% da variância do desempenho nas diferentes provas. Por outro lado, os resultados do estudo de Leclerc et al. (2010) indicaram que o sexo não estava relacionado ao desempenho acadêmico, tampouco influenciava a relação entre inteligência e as crenças de autocompetência.

Além das variáveis do nível do estudante, a literatura apresenta alguns aspectos contextuais que podem influenciar a relação entre a inteligência e o desempenho acadêmico. Nesse sentido, a influência do nível socioeconômico (NSE) é um dos temas mais estudados, principalmente no que se refere à relação com o desempenho acadêmico.

O NSE, a inteligência e o desempenho acadêmico foram estudados por Colom e Flores-Mendonza (2007) numa amostra de 640 estudantes de diferentes escolas brasileiras. Nesse estudo, o NSE foi estudado por meio do salário familiar e nível educacional dos pais. Os resultados indicaram que o NSE não foi capaz de prever as diferenças no desempenho acadêmico ($Beta = 0,04$). Ademais, a inteligência se mostrou como o único preditor das notas escolares ($Beta = 0,37$). Resultados semelhantes foram encontrados por Lloyd e Barenblatt (1984). Esses autores investigaram as relações entre o desempenho, a inteligência, o NSE e a motivação intelectual. As conclusões indicaram que, mediante o controle da variável inteligência, o efeito de predição do NSE sobre o desempenho é muito fraco.

Entretanto, é necessário ressaltar que alguns estudos, que não consideraram a variável inteligência, apresentaram relações significativas entre o NSE e o desempenho acadêmico (Caldas & Bankston-III, 1997). Ao considerar o nível de análise da escola, Laros, Marciano e Andrade (2010) evidenciaram que a inserção do NSE (agregado para a escola) explicou 77% da variância do desempenho acadêmico (do nível da escola). Em outro estudo, os mesmos autores apresentaram um modelo capaz de explicar

aproximadamente 70% da variância do desempenho em língua portuguesa, por meio das variáveis NSE agregado para a escola, escolaridade da mãe e etnia do aluno (Laros, Marciano, & Andrade, 2012).

Ressalta-se, ainda, a relação entre o NSE e a inteligência evidenciada pelas pesquisas que não consideraram o desempenho acadêmico. Laros, Tellegen, Jesus e Karino (no prelo), por exemplo, encontram correlações moderadas e fortes entre os testes de inteligência e o NSE (r entre 0,40 e 0,67). As correlações mais altas referiam-se aos testes verbais de inteligência. Nesse sentido, os aspectos associados ao desenvolvimento do vocabulário e de conhecimentos gerais podem depender mais do acesso à cultura e à educação do que os aspectos relacionados ao raciocínio abstrato e à solução de problemas para os quais não é necessário o conhecimento prévio. Em outras palavras, o desempenho em tarefas de inteligência fluida depende menos do NSE, se comparado às tarefas que exigem conhecimento prévio.

Além do NSE, outros aspectos contextuais podem contribuir para a compreensão da relação entre a inteligência e o desempenho acadêmico, entre esses a cor da pele e etnia (Hattie, 2009). Frederickson e Petrides (2008), por exemplo, investigaram uma amostra de aproximadamente 500 estudantes britânicos. O desempenho dos adolescentes brancos britânicos foi superior se comparado ao dos estudantes negros, bem como superior aos estudantes descendentes de paquistaneses (d entre 0,08 e 0,13). No entanto, quando controlado o efeito da variável inteligência, as diferenças entre os grupos foram minimizadas ou não apresentaram significância estatística. Conclusões semelhantes foram encontradas por Swartwout, Garnaat, Myszka, Fletcher e Dennis (2010). Os resultados desse estudo indicaram que as crianças com o nível socioeconômico baixo e descendentes de hispânicos obtiveram desempenho menor em provas verbais (em inglês), se comparadas às crianças brancas não hispânicas. Todavia,

para as provas não verbais, o desempenho das crianças hispânicas praticamente se igualou ao das crianças brancas não hispânicas. Em outras palavras, as diferenças do desempenho dos estudantes se devem, provavelmente, às questões associadas à linguagem e não à etnia.

Entre os aspectos contextuais, os recursos da escola também exercem papel importante na relação entre inteligência e desempenho acadêmico. Nesse sentido, Gorey (2001) conduziu um estudo meta-analítico para avaliar os efeitos dos programas pré-escolares ao longo da vida dos estudantes. Os autores utilizaram os dados de mais de 30 artigos e compararam as crianças que sofreram a intervenção com as crianças do grupo controle. Os programas pré-escolares apresentaram, na grande maioria dos estudos, um efeito positivo e forte sobre a inteligência e o desempenho acadêmico (médias dos tamanhos do efeito $U = 76,5$ e $78,2$, respectivamente). Depois de cinco anos da intervenção, aproximadamente 70% das crianças que participaram dos programas ainda obtiveram desempenho melhor do que o grupo controle.

Kaufman et al. (2009) também evidenciaram o papel da escola no desempenho acadêmico. Os resultados do estudo indicaram que a inteligência fluida e a inteligência cristalizada (*Gf e Gc*) estão relacionadas ao número de anos que o aluno permanece nas escolas formais. Além disso, o número de anos na escola associou-se ao desempenho nas disciplinas de matemática, leitura e escrita. Para a disciplina de matemática, essa relação foi mais forte ($r = 0,63$) do que para as demais disciplinas ($r = 0,48$ e $0,49$). Em outras palavras, o desempenho em matemática parece depender mais da quantidade de anos nos quais a criança permanece na escola do que o desempenho em leitura e escrita.

Conclusões e Direções para Pesquisas Futuras

De maneira geral, as pesquisas, com objetivos e contextos diversos, apontam para a inteligência como uma importante variável para a compreensão do desempenho acadêmico. Nesse sentido, a variável inteligência parece explicar uma parte importante do desempenho em linguagem e matemática mesmo quando removidos os efeitos de outras variáveis cognitivas, tais como velocidade do processamento e memória de trabalho (Rohde & Thompson, 2007). Ademais, a inteligência parece influenciar também o aumento da proficiência acadêmica no decorrer dos anos (Primi et al., 2010). Assim, as crianças e os adolescentes com escores elevados na inteligência fluida tendem a apresentar aumento mais acentuado no desempenho. Em outras palavras, a aprendizagem também parece ser influenciada pelas habilidades cognitivas.

A personalidade também influencia o desempenho acadêmico (Chamorro-Premuzic & Arteche, 2008; Fabio & Busoni, 2007; Furnham & Monsen, 2009). Por um lado, características de autodisciplina, de senso de dever e de perseverança (fator Realização) associam-se positivamente ao desempenho. Por outro lado, características de festividade, de gregarismo e de assertividade (fator Extroversão) associam-se negativamente ao desempenho. Todavia, as pesquisas utilizam, normalmente, como métodos de análise, correlação e regressão linear. Consequentemente, a personalidade é estudada simplesmente como uma variável independente que, ao lado da inteligência, explica o desempenho acadêmico. O estudo de Chamorro-Premuzic e Arteche (2008), um dos poucos que avaliou mediações e moderações, ampliou a compreensão da personalidade no contexto do desempenho acadêmico. Os autores concluíram que a relação entre inteligência e desempenho depende das características do fator Realização (do modelo *Big Five*). Nesse sentido, pessoas com altos escores de inteligência poderiam ‘relaxar’ na sua organização e autodisciplina para o estudo, o que impactaria

negativamente no desempenho. Ressalta-se, ainda, que o efeito direto entre a inteligência e o desempenho continuou significativo e positivo. Em outras palavras, para as pessoas que não apresentaram esse efeito de ‘relaxamento’ (ou compensação, como os autores denominaram), a inteligência continuou impactando de maneira positiva o desempenho acadêmico.

Ainda que a relação entre a personalidade e o desempenho acadêmico tenha sido amplamente investigada por estudos anteriores, é necessário aumentar as evidências da moderação e/ou mediação que a personalidade exerce sobre a relação entre a inteligência e o desempenho. Nesse sentido, estudos que utilizam métodos experimentais e de análise de equações estruturais poderiam contribuir para a compreensão do tema.

Aspectos da criatividade, da esperança, das crenças de autocompetência e dos pensamentos positivos também apresentam impacto significativo na relação entre a inteligência e o desempenho. Especificamente no que se refere à criatividade, a influência dessa variável parece mais relevante para as disciplinas associadas às ciências humanas, em detrimento às ciências naturais (Freund et al., 2007). Segundo os autores, é possível que as respostas criativas dos estudantes sejam mais valorizadas pelas áreas menos técnicas, cuja estrutura do conhecimento seja mais aberta e flexível. O que explicaria a diferença do impacto da criatividade sobre o desempenho acadêmico.

Ainda no contexto individual, a autopercepção de desempenho, a motivação e o gênero/sexo também foram estudadas em relação ao desempenho acadêmico. Os resultados apresentados por Chamorro-Premuzic et al. (2010) indicam evidências de um sistema de retroalimentação: a autopercepção elevada de desempenho tende a aumentar o real desempenho, que, por sua vez, reforça e aumenta a autopercepção. Ressalta-se, no entanto, que o estudo utilizou a inteligência apenas como uma variável de controle.

Pesquisas que avaliem o efeito de moderação e/ou mediação da inteligência poderiam ampliar a compreensão desse sistema. No que diz respeito às variáveis sexo e gênero, a literatura não apresenta evidências suficientes para uma conclusão robusta.

Sobre os aspectos contextuais, as variáveis nível socioeconômico (NSE), etnia e cor de pele também foram estudadas em relação à inteligência e ao desempenho acadêmico. As relações são fortes ao considerar apenas as variáveis NSE e inteligência (Laros et al., no prelo). Relações fortes também foram encontradas pelas pesquisas que consideraram apenas as variáveis NSE e desempenho acadêmico, principalmente para o nível de análise da escola (Laros, Marciano, et al., 2010, 2012). Contudo, ao controlar o efeito da inteligência, o impacto do NSE sobre o desempenho diminui consideravelmente ou torna-se não significativo (Colom & Flores-Mendoza, 2007; Lloyd & Barenblatt, 1984). Para a variável cor de pele, as conclusões são semelhantes: quando controlado o efeito da variável inteligência, as diferenças de desempenho entre pessoas de cor de pele distintas também são bastante atenuadas. Esses resultados parecem indicar que se o NSE e a cor de pele forem, de fato, relevantes para a compreensão do desempenho acadêmico, o impacto deles é complexo, estrutural e podem influenciar, inclusive, o desenvolvimento das capacidades cognitivas de raciocínio e de resolução de problemas. Nesse sentido, ressalta-se ainda mais a importância de investigar esses aspectos por meio de pesquisas com delineamentos experimentais e longitudinais, bem como por meio de equações estruturais multinível.

A escola parece desempenhar papel importante no desempenho dos estudantes (Gorey, 2001; Hattie, 2009; Kaufman et al., 2009). Crianças com maiores escores de inteligência permanecem mais tempo na escola e apresentam melhor desempenho, principalmente em matemática. Entretanto, é necessário aprofundar os estudos sobre quais aspectos específicos da instituição influenciam a relação entre a inteligência e o

desempenho. Nesse contexto, é relevante investigar, por exemplo, se os professores mais bem preparados teórica e didaticamente influenciam fortemente o desempenho dos estudantes (principalmente daqueles com baixos escores de inteligência).

A presente revisão de literatura apresentou diversas variáveis que foram estudadas em relação à inteligência e ao desempenho acadêmico. Na maior parte das pesquisas, tais variáveis foram investigadas apenas como explicativas do desempenho. Contudo, é possível que esses fenômenos sejam mais complexos e precisem de modelos explicativos mais complexos. Portanto, a partir desta revisão, propõe-se uma agenda de pesquisa na qual as variáveis dos níveis do estudante sejam investigadas como mediadoras e/ou moderadoras do efeito da inteligência sobre o desempenho acadêmico. Ademais, é importante avaliar se tais variáveis influenciam o aumento do desempenho acadêmico no decorrer dos anos (ou seja, a aprendizagem). Espera-se, também, que os aspectos da escola, como a infraestrutura e o desempenho dos professores (relacionado ao conteúdo e à didática) também possam moderar a relação entre inteligência e o desempenho dos estudantes.

Referências

- Almeida, L. S. (1994). *Inteligência: Definição e medida*. Aveiro, PT: CIDInE.
- Berg, C. A. (2000). Intellectual development in adulthood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 117-137). Cambridge, UK: Cambridge University Press.
- Brody, N. (2000). History of theories and measurements of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 16-33). New York, NY: Cambridge University Press.
- Caldas, S. J., & Bankston-III, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, *90*, 269-277.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment (2nd ed.)* (pp. 69-76). New York, NY: The Guilford Press.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153-193
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1-22
- Chamorro-Premuzic, T., & Arteche, A. (2008). Intellectual competence and academic performance: Preliminary validation of a model. *Intelligence*, *36*, 564-573.
- Chamorro-Premuzic, T., Harlaar, N., Goven, C. U., & Plomin, R. (2010). More than just IQ: A longitudinal examination of self-perceived abilities as predictors of academic performance in a large sample of UK twins. *Intelligence*, *38*, 385-392.
- Chen, Z., & Siegler, R. S. (2000). Intellectual development in childhood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 92-116). Cambridge, UK: Cambridge University Press.
- Colom, R., & Flores-Mendoza, C. E. (2007). Intelligence predicts scholastic achievement irrespective of SES factors: Evidence from Brazil. *Intelligence*, *35*, 243-251.
- Davidson, J. E., & Downing, C. L. (2000). Contemporary models of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 34-52). New York, NY: Cambridge University Press.

- Day, L., Hanson, K., Maltby, J., Proctor, C., & Wood, A. (2010). Hope uniquely predicts objective academic achievement above intelligence, personality, and previous academic achievement. *Journal of Research in Personality, 44*, 550-553.
- Fabio, A. D., & Busoni, L. (2007). Fluid intelligence, personality traits and scholastic success: Empirical evidence in a sample of Italian high school students. *Personality and Individual Differences, 43*, 2095-2104.
- Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics across the school-age years. *Psychology in the Schools, 40*, 155-171.
- Frederickson, N., & Petrides, K. V. (2008). Ethnic, gender, and socio-economic group differences in academic performance and secondary school selection: A longitudinal analysis. *Learning and Individual Differences, 18*, 144-151.
- Freund, P. A., Holling, H., & Preckel, F. (2007). A multivariate, multilevel analysis of the relationship between cognitive abilities and scholastic achievement. *Journal of Individual Differences, 28*, 188-197.
- Furnham, A., & Chamorro-Premuzic, T. (2004). Personality and intelligence as predictors of statistics examination grades. *Personality and Individual Differences, 37*, 943-955.
- Furnham, A., & Mosen, J. (2009). Personality traits and intelligence predict academic school grades. *Learning and Individual Differences, 19*, 28-33.
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*, 1539-1552.
- Gorey, K. M. (2001). Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. *School Psychology Quarterly, 16*, 9-30.
- Hakstian, A. R., & Cattell, R. B. (1978). Higher-stratum ability structures on a basis of twenty primary abilities. *Journal of Educational Psychology, 70*, 657-669.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro, RJ: LTC.
- Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 443-451). New York, NY: Macmillan.

- Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment (2nd ed.)* (pp. 41-68). New York, NY: The Guilford Press.
- Kaufman, A. S., Kaufman, J. C., Liu, X., & Johnson, C. K. (2009). How do educational attainment and gender relate to fluid intelligence, crystallized intelligence, and academic skills at ages 22–90 years? *Archives of Clinical Neuropsychology, 24*, 153-163.
- Kuhn, J. T., & Holling, H. (2009). Gender, reasoning ability, and scholastic achievement: A multilevel mediation analysis. *Learning and Individual Differences, 19*, 229-233.
- Laros, J. A., Marciano, J. L. P., & Andrade, J. M. (2010). Fatores que afetam o desempenho na prova de Matemática do SAEB: Um estudo multinível. *Avaliação Psicológica, 9*, 173-186.
- 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., 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.
- Laros, J. A., Tellegen, P. J., Jesus, G. R., & Karino, C. A. (no prelo). *SON-R 2½-7[a] Teste não-verbal de inteligência: Manual com normatização e validação brasileira*. São Paulo, SP: Casa do Psicólogo.
- Leclerc, M., Larivée, S., Archambault, I., & Janosz, M. (2010). Le sentiment de compétence, modérateur du lien entre le QI et le rendement scolaire en mathématiques. *Revue Canadienne de L'Éducation, 33*, 31-56.
- Leeson, P., Ciarrochi, J., & Heaven, P. C. L. (2008). Cognitive ability, personality, and academic performance in adolescence. *Personality and Individual Differences, 45*, 630-635.
- Lloyd, J., & Barenblatt, L. (1984). Intrinsic intellectuality: Its relations to social class, intelligence, and achievement. *Journal of Personality and Social Psychology, 46*, 646-654.

- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment (2nd ed.)* (pp. 136-182). New York, NY: The Guilford Press.
- McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR) - Gc-Gf cross battery assessment*. Boston, MA: Allyn and Bacon.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools, 47*, 651-675.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.
- Primi, R. (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica, 1*, 67-77.
- Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences, 20*, 446-451.
- Rindermann, H., & Neubauer, A. C. (2004). Processing speed, intelligence, creativity, and school performance: Testing of causal hypotheses using structural equation models. *Intelligence, 32*, 573-589.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83-92.
- Rosander, P., Bäckström, M., & Stenberg, G. (2011). Personality traits and general intelligence as predictors of academic performance: A structural equation modelling approach. *Learning and Individual Differences, 5*, 590-596.
- Snijders, J. T., Tellegen, P. J., & Laros, J. A. (1989). *SON-R 5½-17: Manual and research report*. Groningen, ND: Wolters-Noordhoff.
- Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology, 15*, 201-293.
- Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19*, 80-90.
- Swartwout, M. D., Garnaat, S. L., Myszka, K. A., Fletcher, J. M., & Dennis, M. (2010). Associations of ethnicity and SES with IQ and achievement in Spina Bifida Meningomyelocele. *Journal of Pediatric Psychology, 35*, 927-936.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.

Manuscrito 2

Evidências de validade de construto do Teste de Raciocínio Abstrato e Espacial – TRAÉ

Evidence of the construct validity of the Abstract and Spatial Reasoning Test

Resumo

O presente estudo tem como objetivo a construção do Teste de Raciocínio Abstrato e Espacial e a verificação das evidências de validade de construto. A amostra do estudo foi composta por 1.069 estudantes do ensino fundamental com idades entre 11 e 17 anos. Os parâmetros de discriminação (a) e dificuldade (b) dos itens, estimados por meio da TRI, apresentaram médias próximas de 1,00 e 0,70, respectivamente. A curva de informação sugere que o instrumento é mais preciso para avaliar os adolescentes com $thetas$ entre 1 e 2. Em consonância com a teoria CHC de inteligência, as análises fatoriais indicaram que o instrumento possui uma estrutura que permite a estimação de dois fatores de primeira ordem, bem como um fator geral de segunda ordem associado à inteligência fluida. Em suma, os resultados indicaram que o instrumento é adequado ao uso dos psicólogos e, principalmente, dos pesquisadores da área.

Palavras-chave: psicometria; construção de instrumento; validade de construto; raciocínio abstrato; raciocínio espacial.

Abstract

The present study aims to develop the Abstract and Spatial Reasoning Test and obtain evidence of the construct validity. The sample consisted of 1,069 students varying in age from 11 to 17 years. The mean discrimination (a) and difficulty (b) parameter, based on IRT, showed an approximate value of 1.00 and .70, respectively. The information curve suggests that the instrument provides more reliable scores for teenagers with $thetas$ between 1 and 2. Consistent with the CHC model of intelligence, factor analysis indicated that the structure of the instrument allows the estimation of two first-order factors as well as a second-order factor related to fluid intelligence. Concluding, the results indicate that the instrument is appropriate for professionals and in particular for researchers.

Keywords: psychometrics; test development; construct validity; abstract reasoning; spatial reasoning.

O tema da inteligência é um dos aspectos mais investigados na área de psicologia. Destaca-se a importância desse construto psicológico na compreensão do desenvolvimento humano e na predição do desempenho acadêmico (Berg, 2000; Chen & Siegler, 2000; Neisser et al., 1996). Embora a definição de inteligência não seja consensual, o termo pode ser compreendido, de maneira geral, como a capacidade para aprender. Conceitos semelhantes são utilizados por diversos instrumentos psicológicos, tais como o RAVEN, o SON e o WISC. Para a bateria SON, por exemplo, o construto em questão é definido como a habilidade para aprender e a chance de sucesso na escola (Laros, Reis, & Tellegen, 2010; Snijders, Tellegen, & Laros, 1989).

Nas décadas de pesquisa sobre o tema foram propostos diversos modelos de compreensão da inteligência, muitos dos quais apresentaram divergências significativas. Nesse contexto, um dos debates mais importantes diz respeito aos modelos fatoriais da inteligência, que visam responder à questão de quantas e quais dimensões seriam necessárias para descrever, de maneira parcimoniosa, a inteligência. As teorias de Spearman e Thurstone ganharam força nesse debate, principalmente na primeira metade do século passado e podem ser consideradas percussoras de diversos modelos atuais (Brody, 2000; Hogan, 2006).

Spearman (1904) trabalhou com diversos testes de funções sensoriais já existentes na época e reuniu evidências de correlações significativas entre os testes pesquisados. Tais correlações, na perspectiva do autor, poderiam sustentar a existência de uma habilidade geral (fator *g*) subjacente aos testes específicos. Spearman nomeou o modelo de teoria dos dois fatores: um dos fatores composto pela habilidade geral (fator *g*) e o outro, pelas habilidades específicas (*s*). Salienta-se que, nesta perspectiva, a resposta a um determinado item de um teste cognitivo poderia ser explicada, concomitantemente, pelos fatores específico e geral (Almeida, 1994; Brody, 2000).

Thurstone (1938), por sua vez, concluiu que as correlações evidenciadas entre os testes não eram suficientemente altas para sustentar a existência de uma dimensão geral da inteligência. Assim, as dimensões específicas seriam independentes entre si, bem como mais importantes do que a compreensão de uma dimensão única. O modelo englobou sete habilidades específicas: compreensão verbal, fluência verbal, facilidade numérica, raciocínio espacial, velocidade perceptual, indução e memória (Brody, 2000; Davidson & Downing, 2000; Thurstone, 1938).

