

Propriedades psicométricas do conjunto de testes da habilidade visuo espacial

Cristiano Mauro Assis Gomes¹ - Universidade Federal de Minas Gerais, Belo Horizonte, Brasil
Oto Neri Borges - Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

Resumo

Este artigo analisa as propriedades psicométricas de um conjunto de testes elaborados para mensurar a habilidade visuo espacial (Gv), quanto a sua estrutura fatorial, unidimensionalidade, validade convergente e discriminante. Fizeram parte da pesquisa 558 participantes que cursavam ou concluíram o ensino médio. Utilizou-se o Teste de Visualização (Vz), Teste de Flexibilidade de Fechamento (CF), Teste de Memória Visual (MV), Teste de Fluência Figural (FF), Teste de Fluência Ideacional 1 (FI1) e Teste de Fluência Ideacional 2 (FI2). Foram usados procedimentos de análise fatorial exploratória, análise fatorial confirmatória e o ajuste ao modelo Rasch. Os índices de ajuste dos modelos aos dados estão adequados às recomendações da literatura. Os testes possuem uma estrutura fatorial de dois níveis, porém unidimensional. Convergem para mensurar a habilidade visuo espacial (Gv) e divergem na mensuração da fluência, utilizada neste estudo como habilidade contrastante.

Palavras-chave: habilidade visuo espacial; propriedades psicométricas; testes de inteligência.

Psychometric proprieties of visual-spatial ability tests kit

Abstract

This article analyses the psychometrical properties of a test kit developed to measure visual-spatial ability in terms of its factorial structure, unidimensionality, convergent validity and divergent validity. A sample of 558 students that were either attending High School or had just concluded it participated of this research. The Visualization Test (Vz), Flexibility of Closure Test (CF), Visual Memory Test (MV), Figural Fluency Test (FF), Ideational Fluency Test 1 (FI1) and Ideational Fluency Test 2 (FI2) were used. The data analysis used the exploratory factorial analysis, the confirmatory factorial analysis and the fit to the Rasch model procedures. The model fit indexes are in accordance with the cutoff values suggested by the literature. The tests showed a two-level factorial structure, although it can be taken as an unidimensional structure. The tests converge to measure the visual-spatial ability and diverge to measure the fluency ability which was used in this study as contrast ability.

Keywords: visual-spatial ability; psychometric properties; intelligence tests.

Introdução

A Habilidade visuo espacial (Gv) é genérica e ampla, correspondendo a um tipo de processamento geral capaz de produzir, registrar, lembrar transformar imagens e sensações visuais (Primi, 2003; Schelini, 2006). Segundo McGrew (2009) ela é tipicamente aferida por tarefas que envolvem a percepção e transformação, de figuras, formas e imagens visuais bem como por tarefas que envolvem manter a orientação espacial relativa a objetos que se movem no espaço. As habilidades específicas (pertencentes ao estrato I, na teoria de Carroll (1993)) que carregam em Gv são visualização (Vz), relações espaciais (SR), velocidade de fechamento (CS), flexibilidade de fechamento (CF), memória visual (MV), esquadrinamento espacial (SS), entre outros (McGrew, 2009). Ela é uma habilidade de alta ordem, presente nos modelos psicométricos

sobre a inteligência humana (Bickley, Keith & Wolfle, 1995; Carroll, 1993; Cole & Randall, 2003; Johnson & Bouchard Jr., 2005a,b; Johnson, Nijenshuis & Bouchard Jr., 2007; Marañon & Andrés-Pueyo, 2000; McGrew, 2009; Undheim & Gustafsson, 1987). Por seu caráter amplo, esta habilidade se distingue de processos visuo-espaciais específicos, como é o caso da capacidade de girar imagens mentalmente, ou da capacidade de imaginar e transformar mentalmente determinadas figuras bidimensionais em figuras tridimensionais. Em virtude do seu caráter delimitado pelo domínio visuo-espacial, esta habilidade se distingue de qualquer fator geral, o *g* do terceiro estrato da teoria de Carroll (1993), na medida em que um fator geral *strito sensu* necessita ser ativado ou envolvido por qualquer tarefa cognitiva, independentemente do domínio envolvido (Carroll, 1993; Spearrit, 1996).

Alguns aspectos justificam o estudo da habilidade visuo espacial (Gv), enquanto habilidade genérica, de alta ordem. Um deles envolve a identificação apurada dos processos cognitivos que envolvem o domínio visuo espacial (Burton & Fogarty, 2003; Johnson & Bouchard Jr., 2005a,b; Johnson & cols., 2007; Juhel, 1991;

¹ Endereço para correspondência:

Depto. de Psicologia, Universidade Federal de Minas Gerais
Avenida Antônio Carlos, 6627, Campus da Pampulha,
Departamento de Psicologia, Universidade Federal de Minas
Gerais. CEP: 31270-901. e-mail: cgomes@fafich.ufmg.br
Apoio: FAPEMIG e CNPq.

Stumpf & Eliot, 1999). Outro desses aspectos refere-se à capacidade da habilidade visuo espacial de prever certos desempenhos acadêmicos específicos e profissionais (Brandt & Davies, 2006; Hedman & cols., 2006; Nilsson, Hedman & Ahlqvist, 2007; Wanzel & cols., 2003). Um terceiro aspecto advém do argumento de Wechsler e Schelini (2006) sobre uma possível associação entre a inteligência fluída (Gf) e Gv.

O quarto aspecto, bem vinculado ao foco deste trabalho e projeto dos pesquisadores, decorre do fato que Gv é uma importante habilidade ampla na teoria de Carroll (1993), bem como na recente proposta de McGrew (2009) de unificação das teorias de Cattell–Horn e de Carroll em uma teoria e taxionomia geral para o estudo da estrutura das habilidades cognitivas humanas, que denominou Teoria Cattell–Horn–Carroll (CHC). Segundo Primi (2003, p. 71), à época da elaboração de seu trabalho não existia uma bateria que mensurasse todos os fatores amplos (estrato II da Teoria CHC). A bateria Woodcock-Johnson III, em sua opinião, é a que mais se aproxima deste propósito. Ela foi adaptada para o Brasil por Wechsler e Schelini (2006) e possui um teste, composto por 33 itens, para a mensuração de Gv.

Apesar da existência de baterias internacionais elaboradas para mensurar a habilidade visuo espacial (Gv) (Carroll, 2003) e da relevância do estudo das habilidades cognitivas do domínio visuo-espacial, apesar dos progressos em nível nacional na elaboração e validação de testes de inteligência (Primi, 2003), há uma carência de baterias elaboradas especificamente para mensurar esta habilidade, conforme foi constatado em consulta à lista de testes aprovados no SATEPSI (www2.pol.org.Br/satepsi/sistema/admin.cfm?list_a1=sim), atualizada até maio de 2008. O desenvolvimento de baterias visando mensurar de forma equilibrada as habilidades amplas do segundo estrato da teoria CHC é uma agenda importante para a área (McGrew, 2009; Primi, 2003). Nessa direção, destaca-se o trabalho de Schelini e Wechsler (2005, 2006), desenvolvendo uma bateria multidimensional para a inteligência infantil capaz de mensurar quatro habilidades amplas do segundo estrato da teoria CHC. No entanto, essa bateria não mensura Gv. O estudo aqui relatado faz parte de um esforço de desenvolver um conjunto de testes de habilidades do primeiro estrato capaz de identificar seis habilidades amplas do segundo estrato da teoria CHC, incluindo Gv.

Conforme argumentam Sisto, Santos e Noronha (2007), há no Brasil uma limitação de estudos que buscam comprovar as qualidades psicométricas dos testes psicológicos. A unidimensionalidade é uma propriedade psicométrica substancial dentro do conjunto de aspectos fundamentais para a análise da qualidade dos testes psicológicos (Sisto, Rueda & Bartholomeu, 2006). Por outro lado, não há um critério único e bem definido pela literatura no que tange à definição e elaboração de indicadores a respeito da unidimensionalidade (Vitória, Almeida & Primi, 2006). Tennant e Pallant (2006) identificaram na literatura que utiliza o modelo Rasch três abordagens para a investigação da unidimensionalidade: (i) análise prévia utilizando métodos da Teoria Clássica de Testes, (ii) interpretação de que um modelo bem ajustado garante a unidimensionalidade do instrumento, e (iii) testes *post-hoc*, como por exemplo, o uso da análise de componentes principais nos resíduos.

Este artigo relata um estudo das propriedades psicométricas de uma bateria brasileira elaborada especificamente para mensurar a habilidade visuo espacial (Gv). A bateria a ser analisada é o Conjunto de Testes da Habilidade Visuo Espacial, formado pelo Teste de Visualização (Vz), Teste de Flexibilidade de Fechamento (CF) e o Teste de Memória Visual (MV), elaborados pelo primeiro autor deste artigo com base em sua experiência com a tradução, adaptação e validação de 45 testes cognitivos do *Kit of Factor-Referenced Cognitive Tests* do *Educational Testing Service* (Ekstrom & cols., 1976; Gomes, 2005). Ela é uma das baterias específicas desenhadas com o intuito de desenvolver uma bateria brasileira para mensurar seis fatores cognitivos do segundo estrato da teoria de Carroll: inteligência fluída (Gf), a inteligência cristalizada (Gc), habilidade de memória (Gy), a rapidez cognitiva (Gs), a habilidade de recuperação abrangente (Gr) e a habilidade visuo espacial (Gv).

As propriedades psicométricas do Conjunto de Testes de Habilidade Visuo Espacial serão investigadas por meio da unidimensionalidade e da estrutura fatorial de cada um dos testes do conjunto. O Conjunto também será investigado em quanto à sua validade convergente e validade discriminante.

Método

Participantes

Fez parte do estudo uma amostra total de 558 estudantes. Desse total, 384 (68,8%) eram

alunos de uma escola da rede federal de ensino com idade entre 14 e 20 anos, com média de 15,78 e desvio padrão de 1,14; 105 (18,8%) eram alunos de um programa de ensino médio de jovens e adultos da mesma instituição; e 69 (12,4%) já tinham concluído o ensino médio e eram estudantes do primeiro período de uma instituição privada de ensino superior. As idades desses dois últimos grupos não foram coletadas. Todos os alunos estudavam em instituições de ensino do município de Belo Horizonte, Minas Gerais, Brasil. O número de pessoas do gênero feminino foi de 280 (50,2%), e 278 (49,8%) para o gênero masculino. Nem todos os participantes fizeram todos os testes (Vz=558, CF=415, MV=556, FF=292, FI1=292 e FI2=292), de forma que em cada análise foi utilizada a maior amostra possível, sem adoção de procedimento de tratamento de dados faltantes.

Instrumentos

Neste estudo foram utilizados seis instrumentos, Teste de Visualização (Vz), Teste de Flexibilidade de Fechamento (CF), Teste de Memória Visual (MV), Teste de Fluência Figural (FF), Teste de Fluência Ideacional 1 (FI1), Teste de Fluência Ideacional 2 (FI2), todos eles desenvolvidos por Gomes (2005). Cada teste possui o tempo-limite para a sua execução. Tal

limite foi fixado com base em testes similares, porém não foram feitos testes empíricos para estudar sua adequação. Nos quatro primeiros testes, cada um dos itens foi pontuado como correto (1), errado (0) ou faltante (9). Assim, em cada teste o escore mínimo é zero e o máximo corresponde ao total de itens do teste. Nos outros dois testes, cada resposta fornecida pelo respondente, se preenchida a característica solicitada, foi pontuada como correta (1); em caso contrário foi pontuada como errada (0). Nestes dois últimos testes não há dados faltantes, bem como escore máximo.

Teste de Visualização (Vz). Composto por 30 itens, seis figuras em formato bidimensional e seus equivalentes em formato tridimensional, possui um tempo-limite para sua execução de 12 minutos. O respondente deve ser capaz de imaginar a figura bidimensional transformando-se na figura tridimensional. Cada linha da figura bidimensional é marcada com uma letra e cada linha da figura tridimensional é marcada com um número. Cada item, então, apresenta uma letra que indica uma linha da figura bidimensional e o respondente deve escrever o número da linha da figura tridimensional que representa a mesma linha da figura bidimensional. A Figura 1 mostra um exemplo de item deste teste.

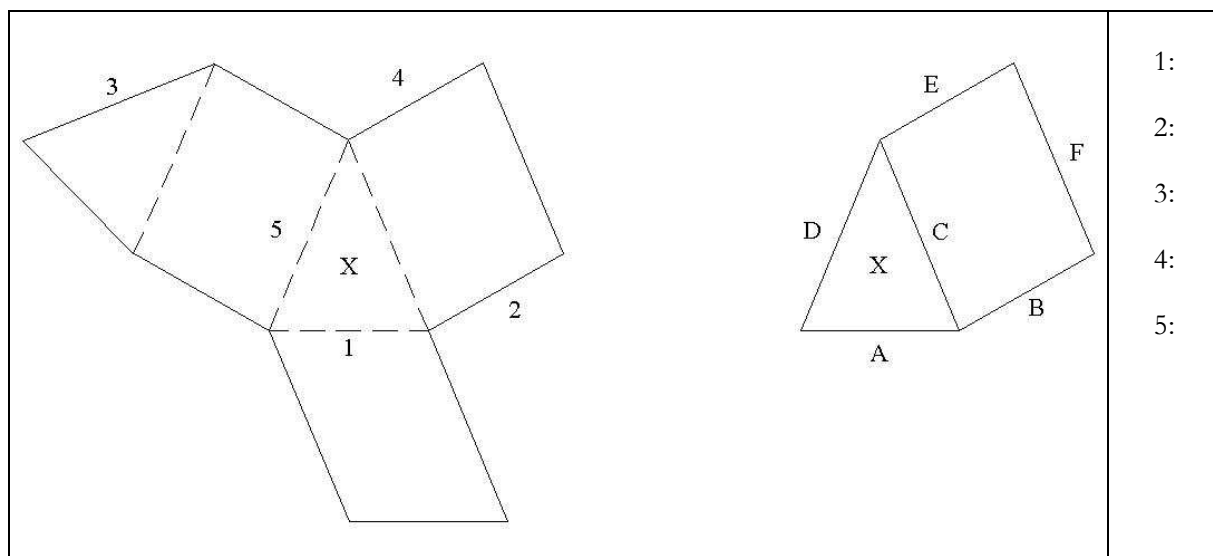


Figura 1. Exemplo de item que compõe o Teste de Visualização (Vz).