O embate entre os modelos unidimensionais (embasados em Spearman) *versus* multidimensionais (embasados em Thurstone) durou muitas décadas e, tampouco é possível afirmar que seja uma discussão superada. Contudo, modelos posteriores ajudaram a conciliar as duas teorias. Um desses modelos foi proposto por Cattell (1943, 1963), que organizou as habilidades cognitivas nas dimensões de inteligência fluída (*Gf*) e cristalizada (*Gc*). A dimensão *Gf* envolve as habilidades de raciocínio e a capacidade para resolução de problemas novos, para os quais a pessoa tem pouco conhecimento prévio. Essa dimensão está amplamente relacionada ao fator *g* de Spearman (Carroll, 2005; McGrew, 2005). A segunda dimensão, *Gc*, refere-se à aquisição e à solidificação (cristalização) de conhecimentos formais e informais, aprendidos por transmissão cultural ou pela escola (Cattell, 1943, 1963). Após alguns estudos de Horn, aluno de doutorado de Cattell, os autores acrescentaram outras sete dimensões básicas: aquisição e recuperação da memória de curto prazo; memória de longo prazo; inteligência visual; inteligência auditiva; velocidade do processamento cognitivo; velocidade de tomada de decisão; e conhecimento quantitativo (Horn, 1994; Horn & Blankson, 2005; McGrew, 2009).

Mais recentemente, o modelo CHC, uma das teorias de inteligência mais utilizadas, também adotou uma posição conciliatória no embate unidimensional *versus*

multidimensional. Esse modelo, proposto McGrew e Flanagan (1998), apresentou uma reorganização e ampliação das teorias já existentes de Cattell, Horn e Carroll (Carroll, 1993, 2005). A teoria CHC organiza as dimensões da inteligência em uma estrutura hierárquica de três níveis, que segue uma ordem de especialização, variando do mais geral - nível III - até as dimensões específicas - nível I. O nível mais amplo (fator de terceira ordem) é composto pela capacidade cognitiva geral. No segundo nível, as habilidades são organizadas em 10 dimensões gerais: inteligência fluída, inteligência cristalizada, conhecimento quantitativo, leitura escrita, memória de curto prazo, processamento visual, processamento auditivo, capacidade da memória de longo prazo, velocidade do processamento e rapidez na decisão. Finalmente, o nível mais baixo (I) é composto por aproximadamente 70 fatores bastante específicos, tais como raciocínio sequencial, raciocínio piagetiano, sensibilidade gramatical, proficiência em línguas estrangeiras e sensibilidade cinestésica (McGrew, 2005; Primi, 2003).

O modelo CHC engloba a inteligência de maneira bastante complexa e, talvez por esse motivo, a construção de um instrumento único, que abarque todo o modelo, seja uma tarefa extremamente difícil. Todavia, no Brasil, alguns instrumentos foram criados ou interpretados com base no modelo CHC, ainda que avaliem apenas uma parte das dimensões propostas pelo modelo em questão. Por exemplo, é possível citar o teste não-verbal de inteligência SON-R 2½-7[a] (Laros, Tellegen, Jesus, & Karino, no prelo). As primeiras versões da família dos testes SON são anteriores às concepções do modelo CHC, mas podem associar-se facilmente à proposta teórica do CHC. Os autores construíram o instrumento de maneira a abranger uma bateria de subtestes de funções específicas, organizando-os numa dimensão geral relacionada à inteligência fluída. Os estudos de validade oferecem suporte à interpretação da estrutura hierárquica do SON-R

2½-7[a]: uma dimensão geral que subjaz as dimensões específicas de raciocínio e de execução (Karino, Laros, & Jesus, 2011; Valentini, Laros, & Jesus, 2010).

Cita-se, ainda, a Bateria de Provas de Raciocínio BPR-5 (Primi & Almeida, 1998) como um exemplo de instrumento brasileiro interpretado com base no modelo CHC. O instrumento é composto por cinco subtestes de raciocínio (abstrato, espacial, numérico, verbal e mecânico) associados a diversas dimensões do nível II do modelo CHC. As correlações evidenciadas entre os cinco subtestes e os estudos de validade também sustentam a estrutura hierárquica e a presença de um fator geral (Primi & Almeida, 1998; Primi, Silva, Santana, Muniz, & Almeida, 2013).

Ainda que existam instrumentos de boa qualidade disponíveis no Brasil para a avaliação da inteligência, esses testes são de aplicação individual ou demandam mais de uma hora de aplicação. Tais aspectos oneram as pesquisas na área, visto que os participantes nem sempre dispõem de tempo suficiente para responder a uma pesquisa, individual ou coletiva, com tamanha duração. Ademais, diversos autores ressaltam a necessidade da elaboração de instrumentos brasileiros, bem como a apresentação de estudos de validade, precisão e normatização (Muñiz, Prieto, Almeida, & Bartram, 1999; Nakano, 2013; Noronha & Alchieri, 2002; Sparta, Bardagi, & Teixeira, 2006).

O presente estudo teve como objetivo geral a construção do Teste de Raciocínio Abstrato e Espacial - TRAE e a obtenção de evidências de validade de construto. O instrumento foi concebido para a avaliação acessível e rápida das habilidades para solucionar problemas abstratos, para os quais o indivíduo tem pouco conhecimento prévio, bem como para avaliação da capacidade de visualização, manipulação mental de imagens e de orientação no espaço tridimensional. Ambas as habilidades estão relacionadas à inteligência fluída.

Especificamente, a presente pesquisa buscou: (a) avaliar a qualidade dos itens da primeira versão do TRAE e consolidar a segunda versão do instrumento; (b) avaliar a qualidade dos itens da segunda versão do TRAE; (c) testar os modelos de estrutura interna; (d) avaliar a precisão do instrumento. Para tanto, a seguir será descrito o método empregado nesta investigação.

Método

Participantes

A presente pesquisa analisou os dados de duas amostras de estudantes. A primeira amostra foi utilizada para avaliação inicial dos itens da TRAE. Participaram desta etapa da pesquisa 149 estudantes (52,3% do sexo masculino) do ensino médio de duas escolas, sendo uma delas da rede pública. Os estudantes tinham idade entre 14 e 19 anos ($M = 16,98$; $DP = 0,87$).

A segunda amostra, utilizada para a consolidação do TRAE e análise da estrutura fatorial, foi composta de 1.295 estudantes do oitavo ano do ensino fundamental. Na análise exploratória dos dados deste estudo foi observado que uma parte dos dados coletados apresentou um padrão muito semelhante ao esperado ao acaso (“chute”). Tais participantes foram excluídos das análises, considerando os seguintes critérios: (a) o número de acertos próximo do esperado ao acaso, incluindo uma margem de tolerância de dois acertos (acima do esperado ao acaso); (b) o padrão de respostas incluía erro nas questões fáceis e acerto nas questões difíceis. Também foram excluídos da amostra inicial os participantes que ofereceram a mesma alternativa de resposta para a maioria das questões. Assim, a amostra final foi composta de 1.069 participantes com idade entre 11 e 17 anos ($M = 13,3$; $DP = 0,7$). Salienta-se que os dados foram coletados

em 42 escolas de uma rede pública de ensino de uma das maiores cidades do Brasil (uma turma de aproximadamente 30 alunos por escola).

Instrumento

Os participantes responderam ao Teste de Raciocínio Abstrato e Espacial (TRAE). O instrumento é dividido em dois subtestes. O primeiro, Raciocínio Abstrato (RA), busca avaliar as habilidades associadas à resolução de problemas de ordem intelectual, para os quais a pessoa tem pouco ou nenhum conhecimento prévio. Nas questões de RA o participante precisa descobrir o princípio da transformação de uma figura A para uma figura B e aplicá-lo na transformação da figura C para a figura D. O grau de dificuldade das questões aumenta conforme o número de partes que compõem a figura e a quantidade de transformações, bem como a plausibilidade das alternativas incorretas. O subteste de RA é composto por 12 questões de múltipla escolha, nas quais são oferecidas quatro alternativas de resposta e o tempo limite de aplicação é de 13 minutos (Anexo II da tese).

O subteste de Raciocínio Espacial (RE) é destinado à avaliação das habilidades visuoespaciais, associadas à sensibilidade para detalhes visuais e à orientação no espaço tridimensional. Para os itens de RE a pessoa deve observar as rotações que ocorrem em três figuras geométricas dispostas em sequência e descobrir como ficaria a próxima figura observando a mesma rotação das anteriores. O grau de dificuldade está relacionado, nesse subteste, ao número e ao tipo de face das figuras, bem como ao tipo de rotação. O subteste é composto por 12 questões de múltipla escolha, nas quais são oferecidas cinco alternativas de resposta. O tempo limite de aplicação é de 17 minutos (Anexo III da tese).

Procedimento

A primeira amostra deste estudo foi coletada em sessões coletivas com duração de aproximadamente 30 minutos. Buscou-se, nessa fase da pesquisa, avaliar a qualidade dos itens, subsidiando, assim, as alterações do instrumento e a exclusão de itens. Para tanto, avaliou-se a dificuldade e discriminação dos itens por meio da Teoria Clássica dos Testes.

Profissionais treinados coletaram a segunda amostra em sessões coletivas de aproximadamente 30 minutos. Os participantes foram voluntários e assentiram a participação. Além disso, foi enviado aos pais um termo de informação, conforme orientação do comitê de ética.

As evidências de validade de estrutura e a qualidade dos itens do TRAE foram avaliadas com base na Teoria Clássica dos Testes e Teoria de Resposta ao Item (Hambleton, Swaminathan, & Rogers, 1991; Nunnally & Bernstein, 1994; Valentini & Laros, 2011). A precisão dos escores do instrumento (fidedignidade) foi estudada por meio do cálculo de Lambda 2 (Cronbach, 1996) e Curva de Informação. As análises foram realizadas com o auxílio dos *softwares* SPSS, Testfact, Bilog-Mg, Noharm e Mplus.

Resultados

Análise dos Itens e Curva de Informação do TRAE

No intuito de consolidar os itens do TRAE, avaliou-se a primeira versão do instrumento composta por 30 itens (15 itens por subtteste). Para tanto, foram analisados os índices de dificuldade e discriminação da Teoria Clássica dos Testes, apresentados na Tabela 1.

Tabela 1. Índices psicométricos da Teoria Clássica dos Testes da primeira versão do TRAE

Item ^a	Dificuldade (p)	Gabarito	Discriminação				
			Correlação item-total por alternativa de resposta (r_{bis})				
			A	B	C	D	E
RA1	0,99	B	0,00	<i>0,52</i>	0,00	0,00	-
RA2	0,94	D	-0,60	-0,61	0,00	<i>0,69</i>	-
RA3	0,91	B	-0,33	<i>0,35</i>	-0,18	-0,22	-
RA4	0,87	B	-0,38	<i>0,55</i>	-0,43	-0,41	-
RA5	0,94	A	<i>0,69</i>	-1,06	-0,33	-0,18	-
RA6	0,93	C	-1,19	-0,47	<i>0,78</i>	-0,54	-
RA7	0,62	D	-0,34	-0,30	<i>0,15</i>	<i>0,33</i>	-
RA8	0,83	B	-0,51	<i>0,57</i>	-0,41	-0,31	-
RA9	0,92	A	<i>0,50</i>	-0,29	-0,17	-0,66	-
RA10	0,75	B	-0,44	<i>0,64</i>	-0,46	-0,43	-
RA11	0,80	C	-0,42	-0,28	<i>0,42</i>	-0,02	-
RA12	0,75	D	-0,80	-0,64	-0,44	<i>0,80</i>	-
RA13	0,67	D	-0,44	-0,40	-0,41	<i>0,60</i>	-
RA14	0,56	A	<i>0,74</i>	-0,10	-0,44	-0,49	-
RA15	0,60	D	-0,27	-0,34	-0,71	<i>0,71</i>	-
RE1	0,88	B	-0,24	<i>0,32</i>	<i>0,13</i>	-0,44	-0,45
RE2	0,90	D	-0,60	-0,51	-0,60	<i>0,67</i>	-0,53
RE3	0,91	C	-0,31	-0,42	<i>0,39</i>	-0,33	-0,02
RE4	0,94	D	-0,60	-0,16	-0,02	<i>0,41</i>	-0,35
RE5	0,91	A	<i>0,43</i>	0,00	-0,39	-0,35	0,00
RE6	0,67	A	<i>0,56</i>	-0,45	-0,53	0,00	-0,20
RE7	0,68	E	-0,16	-0,41	0,00	-0,37	<i>0,50</i>
RE8	0,52	E	-0,26	-0,14	-0,31	-0,22	<i>0,57</i>
RE9	0,66	D	-0,28	-0,35	-0,28	<i>0,56</i>	-0,19
RE10	0,47	C	-0,07	-0,20	<i>0,38</i>	-0,23	-0,02
RE11	0,43	E	-0,11	-0,19	-0,21	-0,14	<i>0,71</i>
RE12	0,39	D	-0,09	-0,28	-0,09	<i>0,65</i>	-0,19
RE13	0,33	A	<i>0,68</i>	-0,10	-0,14	<i>0,22</i>	-0,28
RE14	0,29	C	-0,14	-0,02	<i>0,69</i>	-0,03	-0,30
RE15	0,14	B	<i>0,10</i>	<i>0,67</i>	-0,24	<i>0,06</i>	<i>0,15</i>

Notas. p = proporção de respostas corretas; r_{bis} = correlação bisserial item-total. As correlações bisseriais positivas estão destacadas em itálico.

^a Os itens apresentados nesta tabela fazem parte da primeira versão. A ordem de apresentação não corresponde à segunda versão do instrumento.

Os resultados apresentados na Tabela 1 indicaram que, em geral, os itens do subtteste RA são mais fáceis se comparados aos de RE (a proporção de acertos de RA variou de 0,56 a 0,99 e a de RE variou de 0,14 a 0,94). A média da proporção de acertos dos itens do subtteste RA foi superior a de RE (RA = 0,81 e de RE = 0,61). De maneira geral, as correlações bisseriais foram superiores a 0,30 para a alternativa correta,

indicando que os itens apresentaram discriminação adequada. Alguns itens do subteste RE apresentaram correlações positivas para uma ou mais alternativas incorretas.

Entretanto, mesmo para estes casos, a diferença entre as correlações para a alternativa correta e para as incorretas é alta, o que também respalda a adequação dos índices de discriminação.

Para reduzir o tempo de aplicação do instrumento, excluíram-se três questões por subteste. Buscou-se eliminar os itens mais difíceis de RE e os mais fáceis de RA, deixando os subtestes mais equilibrados no que se refere à média da proporção de acerto dos itens. Neste sentido, eliminaram-se os itens com variância menor. Um motivo adicional para a exclusão dos dois últimos itens de RE refere-se à rotação dupla para a mesma figura, o que os diferenciava dos demais itens e os tornaram mais difíceis.

Após a exclusão de itens, a segunda versão do instrumento foi composta por 24 itens (12 por subteste). Salienta-se, ainda, que os itens da nova versão do instrumento foram renumerados, portanto os números dos itens das versões não são correspondentes. Na Tabela 2 são apresentados os índices de dificuldade, discriminação e acerto ao acaso da segunda versão do TRAE e informação do item por nível de habilidade *theta*. No que diz respeito à TRI, estimou-se, primeiramente, os parâmetros do subteste RA. Após a convergência do modelo e análise dos erros padrões de estimativa, foi realizada uma segunda análise, na qual foram adicionados os itens de RE e fixados os parâmetros dos itens de RA (obtidos na primeira análise). Esse procedimento visou garantir que os parâmetros da TRI dos dois subtestes fossem posicionados na mesma escala.

Tabela 2. Índices de dificuldade, discriminação, acerto ao acaso e informação da segunda versão do TRAE

Item	Parâmetros dos itens					Informação do item por nível de habilidade (TRI)						
	TCT		TRI			-3	-2	-1	0	+1	+2	+3
	p	(r_{bis})	a	b	c^a							
RA1	0,74	0,50	0,43	-0,93	0,25	0,03	0,06	0,08	0,08	0,06	0,04	0,02
RA2	0,70	0,59	0,66	-0,45	0,25	0,01	0,05	0,15	0,19	0,12	0,05	0,02
RA3	0,58	0,68	1,05	0,23	0,25	0,00	0,00	0,07	0,42	0,36	0,09	0,02
RA4	0,41	0,65	1,54	0,97	0,25	0,00	0,00	0,00	0,09	1,04	0,30	0,02
RA5	0,62	0,73	1,20	0,04	0,25	0,00	0,00	0,10	0,61	0,33	0,06	0,01
RA6	0,47	0,53	0,54	1,12	0,25	0,00	0,01	0,03	0,07	0,12	0,12	0,08
RA7	0,44	0,67	1,03	0,89	0,25	0,00	0,00	0,01	0,15	0,47	0,24	0,05
RA8	0,53	0,57	0,68	0,62	0,25	0,00	0,01	0,04	0,14	0,20	0,13	0,06
RA9	0,44	0,58	0,75	1,10	0,25	0,00	0,00	0,02	0,10	0,24	0,21	0,09
RA10	0,37	0,67	1,56	1,12	0,25	0,00	0,00	0,00	0,04	0,95	0,42	0,04
RA11	0,27	0,52	1,63	1,87	0,25	0,00	0,00	0,00	0,00	0,11	1,19	0,23
RA12	0,32	0,59	1,04	1,68	0,25	0,00	0,00	0,00	0,02	0,23	0,47	0,18
M	0,49	0,61	1,04	0,69	-							
DP	0,13	0,07	0,40	0,80	-							
RE1	0,74	0,73	0,96	-0,67	0,20	0,00	0,07	0,37	0,38	0,12	0,03	0,01
RE2	0,62	0,68	0,67	-0,12	0,20	0,01	0,03	0,13	0,23	0,17	0,08	0,03
RE3	0,56	0,67	0,92	0,21	0,20	0,00	0,01	0,10	0,37	0,32	0,10	0,02
RE4	0,60	0,77	1,32	0,01	0,20	0,00	0,00	0,12	0,82	0,34	0,05	0,01
RE5	0,53	0,71	1,10	0,33	0,20	0,00	0,00	0,06	0,47	0,45	0,11	0,02
RE6	0,49	0,69	0,95	0,49	0,20	0,00	0,00	0,05	0,31	0,41	0,15	0,03
RE7	0,35	0,67	1,20	1,12	0,20	0,00	0,00	0,00	0,09	0,65	0,39	0,07
RE8	0,32	0,54	0,74	1,67	0,20	0,00	0,00	0,01	0,05	0,18	0,26	0,16
RE9	0,48	0,68	0,84	0,58	0,20	0,00	0,00	0,05	0,24	0,34	0,16	0,05
RE10	0,30	0,58	0,82	1,70	0,20	0,00	0,00	0,00	0,04	0,20	0,33	0,18
RE11	0,22	0,62	1,53	1,72	0,20	0,00	0,00	0,00	0,00	0,28	1,08	0,18
RE12	0,26	0,58	1,30	1,75	0,20	0,00	0,00	0,00	0,01	0,26	0,81	0,21
M	0,46	0,66	1,03	0,73	-							
DP	0,16	0,07	0,25	0,80	-							

Notas. TCT = Teoria Clássica dos Testes; TRI = Teoria de Resposta ao Item;

p = proporção de acertos (dificuldade do item na TCT); r_{bis} = correlação bisserial item-total (discriminação do item na TCT); a = parâmetro de discriminação da TRI (*slope*); b = parâmetro de dificuldade da TRI (*threshold*); c = parâmetro de acerto ao acaso da TRI (*pseudo-chance*); M = média; DP = desvio padrão.

^a Todos os parâmetros c foram fixados conforme o inverso do número de alternativas (1 / número de alternativas).

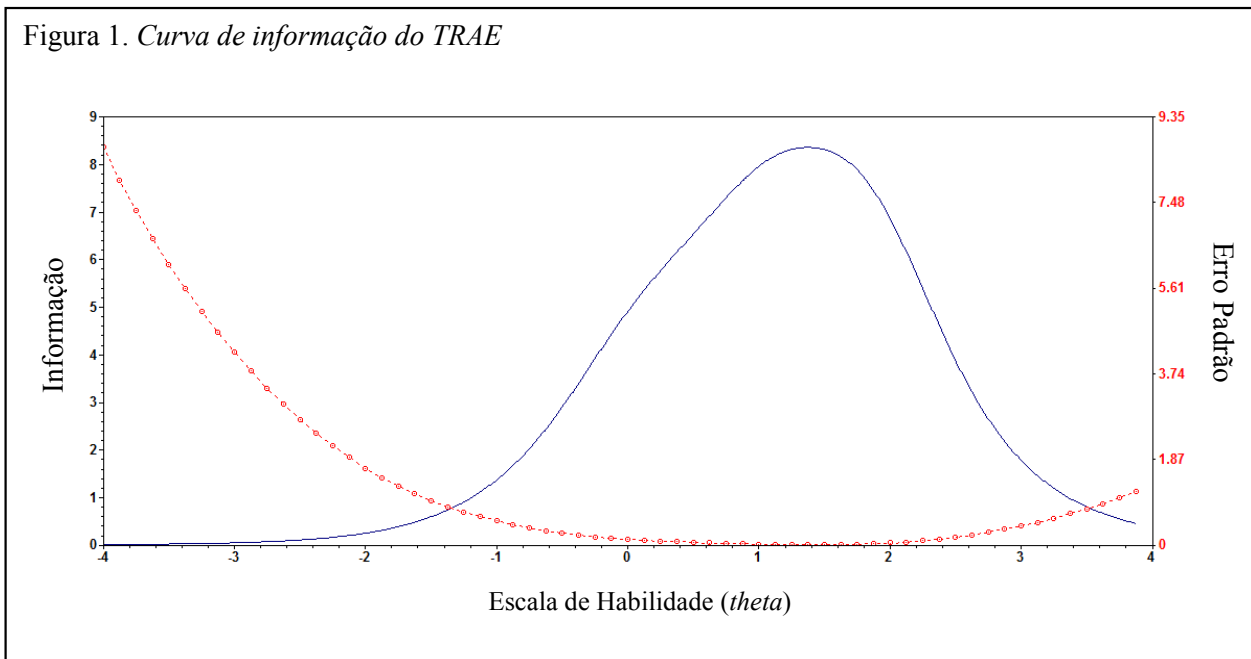
A inspeção da Tabela 2 mostra que ambos os subtestes apresentaram dificuldade mediana (média dos parâmetros b de RA = 0,69 e RE = 0,73). Além disso, os itens se distribuem na proporção aproximada de 13% de itens fáceis, 75% de itens medianos e 13% de itens difíceis. Destaca-se ainda, que as correlações bisseriais item-total foram superiores a 0,30 e que os parâmetros de discriminação da TRI (*slope*) apresentaram

média, aproximada, igual a um. Tais resultados indicam que os itens do instrumento apresentaram capacidade discriminativa minimamente adequada.

Realizaram-se análises de TRI adicionais com a livre estimação do parâmetro c . Contudo, os modelos apresentaram redução nos valores do parâmetro a (discriminação), bem como nas cargas fatoriais. Portanto, optou-se pela estimação do modelo com o parâmetro c fixo. Justifica-se, ainda, a escolha deste modelo em detrimento ao modelo de dois parâmetros (2PL) em função do tipo de resposta do instrumento: os itens são de múltipla escolha, permitindo que alguns participantes com baixa habilidade os respondam de maneira correta.

Na Tabela 2 também são apresentadas a quantidade de informação do item para diferentes níveis de habilidade θ . Ressalta-se que a informação está diretamente relacionada à precisão. Diferentemente da TCT, a informação estimada por meio da TRI possibilita a avaliação da precisão dos escores para cada nível de habilidade. Os resultados da Tabela 2 indicam que a maior quantidade de informação esteve relacionada aos níveis de habilidade entre zero e dois. Além disso, para os níveis de habilidade menos três e menos dois, todos os itens produziram pouca informação precisa.

Buscou-se, ainda, avaliar a curva de informação geral do teste, cujo resultado é apresentado na Figura 1. De maneira semelhante aos dados apresentados na Tabela 2, a curva de informação indica a quantidade de informação associada aos diferentes níveis de escores gerados pelo instrumento.



Por meio da Figura 1 verificou-se que o instrumento é mais preciso para avaliar os adolescentes com habilidade geral entre os θ um e dois. Além disso, para os adolescentes com θ inferior à -1,5 ou superior à 3,5, o instrumento produziu mais erro de estimação do que informação verdadeira.

Análise da Estrutura Interna

Buscou-se explorar a estrutura fatorial do TRAE, tendo como base a divisão dos subtestes de Raciocínio Abstrato (RA) e Raciocínio Espacial (RE), bem como na possibilidade de estruturação do instrumento num fator único relacionado à inteligência fluída. Para tanto, foram realizadas análises fatoriais contemplando uma ou duas dimensões. Considerando que os itens do TRAE são codificados como ‘certo’ ou ‘errado’ (portanto dicotômicos), as análises fatoriais embasadas em matrizes de correlação de Pearson ou de covariâncias poderiam gerar ‘fatores de dificuldade’ não relacionados ao conteúdo do teste (Ayala, 2009). Buscando evitar esse problema, optou-se pela realização de uma análise fatorial de informação plena (FIFA – *Full Information*

Factor Analysis) e uma análise *bi-factor* por se tratarem de modelos estimados a partir de uma matriz de correlações tetracóricas, mais adequadas aos itens dicotômicos.

Ademais, realizaram-se análises não-lineares, que também são robustas para escalas dicotômicas. Os resultados de tais análises são indicados na Tabela 3. Salienta-se que os primeiros itens dos subtestes eram demasiadamente fáceis e prejudicavam a solução fatorial. Contudo, tais itens são importantes para verificar se o participante compreendeu adequadamente a tarefa do teste. Por esses motivos, o primeiro item de cada subteste foi excluído das análises fatoriais, mas foi mantido no subteste.

Tabela 3. Cargas fatoriais dos itens do TRAE para os modelos não-lineares, FIFA e *bi-factor*

Itens	AF Não-Linear			FIFA			
	Unifatorial ($c = 0$)	Dois Fatores (Promax, $c = 0$) ^a		Unifatorial (c livre)	<i>Bi-factor</i> (c livre)		
		RA	RE		Geral	Específico RA	Específico RE
RA2	0,35	0,61		0,30	0,25	0,33	-
RA3	0,60	0,58		0,44	0,44	0,31	-
RA4	0,67	0,48		0,45	0,46	0,25	-
RA5	0,59	0,72		0,43	0,39	0,37	-
RA6	0,29	0,44		0,25	0,20	0,24	-
RA7	0,49	0,65		0,38	0,33	0,32	-
RA8	0,42	0,42		0,35	0,30	0,26	-
RA9	0,47	0,37		0,38	0,35	0,19	-
RA10	0,59	0,56		0,42	0,39	0,32	-
RA11	0,42	0,29		0,35	0,33	0,11	-
RA12	0,46	0,42		0,36	0,26	-0,02	-
RE2	0,55		0,69	0,44	0,43	-	0,02
RE3	0,64		0,54	0,46	0,50	-	0,04
RE4	0,83		0,72	0,50	0,60	-	0,03
RE5	0,71		0,63	0,48	0,51	-	0,01
RE6	0,65		0,59	0,46	0,47	-	0,05
RE7	0,68		0,57	0,46	0,46	-	0,04
RE8	0,39		0,41	0,34	0,30	-	0,01
RE9	0,60		0,60	0,44	0,47	-	-0,01
RE10	0,44		0,46	0,37	0,36	-	0,01
RE11	0,55		0,51	0,41	0,36	-	0,00
RE12	0,43		0,49	0,36	0,31	-	0,00
RMSR	0,01		0,007				
Tanaka	0,96		0,990				

Notas. AF = Análise Fatorial; FIFA = Full Information Factor Analysis; RMSR = Root Mean Square Residual. Tanaka = Indicador de ajuste Tanaka.

^a As cargas fatoriais, relativas a análise não-linear, abaixo de 0,30 foram omitidas na tabela.

No que se refere aos indicadores de ajuste da análise fatorial não-linear, o valor crítico para o indicador de ajuste RMSR é calculado por meio da seguinte equação: $4 \times (1 \div \sqrt{n})$. Assim, para a amostra de 1.069, o valor crítico é de 0,12. Portanto, os valores de RMSR, abaixo do ponto de corte, e o indicador Tanaka acima de 0,95 indicam que os dados se ajustaram adequadamente tanto ao modelo de um fator, quanto ao de dois fatores, usando a análise fatorial não-linear. Percebe-se que o modelo de dois fatores foi levemente mais bem ajustado, embora ambos os modelos tenham explicado aproximadamente 30% da variância dos dados. Ressalta-se, ainda, que os fatores RA e RE apresentaram correlação de moderada a forte ($r = 0,52$).

O fator único da análise FIFA e o fator geral do modelo *bi-factor* explicaram, respectivamente, 16,5% e 15,7% da variância. Ademais, a correlação entre as cargas fatoriais do modelo FIFA unifatorial e do fator geral do *bi-factor* apresentou magnitude igual a 0,96. Percebe-se também que os valores absolutos das cargas fatoriais são muito semelhantes: a diferença entre elas não ultrapassou 0,10. Isso indica que a modelagem dos fatores específicos de RA e RE no modelo *bi-factor* não distorceu as cargas fatoriais do fator geral (do modelo de fator único).

No que se refere aos fatores específicos do modelo *bi-factor* percebe-se que as cargas fatoriais de RA foram maiores do que as de RE. As cargas de RE foram, inclusive, muito próximas de zero. Esses resultados indicam que o subteste RE está quase unicamente medindo a dimensão geral, enquanto RA é mais propenso a apresentar variância específica do subteste.

Os resultados, portanto, sustentam a plausibilidade tanto do modelo de um fator como o de dois fatores. Para testar o ajuste dos modelos foram realizadas análises de equações estruturais. Adicionalmente, testou-se o ajuste de um modelo com fator geral de segunda ordem. Esses resultados são apresentados na Tabela 4.

Tabela 4. *Indicadores de ajuste dos modelos das análises de equações estruturais*

Modelo	χ^2	gl	χ^2 / gl	TLI	CFI	RMSEA (IC 90%)
2 fatores não-correlacionados	1499,28	209	7,17	0,541	0,584	0,076 (0,072-0,080)
1 fator	456,96	209	2,19	0,912	0,920	0,033 (0,029-0,037)
2 fatores correlacionados	322,70	208	1,55	0,959	0,963	0,023 (0,018-0,027)
Segunda ordem	322,70	208	1,55	0,959	0,963	0,023 (0,018-0,027)
<i>Bi-factor</i>	288,48	188	1,53	0,960	0,968	0,022 (0,017-0,027)

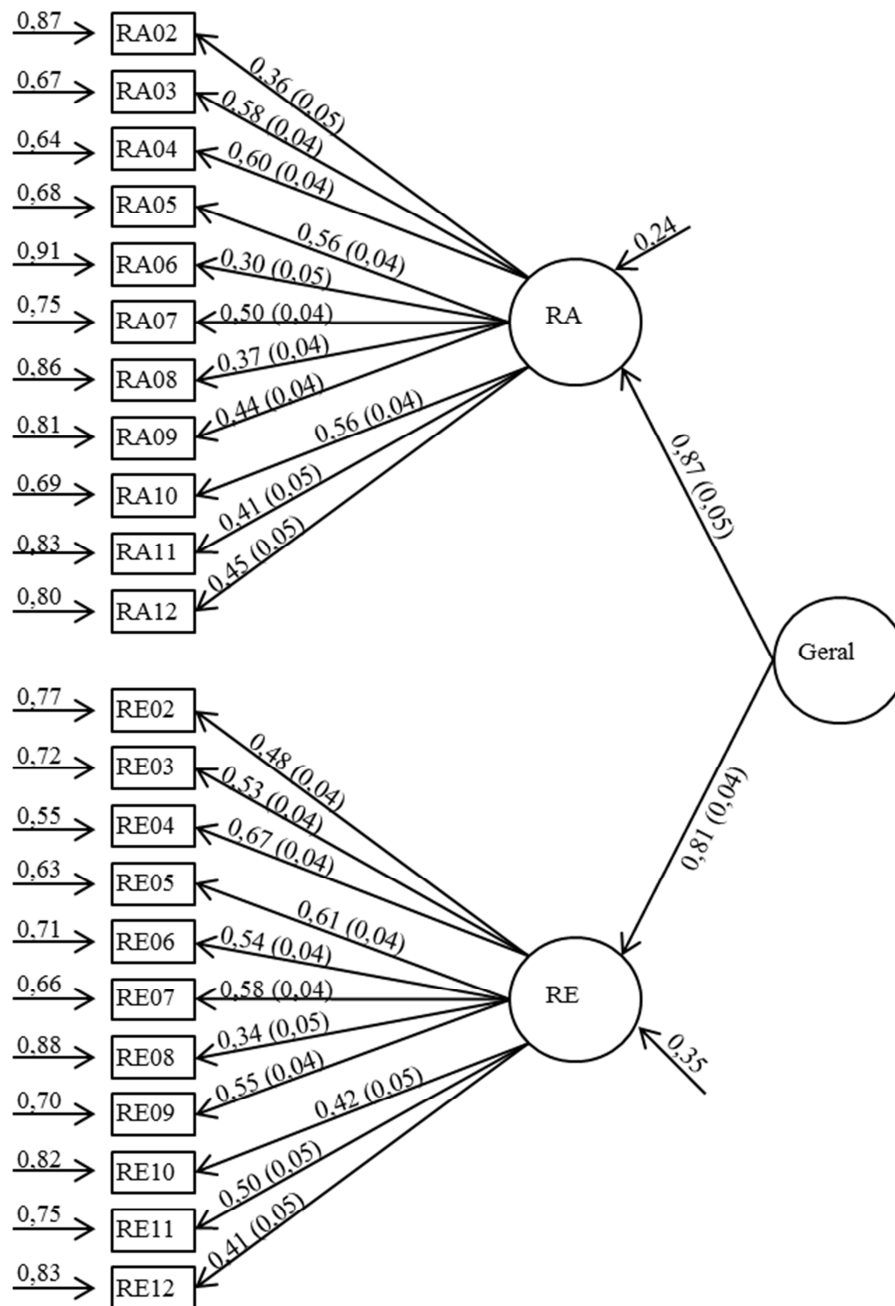
Notas. χ^2 = qui-quadrado; gl = graus de liberdade; TLI = *Tucker Lewis Index*; CFI = *Comparative Fix Index*; RMSEA = *Root Mean Square Error of Approximation*;

Embora o modelo de um fator (dimensão única de primeira ordem) tenha apresentado indicadores de ajuste aceitável (TLI e CFI > 0,90), a inclusão dos fatores específicos correlacionados melhorou significativamente o ajuste do modelo. Salienta-se, no entanto, que o modelo com dois fatores não-correlacionados (correlação fixada em zero) apresentou ajuste extremamente pobre. Ademais, o modelo de dois fatores correlacionados apresentou associação entre RA e RE igual a 0,71.

O modelo *bi-factor* e o de segunda ordem apresentaram indicadores de ajuste muito semelhantes. Esses resultados indicam que ambos os modelos são igualmente plausíveis. Contudo, ao analisar os coeficientes padrão (*pattern* ou cargas fatoriais) dos modelos, percebeu-se que o de segunda ordem apresentou coeficientes mais altos e minimizou os erros padrões de estimativa, se comparado aos resultados do *bi-factor*. Além disso, não foi encontrada significância estatística (razão crítica < 1,96) em seis coeficientes padrão das dimensões específicas do modelo *bi-factor*. Neste sentido, embora os dois modelos sejam plausíveis, adotou-se o modelo de segunda ordem como solução da estrutura do TRAE. Na Figura 2, são indicados os coeficientes padrão

(*pattern*) padronizados, erros padrões de estimação e variâncias residuais do modelo de segunda ordem do TRAE.

Figura 2. Análise fatorial confirmatória do TRAE



Notas. Nos parênteses são indicados os erros padrões de estimação (padronizados); as variâncias residuais são apresentadas acima das flechas pequenas.

TLI = 0,96; CFI = 0,96; RMSEA (IC 90%) = 0,02 (0,018 - 0,027).

Os coeficientes padrão do fator de segunda ordem foram acima de 0,80, o que oferece suporte adicional à necessidade de estimar um fator geral de segunda ordem no modelo. No que se refere aos coeficientes dos fatores de primeira ordem, observa-se que os valores foram acima de 0,30 e estatisticamente significativos. Além disso, os limites inferiores dos intervalos de confiança (IC 95%) desses coeficientes foram sempre superiores a 0,20. Entretanto, aproximadamente 45% desses coeficientes apresentaram magnitude abaixo de 0,50. Consequentemente, as variâncias residuais dos itens são relativamente altas. Isso pode ser parcialmente atribuído à natureza dicotômica dos itens: escalas do tipo ‘certo’ ou ‘errado’ tendem a apresentar restrição de variância, o que prejudica as análises fatoriais.

Finalmente, avaliou-se a precisão dos escores por meio da consistência interna. Para tanto, foi calculado o coeficiente Lambda 2 de Guttman. O fator geral mostrou-se mais preciso ($\lambda_2 = 0,78$) do que os fatores específicos RA ($\lambda_2 = 0,69$) e RE ($\lambda_2 = 0,73$). Também foi possível estimar a precisão relacionada aos fatores específicos por meio da Modelagem por Equações Estruturais, utilizando a variância residual. Neste caso, as variâncias residuais apresentadas para RA e RE na Figura 2 foram, respectivamente, de 0,24 e 0,35. Portanto, a precisão de RA é igual a 0,76 (ou seja, $1 - 0,24 = 0,76$) e a de RE igual a 0,65 (ou seja, $1 - 0,35 = 0,65$).

Discussão

Esta pesquisa buscou evidências de validade de estrutura interna e precisão do Teste de Raciocínio Abstrato e Espacial – TRAE. Ainda que o contexto brasileiro disponha de bons instrumentos para avaliação da inteligência, o TRAE oferece, principalmente aos pesquisadores da área, a possibilidade de realizar uma testagem acessível e rápida de dois tipos de raciocínio: abstrato (RA) e espacial (RE). Além disso, o TRAE pode ser útil como uma ferramenta auxiliar em triagens ou demais

aplicações profissionais nas quais o TRAE possa ser utilizado juntamente com outro instrumento de inteligência.

No que se refere à construção do TRAE, cumpre salientar que o estudo com a primeira versão do instrumento permitiu a consolidação de uma segunda versão, mais sólida. Embora a maioria dos itens da primeira versão tenha apresentado índice de discriminação adequado, a exclusão de alguns itens melhorou o equilíbrio da dificuldade dos subtestes de RA e RE. Esse equilíbrio vai ao encontro das discussões propostas pela literatura em psicometria no que se refere à distribuição da dificuldade dos itens: a maior parte de itens medianos e poucos itens fáceis ou difíceis (Hambleton & Swaminathan, 1985; Pasquali, 2003).

A segunda versão do TRAE apresentou índices de dificuldade e discriminação semelhantes entre os subtestes. No geral, a média do parâmetro de dificuldade e a curva de informação indicaram que o instrumento é mais adequado à população de estudantes levemente acima da média. Ressalta-se, no entanto, que os participantes eram estudantes do oitavo ano do ensino fundamental (antiga sétima série). Assim, pesquisas com estudantes mais jovens poderiam alterar esses resultados, aumentando a precisão do instrumento nas avaliações de adolescentes com escores mais baixos.

No que diz respeito à discriminação dos itens, os parâmetros da TCT e da TRI apresentaram valores adequados. Embora o valor do parâmetro a de 50% dos itens (de 12 dos 24 itens) tenha sido inferior a um, a média desse parâmetro em ambas as escalas foi de aproximadamente um. Sabe-se que discriminação de um item está relacionada à precisão dos escores. Nesse sentido, Primi e Almeida (1998, 2000) também encontraram indicadores de precisão mais baixos na prova de RE da Bateria de Provas de Raciocínio (BPR-5). Em suma, ainda que a discriminação dos itens não seja muito elevada, os índices apresentaram valores esperados e respaldam a qualidade do

instrumento para utilização em pesquisas. Além disso, os resultados de precisão oferecem suporte à utilização do TRAE como uma ferramenta auxiliar no processo de diagnóstico e avaliação psicológica.

Uma das vantagens da TRI sobre a TCT diz respeito à possibilidade de analisar a precisão dos itens e do instrumento para os diferentes níveis de habilidade (Hambleton & Swaminathan, 1985; Kemp, 1955). Ou seja, o instrumento não é igualmente preciso para todos os estudantes. Por um lado, os resultados do TRAE indicaram maior precisão na estimação dos escores dos participantes com habilidades cognitivas medianas ou levemente acima da média. Esse resultado pode ser particularmente interessante aos pesquisadores da área cognitiva que buscam avaliar os estudantes de diferentes níveis de habilidade. Por outro lado, o TRAE mostrou-se impreciso na estimação dos escores dos participantes com baixas habilidades. Principalmente para aqueles com escores de um ou mais desvios padrão abaixo da média, o instrumento parece não produzir muita informação válida. Neste sentido, não se recomenda o uso do TRAE para investigações de pessoas com graves limitações cognitivas.

O estudo também forneceu evidências da estrutura interna do instrumento. As análises fatoriais não-lineares apontaram para a plausibilidade tanto do modelo de um fator como do modelo de dois fatores. Ainda que os indicadores de ajuste do modelo de dois fatores tenham sido levemente superiores aos do modelo unifatorial, a média das cargas fatoriais foi de 0,53 para ambos os modelos. Ademais, os modelos explicaram praticamente a mesma quantidade da variância dos itens. Ou seja, ainda que a estimação separada dos escores de RA e de RE seja possível, os itens também são explicados por uma dimensão geral. Tais resultados oferecem suporte à unidimensionalidade do modelo e justificam o cálculo de um escore geral para os dois subtestes. Essa discussão

também é relevante para a estimação dos parâmetros da TRI, cujos modelos utilizados nesta pesquisa têm como pressuposto a unidimensionalidade.

As análises fatoriais por meio da TRI apresentaram conclusões semelhantes às análises não-lineares. As cargas fatoriais da análise FIFA apresentaram valores esperados e a estimação dos fatores específicos do modelo *bi-factor* não distorceu o fator geral. Esse resultado também oferece sustentação à unidimensionalidade do TRAE (Reise, Morizot, & Hayes, 2007). No entanto, o fator específico RE do modelo FIFA *bi-factor* praticamente não acrescentou explicação da variância dos itens. Ademais, apenas cinco itens (42%) apresentaram cargas fatoriais acima de 0,30 no fator específico de RA. Tais resultados divergem dos indicados com a Bateria de Provas de Raciocínio (BPR-5). No estudo de Primi et al. (2013) o modelo *bi-factor* para a forma A da BPR-5 (aplicada a estudantes do ensino fundamental) apresentou 48% de cargas fatoriais superiores a 0,29 no fator específico de RA, bem como 85% no fator RE. Cabe destacar que o TRAE e a BPR-5 são instrumentos que oferecem tarefas semelhantes, mas que apresentam dimensões e número de itens diferentes. O instrumento de Primi e Almeida oferece, ainda, a possibilidade de avaliar aspectos de raciocínio verbal, mecânico e numérico, o que difere da proposta do TRAE. Tais características podem explicar as diferenças encontradas nos estudos.

As análises fatoriais confirmatórias sustentaram a possibilidade de estimar os fatores específicos de RA e RE. Por um lado, o modelo de dois fatores não-correlacionados apresentou ajuste aos dados muito aquém do esperado, indicando que a estimação de fatores específicos, sem considerar a relação entre eles, é pouco plausível. Por outro lado, a estimação de um modelo com um fator único apresentou indicadores de ajuste mais pobres se comparados aos modelos de dois fatores correlacionados, de segunda ordem e o *bi-factor*.

O modelo de segunda ordem apresentou ajuste idêntico ao de dois fatores correlacionados. Isso era esperado e pode ser atribuído a uma das características da Modelagem por Equações Estruturais: ao estimar um fator de segunda ordem (fator geral) para apenas dois fatores de primeira ordem (RA e RE) o modelo é não-identificado, a menos que se fixe um parâmetro adicional (Keith, 2005; Thompson, 2004). Neste caso, fixaram-se as cargas fatoriais como iguais entre os fatores de primeira e segunda ordem. Considerando que os fatores de primeira ordem apresentaram alta correlação, era esperado que a estimação de uma dimensão de segunda ordem não aumentasse o ajuste (se comparado ao modelo de dois fatores correlacionados). Entretanto, o fator de segunda ordem deixa o modelo mais parcimonioso e permite a estimação de uma habilidade geral, além das habilidades específicas. Assim, o pesquisador tem a informação de três habilidades (uma geral e duas específicas).