Teste de Flexibilidade de Fechamento (CF). Composto por 32 itens e um tempo-limite de 12 minutos para sua execução, cada item possui uma figura modelo e um espaço de resolução. O modelo é uma figura bidimensional formada de pontos e traços entre esses pontos. O espaço de

resolução, por sua vez, é um conjunto de pontos. A tarefa do respondente é fazer os traços da figura modelo no conjunto de pontos do espaço de resolução, de modo a reproduzir fielmente a figura modelo. A Figura 2 mostra um dos itens deste teste.

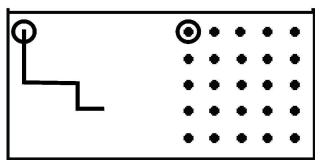


Figura 2. Exemplo de item que compõe o Teste de Flexibilidade de Fechamento (CF).

Teste de Memória Visual (MV). Composto por 12 itens, uma folha de memorização e uma folha de resposta, há um tempo-limite de sete minutos para a execução. A folha de memorização apresenta 12 mapas. A folha de resposta apresenta também 12 mapas e duas opções de resposta, onde o respondente deve indicar se o mapa da folha de resposta é ou não é exatamente o mesmo da folha de memorização. A Figura 3 mostra um item do teste.

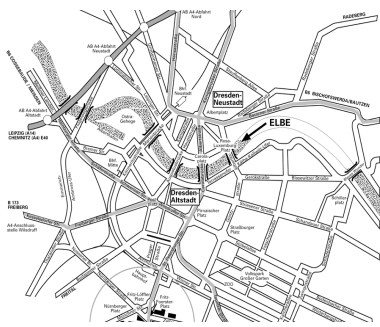


Figura 3. Exemplo de item que compõe o Teste de Memória Visual (MV).

Teste de Fluência Figural (FF). É composto por 20 itens e um tempo limite de um minuto e meio para sua execução. Cada item é composto por uma camiseta em branco. O respondente deve desenhar detalhes na camiseta de modo que cada uma seja única. O escore é formado pelo número de camisetas desenhadas.

Teste de Fluência Ideacional 1 (FI1). Não possui um número definido de itens, mas há um tempo-limite de quatro minutos para sua execução. O respondente deve ser capaz de escrever o maior número de ideias relacionadas a um tópico pre-determinado. Um exemplo de tópico é "Uma viagem de trem". O escore é formado pelo número de ideias escritas relacionadas com o tópico.

Teste de Fluência Ideacional 2 (FI2). Não possui um número definido de itens, mas há um tempo-limite de três minutos para sua execução. O respondente deve ser capaz de escrever o maior número de objetos totalmente ou

predominantemente relacionados a uma categoria pre-determinada. Um exemplo de categoria é "Objetos vermelhos". O escore é formado pelo número de objetos escritos relacionados com a categoria.

Procedimento

Os testes foram aplicados conjuntamente, de forma coletiva, e sempre por psicólogos ou estudantes de psicologia devidamente treinados. Foram tomados todos os cuidados éticos em relação aos participantes e a pesquisa contou com a aprovação do Comitê de Ética da UFMG (n. ETIC 181/06).

Para o tratamento dos dados, os dados faltantes foram repontuados como zero. Na identificação da estrutura fatorial de cada teste do Conjunto de Testes da Habilidade Visuo Espacial foram utilizados os procedimentos fatoriais exploratórios de retenção de fatores do autovalor-maior-do-que-um, o *scree test*, a análise paralela por permutação e a máxima verossimilhança. Foram selecionados vários procedimentos de retenção, na medida em que eles tendem a gerar diferentes soluções fatoriais (Beauducel, 2001; Frazier & Youngstrom, 2007; Guadagnoli & Velicer, 1988; Henson & Roberts, 2006).

A partir da identificação de diferentes soluções fatoriais pela análise fatorial exploratória (EFA – *exploratory factor analysis*), uma solução era identificada como a mais adequada através da análise fatorial confirmatória (CFA – *confirmatory factor analysis*), usando um índice comparativo de ajuste CFI (*comparative fit index*) (Bentler, 1990) e a estatística raiz quadrática média residual estandarizada, RMR (*root mean square residual*), e a comparação dos qui-quadrados e graus de liberdade provenientes das diferentes soluções (Byrne, 2001).

O CFI foi proposto por Bentler (1990) como uma estatística que varia de zero a um e é adequada para mensurar o grau de ajuste de um modelo aos dados, mesmo em amostras pequenas (Byrne, 2001). Segundo Byrne, originalmente o critério de bom ajuste era um $CFI > 0,90$, porém mais tarde Hu e Bentler (1999) sugeriram que $CFI > 0,95$ é um critério mais estrito de um bom ajuste do modelo aos dados. Foi adotado neste trabalho o critério menos restrito ($CFI > 0,90$). O RMR corresponde ao valor médio dos resíduos obtidos pela diferença entre a matriz de variância-covariância dos dados da amostra e a matriz de variância-covariância advinda do ajuste do modelo proposto. Segundo Byrne (2001) estes resíduos são

de difícil interpretação por serem relativos aos valores das variâncias e covariâncias observadas. Ela sugere que o RMR estandarizado, por variar de zero a um, é de mais fácil interpretação e que um modelo está bem ajustado aos dados se o RMR estandarizado é igual ou menor do que 0,05 (Hu & Bentler, 1995).

Tennant e Pallant (2006) identificaram, na literatura que utiliza o modelo Rasch, três abordagens para a investigação da unidimensionalidade: (i) análise prévia utilizando métodos da Teoria Clássica de Testes, (ii) a interpretação de que um modelo bem ajustado garante a unidimensionalidade do instrumento, e (iii) testes *post-hoc*, como, por exemplo, o uso da análise de componentes principais nos resíduos. Neste estudo utilizou-se a primeira abordagem, por meio da EFA e da CFA, e a segunda abordagem, estudando o ajuste do modelo Rasch aos dados, usando índice *infit*, escolhido por ser um indicador sensível da relação entre o nível de desempenho atribuído às pessoas e o nível de dificuldade atribuído aos itens (Linacre, 2002).

Os itens foram considerados bem ajustados se seus *infits* estavam na faixa entre 0,7 e 1,3 (Wright & Linacre, 1994). O teste foi considerado bem ajustado se todos *infits* dos itens estavam na faixa em que eram produtivos para a medida, $0,5 < \textit{infit} < 1,5$ (Wright & Linacre, 1994).

Com relação à validade convergente e discriminante, foi realizada uma análise fatorial confirmatória do Conjunto de Testes da Habilidade Visuo-Espacial, acrescentada de três testes elaborados especificamente para mensurar a fluência (Gr), uma outra habilidade cognitiva alta ordem do campo psicométrico. Como critério de ajuste do modelo, além do índice CFI, foi usada a estatística erro de aproximação quadrático médio, RMSEA (*root mean square error of approximation*), considerado dos mais informativos na modelagem da estrutura de covariância (Byrne, 2001). Esse índice leva em conta o erro de aproximação na população e tenta responder à questão de quão bem pode o modelo reproduzir a matriz de covariância da população, caso ela estivesse disponível, se os valores dos parâmetros, apesar de desconhecidos, forem escolhidos de forma otimizada. Browne e Cudeck (1993) sugeriram que valores de RMSEA menores do que 0,05 são indicadores de bom ajuste. Um ajuste perfeito corresponde a RMSEA=0,00.