Deve-se considerar, ainda, a comparação entre os modelos de segunda ordem e o *bi-factor*, que também apresentaram indicadores de ajuste semelhantes. Mulaik e Quartetti (1997) observaram que os indicadores de ajustes desses tipos de modelos são parecidos na maioria das pesquisas, embora eles apresentem diferenças teóricas substanciais. O modelo *bi-factor*, por exemplo, impõe uma restrição teórica relacionada à ausência de correlações entre os fatores latentes específicos. Tal restrição, na presente pesquisa, alterou consideravelmente as cargas fatoriais dos itens. No modelo *bi-factor* confirmatório, 17% das cargas fatoriais não apresentaram significância estatística no respectivo fator específico. Ademais, tal restrição impõe ao modelo uma estrutura não-hierárquica, semelhante ao modelo teórico proposto por Spearman (Almeida, 1994; Brody, 2000; Spearman, 1904). Por outro lado, o modelo de segunda ordem oferece suporte interpretativo da moderna teoria de CHC (McGrew & Flanagan, 1998; McGrew

& Wendling, 2010). Nesse sentido, as dimensões RA, RE e geral são organizadas em dois níveis hierárquicos, nos quais a dimensão geral organiza e sustenta as dimensões específicas. Teoricamente, a dimensão geral estaria mais associada à inteligência fluida (*Gf* do segundo nível hierárquico), pois as soluções dos itens do TRAE não exigem conhecimento prévio, mas estão associadas à capacidade de raciocínio indutivo e dedutivo. O subteste de RE também pode associar-se à dimensão de Processamento Visual (*Gv*) do modelo CHC por envolver tarefas nas quais o participante deve visualizar, reter, recuperar e transformar mentalmente as figuras apresentadas (McGrew, 2005).

De maneira geral, a presente pesquisa oferece suporte às evidências de validade da estrutura interna, também denominada como validade de construto, do Teste de Raciocínio Abstrato e Espacial (TRAE). Os parâmetros de dificuldade e discriminação da TCT e TRI também indicam que o instrumento é adequado ao uso pelos profissionais e, principalmente, pelos pesquisadores da área. Além disso, em consonância com a teoria CHC de inteligência, o TRAE parece apresentar uma estrutura robusta que permite a estimação de fatores de primeira ordem (RA e RE), bem como um fator geral de segunda ordem associado à inteligência fluida. Espera-se que o instrumento construído possa ser útil como ferramenta auxiliar no processo de avaliação psicológica e principalmente para o uso em pesquisas cujo tempo de aplicação é pequeno.

Referências

- Almeida, L. S. (1994). *Inteligência: Definição e medida*. Aveiro, PT: CIDInE.
- Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.
- Berg, C. A. (2000). Intellectual development in adulthood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 117-137). Cambridge, UK: Cambridge University Press.
- Brody, N. (2000). History of theories and measurements of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 16-33). New York, NY: Cambridge University Press.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment* (2nd ed., pp. 69-76). New York, NY: The Guilford Press.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40, 153-193
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1-22
- Chen, Z., & Siegler, R. S. (2000). Intellectual development in childhood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 92-116). Cambridge, UK: Cambridge University Press.
- Cronbach, L. J. (1996). *Fundamentos da testagem psicológica* (5 ed.). Porto Alegre, RS: Artes Médicas.
- Davidson, J. E., & Downing, C. L. (2000). Contemporary models of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 34-52). New York, NY: Cambridge University Press.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Norwell, MA: Kluwer.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. London, UK: Sage.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro, RJ: LTC.

- Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 443-451). New York, NY: Macmillan.
- Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (2nd ed., pp. 41-68). New York, NY: The Guilford Press.
- Karino, C. A., Laros, J. A., & Jesus, G. R. (2011). Evidências de validade convergente do SON-R 2 ½ - 7 [a] com o WPPSI-III e WISC-III. *Psicologia: Reflexão e Crítica*, 24, 621-629.
- Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment* (2 ed., pp. 581-614). New York, NY: The Guilford Press.
- Kemp, L. C. D. (1955). Environmental and other characteristics determining attainments in primary schools. *British Journal of Educational Psychology*, 25, 66-77.
- 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.
- Laros, J. A., Tellegen, P. J., Jesus, G. R., & Karino, C. A. (no prelo). *SON-R 2½-7[a] Teste não-verbal de inteligência: Manual com normatização e validação brasileira*. São Paulo, SP: Casa do Psicólogo.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment* (2nd ed., pp. 136-182). New York, NY: The Guilford Press.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1-10.
- McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR) - Gc-Gf cross battery assessment*. Boston, MA: Allyn and Bacon.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, 47, 651-675.

- Mulaik, S. A., & Quartetti, D. A. (1997). First-order or higher-order general factor? *Structural Equation Modeling, 4*, 193-211.
- Muñiz, J., Prieto, G., Almeida, L., & Bartram, D. (1999). Test use in Spain, Portugal and Latin American Countries. *European Journal of Psychological Assessment, 15*, 151-157.
- Nakano, T. (2013). Problemas apresentados pelos instrumentos com parecer desfavorável no SATEPSI. *Avaliação Psicológica, 12*, 121-130.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77 - 101.
- Noronha, A. P. P., & Alchieri, J. C. (2002). Reflexões sobre os instrumentos de avaliação psicológica. In R. Primi (Ed.), *Temas em avaliação psicológica* (pp. 7-16). Campinas, SP: Imprensa Digital do Brasil Gráfica e Editora.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3 ed.). New York, NY: McGraw-Hill.
- Pasquali, L. (2003). *Psicometria: Teoria dos testes na psicologia e na educação*. Petrópolis, RJ: Editora Vozes.
- Primi, R. (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica, 1*, 67-77.
- Primi, R., & Almeida, L. (1998). *BPR-5: Manual Técnico*. São Paulo, SP: Casa do Psicólogo.
- Primi, R., & Almeida, L. (2000). Estudo de validação da Bateria de Provas de Raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa, 16*, 165-173.
- Primi, R., Silva, M. C. R., Santana, P. R., Muniz, M., & Almeida, L. S. (2013). The use of the bi-factor model to test the uni-dimensionality of a battery of reasoning tests. *Psicothema, 25*, 115-122.
- Reise, S. P., Morizot, J., & Hayes, R. D. (2007). The use of the bi-factor model to test the uni-dimensionality of a battery of reasoning tests. *Quality Life Research, 16*, 19-32.
- Snijders, J. T., Tellegen, P. J., & Laros, J. A. (1989). *SON-R 5½-17: Manual and research report*. Groningen, ND: Wolters-Noordhoff.
- Sparta, M., Bardagi, M. P., & Teixeira, M. A. P. (2006). Modelos e instrumentos de avaliação em orientação profissional: perspectiva histórica e situação no Brasil. *Revista Brasileira de Orientação Profissional, 7*, 19-32.

- Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis[*]: Understanding Concepts and Applications**. Washington, DC: American Psychologist Association.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.
- Valentini, F., & Laros, J. A. (2011). Teoria de Resposta ao Item na Avaliação Psicológica. In R. A. M. Ambiel, I. S. Rabelo, S. V. Pacanaro, G. A. S. Alves & I. F. A. S. Leme (Eds.), *Avaliação psicológica: guia de consulta para estudantes e profissionais de psicologia*. São Paulo, SP: Casa do Psicólogo.
- Valentini, F., Laros, J. A., & Jesus, G. R. (2010). *Invariância da estrutura fatorial do teste de inteligência SON-R 2½ -7[a] por meio da análise fatorial confirmatória*. Paper presented at the IX Encontro Mineiro de Avaliação Psicológica, Belo Horizonte, MG.

Manuscrito 3

Evidências de validade do Teste de Raciocínio Abstrato e Espacial com base na abordagem multitraço-multimétodo

Validity evidence of the Abstract and Spatial Reasoning Test based on the multitrait-multimethod approach

Relato Breve de Pesquisa

Resumo

O objetivo deste estudo foi obter indicações da fidedignidade e da validade convergente e discriminante do teste de Raciocínio Abstrato e Espacial (TRAE). O TRAE e a BPR-5 foram aplicados em 149 estudantes do ensino médio. Os escores da escala geral do TRAE apresentaram coeficiente de fidedignidade adequado (0,76), entretanto os escores dos quatro subtestes foram menos precisos. Por meio da abordagem multitraço-multimétodo, no contexto da modelagem por equações estruturais, 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. No entanto, um modelo alternativo de correlação perfeita entre RA e RE também é plausível, o que indica restrições quanto à validade discriminante. Portanto, os resultados respaldam a precisão e validade convergente do TRAE, principalmente para a escala geral. Todavia, sugere-se que os subtestes sejam interpretados com cautela.

Palavras-chave: raciocínio; validade; multitraço-multimétodo.

Abstract

In the present study evidence was obtained of the reliability and of the convergent and discriminant validity of the Abstract and Spatial Reasoning Test (TRAE, in Portuguese). The TRAE and the BPR-5 were both administered to 149 high school students. 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 ability; validity; multitrait-multimethod.

O estudo da inteligência é um dos empreendimentos mais bem sucedidos da psicologia moderna. Por meio das pesquisas sobre o tema, foi possível reunir evidências sobre o funcionamento de diversos aspectos genéticos e ambientais que embasam a inteligência. Neste sentido, a construção de instrumentos para avaliação cognitiva é de suma importância para a prática e para a pesquisa psicológica.

Além disso, o desenvolvimento de teorias sobre a inteligência é premissa da pesquisa e da prática na área. Neste contexto, destaca-se o modelo CHC (Cattell-Horn-Carroll), teorizado por McGrew e Flanagan (McGrew, 2005; McGrew & Flanagan, 1998). 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, sendo essa dimensão equivalente ao fator *g* de Spearman; o segundo nível envolve dez dimensões, nas quais estão inclusas a inteligência fluída e a inteligência cristalizada (*Gf-Gc*); e o primeiro nível envolve aproximadamente 70 habilidades específicas.

Dentre as habilidades específicas, destacam-se, na literatura científica, os raciocínios abstrato e espacial. 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. Ambos os tipos de raciocínio estão mais associados à inteligência fluída (McGrew, 2005; Primi & Almeida, 1998; Tellegen & Laros, 2011).

Para ampliar a compreensão teórica da inteligência, bem como subsidiar a prática psicológica é necessário construir instrumentos adequados à população brasileira. Este estudo insere-se nesse contexto e busca, especificamente, avaliar os

indicadores de fidedignidade do Teste de Raciocínio Abstrato e Espacial (TRAE), bem como as evidências de validade convergente e discriminante desse instrumento.

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. Tais participantes tinham idades entre 14 e 19 anos ($M = 16,98$; $DP = 0,87$). Adotou-se o procedimento de amostragem não probabilística.

Instrumentos

Os estudantes responderam ao Teste de Raciocínio Abstrato e Espacial – TRAE e à Bateria de Provas de Raciocínio 5 – BPR-5 (Primi & Almeida, 1998). 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 (Anexos II e III da tese). O segundo instrumento, a BPR-5, é utilizado para avaliar a habilidade cognitiva geral e as habilidades cognitivas em cinco áreas específicas: raciocínio abstrato, raciocínio verbal, raciocínio espacial, raciocínio numérico e raciocínio mecânico. Cada subteste contém entre 20 e 25 questões e 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 e o terceiro ano do ensino médio.

Procedimentos

Os instrumentos foram aplicados em sessões coletivas, realizadas nas salas de aula. Os estudantes foram voluntários e assentiram a participação. Além disso, foi enviado aos pais um termo de informação, conforme orientação do comitê de ética. Os dados de precisão foram analisados por meio do indicador Lambda 2 de Guttman; e as evidências da validade convergente e discriminante foram estudadas por meio correlações de Pearson e modelagem multitraço-multimétodo (Byrne, 2010; Campbell & Fiske, 1959; Widaman, 1985).

Resultados

No intuito de responder ao objetivo específico referente à consistência interna do instrumento, calculou-se o coeficiente lambda 2 de Guttman (λ_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 *
TRAE	RA	4	15	11,91	2,31	0,67
	RE	3	15	8,91	2,63	0,67
	Geral	8	30	20,83	4,20	0,76
BPR-5 (amostra deste estudo)	RV	7	23	15,97	3,09	0,61
	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
BPR-5 (manual)	RV	0	23	15,66	3,81	0,82
	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

Notas. * Para a amostra do presente estudo, a precisão do TRAE e da BPR-5 foi estimada pelo coeficiente lambda 2 de Guttman (λ_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, primeiramente, 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, por meio da seguinte equação (Hogan, 2006):

$$r \text{ corrigida} = r_{xy} \div \sqrt{r_{xx} \cdot r_{yy}}$$

Na qual, r_{xy} = correlação bruta sem correção; r_{xx} = fidedignidade do teste X; e r_{yy} = fidedignidade do teste Y.

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)

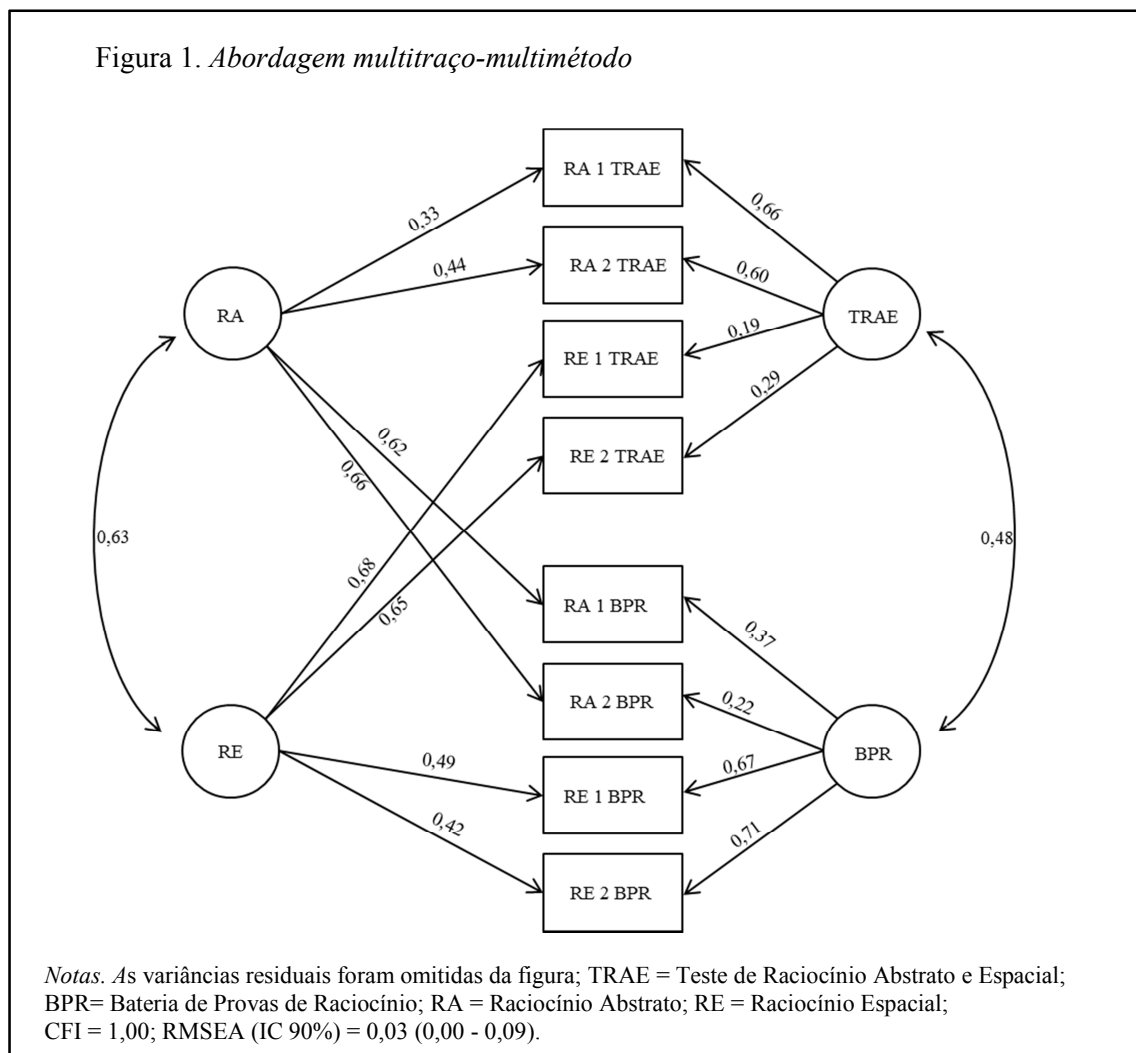
Notas. n = 149; Na tabela, são apresentados os coeficientes das correlações de Pearson; entre parênteses são apresentadas as correlações atenuadas para a ausência de fidedignidade perfeita.

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.

Além das análises de correlação, buscou-se reunir evidências de validade convergente e discriminante do TRAE 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 (dentro 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. 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).

No que se refere à MTMM, utilizou-se a adaptação do método para a modelagem por equações estruturais (MEE) 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 MTMT versa a comparação de modelos aninhados: o primeiro modelo é utilizado como base de comparação e consiste na livre estimação das correlações entre os fatores latentes e os fatores dos métodos; no segundo modelo, os fatores latentes são excluídos e os fatores dos métodos são estimados livremente; para o terceiro modelo, as correlações entre os fatores latentes são fixadas em 1 (correlações perfeitas); no quarto modelo, as correlações entre os fatores latentes voltam a ser livremente estimadas, mas as correlações entre os fatores dos métodos são fixadas em 0 (correlações nulas). Os modelos são, então, comparados por meio dos indicadores de ajuste χ^2 , CFI, RMSEA e AIC. Neste sentido, espera-se que o modelo 1 se ajuste melhor aos dados do que os modelos 2 e 3. Espera-se, também, que o modelo 4 apresente indicadores de ajuste semelhantes ao do modelo 1 (Byrne, 2010). Esses resultados são apresentados na Tabela 3.

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

Modelo	χ^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	
			Diferença (Δ)		
Comparação dos Modelos			χ^2 (gl)	CFI	AIC

Teste da Validade Convergente					
	Modelo 1 <i>versus</i> Modelo 2		46,45 (9)	0,113	28,454
Teste da Validade Discriminante					
	Modelo 1 <i>versus</i> Modelo 3		3,857 (1)	0,009	1,857
	Modelo 1 <i>versus</i> Modelo 4		4,418 (1)	0,011	2,418

Notas. Modelo 1 = livre estimação das correções 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.

Os resultados apresentados na Tabela 3 indicam que o modelo 1 se ajusta significativamente melhor aos dados do que o modelo 2. Ademais, a diferença de ajuste entre os modelos 1 e 4 é pequena, conforme o esperado. No entanto as diferenças dos indicadores de ajuste entre os modelos 1 e 3 são pequenas, embora estatisticamente significativas ($p < 0,05$).

Discussão

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 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, principalmente para a escala geral, cuja correlação atenuada entre os instrumentos foi de 0,73. Os resultados da modelagem multitraço-multimétodo (MTMM) também oferecem suporte à validade convergente dos instrumentos. O modelo apresentou significativa diminuição dos indicadores de ajuste frente à remoção dos fatores latentes 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 esses 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 (Byrne, 2010; Kline, 2011).

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, conforme o esperado, RA e RE 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 et al. 2010). 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, esperava-se que, ao estimar uma correlação perfeita entre RA e RE, o modelo 3 apresentasse significativa redução de ajuste em comparação ao modelo 1. No entanto, a semelhança dos valores de χ^2 , CFI e AIC indica que a colinearidade de RA e RE é bastante plausível. Em outras palavras, RA e RE não parecem apresentar diferenças muito relevantes. Embora não esperados, esses resultados são condizentes com outros estudos. Tellegen e Laros (2011) tampouco evidenciaram clara distinção entre esses dois tipos de raciocínio nos estudos do teste SON-R 6-40. Almeida e Primi (1998, 2000) também encontraram correlações moderadas entre RA e RE ($r > 0,45$) nas pesquisas da BPR-5. Portanto, a distinção entre RA e RE tampouco é clara na literatura científica, 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. Esse 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).

Ressalta-se, como limitação da pesquisa, que a amostra é relativamente pequena e selecionada por conveniência, o que pode gerar instabilidade nos modelos apresentados. Sugere-se, portanto, a condução de estudos que utilizem amostras mais representativas.

De maneira geral, os resultados discutidos nesta pesquisa, respaldam a precisão e 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 precisão das escalas de RA e RE, sugere-se que a interpretação dos resultados dos subtestes seja feita com cautela.

Referências

- Almeida, L. S., Nascimento, E., Lima, A. O. F., Vasconcelos, A. G., Akama, C. T., Santos, M. T. (2010). Bateria de Provas de Raciocínio (BPR-5): Estudo exploratório em alguns alunos universitários. *Avaliação Psicológica*, 9, 155-162.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.
- Campbell, D. T., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro, RJ: LTC.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment* (2nd ed., pp. 136-182). New York, NY: The Guilford Press.
- McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR) - Gc-Gf cross battery assessment*. Boston, MA: Allyn and Bacon.
- Primi, R., & Almeida, L. (2000). Estudo de validação da Bateria de Provas de Raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa*, 16, 165-173.
- Primi, R., & Almeida, L. (1998). *BPR-5: Manual Técnico*. São Paulo, SP: Casa do Psicólogo.
- Tellegen, P. J., & Laros, J. A. (2011). *SON-R 6-40: Niet-verbale Intelligentie Test. Verantwoording, Instructies & Normen [SON-R 6-40: Nonverbal Intelligence Test. Research report, Instructions & Norms]*. Amsterdam, ND: Hogrefe.
- Widaman, F. K. (1985). Hierarchically tested covariance structure models for multitrait-multimethod data. *Applied Psychology Measurement*, 9, 1-26.

Manuscrito 4

Análise multinível das relações entre inteligência fluida,
desempenho acadêmico e aprendizagem

The relationship among fluid intelligence, academic achievement,
and learning using multilevel analysis

Sugestão de título abreviado: Desempenho acadêmico e inteligência

Resumo

O objetivo deste estudo foi construir e testar modelos multinível do desempenho acadêmico e da aprendizagem, considerando as variáveis explicativas inteligência fluida (*Gf*), infraestrutura na escola e nível socioeconômico agregado para a escola (NSEA). O primeiro estudo avaliou 1.295 estudantes com idades entre 11 e 17 anos. No segundo, a aprendizagem desses estudantes foi avaliada pela comparação do desempenho acadêmico em dois anos consecutivos. Os resultados apontaram para relação moderada entre a *Gf* e o desempenho acadêmico, bem como entre a *Gf* e a aprendizagem. A infraestrutura na escola influenciou ambas as relações; o NSEA, apenas a relação entre a *Gf* e o desempenho acadêmico. Nesse sentido, uma infraestrutura melhor e um NSEA mais alto diminuíram o efeito da *Gf* sobre o desempenho e a aprendizagem. Isso sugere que a infraestrutura e o NSEA reduzem as diferenças do desempenho e da aprendizagem entre os estudantes com baixos e altos escores de inteligência.