Resultados e discussão

Todos os critérios de retenção de fatores indicaram para o Teste de Visualização (Vz) uma solução fatorial de três fatores de primeiro nível e um fator geral (g), excetuando apenas o procedimento de máxima verossimilhança que apresentou caso *heywood*. A presença de uma única solução nos diferentes critérios de retenção indica que a solução encontrada é robusta. Em relação ao Teste de Flexibilidade de Fechamento (CF), a análise paralela por permutação indicou uma solução fatorial de três fatores de primeiro nível, o *scree test* e o autovalor-maior-do-que-um indicaram uma solução de cinco fatores de primeiro nível e o procedimento da máxima verossimilhança indicou 10 fatores de primeiro nível. O Teste de Memória Visual (MV) também apresentou diferentes soluções fatoriais a partir dos critérios de retenção de fatores. O *scree test* indicou uma solução de um único fator, a análise paralela por permutação, dois fatores de primeiro nível, o autovalor-maior-do-que-um, quatro fatores de primeiro nível e a máxima verossimilhança, cinco fatores de primeiro nível.

A análise fatorial confirmatória foi utilizada para que se pudessem comparar as diferentes soluções fatoriais de cada teste, obtidas pela análise fatorial exploratória. As diferentes soluções de cada teste foram modeladas, de forma a se obter o seu grau de ajuste aos dados. A solução com melhor grau de ajuste aos dados foi selecionada como a mais adequada de cada teste.

A melhor solução de cada um dos testes apresentou um CFI superior a 0,90 e um RMR inferior a 0,05 (Teste de Visualização, CFI=0,93, RMR=0,01, $\chi^2=1145,97$, gl=361; Teste de Fechamento Figural, CFI=0,94, RMR=0,01, $\chi^2=839,65$, gl=415; Teste de Memória Visual, CFI=0,96, RMR=0,01, $\chi^2=81,51$, gl=41), mostrando bom ajuste aos dados. Os três testes apresentaram a mesma estrutura fatorial (Figura 4). Em todos há a presença de alguns fatores de primeiro nível e a presença de um fator geral que interage diretamente com os itens. Interpretou-se o fator geral identificado em cada um dos testes como a habilidade específica do teste em questão. No caso do Teste de Visualização (Vz), por exemplo, o fator geral é a habilidade de Visualização (Vz). Por sua vez, interpretou-se que os fatores primários obtidos em cada um dos testes podem ser tanto processos cognitivos específicos que envolvem a habilidade cognitiva do próprio teste, aspectos contextuais envolvidos na resolução dos itens do teste, como outros processos cognitivos concorrentes e não

relacionados à habilidade cognitiva do próprio teste.

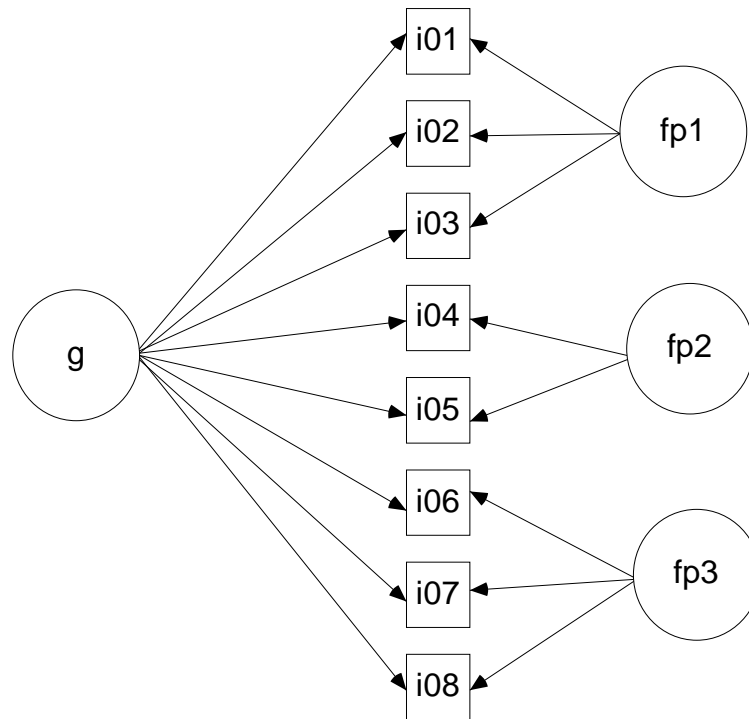


Figura 4. Estrutura fatorial comum encontrada em todos os testes do conjunto de testes.
 Legenda: g=fator geral; fp1=fator primário 1; fp2=fator primário 2; fp3=fator primário 3; i=item.

Todos os três testes do Conjunto mostraram-se unidimensionais. Há evidências favoráveis a essa afirmação obtidas a partir da análise fatorial e da análise do modelo Rasch. O Teste de Visualização apresenta características de pronunciada unidimensionalidade. Todos os seus itens apresentaram uma carga superior a 0,3 em relação ao fator geral, que explicou 74,38% da variância comum (Tabela 1). Em relação à escala unidimensional gerada pelo modelo Rasch, ela apresentou boa fidedignidade das estimações dos itens e das pessoas, 0,99 e 0,89, respectivamente. O índice *infít* médio foi de 0,98 e desvio padrão de 0,21. Apenas os itens nove, 23 e 25 (1,38, 1,40 e 0,56, respectivamente) não estiveram entre os valores de 0,7 a 1,3 do *infít*, não se adequando

muito bem à escala. Apesar disso, por estarem na faixa produtiva para medida, $0,5 < infit < 1,5$ (Wright & Linacre, 1994), contribuindo para a mensuração da habilidade-alvo, e por apresentarem cargas fatoriais relevantes junto ao fator geral, eles foram mantidos no teste. Houve boa distribuição dos itens quanto ao seu grau de dificuldade (Figura 5): a média da dificuldade dos itens foi 0,00, o desvio padrão 1,55, a dificuldade máxima foi 3,22 e a mínima -3,01. O Teste de Visualização foi fácil para a amostra utilizada, sendo a média da habilidade 2,03. Apesar disso, as pessoas foram bem discriminadas: o desvio padrão 2,50, enquanto a habilidade máxima alcançou 5,71 e a mínima -5,58. Apenas 75 pessoas alcançaram o escore máximo e três tiveram o escore mínimo.

Tabela 1. Estrutura fatorial do Teste de Visualização

	g	fp1	fp2	fp3
i01	0,35	-0,02	0,65	0,00
i02	0,46	0,00	0,58	0,00
i03	0,59	0,00	0,35	0,00
i04	0,47	0,00	0,55	0,00
i05	0,56	0,00	0,45	0,00
i06	0,45	0,00	0,41	0,00
i07	0,47	0,13	0,00	0,28
i08	0,60	0,00	0,30	0,00
i09	0,46	0,00	0,25	0,15
i10	0,46	0,00	0,47	0,00
i11	0,60	0,10	0,22	0,00
i12	0,59	-0,07	0,23	0,00
i13	0,34	0,00	0,00	0,69
i14	0,38	0,00	0,00	0,74
i15	0,59	0,00	0,18	0,22
i16	0,79	-0,04	0,15	0,00
i17	0,66	0,07	0,00	0,25
i18	0,81	-0,03	0,10	0,00
i19	0,82	-0,07	0,00	0,00
i20	0,58	0,04	0,00	0,17
i21	0,62	-0,01	0,00	0,00
i22	0,85	-0,07	0,00	0,00
i23	0,60	0,10	0,00	0,00
i24	0,78	0,05	0,00	0,00
i25	0,89	-0,10	0,00	-0,13
i26	0,73	0,36	0,00	0,00
i27	0,68	0,32	0,00	0,18
i28	0,82	0,35	0,00	0,00
i29	0,76	0,47	0,00	0,00
i30	0,78	0,39	0,00	0,00
% comun.	74,38	4,88	12,75	7,98

Legenda: i=item, g=fator geral, fp=fator primário, %comun.=percentual da comunalidade.

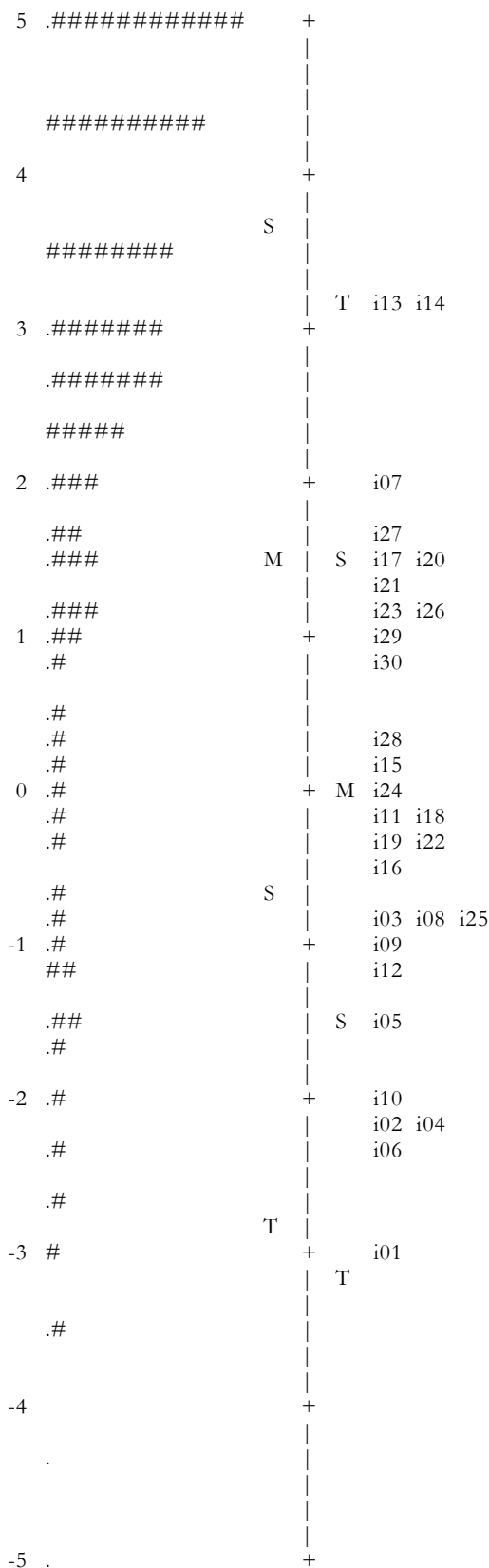


Figura 5. Distribuição dos itens e pessoas do Teste de Visualização pelo Modelo Rasch. Legenda: # = seis pessoas.

Semelhante ao padrão encontrado no Teste de Visualização (Vz), todos os itens do Teste de Flexibilidade de Fechamento (CF) apresentaram

uma carga superior a 0,3 em relação ao fator geral, que explicou 68,98% da variância comum (Tabela 2). A escala unidimensional gerada pelo modelo

Rasch apresentou boa fidedignidade das estimações dos itens e das pessoas, 0,98 e 0,89, respectivamente. O índice *infít* médio dos itens em relação à escala foi de 1,00 e desvio padrão de 0,12. Todos os itens estiveram entre os valores de 0,7 a 1,3, adequando-se à escala. Houve distribuição dos itens aceitável quanto ao seu grau de dificuldade (Figura 6): a média da dificuldade dos itens foi 0,00, o desvio padrão, 1,22, a dificuldade máxima foi 1,55 e a mínima -2,58. O teste de CF foi fácil para a amostra utilizada, sendo a média da habilidade das pessoas 1,57. No entanto, as pessoas foram bem discriminadas: o desvio padrão foi 2,40, enquanto a habilidade máxima alcançou 5,19 e a mínima, -5,44. Apenas cinco pessoas obtiveram o escore mínimo,

enquanto 69 pessoas atingiram o escore máximo. Apesar de todas as evidências favoráveis à unidimensionalidade, a distribuição do grau de dificuldade dos itens indica a necessidade de elaboração de itens mais difíceis para a amostra desta pesquisa, tendo em vista incorporar mais itens pertencentes ao intervalo dois e três desvios padrão da escala. Uma estratégia possível para a elaboração de itens mais difíceis diz respeito à elaboração de itens com figuras formadas com mais pontos. Todos os itens do teste apresentam uma figura-modelo formada mediante a ligação de cinco pontos. Itens com figuras-modelo formadas por seis ou sete pontos podem ampliar o grau de complexidade e aumentar a dificuldade do teste.

Tabela 2. Estrutura fatorial do Teste de Fechamento Figural

	g	fp1	fp2	fp3	fp4	fp5
i01	0,45	0,00	0,00	0,00	0,51	0,00
i02	0,47	0,00	0,00	0,00	0,41	0,00
i03	0,43	0,00	0,00	0,45	0,00	0,00
i04	0,62	0,00	0,00	0,00	0,20	0,00
i05	0,68	0,00	0,27	0,00	0,02	0,00
i06	0,62	0,00	0,16	0,00	0,00	0,32
i07	0,65	0,25	0,17	-0,24	0,00	0,00
i08	0,74	0,32	0,00	0,00	0,00	0,00
i09	0,48	0,00	0,00	0,00	0,54	0,00
i10	0,43	0,00	0,11	0,00	0,34	0,00
i11	0,37	0,00	0,00	0,66	0,00	0,00
i12	0,60	0,00	0,32	0,00	0,00	0,00
i13	0,63	0,00	0,17	0,00	0,00	0,00
i14	0,49	0,22	0,00	0,29	0,00	0,38
i15	0,57	0,27	0,00	0,00	0,00	0,00
i16	0,73	0,39	-0,33	0,00	0,00	0,00
i17	0,66	0,00	0,09	0,00	0,12	0,00
i18	0,31	0,00	0,00	0,60	0,00	0,00
i19	0,68	0,00	0,14	0,00	0,00	-0,31
i20	0,66	0,00	0,28	0,00	0,00	0,00
i21	0,58	0,00	0,28	0,00	0,00	0,00
i22	0,71	0,28	0,14	-0,20	0,12	0,00
i23	0,62	0,59	0,31	0,00	0,00	0,00
i24	0,58	0,19	0,00	0,00	0,00	0,00
i25	0,48	0,00	0,19	0,40	0,00	0,00
i26	0,53	0,00	0,00	0,36	0,18	0,00
i27	0,63	0,00	0,19	0,00	0,00	0,00
i28	0,53	0,00	0,38	0,00	0,00	0,00
i29	0,68	0,00	0,25	0,00	0,00	0,00
i30	0,63	0,25	0,18	0,00	0,00	0,00
i31	0,64	0,46	0,00	0,00	0,00	0,00
i32	0,66	0,31	0,00	0,00	0,00	0,00
%comun.	68,98	7,88	6,13	9,13	5,76	2,12

Legenda: i=item, g=fator geral, fp=fator primário, %comun.=percentual da comunalidade.