Palavras-Chave: inteligência fluida; desempenho acadêmico; aprendizagem; infraestrutura na escola; nível socioeconômico.

Abstract

The present study aims to develop and test multilevel models explaining academic achievement and learning considering fluid intelligence (*Gf*) of the students, infrastructure and socioeconomic status of the school (SESS) as predictor variables. The first study examined 1,295 students varying in age from 11 to 17 years. In the second study the learning process of these students was assessed by comparing their school achievements of two consecutive years. The results indicated a moderate relationship between *Gf* and academic achievement, and between *Gf* and learning as well. Infrastructure influenced both relationships, while SESS influenced only the relationship between *Gf* and achievement, in the sense that a better infrastructure and a higher SESS had a diminishing effect on the relationships. This effect suggests that school infrastructure and SESS decreases existing differences in academic achievement and learning among the students with low and high intelligence scores.

Keywords: fluid intelligence; academic achievement; learning; school infrastructure; socioeconomic status.

O estudo da inteligência é um dos construtos psicológicos mais estudados ao longo da história da psicologia moderna. Por meio da inteligência é possível prever diversos construtos, tais como desempenho ocupacional, econômico e social, entre outros (Neisser et al., 1996). Além disso, a inteligência prediz, principalmente, uma parte expressiva da variabilidade no desempenho acadêmico dos estudantes (Colom & Flores-Mendoza, 2007; Hattie, 2009; McGrew & Wendling, 2010; Primi, Ferrão, & Almeida, 2010; Rohde & Thompson, 2007).

Nas décadas de pesquisa sobre a inteligência foram propostos diversos modelos teóricos. Um dos mais importantes foi proposto por Cattell (1943, 1963), que organizou a inteligência em duas dimensões gerais: inteligência fluída (*Gf*) e cristalizada (*Gc*). A dimensão *Gf* refere-se às habilidades de raciocínio e à capacidade para resolução de problemas novos, para os quais a pessoa tem pouco conhecimento prévio. Tais habilidades envolvem o processamento de informações de maneira lógica, a classificação e a identificação das relações entre conceitos concretos e abstratos, a elaboração e teste de hipóteses (raciocínio hipotético-dedutivo), entre outros (Horn, 1994; McGrew, 2005, 2009; Primi et al., 2010). A dimensão *Gc* diz respeito à aquisição e à solidificação de conhecimentos formais e informais, apreendidos por transmissão cultural ou por meio da escola (Cattell, 1943, 1963).

Na década de 1990, as dimensões *Gf* e *Gc* foram abarcadas pela moderna teoria CHC (Cattell-Horn-Carroll). A teoria CHC organizou as dimensões da inteligência em uma estrutura hierárquica de três níveis, que segue uma ordem de especialização, variando do mais geral - nível III - até as dimensões específicas - nível I. Nesse contexto, as dimensões *Gf* e *Gc* são compreendidas, por um lado, como fatores intermediários (nível II) que organizam diversas habilidades específicas e, por outro lado, são organizadas por um grande fator geral (nível III). Ressalta-se, ainda, que para

o modelo CHC, a dimensão *Gf* é a que mais se aproxima do fator geral (Carroll, 2005; McGrew, 2005).

Entre as habilidades cognitivas, a inteligência fluida e o fator geral são as dimensões mais estudadas na relação com o desempenho acadêmico e a aprendizagem. Diversos estudos apontam para as evidências que os estudantes com maiores níveis de inteligência também apresentam melhor desempenho na escola (Colom & Flores-Mendoza, 2007; Floyd, Evans, & McGrew, 2003; Kemp, 1955; Primi et al., 2010). De maneira geral, as pesquisas apontam para valores de *r* e *Beta* entre 0,30 e 0,70 (Neisser et al., 1996; Valentini & Laros, 2013).

A literatura científica sobre o tema é ampla, o que permitiu o estudo de meta-análise de McGrew e Wendling (2010) sobre a relação entre o modelo CHC e o desempenho dos estudantes. Os autores reuniram 19 pesquisas e concluíram que o fator *g* (geral) e a *Gf* relacionaram-se positivamente ao desempenho acadêmico em todos os estudos ($p < 0,05$). No Brasil, Colom e Flores-Mendonza (2007) investigaram os dados de 640 estudantes e observaram um coeficiente de regressão moderado ($Beta = 0,37$) entre a inteligência e as notas escolares dos estudantes.

Existem evidências que a inteligência também influencia o nível de aprendizagem dos estudantes, além do desempenho (Geary, 2011; Hambrick, Pink, Meinz, Pettibone, & Oswald, 2008; Primi et al., 2010; Swanson, Jerman, & Zheng, 2008). Primi et al. (2010) associaram os escores de inteligência fluida a quatro avaliações longitudinais do desempenho em matemática. Os pesquisadores encontraram relações entre o desempenho em matemática e as dimensões da inteligência fluida, raciocínio abstrato, numérico, verbal e espacial ($r \geq 0,27$, para a dimensão geral). Ademais, o maior crescimento da proficiência em matemática foi observado nos estudantes com os maiores escores de inteligência. Ou seja, os estudantes com altos

escores de inteligência parecem adquirir mais conhecimento em matemática. Resultados semelhantes foram encontrados por Geary (2011), que investigou uma amostra de estudantes do ensino fundamental. As conclusões do estudo apontaram para a importância das variáveis inteligência e velocidade de processamento na compreensão da evolução da proficiência em matemática após cinco anos da realização da primeira avaliação. Ressalta-se ainda que a velocidade do processamento também é considerada uma dimensão específica da inteligência, segundo o modelo CHC. No entanto, controvérsias foram encontradas na literatura. Zhang, Davis, Salthouse e Tucker-Drob (2007) realizaram um estudo transversal com 2.453 pessoas separadas em cinco grupos de idade. Os dados foram avaliados por meio da curva de crescimento e indicaram uma relação significativa entre a inteligência fluida e o desempenho, mas não entre a inteligência e a aprendizagem.

A literatura científica também apresenta estudos que se referem à influência dos aspectos contextuais sobre a inteligência e sobre o desempenho acadêmico (Valentini & Laros, 2013). No que se refere ao nível de análise da escola, o nível socioeconômico da família e do aluno (NSE) tem sido agregado para a escola (NSEA) e associado fortemente ao desempenho acadêmico (Caldas & Bankston-III, 1997). Laros, Marciano e Andrade (2010), utilizando a técnica de Análise Multinível, evidenciaram que o NSEA explicou aproximadamente 70% da variância do desempenho em matemática (do nível da escola). Os mesmos autores verificaram que o desempenho em língua portuguesa também pode ser explicado, em aproximadamente 70%, pelas variáveis NSEA, escolaridade da mãe e etnia do aluno (Laros, Marciano, & Andrade, 2012). Convém ressaltar que, nesses três estudos, os autores não consideraram a variável inteligência na predição do desempenho acadêmico.

Entretanto, no que se refere ao nível de análise do estudante, a literatura parece indicar para relações fracas ou, no máximo, moderadas entre NSE e desempenho, após o controle da variável inteligência, conforme ressaltado por Valentini e Laros (2013). O estudo de Kemp (1955), um dos primeiros sobre o tema, indicou que, após o controle da variável inteligência, a correlação entre NSE e desempenho acadêmico sofreu redução de 0,56 para 0,30. Esse resultado indica que uma parte significativa da relação entre o NSE e o desempenho pode ser explicada pela covariância com a variável inteligência. Mais recentemente, a pesquisa de Colom e Flores-Mendonza (2007) tampouco encontrou relação significativa entre o NSE e o desempenho acadêmico. Para esta relação, após a inserção da variável inteligência, o valor de *Beta* foi igual a 0,04. Além disso, a inteligência se mostrou como o único preditor das notas escolares ($Beta = 0,37$). Resultados semelhantes foram encontrados por Lloyd e Barenblatt (1984), que investigaram as relações entre o desempenho, a inteligência, o NSE e a motivação intelectual. As conclusões apontaram para um efeito preditivo muito fraco do NSE sobre o desempenho, mediante o controle da variável inteligência.

Ressalta-se ainda a relação entre o NSE e a inteligência. O estudo já citado de Kemp (1955), por exemplo, encontrou correlação igual a 0,52 entre a inteligência e o NSE. Relações parecidas foram encontradas num estudo recente de Laros, Tellegen, Jesus e Karino (no prelo). Os autores evidenciaram correlações moderadas e fortes entre os escores dos testes de inteligência e o NSE (r entre 0,40 e 0,67). As correlações mais altas referiam-se aos testes verbais de inteligência. Em outras palavras, o desempenho nas tarefas que envolviam a manipulação de conhecimentos gerais já adquiridos, principalmente aqueles referentes à linguagem, esteve mais fortemente associado ao NSE do que o desempenho em tarefas que envolviam mais a inteligência fluida.

Em suma, conforme destacado por Valentini e Laros (2013), a literatura científica parece indicar relações fortes para os estudos que consideraram apenas as variáveis inteligência e NSE (Laros et al., no prelo). Relações fortes também foram encontradas entre o NSE e o desempenho acadêmico pelas pesquisas que não consideraram a variável inteligência (Laros et al., 2010). No entanto, ao controlar o efeito da inteligência, o impacto do NSE sobre o desempenho diminui consideravelmente ou torna-se não significativo (Colom & Flores-Mendoza, 2007; Lloyd & Barenblatt, 1984). Essas discussões parecem indicar que se o NSE for, de fato, relevante para a compreensão do desempenho acadêmico, o impacto dele é mais complexo. Nesse sentido, ainda que a literatura apresente diversos estudos sobre o tema, não foi encontrada nenhuma pesquisa multinível que avaliasse o efeito de moderação do NSE agregado para o nível da escola (NSEA) sobre a relação entre inteligência e desempenho acadêmico, ou sobre o aumento do desempenho de um ano para outro. Além do NSEA, a literatura carece de estudos que avaliem a relevância da infraestrutura na escola na moderação da relação entre a inteligência e o desempenho acadêmico.

O presente estudo teve como objetivo geral a construção de modelos explicativos do desempenho acadêmico e do nível de aprendizagem do estudante a partir de variáveis dos níveis de análise do aluno e da escola. Definiu-se nível de aprendizagem do aluno por meio do aumento da proficiência em língua portuguesa e matemática entre duas medidas repetidas. Especificamente, buscou-se: (a) testar um modelo simples explicativo do desempenho acadêmico em língua portuguesa, matemática e ciências naturais a partir da inteligência fluida (*Gf*) do estudante; (b) testar um modelo multinível explicativo do desempenho acadêmico a partir da *Gf*, infraestrutura na escola, nível socioeconômico agregado (NSEA) e outras variáveis dos níveis de análise do estudante e da escola; (c) testar um modelo multinível explicativo

do nível de aprendizagem a partir da *Gf* do estudante e da infraestrutura na escola. As principais hipóteses do estudo envolvem (1) relação forte entre a *Gf* e o desempenho acadêmico, (2) efeito de moderação da infraestrutura na escola sobre a relação entre a *Gf* e o desempenho, (3) relação moderada entre a *Gf* e o nível de aprendizagem e (4) efeito de moderação da infraestrutura na escola e do NSEA sobre a relação entre a *Gf* e o nível de aprendizagem.

Método

Participantes

A amostra do presente estudo foi composta de 1.295 estudantes do ensino fundamental. Na análise exploratória dos dados, observou-se que uma parte das respostas dos participantes foi muito semelhante ao padrão de acertos ao acaso (“chute”). Acerto ao acaso acontece quando uma pessoa com uma habilidade muito baixa marca a resposta certa devido à sorte. Os participantes com esse tipo de padrão de respostas foram excluídos das análises, considerando os seguintes critérios: (a) quando o número de acertos do participante foi próximo ao esperado ao acaso, incluindo uma margem de tolerância de dois acertos (acima do esperado ao acaso); (b) quando o padrão de respostas incluía erro nas questões fáceis e acerto nas questões difíceis. Além disso, foram excluídos da amostra inicial os participantes que ofereceram a mesma alternativa de resposta para mais de 50% das questões. Assim, a amostra final do estudo de delineamento relacional foi composta de 1.069 participantes com idade entre 11 e 17 anos ($M = 13,3$; $DP = 0,70$).

Para o estudo de delineamento de medidas repetidas os estudantes que não responderam às provas em algum dos dois momentos da coleta foram excluídos das análises. Assim, os dados de 943 participantes foram considerados no estudo de

delineamento de medidas repetidas. Os dados foram coletados em 42 escolas (aproximadamente uma turma de 30 estudantes por escola) de uma rede pública de ensino de uma das maiores cidades do Brasil.

Instrumentos

Para avaliação da inteligência fluida, os participantes responderam ao Teste de Raciocínio Abstrato e Espacial – TRAE (Valentini et al., 2013). O instrumento é dividido em dois subtestes. O primeiro, Raciocínio Abstrato (RA), busca avaliar as habilidades associadas à resolução de problemas cognitivos, para os quais a pessoa tem pouco ou nenhum conhecimento prévio (Anexo II da tese). Nas questões de RA o sujeito precisa descobrir o princípio da transformação de uma figura A para uma figura B e aplicá-lo na transformação da figura C para a figura D. O subteste de RA é composto por 12 itens de múltipla escolha, nos quais são oferecidas quatro alternativas de resposta, e o tempo limite de aplicação é de 13 minutos. O segundo subteste, Raciocínio Espacial (RE), é destinado à avaliação das habilidades visuoespaciais, associadas à sensibilidade para detalhes visuais e à orientação no espaço tridimensional (Anexo III da tese). Para os itens de RE o sujeito deve observar as rotações que ocorrem em três figuras geométricas dispostas em sequência e descobrir como ficaria a próxima figura observando a mesma rotação das anteriores. O subteste é composto por 12 itens de múltipla escolha, nos quais são oferecidas cinco alternativas de resposta. O tempo limite para a sua aplicação é de 17 minutos.

Para avaliação do desempenho acadêmico dos alunos, foram analisados os bancos de dados da avaliação educacional de uma das maiores cidades do Brasil. Tais bancos de dados também foram utilizados para analisar as variáveis contextuais retiradas dos questionários de alunos, professores e diretores. Vale esclarecer que a

secretaria de educação dessa cidade aplica anualmente uma prova de língua portuguesa (20 itens) e uma prova de matemática (20 itens). No ano de 2012 também foi aplicada uma prova de conhecimento em ciências naturais (40 itens). Na realidade, os estudantes respondem a diferentes cadernos de provas, que podem conter diferentes questões. No entanto, os parâmetros dos itens e as proficiências dos estudantes são equalizados, por meio da Teoria de Resposta ao Item (TRI), em uma única escala. Essa escala também é comum para as diferentes séries escolares e ano de aplicação, o que torna factível a comparação, por exemplo, entre o desempenho médio de alunos da quarta série que realizaram a prova em 2011 e o desempenho médio de alunos da oitava série que realizaram a prova em 2012. As provas de língua portuguesa e de matemática foram equalizadas com a escala do Sistema de Avaliação da Educação Básica (Saeb) do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP/MEC), o que permite, adicionalmente, a comparação dos estudantes da presente pesquisa com os demais estudantes do país.

Buscou-se ainda estimar os escores de infraestrutura nas escolas por meio da escala e dos parâmetros disponibilizados pelo estudo de Neto, Jesus, Karino e Andrade (2013). Essa escala é composta por 25 questões dicotômicas que avaliam a presença de itens relacionados à estrutura da escola, tais como abastecimento de água e sala de professores, bem como itens que se referem à presença de equipamentos, tais como computadores, copiadora e internet. Os itens relacionados à educação infantil (por exemplo, berçário) não foram utilizados na presente pesquisa.

Procedimentos

Os dados concernentes ao TRAE foram coletados em sessões coletivas de aproximadamente 30 minutos. Os participantes foram voluntários e assentiram a

participação. Além disso, foi enviado aos pais um termo de informação, conforme orientação do comitê de ética.

Os escores do TRAE foram gerados por meio do Modelo de TRI de Três Parâmetros Logístico (Lord, 1974, 1980). O nível socioeconômico (NSE) dos estudantes foi estimado por meio do Modelo de Resposta Gradual (Samejima, 1974). O NSE agregado (NSEA) foi calculado a partir da média do NSE dos estudantes da escola.

Os resultados do estudo relacional foram analisados por meio da Modelagem por Equações Estruturais Multinível (Heck & Thomas, 2009), no qual se buscou analisar as variáveis do aluno no nível intragrupo e as variáveis da escola no nível entregupo. O delineamento de medidas repetidas foi analisado por meio da Modelagem Multinível Longitudinal, considerando os seguintes níveis de análise: o primeiro nível (intrasujeito) consiste nas medidas repetidas do desempenho dos estudantes; o segundo nível diz respeito às variáveis do estudante; o terceiro nível refere-se às variáveis da escola. As análises foram realizadas com o auxílio dos *softwares* Bilog-Mg 3.0, Multilog, Mplus 6.11 e MLwiN 2.22. Os coeficientes padronizados dos modelos multinível foram calculados por meio da equação (Hox, 2010):

$$\text{coeficiente padronizado} = \text{coeficiente não padronizado} \times \frac{DP_{VI}}{DP_{VD}}$$

No qual, DP_{VI} refere-se ao desvio-padrão da variável independente, e DP_{VD} refere-se ao desvio-padrão da variável dependente.

Resultados

Delineamento Relacional

As análises exploratórias indicaram que as distribuições dos escores das variáveis dependentes foram semelhantes à curva de distribuição normal. Além disso, o valor de assimetria (*skewness*) foi inferior a 0,50 para as variáveis relacionadas ao

desempenho acadêmico e à inteligência, e o coeficiente de normalidade multivariada de Mardia foi igual a 1,64 (razão crítica = 3,20). O critério para considerar um item com distribuição normal foi assimetria inferior a 1,00 (Miles & Shevlin, 2001) e coeficiente Mardia inferior a 5,00 (Byrne, 2010).

Avaliou-se, primeiramente, um modelo simples que explicasse o desempenho acadêmico por meio da inteligência fluida, considerando apenas o nível de análise do aluno. O modelo, conforme Tabela 1 e Figura 1, contém os escores dos subtestes de RA e RE que são explicados pelo fator latente inteligência fluida (*Gf*); bem como os escores de língua portuguesa e de ciências naturais explicados por um fator de desempenho verbal (de primeira ordem), que, juntamente com os escores de matemática, são explicados por um fator de desempenho geral (de segunda ordem). Os coeficientes não-padronizados estimados por meio da modelagem por equações estruturais são apresentado na Tabela 1.

Tabela 1. *Parâmetros do modelo explicativo do desempenho do aluno por meio da inteligência fluida*

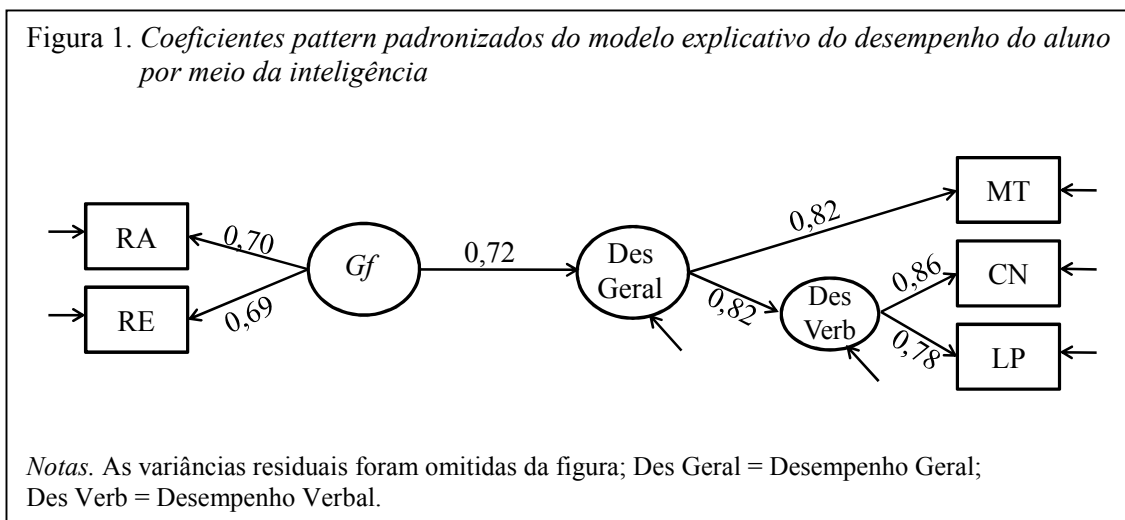
Parâmetros		Coeficientes não-padronizados		
		Efeito (IC 95%)	EP	RC
<i>Fatores</i>				
<i>Gf</i>	→ RA	0,54 (0,51 - 0,58)	0,02	27,2
<i>Gf</i>	→ RE	0,54 (0,51 - 0,58)	0,02	27,2
Desempenho Verbal	→ LP	0,84 (0,76 - 0,92)	0,04	21,0
Desempenho Verbal	→ CN	1,00 (fixo)	-	-
Desempenho Geral	→ MT	1,00 (fixo)	-	-
Desempenho Geral	→ Desempenho Verbal	1,04 (0,90 - 1,18)	0,07	14,9
<i>Regressão</i>				
<i>Gf</i>	→ Desempenho Geral	0,44 (0,38 - 0,50)	0,03	14,7
<i>Ajuste do Modelo</i>				
χ^2	2,67		TLI	1,00
gl	4		CFI	1,00
χ^2 /gl	0,67		RMSEA (IC 90%)	0,00 (0,00 – 0,04)
<i>p</i>	0,61			

Notas. As setas (→) indicam a direção do parâmetro; *Gf* = inteligência fluida; RA = raciocínio abstrato; RE = raciocínio espacial; LP = língua portuguesa; CN = ciências naturais; MT = matemática; EP = erro padrão de estimação; RC = razão crítica; χ^2 = qui-quadrado; gl = graus de liberdade; TLI = *Tucker-Lewis Index*; CFI = *Comparative Fix Index*; RMSEA = *Root Mean Square Error of Approximation*; IC = Intervalo de Confiança.