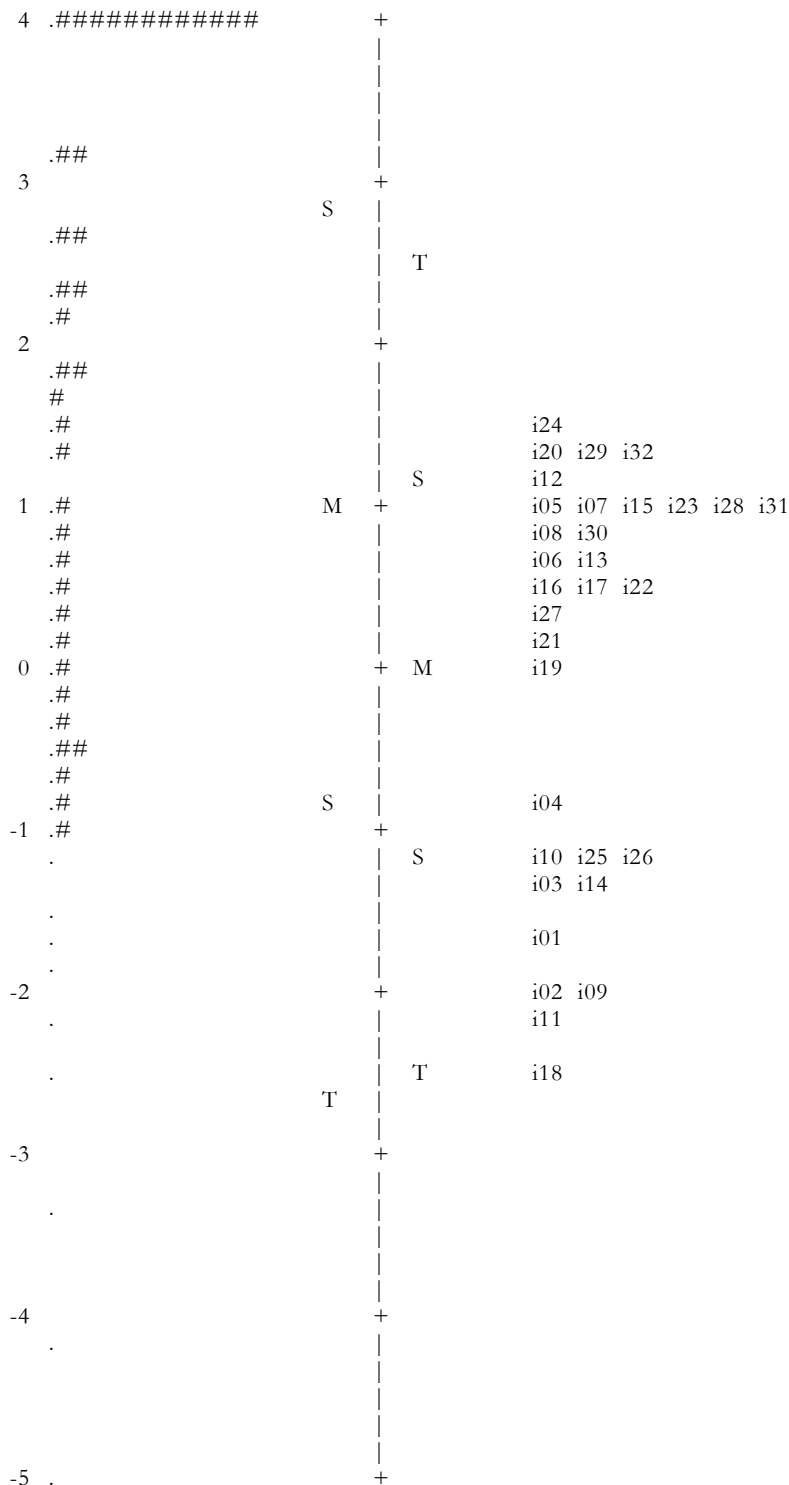


Figura 6. Distribuição dos itens e pessoas do Teste de Flexibilidade de Fechamento pelo Modelo Rasch. Legenda: #=oito pessoas.

O Teste de Memória Visual (MV) também é explicado preponderantemente pelo fator geral (53,32% da comunalidade). Excetuando os itens seis e oito, todos os outros apresentam carga igual ou superior a 0,3 em relação ao fator geral (Tabela 3). A escala gerada pelo modelo Rasch apresenta boa fidedignidade da estimação dos itens (0,96) e

uma fidedignidade da estimação das pessoas insatisfatória (0,41). A fidedignidade apurada pelo coeficiente KR-20 foi 0,77. A diferença entre os dois resultados pode ser devida ao alto percentual de escores extremos na amostra: 37,95% dos sujeitos obtiveram o escore máximo no teste. Segundo Linacre (2006), grandes diferenças entre

os valores da fidedignidade apurada pelos escores brutos (KR-20) e pelas medidas devem-se à presença de casos extremos ou de dados faltantes. O índice *infit* médio dos itens em relação à escala foi de 1,00 e desvio-padrão de 0,15. Todos os itens estiveram entre os valores de 0,7 a 1,3, adequando-se à escala. Houve distribuição dos itens aceitável quanto ao seu grau de dificuldade (Figura 7): a média da dificuldade dos itens foi 0,00, o desvio padrão, 0,70, a dificuldade máxima foi 1,0 e a mínima, -1,7. O teste MV foi fácil para a amostra utilizada: a média da habilidade foi 2,33 e o desvio

padrão, 1,52, enquanto a habilidade máxima alcançou 3,84 e a mínima, -1,77. O Teste de Memória Visual possui uma distribuição dos itens que aponta para a necessidade de elaboração de itens mais difíceis (Figura 7). Uma estratégia possível para a resolução desse problema é a elaboração de mapas que possuam mais detalhes a serem percebidos e memorizados. Outra possibilidade é a elaboração de mapas na página de resposta que possuam uma forte semelhança com os mapas da página de memorização, mas que não sejam iguais aos mesmos.

Tabela 3. Estrutura fatorial do Teste de Memória Visual

	g	fp1	fp2
i01	0,45	0,00	-0,13
i02	0,60	0,23	0,12
i03	0,33	0,00	0,95
i04	0,37	0,66	0,00
i05	0,65	0,19	0,00
i06	0,24	0,00	0,11
i07	0,58	0,28	0,13
i08	0,15	0,00	0,07
i09	0,33	0,00	0,10
i10	0,34	0,08	0,00
i11	0,45	0,55	0,00
i12	0,41	0,28	0,00
%comun.	53,32	23,54	23,13

Legenda: i=item, g=fator geral, fp=fator primário, %comun.=percentual da comunalidade.

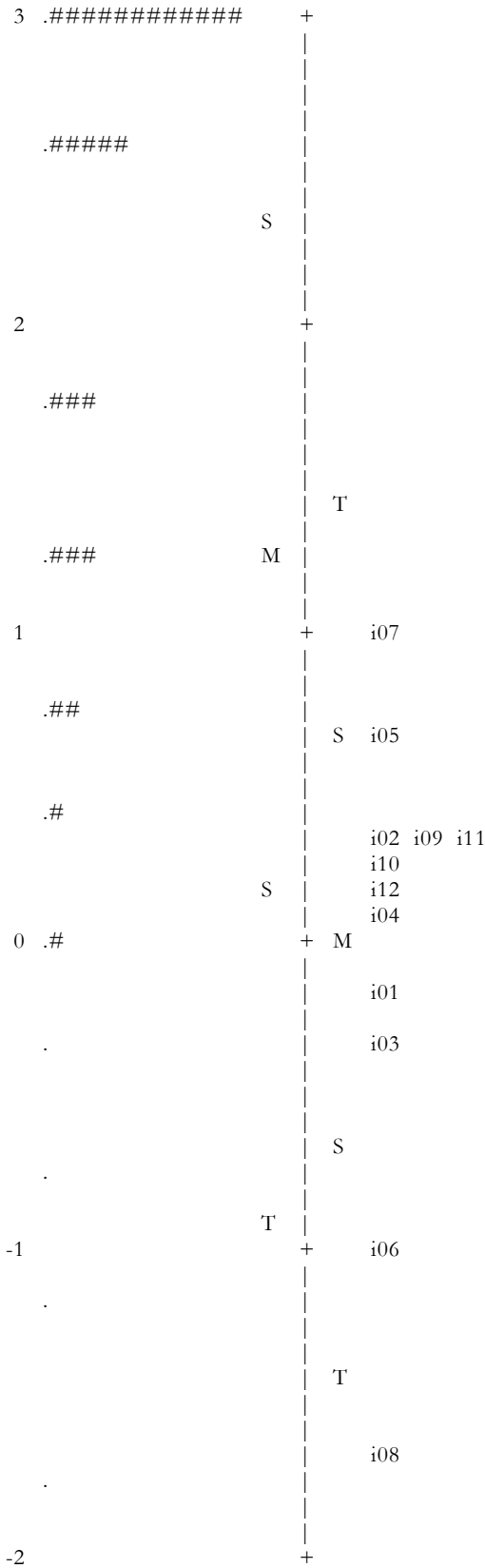


Figura 7. Distribuição dos itens e pessoas do Teste de Memória Visual pelo Modelo Rasch. Legenda: #=dezessete pessoas.

Com relação à análise da validade convergente e à validade discriminante do Conjunto de Teste da Habilidade Visuo Espacial, foi realizada uma análise fatorial confirmatória abrangendo os três testes da Habilidade Visuo Espacial e três testes de Fluência (Gr). A análise

fatorial confirmatória foi feita a partir da pontuação total dos participantes nos seis testes, e a solução encontrada apresentou um ótimo ajuste em relação aos dados, com um CFI de 1,00 e um RMSEA de 0,00.

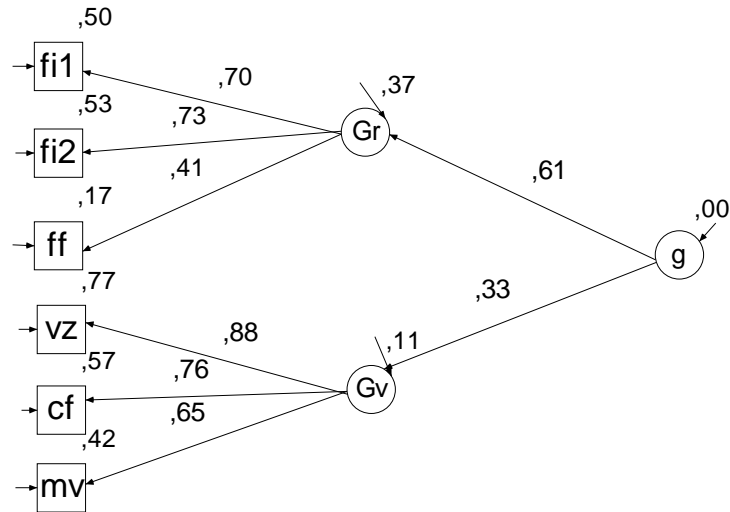


Figura 8. Solução fatorial dos seis testes.

Conforme pode ser observado na Figura 8, todos os testes do Conjunto de Testes da Habilidade Visuo espacial convergem na mensuração da habilidade-alvo, com uma carga fatorial considerável. Além disso, nenhum dos testes da Habilidade Visuo espaciais e relaciona com a fluência, de modo que se pode argumentar favoravelmente a respeito da validade convergente e da validade discriminante do Conjunto de Testes da Habilidade Visuo Espacial (Gv).

Considerações finais

Neste trabalho foram avaliadas as propriedades psicométricas do Conjunto de Testes da Habilidade Visuo Espacial (Gv). As principais limitações do trabalho advêm da forma de seleção da amostra utilizada. Formada por três grupos, a amostra foi delimitada pela oportunidade, acarretando: (i) um certo viés amostral e (ii) uma limitação de dados demográficos, socioeconômicos e culturais. Quanto à primeira limitação apontada, há um grupo predominante de participantes formado por estudantes de ensino médio de uma escola pública federal, com um processo seletivo de ingresso muito competitivo. Os estudantes desse grupo possuem, na média, habilidades maiores do que a média do conjunto

de adolescentes brasileiros. Esta limitação impõe cautela tanto na generalização dos resultados para o conjunto da população brasileira quanto à segunda limitação descrita. Em virtude de certas circunstâncias da coleta de dados, não foi possível obter dados demográficos, socioeconômicos e culturais de todos os participantes. No entanto, apesar dessas limitações, o trabalho apresenta contribuições substantivas.

A avaliação das propriedades psicométricas do Conjunto de Testes da Habilidade Visuo Espacial (Gv) realizada, enfatizou a análise da unidimensionalidade de cada teste e a validade convergente e discriminante do conjunto de testes. Analisou-se também a distribuição dos itens de cada teste, quanto ao seu grau de dificuldade. Ao utilizarem-se diferentes procedimentos estatísticos, foi possível triangular dados distintos e gerar evidências que corroboravam ou amenizavam evidências de outros procedimentos. Uma contribuição deste estudo refere-se à utilização combinada da análise fatorial exploratória e da análise fatorial confirmatória para a identificação da solução fatorial adequada de um teste. Outra contribuição refere-se ao uso combinado da análise fatorial e do modelo Rasch para análise da unidimensionalidade. Os valores do *infit* do modelo Rasch indicam se o item de algum teste

funciona bem dentro da escala gerada. Por outro lado, a carga fatorial de cada item relacionada ao fator geral encontrado, assim como o percentual da comunalidade explicado pelo fator geral, indicam o quanto o teste é explicado pela dimensão-alvo, assim como a magnitude da relação entre o item e essa dimensão. Esses dados se complementam e informam aspectos distintos da unidimensionalidade, de modo que a sua combinação favorece uma análise mais ampla sobre o comportamento dos itens e sua adequação.

O Conjunto de Testes da Habilidade Visuo Espacial (Gv) é válido para a amostra deste estudo. Seus testes são pronunciadamente unidimensionais, mensuram de forma convergente a habilidade-alvo, assim como não participam da mensuração de outra habilidade cognitiva. Apesar disso, novos estudos são necessários, pois a distribuição dos itens de alguns testes não alcança a amplitude necessária quanto ao grau de dificuldade. Ao mesmo tempo, é necessário que as evidências de unidimensionalidade possam ser também identificadas em amostras mais amplas e diversificadas através de novos estudos.