De maneira geral, o modelo apresentou um ajuste adequado aos dados. O valor *p* do qui-quadrado do modelo foi acima de 0,05, portanto não foi encontrada diferença estatisticamente significativa entre a matriz original dos dados e a matriz estimada pelo modelo. Os valores de TLI, CFI e RMSEA apontaram para um modelo superajustado (ou ajuste sem resíduo).

Por meio da Tabela 1, verifica-se que no intervalo de confiança de 95% dos parâmetros não-padronizados não estão incluídos o valor 0. Além disso, todos os valores de razão crítica foram bem superiores a 1,96, indicando que os parâmetros avaliados foram significativamente diferentes de 0. Salienta-se que os parâmetros apresentados na Tabela 1 não são padronizados e, portanto, não podem ser comparados

entre si. Para compará-los, a Figura 1 apresenta o modelo simples e os coeficientes *pattern* padronizados.



Os coeficientes *pattern* indicados na Figura 1 foram acima de 0,69, o que oferece suporte à estimação das variáveis latentes *Gf* e desempenho acadêmico geral. Ademais, o coeficiente estimado entre essas duas variáveis foi forte ($Beta = 0,72$), portanto, no modelo simples, aproximadamente 50% da variância do desempenho geral pode ser explicada pela inteligência fluida do estudante.

Além do modelo simples, avaliou-se a relação entre variáveis do aluno, incluindo a *Gf*, o desempenho acadêmico e as variáveis da escola, por meio de cinco modelos multinível. No primeiro modelo (o modelo vazio) foram inseridas apenas as três variáveis dependentes, sem nenhuma variável independente, o que proporciona uma estimativa de correlação intraclasse (ICC). Considerando que as três variáveis dependentes foram inseridas concomitantemente no modelo, estimou-se o modelo vazio considerando as covariâncias entre as variáveis dependentes. O modelo 1 é apresentado na Tabela 2.

Tabela 2. *Parâmetros não-padronizados do modelo 1 (modelo vazio)*

Parâmetros	Efeito	EP	RC
<i>Intragrupo (Nível Estudante)</i>			
<i>Variância</i>			
LP	0,58	0,03	21,09
MT	0,47	0,02	19,94
CN	0,68	0,03	21,98
<i>Covariância</i>			
LP ↔ MT	0,23	0,02	11,15
CN ↔ MT	0,29	0,02	15,04
CN ↔ LP	0,39	0,03	15,78
<i>Entregrupo (Nível Escola)</i>			
<i>Intercepto</i>			
LP	-0,20	0,06	
MT	-0,12	0,05	
CN	0,11	0,06	
<i>Variância</i>			
LP	0,10	0,03	3,48
MT	0,08	0,03	2,94
CN	0,13	0,04	3,62
<i>Covariância</i>			
LP ↔ MT	0,09	0,02	3,17
CN ↔ MT	0,09	0,03	2,92
CN ↔ LP	0,11	0,03	3,33
Correlação intraclasse (ICC)	LP=0,15 / MT=0,15 / CN=0,16		

Notas. As setas (↔) indicam a direção do parâmetro; LP = língua portuguesa; CN = ciências naturais; MT = matemática; EP = erro padrão; RC = razão crítica.

Conforme descrito na Tabela 2, as correlações intraclasse (ICC) indicaram que aproximadamente 14, 15 e 16% da variância do desempenho em língua portuguesa, matemática e ciências naturais, respectivamente, podem ser atribuídas ao nível da escola. Estes resultados sugerem que as médias do desempenho acadêmico variaram entre as escolas, justificando o uso de modelos multinível. Os interceptos para o modelo vazio apontaram para as médias de desempenho próximas de 0. Tal resultado era esperado, uma vez que as proficiências foram estimadas por meio da TRI na escala *theta* que assume média igual a 0 e desvio-padrão igual a 1.

Nas modelagens subsequentes, buscou-se testar os modelos multinível estimando os fatores latentes de inteligência e de desempenho acadêmico. Todavia, a presença desses fatores latentes inviabilizava a estimação dos modelos com a randomização dos coeficientes de regressão, ora por não identificar o modelo, ora pela limitação de recursos computacionais. Nesse sentido, optou-se pela retirada desses fatores latentes, tratando os escores de inteligência fluida e de língua portuguesa, de matemática e de ciências naturais como variáveis observadas. Ademais, permitiu-se a covariância entre as variâncias residuais das variáveis dependentes (desempenho acadêmico). O modelo 2 contém as variáveis explicativas do nível do aluno, e o modelo 3, as variáveis do nível do aluno e da escola, ambos são apresentados na Tabela 3. No nível de análise do estudante, inseriu-se uma variável latente relacionada à leitura. Essa variável foi estimada com base em três itens do questionário de alunos, cujos conteúdos dizem respeito à quantidade de livros que o estudante leu por iniciativa própria ou por indicação da escola, bem como se referem à quantidade de livros disponíveis em sua casa.

Tabela 3. Parâmetros não-padronizados do modelo 2 e do modelo 3

Parâmetros			Modelo 2			Modelo 3			
			Efeito	EP	RC	Efeito	EP	RC	
<i>Intragruppo (Nível Estudante)</i>									
Fator	Leitura	→	Item1	1,00	-	-	1,00	-	-
	Leitura	→	Item 2	0,23	0,10	2,79	0,29	0,10	2,78
	Leitura	→	Item 3	0,18	0,08	2,42	0,19	0,08	2,41
	Fator Leitura	→	LP	0,10	0,02	4,15	0,10	0,02	4,16
	Fator Leitura	→	CN	0,08	0,03	3,35	0,08	0,03	3,35
	<i>Gf</i>	→	LP	0,30	0,03	9,27	0,30	0,03	9,28
	<i>Gf</i>	→	MT	0,36	0,03	11,08	0,35	0,03	11,40
	<i>Gf</i>	→	CN	0,38	0,03	13,72	0,37	0,03	13,76
	Variância e Residual	Fator Leitura			1,34	0,47	2,84	1,32	0,47
LP				0,51	0,02	22,53	0,51	0,02	22,54
MT				0,39	0,02	17,13	0,39	0,02	17,14
CN				0,58	0,03	19,08	0,58	0,03	19,05
Covariância Residual	LP	↔	MT	0,16	0,02	10,31	0,16	0,02	10,11
	CN	↔	MT	0,20	0,02	11,66	0,20	0,02	11,65
	CN	↔	LP	0,31	0,03	12,10	0,31	0,03	12,09
<hr/>									
<i>Entreguppo (Nível Escola)</i>									
Expectativa do professor	→	LP		-	-	-	0,48	0,15	3,17
Expectativa do professor	→	MT		-	-	-	0,47	0,10	4,54
Expectativa do professor	→	CN		-	-	-	0,60	0,17	3,51
Interação (INF x NSEA)	→	MT		-	-	-	0,23	0,10	2,39
Médias e Interceptos	LP			-0,19	0,04		-0,18	0,04	
	MT			-0,10	0,04		-0,10	0,03	
	CN			0,13	0,05		0,14	0,04	
Variância e Residual	LP			0,05	0,02	3,57	0,04	0,01	3,88
	MT			0,04	0,01	3,37	0,02	0,01	2,97
	CN			0,08	0,02	4,37	0,05	0,01	4,10
Covariância e Residual	LP	↔	MT	0,04	0,01	3,34	0,03	0,01	3,79
	CN	↔	MT	0,04	0,01	3,00	0,02	0,01	2,60
	CN	↔	LP	0,06	0,02	3,72	0,04	0,01	3,67
Indicadores de ajuste				AIC=14.036,71 BIC (ajustado) = 14.088,50 CFI=0,985 / TLI=0,963			AIC=14.026,15 BIC (ajustado) = 14.085,09 CFI=0,986 / TLI=0,965		
Correlação intraclasse (ICC)				LP=0,084 / MT=0,069/ CN = 0,098			LP=0,084 / MT=0,071/ CN = 0,099		

Notas. As setas (→) indicam a direção do parâmetro; *Gf* = inteligência fluida; LP = língua portuguesa; CN = ciências naturais; MT = matemática; Expectativa do professor = expectativa sobre a quantidade de alunos que irão terminar o ensino fundamental; INF x NSEA = interação entre a infraestrutura na escola e o NSE agregado para a escola; EP = erro padrão; RC = razão crítica.

Os resultados apresentados na Tabela 3 apontaram para efeitos estatisticamente significativos ($RC > 1,96$) para todas as estimativas. Os valores de CFI e TLI, acima de 0,95, indicaram que os modelos 2 e 3 se ajustaram de maneira adequada aos dados. Os indicadores de AIC e BIC estão embasados no valor de ajuste qui-quadrado e ponderam os graus de liberdade e tamanho da amostra. O qui-quadrado reflete a falta de ajuste, assim, quanto menor for o seu valor, maior será o ajuste do modelo. Portanto, a diminuição dos indicadores AIC e BIC sugere que o modelo 3 foi mais bem ajustado do que o modelo 2.

Por meio das Tabelas 2 e 3 verificou-se a diminuição da variância do nível do estudante entre os modelos 1 e 2, bem como redução da variância do nível da escola entre os modelos 2 e 3. Tal redução é atribuída à explicação da variância proporcionada pelas variáveis independentes. Contudo, não é possível realizar o cálculo de R^2 (variância explicada), pois as variâncias residuais de LP, MT e CN foram correlacionadas nos modelos. Assim, uma parte da redução da variância de LP, por exemplo, pode ser atribuída à correlação com MT e CN. Caso as covariâncias entre LP, MT e CN fossem desconsideradas, o modelo explicaria, aproximadamente, 12, 17 e 15% das variâncias do desempenho do aluno em LP, MT e CN, respectivamente.

Verificou-se também que a inserção das variáveis independentes do nível do aluno reduziu a variância do nível da escola. Embora não esperado, isso pode indicar que a variância do fator leitura e da inteligência fluida (variáveis independentes) não se distribuem igualmente entre as escolas. Em outras palavras, a redução de variância do nível entre grupo não parece ser um efeito real, mas apenas um reflexo da distribuição desigual, entre as escolas, das variáveis independentes do nível do estudante (Hox, 2010).

De maneira geral, o modelo 2 indicou que a *Gf* e a dimensão latente relacionada à leitura do estudante são importantes para a compreensão do desempenho acadêmico. Os coeficientes não-padronizados apresentados nas tabelas podem ser interpretados como a quantidade de alteração da variável dependente, caso uma unidade da variável independente fosse alterada (consulte a tabela em anexo para obter a média e desvio-padrão das variáveis). Por exemplo, caso a variável *Gf* fosse alterada em uma unidade, tal alteração provocaria a mudança em 0,30 pontos do desempenho em LP e 0,36 em MT. Esses coeficientes não podem ser comparados entre si, uma vez que possuem métricas distintas. Para comparação do efeito das variáveis explicativas, é possível padronizá-los por meio da equação disponibilizada no método deste manuscrito. Os coeficientes padronizados (não apresentados nas tabelas) foram acima de 0,36 para a relação entre a *Gf* e o desempenho e acima de 0,11 para a relação entre a leitura e o desempenho.

No que se refere ao nível de análise da escola, a expectativa do professor sobre a quantidade de alunos que iriam terminar o ensino fundamental e a interação entre o nível socioeconômico agregado (NSEA) e a infraestrutura na escola também são variáveis importantes para a predição do desempenho acadêmico. Os coeficientes padronizados foram acima de 0,40 para a expectativa do professor e igual a 0,16 para o efeito de interação.

Nos modelos 4 e 5 testaram-se os coeficientes randômico das variáveis do nível do aluno, assumindo efeitos diferenciados entre as escolas, bem como interações entre os níveis de análise que pudessem explicar a randomização dos parâmetros. Tais resultados são apresentados na Tabela 4.

Tabela 4. *Parâmetros não padronizados dos modelos randômicos*

Parâmetros	Modelo 4			Modelo 5			
	Efeito	EP	RC	Efeito	EP	RC	
<i>Intragrupo (Nível Estudante)</i>							
Fator Leitura → CN	0,08	0,03	3,35	0,08	0,02	3,38	
Fator Leitura → LP	0,10	0,02	4,17	0,10	0,02	4,22	
<i>Gf</i> → CN	0,37	0,03	13,74	0,37	0,03	13,71	
<i>Variância e Residual</i>	Fator Leitura	1,33	0,48	2,80	1,33	0,47	2,83
	LP	0,51	0,02	22,61	0,50	0,02	22,64
	MT	0,38	0,02	17,59	0,38	0,02	17,71
	CN	0,58	0,03	19,08	0,58	0,03	19,09
<i>Entregrupo (Nível Escola)</i>							
Expectativa professor → LP	0,47	0,15	3,09	0,47	0,16	2,99	
Expectativa professor → MT	0,46	0,11	4,36	0,43	0,11	3,90	
Expectativa professor → CN	0,60	0,17	3,50	0,60	0,17	3,48	
Interação (INF x NSEA) → MT	0,23	0,09	2,39	0,23	0,09	2,42	
Interação (INF x NSEA) → <i>Slope1 (Gf→ LP)</i>	-	-	-	-0,23	0,10	-2,32	
Expectativa professor → <i>Slope2 (Gf→ MT)</i>	-	-	-	0,19	0,06	3,14	
<i>Média e Intercepto</i>	<i>Slope1 (Gf→ LP)</i>	0,29	0,03	0,30	0,03		
	<i>Slope2 (Gf→ MT)</i>	0,36	0,03	0,35	0,03		
	LP	-0,18	0,04	-0,18	0,04		
	MT	-0,10	0,03	-0,11	0,03		
	CN	0,14	0,04	0,14	0,04		
<i>Variância e Residual</i>	<i>Slope1 (Gf→ LP)</i>	0,003	0,004	0,84	0,002	0,004	0,47
	<i>Slope2 (Gf→ MT)</i>	0,005	0,004	1,42	0,004	0,004	0,93
	LP	0,04	0,01	3,84	0,04	0,01	3,89
	MT	0,02	0,01	2,88	0,02	0,01	2,87
	CN	0,05	0,01	4,11	0,05	0,01	4,11
Indicadores de ajuste	AIC= 14.027,59 BIC (ajustado) = 14.090,11			AIC = 14.023,11 BIC (ajustado) = 14.089,19			
Correlação intraclasse (ICC)	LP=0,084 / MT=0,071/ CN= 0,099			LP=0,084 / MT=0,071 / CN= 0,099			

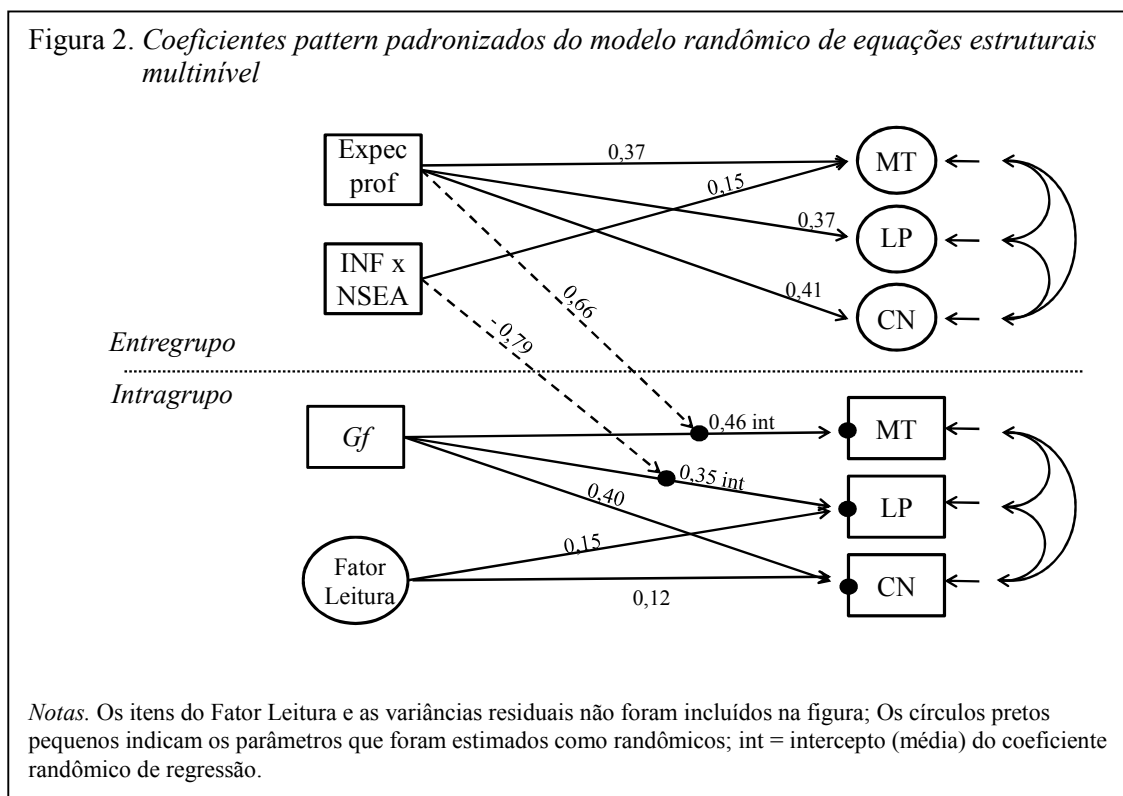
Notas. As setas (→) indicam a direção do parâmetro; foram omitidos da tabela os itens do fator leitura, bem como as covariâncias residuais entre as variáveis de desempenho; *Gf*= inteligência fluida; INF x NSEA= interação entre a infraestrutura na escola e o NSE agregado para a escola; *Slope* = coeficiente de inclinação; EP = erro padrão; RC = razão crítica.

Observa-se na Tabela 4 que o modelo 4 apresentou leve aumento dos indicadores AIC e BIC, se comparado ao modelo 3. Ou seja, o modelo teve leve diminuição do ajuste. Percebe-se, ainda, que as variâncias dos dois *slopes* randômicos

foram pequenas, e suas razões críticas abaixo do valor de 1,64 (valor de significância de 5% unicaudal) mostraram que elas não foram significativamente diferentes de zero.

Portanto, no modelo 4, a relação entre a *Gf* e o desempenho acadêmico do estudante não parece sofrer grande alteração entre as escolas.

Entretanto, o modelo 5 apresentou melhores indicadores de AIC e BIC em comparação com os modelos 3 e 4. Ademais, os *slopes* randomizados foram explicados, em parte, pela interação entre infraestrutura na escola e nível socioeconômico agregado para o nível da escola, bem como pela expectativa do professor sobre a quantidade de alunos que iriam terminar o ensino fundamental. Tais resultados indicam que, embora a variância dos *slopes* seja pequena, a relação entre *Gf* e desempenho acadêmico pode ser influenciada pelas variáveis da escola. Salienta-se que foi testado um modelo adicional, no qual as variáveis NSEA e infraestrutura na escola foram consideradas como variáveis independentes (sem a interação entre elas). Nesse modelo alternativo, o NSEA apresentou efeito significativo, mas o efeito da infraestrutura sobre o desempenho não foi significativo. Esse resultado indica que a infraestrutura, por si, não reduziu a relação entre *Gf* e desempenho, mas apresentou um efeito apenas quando considerada em conjunto com o NSEA. Nesse sentido, as escolas que mais reduziram o efeito da *Gf* sobre o desempenho acadêmico foram aquelas com maior NSEA e melhor infraestrutura. Para facilitar a interpretação do modelo 5, a Figura 2 apresenta as estimativas padronizadas dos coeficientes *pattern*.



Conforme a Figura 2, o desempenho do estudante no teste de inteligência fluida apresentou relação moderada (variando de 0,35 a 0,46) com as três provas de desempenho acadêmico. Para o nível da escola, a expectativa do professor apresentou relação moderada com o desempenho acadêmico (variando de 0,37 a 0,41). Destaca-se, ainda o papel da infraestrutura na escola e do NSEA na moderação do efeito da *Gf* sobre o desempenho acadêmico. Nesse sentido, a interação entre estas duas variáveis da escola diminuiu a relação entre a *Gf* e o desempenho em língua portuguesa. Ressalta-se, no entanto, que a variabilidade dos *slopes* entre a *Gf* e a língua portuguesa foi pequena, podendo superestimar o efeito moderador da infraestrutura e do NSEA.

Para complementar as análises, conduziram-se regressões simples (não apresentadas nas tabelas) entre a inteligência fluida e o desempenho em língua portuguesa separadas por escolas. Os *Betas*, na média, foram mais elevados para as escolas com menor NSEA. Para as cinco escolas de menor NSEA, a média dos *Betas* da

relação entre a *Gf* e a língua portuguesa foi de 0,53, e para as cinco escolas de maior NSEA a média dos *Betas* foi de 0,24. Ou seja, a relação entre a inteligência e a língua portuguesa parece diminuir conforme aumenta o NSEA.

Delineamento de Medidas Repetidas

Para a análise de medidas repetidas foram considerados os escores de proficiência das provas de língua portuguesa (LP) e de matemática (MT) nos anos de 2011 e 2012. A prova de ciências naturais não pode ser utilizada nesse delineamento, pois foi aplicada apenas uma única vez. Na análise, foram considerados três modelos para cada prova, separadamente: o *primeiro* modelo continha apenas a variável dependente (também chamado de modelo vazio); no *segundo* modelo foi inserida a variável explicativa tempo para avaliar se a alteração na proficiência do estudante entre os anos de 2011 e 2012 era estatisticamente significativa; no *terceiro* foram inseridas as variáveis explicativas do nível do aluno (2) e da escola (3) e testados os efeitos randômicos. Os resultados do primeiro modelo são apresentados na Tabela 5.

Tabela 5. *Parâmetros não-padronizados do modelo multinível de medidas repetidas 1 (modelo vazio)*

Parâmetros	Língua Portuguesa			Matemática		
	Efeito	EP	RC	Efeito	EP	RC
<i><u>Intrasujeito (Medida Repetida)</u></i>						
<i>Variância</i>	0,27	0,01	20,38	0,26	0,01	20,30
<i><u>Intragrupo (Nível Estudante)</u></i>						
<i>Variância</i>	0,33	0,02	14,22	0,25	0,02	12,89
<i><u>Entregrupo (Nível Escola)</u></i>						
<i>Intercepto do modelo</i>	-0,27	0,05		-0,32	0,05	
<i>Variância</i>	0,09	0,03	3,72	0,08	0,02	3,67
<i>Deviance (gl)</i>	4.126,98 (4)			3.927,17 (4)		
Correlação intraclasses aluno	0,48			0,42		
Correlação intraclasses escola	0,13			0,14		

Notas. EP = erro padrão; RC = razão crítica; gl = graus de liberdade.