A habilidade visuo espacial também é mensurada pela bateria Woodcock-Johnson III (Wechsler & Schelini, 2006). Enquanto essa bateria mensura Gv através de um único teste, o conjunto aqui analisado permite mensurar Gv por meio de três habilidades do primeiro estrato da teoria CHC. Esta característica, está de acordo com a tradição de que um fator seja identificado por, no mínimo, três indicadores (Carroll, 1993). Além disso, o conjunto de testes proposto para mensurar Gv pode ser aplicado em uma sessão da ordem de 45 minutos, o que é adequado para propósitos de pesquisa e aplicados, por exemplo, na psicologia escolar. Por ter tais características, o Conjunto de Testes da Habilidade Visuo Espacial (Gv) poderá vir a se tornar uma ferramenta para as pesquisas brasileiras no campo dos estudos sobre a habilidade visuo espacial.

Referências

- Beauducel, A. (2001). Problems with parallel analysis in data sets with oblique simple structure. *Methods of Psychological Research Online*, 6 (2). Obtido do World Wide Web em 25 de janeiro de 2007, <http://www.dgps.de/fachgruppen/methoden/mpr-online/issue14/art2/beauducel.pdf>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bickley, P. G., Keith, T. Z. & Wolfle, L. M. (1995). The three-stratum theory of cognitive abilities: test of the structure of intelligence across the life span. *Intelligence*, 20, 309-328.
- Brandt, M. G. & Davies, E. T. (2006). Visual-spatial ability, learning modality and surgical knot tying. *Canadian Journal of Surgery*, 49(6), 412-416.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. Em K. A. Bollen & J. S. Long (Orgs.). *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Burton, L. J. & Fogarty, G. J. (2003). The factor structure of visual imagery and spatial abilities. *Intelligence*, 31, 289-318.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Carroll, J. B. (1993). *Human cognitive abilities: a survey of factor analytic studies*. Nova Iorque: Cambridge University Press.
- Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: current evidence supports g and about ten broad factors. Em H. Nyborg (Org.). *The scientific study of general intelligence: tribute to Arthur R. Jensen* (pp. 1-20). Oxford, UK: Elsevier Science/Pergamon Press.
- Cole, J. C. & Randall, M. K. (2003). Comparing the cognitive ability models of Spearman, Horn and Cattell, and Carroll. *Journal of Psychoeducational Assessment*, 21, 160-179.
- Ekstrom, R. B., French, J. W., Harman, H. H. & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Frazier, T. W. & Youngstrom, E. A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: are we overfactoring? *Intelligence*, 35, 169-182.
- Gomes, C. M. A. (2005). *Uma análise dos fatores cognitivos mensurados pelo Exame Nacional do Ensino Médio (ENEM)*. Tese de doutorado.

- Belo Horizonte, MG: Universidade Federal de Minas Gerais – Faculdade de Educação.
- Guadagnoli, E. & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103* (2), 265-275.
- Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179-203.
- Hedman, L., Ström, P., Andersson, P., Kjellin, A., Wredmark, T. & Felländer-Tsai, L. (2006). High-level visual-spatial ability for novices correlates with performance in a visual-spatial complex surgical simulator task. *Surgical Endoscopy and Other Interventional Techniques*, *20* (8), 1275-1280.
- Henson, R. K. & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*(3), 393-416.
- Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. Em R. H. Hoyle (Org.). *Structural equation modeling: concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55.
- Johnson, W. & Bouchard Jr., T. J. (2005a). Constructive replication of the visual perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability. *Intelligence*, *33*, 417-430.
- Johnson, W. & Bouchard Jr., T. J. (2005b). The structure of human intelligence: it is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393-416.
- Johnson, W., Nijenshuis, J. & Bouchard Jr., T. J. (2007). Replication of the hierarchical visual-perceptual-image rotation model in de Wolff and Buitens's (1963) battery of 46 tests of mental ability. *Intelligence*, *35*, 69-81.
- Juhel, J. (1991). Spatial abilities and individual differences in visual information processing. *Intelligence*, *15*(1), 117-137.
- Linacre J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16* (2), 878.
- Marañón, R. C. & Andrés-Pueyo, A. (2000). The study of human intelligence: a review at the turn of the millennium. *Psychology in Spain*, *4* (1), 167-182.
- 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), 1-10.
- Nilsson, T., Hedman, L. & Ahlqvist, J. (2007). Visual-spatial ability and interpretation of three-dimensional information in radiographs. *Dentomaxillofacial Radiology*, *36*(2), 86-91.
- Primi, R. (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica*, *1*, 67-77.
- Schelini, P. W. (2006). Teoria das inteligências fluida e cristalizada: início e evolução. *Estudos de Psicologia (Natal)*, *11*(3), 323-332.
- Schelini, P. W. & Wechsler, S. M. (2005). Bateria multidimensional de inteligência infantil: desenvolvimento de instrumento. *Psico-USF*, *10* (2), 129-139.
- Schelini, P. W. & Wechsler, S. M. (2006). Estudo da estrutura fatorial da Bateria Multidimensional de Inteligência Infantil. *Estudos de Psicologia (Campinas)*, *23*, 105-112.
- Sisto, F. F., Marin Rueda, F. J. & Bartholomeu, D. (2006). Estudo sobre a unidimensionalidade do teste Matrizes Progressivas Coloridas de Raven. *Psicologia: Reflexão e Crítica*, *19*(1), 66-73.
- Sisto, F. F., Santos, A. A. A. & Noronha, A. P. P. (2007). Estudo sobre a dimensionalidade do teste R1 – Forma B. *Psicologia em Estudo*, *12*(1), 185-193.
- Spearrit, D. (1996). Carroll's model of cognitive abilities: educational implications. *International Journal of Educational Research*, *25*(2), 107-198.
- Stumpf, H. & Eliot, J. (1999). A structural analysis of visual spatial ability in academically talented students. *Learning and Individual Differences*, *11* (2), 137-151.
- Tennant, A. & Pallant, J. F. (2006). Unidimensionality matters! (A tale of two Smiths?). *Rasch Measurement Transactions*, *20*(1), 1048-51.
- Undheim, J. O. & Gustafsson, J. E. (1987). The hierarchical organization of cognitive abilities: restoring general intelligence through the use

- of Linear Structural Relations (LISREL). *Multivariate Behavioral Research*, 22, 149-171.
- Vitória, F., Almeida, L. S. & Primi, R. (2007). Unidimensionalidade em testes psicológicos: conceito, estratégias e dificuldades na sua avaliação. *Revista de Psicologia da Vetor Editora*, 7 (1), 1-7.
- Wanzel, K. R., Hamstra, S. J., Caminiti, M. F., Anastakis, D. J., Grober, E. D. & Reznick, R. K. (2003). Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery*, 134(5), 750-757.
- Wechsler, S. M. & Schelini, P. W. (2006). Bateria de habilidades cognitivas Woodcock-Johnson III: validade de construto. *Psicologia: Teoria e Pesquisa*, 22, 287-295.
- Wright, B. D. & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8 (3), 370.

Recebido em junho de 2008
Reformulado em dezembro de 2008
Aprovado em janeiro de 2009

Sobre os autores:

Cristiano Mauro Assis Gomes é psicólogo, doutor em Educação, professor adjunto do Departamento de Psicologia da Universidade Federal de Minas Gerais e trabalha no campo da psicologia educacional e do desenvolvimento humano, avaliação psicológica e educacional.

Oto Neri Borges é físico, doutor em Física, professor associado do Departamento de Física da Universidade Federal de Minas Gerais, professor do Colégio Técnico de Minas Gerais e Coordenador do Programa de Pós-Graduação em Educação da Universidade Federal de Minas Gerais. Trabalha na área do ensino de ciências, cognição e avaliação educacional.