Os interceptos para os modelos vazios de língua portuguesa (LP) e de matemática (MT) apontaram para as médias de desempenho próximas de 0. Ressalta-se que as proficiências foram estimadas por meio da TRI na escala *theta* ($M = 0$, $DP = 1$). As variâncias dos níveis de análise do aluno (intragrupo) e da escola (entregrupo) indicam que os interceptos podem variar entre os alunos e as escolas, justificando a utilização da modelagem multinível longitudinal. As correlações intraclasse, separadas por nível de análise, sugeriram que a variância do desempenho do aluno depende mais dos aspectos associados ao próprio aluno do que aos da escola. Contudo, as correlações intraclasse para a escola (0,13 e 0,14) foram relevantes e fortes o suficiente para justificar a inclusão deste nível de análise. No modelo 2, inseriu-se a variável momento de aplicação, cujo resultado é indicado na Tabela 6.

Tabela 6. *Parâmetros não-padronizados dos modelos multinível de medidas repetidas 2*

Parâmetros	Língua Portuguesa			Matemática		
	Efeito	EP	RC	Efeito	EP	RC
<i>Intrasujeito (Medida Repetida)</i>						
Momento de Aplicação ¹	0,16	0,02	6,83	0,25	0,02	10,87
Variância	0,25	0,01	21,00	0,23	0,01	21,28
<i>Intragrupo (Nível Estudante)</i>						
Variância	0,33	0,02	15,14	0,26	0,02	13,64
<i>Entregrupo (Nível Escola)</i>						
Intercepto do modelo	-0,28	0,05		-0,24	0,05	
Variância	0,10	0,03	3,80	0,08	0,02	3,81
Deviance (gl)	4.081,15 (5)			3.815,04 (5)		
Δ Deviance (Δ gl) entre modelos 1 e 2	45,83 (1)			112,13 (1)		
Δ Deviance \div Δ gl	45,83			112,13		
Correlação intraclasse aluno	0,49			0,46		
Correlação intraclasse escola	0,15			0,14		

Notas. ¹ = O momento de aplicação foi codificado como segue: 2011 = 0; 2012 = 1; EP = erro padrão; RC = razão crítica; gl = graus de liberdade; Δ Deviance (Δ gl) = diferenças dos deviances e graus de liberdade entre os modelos 1 e 2.

O momento de aplicação da prova apresentou efeito estaticamente significativo sobre o desempenho do estudante. O primeiro ano de aplicação das provas (2011) foi codificado como 0; assim, o efeito do momento de aplicação pode ser interpretado como o aumento médio do desempenho dos estudantes. Por exemplo, entre os anos de 2011 e 2012 os estudantes avaliados aumentaram, em média, 0,16 pontos da proficiência de língua portuguesa, e 0,25 pontos da proficiência de matemática. Percebe-se, por meio da comparação dos *deviances*, que o modelo 2 apresentou ajuste significativamente melhor se comparado ao modelo 1 ($\Delta \text{Deviance} \div \Delta \text{gl} > 1,96$). Buscou-se, ainda, avaliar a capacidade das variáveis inteligência fluida e infraestrutura na escola na explicação do aumento da proficiência, por meio do modelo 3, apresentado na Tabela 7.

Tabela 7. *Parâmetros não-padronizados dos modelos multinível de medidas repetidas 3 (inclusão dos efeitos randômicos)*

Parâmetros	Língua Portuguesa			Matemática		
	Efeito	EP	RC	Efeito	EP	RC
<i>Intrasujeito (Medida Repetida)</i>						
Slope1: Momento de Aplicação (rand) ¹	0,18	0,03	5,74	0,26	0,03	8,67
<i>Variância Residual</i>	0,00	0,00		0,00	0,00	

<i>Intragrupo (Nível Estudante)</i>						
Slope2: Interação Gf x Momento da Aplicação (rand)	0,46	0,05	9,04	0,62	0,05	13,82
<i>Variância Residual</i>	0,55	0,03	18,33	0,29	0,01	20,79

<i>Entregrupo (Nível Escola)</i>						
Interação Gf x Momento da Aplicação x Infraestrutura na Escola	-0,44	0,05	-8,94	-0,60	0,05	-13,33
<i>Intercepto do modelo</i>	-0,35	0,05		-0,23	0,04	
<i>Variância Residual</i>						
Intercepto	0,06	0,02	3,37	0,04	0,01	3,33
Slope1	0,01	0,01	1,56	0,01	0,01	1,29
Slope2	0,03	0,01	2,50	0,03	0,01	2,80

<i>Deviance (gl)</i>	3.979,26 (14)			3.576,29 (14)		
$\Delta \text{Deviance} (\Delta \text{gl})$ entre modelos 2 e 3	102,89 (9)			238,75 (9)		
$\Delta \text{Deviance} \div \Delta \text{gl}$	11,43			26,52		

Notas. ¹ = O momento de aplicação foi codificado como segue: 2011 = 0; 2012 = 1; Gf = inteligência fluida; (rand) = coeficiente randômico; EP = erro padrão; RC = razão crítica; gl = graus de liberdade; $\Delta \text{Deviance} (\Delta \text{gl})$ = diferenças dos *deviances* e graus de liberdade entre os modelos 2 e 3.

Por meio da comparação dos *deviances*, verifica-se que o modelo 3 apresentou ajuste significativamente melhor se comparado ao modelo 2 ($\Delta \text{Deviance} \div \Delta \text{gl} > 1,96$). Todos os *slopes* estimados foram estatisticamente diferentes de 0, portanto contribuíram para a explicação do desempenho do aluno. A interação entre o momento de aplicação e a variável inteligência fluida apresentou efeito moderado sobre o desempenho. Para o aumento da proficiência em LP e MT a *Gf* apresentou coeficiente padronizado (não disponibilizado nas tabelas) igual a 0,35 e 0,51, respectivamente. Estes resultados indicaram que a inteligência fluida não somente contribuiu para o desempenho do estudante, mas acentuou o aumento da proficiência (ou seja, da aprendizagem).

Para o nível da escola, a interação entre infraestrutura, *Gf* do estudante e momento de aplicação apresentou um efeito negativo. Para as provas de LP e MT, o coeficiente padronizado da interação foi igual a -0,25 e -0,36, respectivamente. Ou seja, se a *Gf* acentuou o aumento da proficiência (ou aprendizagem), a infraestrutura na escola alterou esse efeito: quanto melhor a infraestrutura na escola, menor é a relação entre a inteligência e a aprendizagem. Adicionalmente, testou-se um modelo alternativo considerando a variável nível socioeconômico agregado (NSEA) como preditor da aprendizagem. Todavia o NSEA não apresentou coeficiente de regressão estatisticamente significativo.

Discussão

O presente estudo buscou desenvolver e testar modelos explicativos do desempenho acadêmico e da aprendizagem (operacionalizado pelo aumento da proficiência em língua portuguesa e matemática entre os anos de 2011 e 2012). Ainda que a literatura sobre esses temas seja vasta, a presente pesquisa visou contribuir com uma perspectiva multinível de compreensão do fenômeno. Para tanto, foram

consideradas as variáveis explicativas do nível do estudante (principalmente inteligência fluida - *Gf*) e do nível da escola (principalmente infraestrutura e nível socioeconômico agregado no nível da escola - NSEA).

O primeiro resultado a ser destacado diz respeito às relações entre a inteligência fluida e o desempenho acadêmico evidenciadas pelo modelo de equações estruturais de um único nível de análise, o nível do estudante. O modelo apresentou ajuste adequado aos dados, o que oferece suporte à possibilidade de avaliar a *Gf* por meio dos testes de Raciocínio Abstrato (RA) e Raciocínio Espacial (RE), bem como avaliar uma dimensão geral de desempenho acadêmico (ou proficiência) por meio das provas de língua portuguesa, de matemática e de ciências naturais. Ainda que as tarefas cognitivas envolvidas na resolução de problemas de língua portuguesa sejam diferentes daquelas para resolução de ciências naturais, por exemplo, as correlações entre os escores das provas parecem sustentar a possibilidade de estimação de uma dimensão latente geral relacionada à proficiência geral (ou conhecimento geral).

Os resultados do modelo simples indicaram relação forte entre a *Gf* e o desempenho acadêmico ($Beta = 0,72$). Portanto, o estudo corrobora a primeira hipótese da pesquisa, que previa uma relação forte entre essas variáveis. Além disso, esse resultado é condizente com a literatura que também aponta para as associações moderadas e fortes entre esses dois construtos (Colom & Flores-Mendoza, 2007; Floyd et al., 2003; McGrew & Wendling, 2010; Neisser et al., 1996; Primi et al., 2010). O estudo de Primi et al. (2010), por exemplo, encontrou correlações de aproximadamente 0,60 entre a *Gf* e o desempenho em matemática. Nesse contexto, o coeficiente de regressão evidenciado pela presente pesquisa é um dos maiores entre os diversos estudos sobre o tema. Cabe destacar que grande parte das pesquisas anteriores utilizou correlações simples ou regressões lineares como métodos de análise. Tais métodos,

normalmente, subestimam as relações entre variáveis por não considerarem o erro de medida (ou imprecisão). Ainda que estejam disponíveis equações para atenuação da correlação para a falta de fidedignidade perfeita de um ou ambos os instrumentos (Hogan, 2006), na maioria dos estudos elas não foram utilizadas. Na presente pesquisa, os erros de medida foram modelados, por meio das equações estruturais, juntamente com as variáveis independentes e dependentes. Esse procedimento permitiu que o efeito da relação entre a *Gf* e o desempenho acadêmico fosse isolado do erro de medida, o que pode explicar o alto valor do *Beta* evidenciado pela presente pesquisa. Essa conclusão também enfatiza a relevância da modelagem do erro de medida (ou o uso de equações para corrigir correlações brutas) para uma compreensão mais adequada dos efeitos da inteligência sobre o desempenho acadêmico.

O segundo aspecto a ser destacado dos resultados diz respeito ao modelo de equações estruturais multinível. Essa modelagem pode oferecer um avanço na literatura, uma vez que se tem pouco conhecimento dos efeitos de moderação das variáveis do nível da escola sobre relação entre a *Gf* e o desempenho acadêmico. Ademais, as correlações intraclasse deste estudo, embora não sejam altas (por volta de 0,15), são suficientes para que a variabilidade do desempenho entre as escolas também seja considerada na modelagem (Hox, 2010).

Os modelos multinível apresentaram ajuste adequado aos dados e também indicaram a variável *Gf* como um importante preditor do desempenho acadêmico. Contudo, os *Betas* do modelo multinível foram inferiores ao do modelo simples de equações estruturais. Essa redução pode ser explicada pela dificuldade computacional que se enfrentou ao estimar um fator geral para o desempenho acadêmico nos modelos multiníveis. Em algumas tentativas, o modelo resultava em uma matriz não positivamente definida e, em outras, não havia recurso computacional (memória ram)

suficiente para estimar o modelo. Ainda assim, as relações entre *Gf* e desempenho acadêmico, para os modelos multinível, foram moderadas e são coerentes com as encontradas na literatura científica (Valentini & Laros, 2013). Além disso, o efeito mais forte ($Beta = 0,46$) foi entre a *Gf* e o desempenho em matemática. Esse resultado era esperado uma vez que a capacidade cognitiva para solução de problemas novos de ordem abstrata e espacial (inteligência fluida) parece ser um aspecto importante para a solução de tarefas matemáticas, que normalmente envolvem a representação e manipulação abstrata de variáveis. O desempenho em língua portuguesa, por sua vez, parece estar associado mais aos aspectos de raciocínio e manipulação de conceitos verbais, o que poderia estar mais associado à inteligência cristalizada (*Gc*).

Os resultados da análise multinível revelaram que o fator leitura do aluno influenciou positivamente o desempenho em língua portuguesa e ciências naturais, mas não em matemática. A leitura envolve a quantidade de livros disponíveis em casa, bem como a quantidade de livros efetivamente lidos por iniciativa própria ou por indicação da escola. Tais aspectos parecem estar mais relacionados às tarefas que envolvam a abstração de conceitos verbais, sinônimos e antônimos do que àquelas que envolvem a manipulação de números, variáveis e equações.

No que diz respeito às variáveis da escola, o modelo evidenciou a importância do nível socioeconômico agregado (NSEA), da infraestrutura na escola e da expectativa do professor como importantes variáveis para a predição do desempenho acadêmico. Ainda que a interação entre NSEA e a infraestrutura tenha apresentado apenas um coeficiente de regressão pequeno ($Beta = 0,15$) sobre o desempenho em matemática, o efeito de moderação dessas variáveis sobre a relação entre *Gf* e desempenho em língua portuguesa foi forte ($-0,79$). Portanto, a segunda hipótese da pesquisa, que previa o efeito de moderação, também foi corroborada. Esse resultado indica que a escola pode

diminuir o efeito da inteligência fluida sobre o desempenho acadêmico. Em outras palavras, a escola parece exercer um ‘nivelamento’ dos alunos quanto à língua portuguesa: as escolas com maior NSEA e melhor infraestrutura parecem ‘compensar’ a dificuldade cognitiva dos alunos, aproximando o desempenho em língua portuguesa dos alunos com os menores escores de inteligência com o desempenho dos alunos com os maiores escores de inteligência. Esse resultado também pode proporcionar um avanço na literatura científica. Por um lado, estudos que desconsideraram a inteligência do estudante apontaram para relações fortes entre o NSE (e NSEA) e o desempenho acadêmico (Caldas & Bankston-III, 1997; Laros et al., 2010, 2012). Por outro lado, as pesquisas que desconsideraram o nível de análise da escola evidenciaram apenas relações fracas entre o status socioeconômico e o desempenho acadêmico, após o controle da variável inteligência (Colom & Flores-Mendoza, 2007; Lloyd & Barenblatt, 1984). Essa ambiguidade foi apontada pela revisão teórica de Valentini e Laros (2013), que apresentou a hipótese da existência de uma relação complexa entre as variáveis. Os resultados do presente estudo oferecem suporte à tal hipótese ao apontar para as evidências do efeito de moderação das variáveis infraestrutura e NSEA sobre a relação entre a inteligência e o desempenho em língua portuguesa. Ressalta-se, no entanto, que o efeito só foi encontrado mediante a interação entre a infraestrutura e o NSEA. Ou seja, ao considerar apenas a infraestrutura, não se encontrou efeito significativo sobre a relação (*slope*) entre a inteligência e o desempenho em língua portuguesa. Portanto, as alterações dos *slopes* da relação entre inteligência e língua portuguesa ocorreram da seguinte maneira: (1) grande alteração dos *slopes* para as escolas com o maior NSEA e melhor infraestrutura; (2) pouca alteração dos *slopes* entre as escolas com baixo NSEA e infraestrutura ruim; (3) alteração mediana dos *slopes* para aquelas escolas com baixo NSEA e melhor infraestrutura. Em outras palavras, a infraestrutura adequada na escola

foi capaz de aproximar o desempenho dos estudantes com os maiores e os menores escores de inteligência, ainda que essas escolas sejam compostas pelos estudantes com os indicadores de nível socioeconômico mais baixos.

Ainda no nível de análise da escola, a expectativa média dos professores sobre quantos estudantes iriam concluir o ensino fundamental foi uma variável importante para o modelo multinível. As relações entre esse aspecto e o desempenho acadêmico foram moderadas. Tais resultados são compatíveis com a literatura. O estudo de meta-análise de Hattie (2009) indicou um tamanho de efeito moderado ($d = 0,43$) entre a expectativa do professor e o desempenho dos estudantes. Ressalta-se, no entanto, que a expectativa do professor pode ser apenas um reflexo do desempenho médio dos estudantes. Nesse caso, a expectativa do professor estaria associada à sua avaliação adequada sobre o desempenho e o potencial dos estudantes e não, exatamente, ao desempenho dos alunos. Em outras palavras, a expectativa do professor poderia ser um reflexo do real desempenho dos alunos, em vez do desempenho dos estudantes ser influenciado pela expectativa do professor. Obviamente, essa hipótese precisaria ser mais bem estudada por meio de métodos experimentais ou longitudinais. A expectativa do professor também apresentou efeito de moderação positivo sobre a relação entre *Gf* e matemática. Assim, os professores podem acentuar a relação entre a inteligência e o desempenho do estudante. Em outras palavras, a expectativa elevada do professor respalda um aumento, ainda maior, do desempenho acadêmico dos estudantes com os maiores escores de inteligência e reduz, ainda mais, o desempenho dos estudantes com os menores escores de inteligência.

O terceiro aspecto a ser destacado dos resultados diz respeito aos modelos de medidas repetidas. A presente pesquisa encontrou evidência de que a *Gf* está relacionada à aprendizagem do aluno. Em outras palavras, o aumento da proficiência em

língua portuguesa e matemática de um ano para o outro depende, em parte, da inteligência fluida. Tais resultados corroboram a terceira hipótese da presente pesquisa, que previa relação moderada entre a *Gf* e a aprendizagem, bem como estão de acordo com a maior parte da literatura científica (Geary, 2011; Hambrick et al., 2008; Primi et al., 2010; Swanson et al., 2008). Contudo, esse efeito é menor do que o encontrado nas pesquisas de Primi et al (2010) e de Geary (2011). Essa diferença pode ser explicada pela limitada variância da aprendizagem entre os estudantes da presente pesquisa. Ou seja, os estudantes que participaram da pesquisa apresentaram níveis parecidos de aprendizagem. Entretanto, as diferenças entre os estudantes, ainda que pequenas, podem ser explicadas, em parte, pela inteligência fluida.

Geary (2011), no seu estudo sobre a relação entre a inteligência e a aprendizagem, havia apontado como limitação e sugestão para pesquisas posteriores a avaliação de variáveis da escola sobre o desempenho acadêmico. Portanto, os resultados da presente pesquisa podem oferecer outro avanço à literatura científica: além da inteligência, a infraestrutura na escola também apresentou efeito sobre a aprendizagem do estudante. Especificamente, a infraestrutura na escola moderou, de maneira negativa, a relação entre a *Gf* e a aprendizagem. Essa conclusão corrobora, em parte, a quarta hipótese desta pesquisa: embora se tenha encontrado efeitos mais baixos do que o esperado (*Beta* de -0,25 para língua portuguesa e -0,36 para matemática), esses efeitos apontam para uma variável contextual importante na compreensão da aprendizagem. Isso sugere que a infraestrutura na escola reduz a diferença da aprendizagem entre os estudantes com baixos e altos escores de inteligência fluida. Nesse sentido, se, por um lado, os alunos que mais aprenderam foram aqueles que apresentaram os maiores escores de *Gf*, por outro lado, a infraestrutura na escola exerceu um papel importante, aproximando a aprendizagem em língua portuguesa e matemática dos alunos com os

menores escores de inteligência com a aprendizagem dos alunos com os maiores escores de inteligência.

Como principal limitação da pesquisa, aponta-se para o número relativamente pequeno de escolas nas quais os dados foram coletados (42 escolas). Ademais, as escolas apresentaram um perfil de alunos, de desempenho e de aprendizagem relativamente parecido, diminuindo a variabilidade dos dados nesse nível de análise. Portanto, sugere-se que novas pesquisas selecionem amostras de escolas mais heterogêneas, principalmente no que diz respeito à infraestrutura. Outra limitação refere-se à avaliação da aprendizagem em apenas dois momentos. Ainda que os alunos tenham demonstrado aumento da proficiência em língua portuguesa e matemática entre os anos de 2011 e 2012, esse intervalo de tempo é relativamente curto, podendo acarretar instabilidade nas medidas de aprendizagem. No intuito de evitar o problema, sugere-se que novas pesquisas multiníveis de medidas repetidas avaliem os estudantes em mais de dois momentos ou em intervalos de tempo maiores.

Em que pese às limitações do estudo, aponta-se, em resumo, para as evidências do efeito de moderação da infraestrutura da escola e do NSEA na relação entre a inteligência fluida e o desempenho acadêmico, bem como entre a inteligência fluida e a aprendizagem. Tais evidências sustentam a possibilidade e a importância de compreender o desempenho acadêmico e a aprendizagem de maneira complexa e multinível.

Referências

- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.
- Caldas, S. J., & Bankston-III, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, *90*, 269-277.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment* (2nd ed., pp. 69-76). New York, NY: The Guilford Press.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153-193
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1-22
- Colom, R., & Flores-Mendoza, C. E. (2007). Intelligence predicts scholastic achievement irrespective of SES factors: Evidence from Brazil. *Intelligence*, *35*, 243-251.
- Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics across the school-age years. *Psychology in the Schools*, *40*, 155-171.
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*, *47*, 1539-1552.
- Hambrick, D. Z., Pink, J. E., Meinz, E. J., Pettibone, J. C., & Oswald, F. L. (2008). The roles of ability, personality, and interests in acquiring current events knowledge: A longitudinal study. *Intelligence* *36*, 261-278.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- Heck, R. H., & Thomas, S. L. (2009). *An introduction to multilevel modeling techniques* (2nd ed.). New York, NY: Routledge.
- Hogan, T. P. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro, RJ: LTC.
- Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 443-451). New York, NY: Macmillan.
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

- Kemp, L. C. D. (1955). Environmental and other characteristics determining attainments in primary schools. *British Journal of Educational Psychology*, *25*, 66–77.
- Laros, J. A., Marciano, J. L. P., & Andrade, J. M. (2010). Fatores que afetam o desempenho na prova de Matemática do SAEB: Um estudo multinível. *Avaliação Psicológica*, *9*, 173-186.
- 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., Tellegen, P. J., Jesus, G. R., & Karino, C. A. (no prelo). *SON-R 2½-7[a] Teste não-verbal de inteligência: Manual com normatização e validação brasileira*. São Paulo, SP: Casa do Psicólogo.
- Lloyd, J., & Barenblatt, L. (1984). Intrinsic intellectuality: Its relations to social class, intelligence, and achievement. *Journal of Personality and Social Psychology*, *46*, 646-654.
- Lord, F. M. (1974). Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, *39*, 247-267.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment* (2nd ed., pp. 136-182). New York, NY: Guilford Press.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, *47*, 651-675.
- Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. London, UK: Sage.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77 - 101.

- Neto, J. J. S., Jesus, G. R., Karino, C. A., & Andrade, D. F. (2013). Uma escala para medir a infraestrutura da escola. *Estudos em Avaliação Educacional*, 24, 78-99.
- Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences*, 20, 446-451.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35, 83-92.
- Samejima, F. (1974). Normal ogive model on the continuous response level in the multi-dimensional latent space. *Psychometrika*, 39, 111-121.
- Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in Working Memory and Mathematical Problem Solving in Children at Risk and Not at Risk for Serious Math Difficulties. *Journal of Educational Psychology*, 100, 343-379.
- Valentini, F., & Laros, J. A. (2013). *Inteligência e desempenho acadêmico: Revisão de literatura*. Manuscrito não publicado. Brasília, DF.
- Valentini, F., Laros, J. A., Lima, R. M. F., Fonseca, W. R., Morais, L. S., & Tavares, R. L. J. (2013). *Evidências de validade de construto do Teste de Raciocínio Abstrato e Espacial – TRAE*. Manuscrito não publicado. Brasília, DF.
- Zhang, Z., Davis, H. P., Salthouse, T. A., & Tucker-Drob, E. M. (2007). Correlates of individual, and age-related, differences in short-term learning. *Learning and Individual Differences*, 17, 231-240.

Anexo do manuscrito 4

Tabela Anexo. *Estatísticas descritivas*

Variáveis	Mínimo	Máximo	Média	DP
<i><u>Delineamento Relacional</u></i>				
Matemática	-1,76	1,93	-0,08	0,74
Língua Portuguesa	-2,77	1,79	-0,15	0,82
Ciências Naturais	-2,63	2,72	0,16	0,90
Inteligência	-1,69	2,70	0,00	0,90
Fator Leitura				
Item 1	0,00	4,00	1,98	1,21
Item 2	0,00	4,00	2,03	1,05
Item 3	0,00	4,00	1,44	0,96
Interação Infraestrutura * NSEA	-0,89	0,02	-0,36	0,19
Expectativa do Professor	2,57	4,00	3,00	0,25
<i><u>Delineamento de medidas repetidas</u></i>				
Matemática (geral)	-2,72	1,93	-0,20	0,77
Matemática (2011)	-2,72	1,86	-0,32	0,78
Matemática (2012)	-1,76	1,93	-0,08	0,74
Língua Portuguesa (total)	-2,77	1,79	-0,23	0,83
Língua Portuguesa (2011)	-2,74	1,68	-0,31	0,83
Língua Portuguesa (2012)	-2,77	1,79	-0,15	0,82
Interação tempo * inteligência	-1,67	2,70	0,00	0,63
Interação tempo * inteligência * infraestrutura na escola	-2,96	2,96	0,00	0,46

Nota. DP = desvio-padrão.

Considerações Finais

A presente tese teve como objetivo principal avaliar o impacto de variáveis pessoais (principalmente inteligência fluida) e de variáveis contextuais (principalmente infraestrutura na escola) sobre o desempenho acadêmico e a aprendizagem. Como resultado dos estudos, foram produzidos quatro manuscritos, que, em conjunto, buscaram responder ao objetivo principal desta tese.

O primeiro manuscrito apresentou uma revisão crítica da literatura sobre a relação entre a inteligência fluida (*Gf*) e o desempenho acadêmico dos estudantes. Cumpre salientar que não há um consenso na literatura sobre o conceito de inteligência e o modelo teórico mais adequado para a compreensão das habilidades cognitivas. A revisão de literatura focou os modelos fatoriais de inteligência, uma vez que tais modelos são os mais utilizados nos estudos que testaram a relação entre a inteligência e o desempenho acadêmico. Sobre os modelos fatoriais, a literatura também apresenta divergências sobre quantas e quais as dimensões que seriam necessárias para definir a estrutura da inteligência. As soluções multifatoriais podem ser muito atrativas ao ‘senso comum’, uma vez que permitem, dentro do desenvolvimento normal, uma pessoa ser um ‘gênio’ em raciocínio espacial, por exemplo, e ter um desempenho muito abaixo da média em raciocínio verbal. Contudo, as pesquisas na área, que contam com mais de 100 anos de história, parecem indicar para algo diferente do ‘senso comum’: ainda que existam dimensões específicas das habilidades cognitivas, há evidências que sustentam a possibilidade de agrupá-las em algumas poucas dimensões gerais (ou apenas uma dimensão geral). Nesse contexto, o modelo Cattell-Horn-Carroll (CHC) destaca-se como uma moderna teoria ao permitir a estimação de aproximadamente 70 dimensões específicas e, ao mesmo tempo, agregá-las em uma dimensão geral única. Em outras palavras, as dimensões específicas não são completamente independentes entre si.

Grande parte dos estudos revisados sobre a inteligência e o desempenho acadêmico utilizou alguma das dimensões definidas pelo modelo CHC. Tais dimensões podem prever grande parte da variância do desempenho acadêmico. Juntamente com a inteligência, outras dimensões do nível de análise do estudante têm sido associadas ao desempenho, por exemplo, a personalidade, a motivação, a autopercepção de desempenho, entre outras. No nível de análise da escola, destaca-se o papel da infraestrutura e do nível socioeconômico agregado na relação com o desempenho acadêmico, embora algumas pesquisas tenham apontado para efeitos apenas fracos entre essas variáveis. Salienta-se que a maior parte dos artigos revisados tratou de pesquisas que consideraram apenas um nível de análise (normalmente o do estudante). Outros artigos que utilizaram análises multinível não investigaram o efeito de moderação das variáveis contextuais sobre a relação entre a inteligência e o desempenho acadêmico.

Nesse contexto, os manuscritos dois, três e quatro buscaram suprir, em parte, as lacunas apontadas na revisão teórica. Primeiramente, buscou-se construir um instrumento que pudesse avaliar a inteligência fluida de maneira acessível, rápida e coletiva. Nesse sentido, o segundo manuscrito apresentou dois estudos relacionados à construção do Teste de Raciocínio Abstrato e Espacial (TRAÉ). A primeira versão do instrumento (de 15 itens) foi aplicada numa amostra pequena para obter os indicadores de qualidade dos itens. Ainda que os itens da primeira versão do TRAÉ tenham demonstrado boa capacidade discriminativa e precisão, a exclusão de seis itens melhorou o equilíbrio da dificuldade entre os subtestes. A segunda versão do TRAÉ (de 12 itens) foi, então, aplicada em uma amostra de aproximadamente 1.300 estudantes. O estudo indicou que os itens apresentaram equilíbrio na distribuição do parâmetro de dificuldade, bem como discriminação adequada. Ademais, a curva de informação indicou que o TRAÉ é mais preciso ao estimar as habilidades levemente acima da

média. No que se refere à estrutura interna, as análises fatoriais exploratórias e confirmatórias sustentaram uma solução hierárquica, possibilitando a estimação de duas dimensões específicas da inteligência (raciocínio abstrato e raciocínio espacial), bem como a estimação de uma habilidade geral associada à inteligência fluida. Os indicadores de ajuste dos modelos de segunda ordem e do *bi-factor* ofereceram suporte à adequação dos modelos.

No entanto, ainda era necessário aprofundar os estudos da validade de construto do TRAE. Nesse intuito, buscou-se aprofundar as evidências de validade convergente e discriminante, relacionando os escores do teste TRAE com os escores de outro instrumento que já havia apresentado boas evidências de validade de construto, nesse caso a BPR-5. Os resultados desse estudo foram descritos no terceiro manuscrito. As correlações entre o TRAE e a BPR-5 foram fortes o suficiente para respaldar as evidências de validade convergente do TRAE, principalmente após a atenuação para ausência de fidedignidade perfeita dos instrumentos. A maior relação encontrada foi entre os escores gerais dos instrumentos. Esse resultado era esperado uma vez que o TRAE e a BPR-5 visam mensurar a inteligência fluida. Tais conclusões oferecem suporte à validade convergente do TRAE. As menores correlações foram entre os escores dos subteste do TRAE e os escores de raciocínio verbal e numérico da BPR-5. Estes últimos resultados também eram esperados, uma vez que a manipulação de conceitos linguísticos (raciocínio verbal) e as operações aritméticas básicas (raciocínio numérico) exigem conhecimento prévio mínimo, e podem ser explicadas pela dimensão latente de inteligência cristalizada (*Gc*). Esse resultado oferece suporte à validade discriminante do TRAE.

Os resultados da análise multitraço-multimétodo também ofereceram evidências adicionais à validade convergente do TRAE. Os coeficientes *pattern* (cargas fatoriais)

para a relação entre os fatores latentes (RA e RE) e as parcelas de itens foram minimamente adequadas. Ademais, a exclusão dos fatores latentes de RA e RE no modelo 2 foi acompanhada por uma significativa redução no ajuste do modelo. Esse resultado também oferece suporte à validade convergente do TRAE. Contudo, a correlação estimada entre os fatores latentes de RA e RE foi alta, e o modelo no qual essa correlação foi fixada em 1 (modelo 3) também foi plausível. Esse resultado indica limitação no que se refere à capacidade do TRAE discriminar o raciocínio abstrato do raciocínio espacial. Embora isso não fosse esperado, pesquisas anteriores também apontaram para correlações moderadas e fortes entre esses dois tipos de raciocínio. Além disso, RA e RE estão amplamente associados à inteligência fluida, o que também pode explicar a correlação entre eles.

Após a consolidação da segunda versão do TRAE, pôde-se, então, avaliar a relação entre a inteligência fluida (*Gf*), o desempenho acadêmico e a aprendizagem. Tais resultados foram descritos no quarto manuscrito. Para a realização dos estudos utilizou-se diversos bancos de dados da Secretaria de Educação e do Ministério da Educação. Além disso, realizou-se uma coleta de dados do teste TRAE em parceria com a Secretaria de Educação e o Centro de Seleção e Promoção de Eventos (CESPE-UnB). O modelo simples, de apenas um nível de análise, sugeriu forte relação entre a *Gf* e o desempenho acadêmico. Esse resultado é condizente com a maior parte da literatura científica. Ainda que os pesquisadores tenham buscado outros aspectos que pudessem influenciar o desempenho, a inteligência é um dos construtos psicológicos com maior efeito preditivo do desempenho do estudante.

Os modelos multinível aprofundaram a discussão sobre o tema. Por meio desses modelos, evidenciaram-se relações moderadas entre a *Gf* e o desempenho nas provas de língua portuguesa, de matemática e de ciências naturais. Tal relação é levemente distinta

entre os estudantes das diferentes escolas. Contudo, essas pequenas diferenças podem ser amplamente explicadas pela infraestrutura na escola e pelo nível socioeconômico agregado para a escola (NSEA). A infraestrutura diz respeito às instalações de salas de aula, banheiros, bibliotecas e à presença de itens como computadores e internet, entre outros. O NSEA diz respeito à média da quantidade de televisores, rádios, refrigeradores, automóveis, entre outros, que as famílias dos estudantes da escola possuem. Portanto, tais aspectos podem reduzir a diferença do desempenho entre os estudantes com os maiores e menores escores de inteligência. Ou seja, a infraestrutura na escola e o NSEA parecem diminuir o efeito da inteligência sobre o desempenho. Nesse caso, é possível que a qualidade das instalações da escola, presença de materiais de trabalho para os professores e alunos, entre outros, ofereçam subsídios importantes para o desempenho acadêmico, principalmente para os estudantes com menores escores de inteligência.

Resultados semelhantes foram encontrados no estudo de medidas repetidas. Comparou-se o desempenho acadêmico dos estudantes em dois anos consecutivos tendo como objetivo estimar o nível de aprendizagem. Os escores das provas nos diferentes anos foram equalizados por meio da TRI, o que permitiu a comparação entre as notas. Conforme previsto, a inteligência associou-se à aprendizagem (ou aumento da proficiência). Em outras palavras, os alunos com os maiores escores de *Gf*, não somente demonstraram melhor desempenho, como foram os que mais aprenderam. Novamente, a infraestrutura na escola alterou esse efeito, diminuindo a diferença da aprendizagem entre os estudantes com os maiores e os menores escores de inteligência. Contudo, o NSEA não apresentou efeito sobre a relação entre inteligência e aprendizagem.

Os estudos desta tese apresentaram limitações: (1) A revisão teórica careceu de uma melhor sistematização dos resultados. Portanto, os conceitos abarcados podem

estar sub ou super-representados na revisão de literatura. Nesse sentido, uma revisão mais sistemática, que adotasse critérios mais rígidos poderia melhorar o equilíbrio da revisão teórica. (2) Ainda que o TRAE tenha apresentado boas evidências de validade, a precisão desse instrumento foi apenas aceitável. O pequeno número de itens e a baixa variabilidade dos escores entre os estudantes podem explicar os índices abaixo do esperado. Novos estudos com amostras mais heterogêneas, no que se refere à inteligência fluida, poderiam melhorar a precisão dos escores gerados pelo instrumento. (3) As escolas também apresentaram perfis de alunos com desempenho e aprendizagem semelhantes entre si. Tais semelhanças entre os alunos reduziu a variabilidade dos dados no nível de análise da escola. Portanto, sugere-se que novas pesquisas colem amostras de escolas mais heterogêneas, principalmente no que diz respeito à infraestrutura. Além disso, seria importante avaliar as escolas de diferentes locais (capital e interior; urbanas e rurais) e de diferentes tipos (privadas e públicas).

Em que pese às limitações da presente tese, os estudos descritos podem promover um pequeno avanço na literatura científica: (1) Disponibilizou-se um instrumento com boas características psicométricas para a avaliação de dois tipos de raciocínio e de uma habilidade geral. (2) Apontou-se para as evidências do efeito de moderação da infraestrutura na escola e do nível socioeconômico agregado sobre a relação entre a inteligência e o desempenho acadêmico, bem como entre a inteligência e a aprendizagem. Tais resultados salientam a importância de compreender a relação entre a inteligência fluida, o desempenho acadêmico e a aprendizagem como um fenômeno complexo e multinível.

A presente pesquisa evidenciou a importância de considerar as variáveis da escola nos modelos explicativos do desempenho acadêmico e da aprendizagem. No entanto, é necessário ampliar os estudos sobre quais os aspectos escolares são, de fato,

relevantes para o desempenho do aluno e como esses aspectos podem influenciar a relação entre a inteligência fluida e o desempenho. Como agenda de pesquisa, ressalta-se a importância de avaliar os temas relacionados à qualidade didática e à proficiência dos professores. Nesse sentido, os professores com altos níveis de conhecimento sobre o conteúdo que lecionam conseguem aumentar o desempenho dos seus alunos, mesmo daqueles com baixos escores de inteligência? A didática ao lecionar pode diminuir a diferença de desempenho e aprendizagem entre os alunos com os maiores e menores escores de inteligência? Essas perguntas poderiam ser investigadas em pesquisas futuras.

Anexos da Tese

Anexo I *Comitê de Ética*



Comitê de Ética em Pesquisa
Instituto de Ciências Humanas
Universidade de Brasília

Universidade de Brasília
Instituto de Ciências Humanas
Campus Universitário Darcy Ribeiro

ANÁLISE DE PROJETO DE PESQUISA

Título do Projeto: RELAÇÕES ENTRE HABILIDADE COGNITIVA E DESEMPENHO ACADÊMICO

Número do projeto: 03-06/2012

Com base nas Resoluções 196/96, do CNS/MS, que regulamenta a ética da pesquisa em seres humanos, o Comitê de Ética em Pesquisa com Seres Humanos do Instituto de Ciências Humanas da Universidade de Brasília, após análise dos aspectos éticos, resolveu **APROVAR** o projeto intitulado "RELAÇÕES ENTRE HABILIDADE COGNITIVA E DESEMPENHO ACADÊMICO".

O pesquisador responsável fica notificado da obrigatoriedade da apresentação de um relatório final sucinto e objetivo sobre o desenvolvimento do Projeto, no prazo de 1 (um) ano a contar da presente data (itens VII.13 letra "d" e IX.2 letra "c" da Resolução CNS 196/96).

Brasília, 01 de agosto de 2012.

Debora Diniz
Coordenadora Geral – CEP/IH

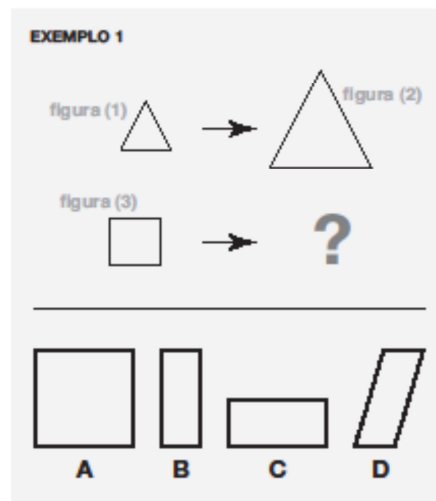
Anexo II

Raciocínio Abstrato

A seguir serão apresentadas algumas figuras que sofrem alguma mudança. A figura (1) se transforma na figura (2). A figura (3) deve mudar do mesmo jeito. Depois de mudar a figura (3), como ficaria a próxima figura?

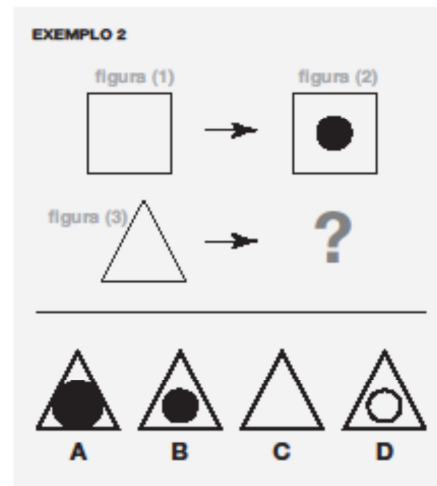
Qual das figuras (alternativas A, B, C ou D) deve vir no lugar do ponto de interrogação?

Veja o exemplo 1.



O triângulo (1) muda e fica maior (2). O quadrado (3) deve mudar do mesmo jeito, ficando maior. A alternativa "A" é a única em que o quadrado fica maior. Ou seja, a alternativa "A" é a resposta correta! Veja que na sua folha de respostas a alternativa "A" está marcada para o exemplo 1. Você não precisa marcar nada ainda.

Agora, veja o exemplo 2.



A figura (1) se transforma na figura (2), pois uma bolinha pequena totalmente preta aparece dentro do quadrado. O triângulo (3) deve mudar do mesmo jeito. A alternativa "B" é a única que contém um triângulo com uma bolinha pequena totalmente preta dentro. Ou seja, a alternativa "B" é a resposta correta! Note que, na sua folha de respostas, a alternativa "B" está marcada para o exemplo 2.

As próximas questões são semelhantes a esses exemplos. Trabalhe o mais rápido que puder.

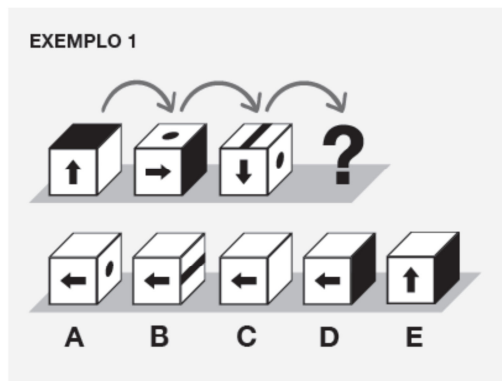
Suas respostas para as questões devem ser marcadas na folha de respostas.

Anexo III

Raciocínio Espacial

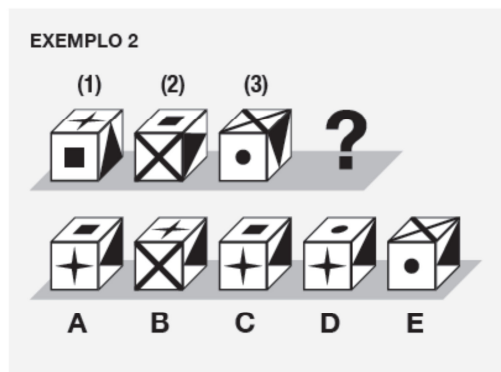
A seguir serão apresentadas algumas figuras que estão em movimento e giram conforme a sequência apresentada na parte superior da questão. Sua tarefa é verificar como a figura (um cubo) está girando e descobrir como ficaria a próxima figura.

Veja o Exemplo 1.



Se você olhar o exemplo 1, a figura gira no sentido horário (para a direita). Note que na primeira figura, a face totalmente preta está na parte superior. Após o giro, a face preta fica na parte direita (segunda figura), o que indica que a figura girou no sentido horário. O mesmo acontece com a bolinha. Na segunda figura, a bolinha está na parte superior e, na terceira figura está na parte da direita. Ou seja, a figura girou de novo no sentido horário. Portanto, a próxima figura deveria ter uma linha na face direita. A alternativa "B" é a única que contém uma linha na face direita. Ou seja, a alternativa "B" é a resposta correta. Note que, na sua folha de respostas, a alternativa "B" está marcada para o exemplo 1. Você não precisa marcar nada ainda.

Preste atenção, porque as figuras podem girar para outros sentidos.



Agora, as figuras giram da frente para cima: o quadrado preto passa da parte da frente para cima (2); e, depois, a cruz também passa da parte da frente para cima (3). Portanto a alternativa "D" completa o movimento, pois a bolinha que estava na parte da frente deve passar para cima.

As próximas questões são semelhantes a esses exemplos. Você deve verificar como a figura gira e escolher a alternativa (A, B, C, D ou E) que completa o movimento.

Suas respostas para as questões devem ser marcadas na folha de respostas.

Anexo IV

Sintaxe do Modelo de Equação Estrutural Multinível (Mplus)

```
USEVARIABLE ARE codEsc LP_TRI MT_TRI CN_TRI
RAE_TRI A11Inv A12Inv A30Inv p11 NSErec;
```

```
between = p11 NSErec;
within = RAE_TRI A11Inv A12Inv A30Inv;
centering = grandmean(RAE_TRI A11Inv A12Inv A30Inv p11 NSErec);
MISSING ARE all (999);
CLUSTER IS codEsc;
```

```
Define: NSErec=(INF_TRI*NSEA_Rtri);
Analysis: TYPE=TWOLEVEL random;
estimator=MLR;
ITERATIONS=2000;
CONVERGENCE=0.01;
```

```
Model: %BETWEEN%
LP_TRI on p11;
MT_TRI on NSErec p11;
CN_TRI on p11;
```

```
S1 on NSErec;
S2 on p11;
```

```
LP_TRI with MT_TRI;
MT_TRI with CN_TRI;
CN_tri with LP_TRI;
```

```
%WITHIN%
Liv by A12Inv A11Inv A30Inv;
```

```
S1|LP_TRI on RAE_TRI;
S2|MT_TRI on RAE_TRI;
CN_TRI on RAE_TRI;
```

```
LP_TRI on Liv;
CN_TRI on Liv;
```

```
LP_TRI with MT_TRI;
MT_TRI with CN_TRI;
CN_tri with LP_TRI;
```

```
output: sampstat;
```