

QUALIDADES PSICOMÉTRICAS DO CONJUNTO DE TESTES DE INTELIGÊNCIA FLUIDA

Cristiano Mauro Assis Gomes - Universidade Federal de Minas Gerais
Oto Borges - Universidade Federal de Minas Gerais

RESUMO

Este artigo analisa as qualidades psicométricas do Conjunto de Testes de Inteligência Fluida, em termos de sua estrutura fatorial, unidimensionalidade, validade convergente e discriminante. Fizeram parte da pesquisa 558 participantes que cursavam ou concluíram o ensino médio. Foram utilizados procedimentos de análise fatorial exploratória e confirmatória para identificação da estrutura fatorial de cada teste, assim como para análise da validade convergente e validade discriminante. O modelo Rasch foi utilizado para análise da unidimensionalidade de cada teste. Os resultados apontaram que os testes possuem uma estrutura fatorial complexa, embora ela possa ser considerada uma estrutura unidimensional. Os índices de ajuste dos modelos fatoriais foram adequados aos dados, e os testes apresentaram boa distribuição quanto à dificuldade dos itens. Quanto à validade convergente e validade discriminante, os testes convergem para mensurar a Inteligência Fluida e divergem na mensuração de uma habilidade cognitiva contrastante, a Fluência, utilizada neste estudo como uma habilidade contrastante.

Palavras-chave: validade do teste; psicométrica; medidas de inteligência.

PSYCHOMETRICAL PROPERTIES ANALYSIS OF FLUID INTELLIGENCE TESTS KIT

ABSTRACT

The article analyses psychometrical properties of the Fluid Intelligence Tests Kit, by means of its factorial structure, unidimensionality, convergent validity and divergent validity. The sample was formed by 558 students that were either attending High School or had just concluded it. Exploratory and confirmatory factorial analyses were used to identify the factorial structure of the tests and to analyze their convergent and divergent validities. The Rasch model was used to analyze the unidimensionality of each test. The results showed that the tests have a complex factorial structure although it can be take as an unidimensional structure. The factorial model fit was good and tests showed good item difficulties distribution. As for the convergent and divergent validity, the tests converge to measure the Fluid Intelligence and diverge to measure a diverse cognitive ability, the Fluency ability, used here as a contrast ability.

Keywords: test validity; psychometrics; intelligence measurement.

INTRODUÇÃO

Desde meados da década de 1990 tem crescido a aceitação da idéia de que as habilidades cognitivas são dispostas em uma arquitetura hierárquica de níveis distintos, cada um deles correspondendo a diferentes graus de abrangência e especialização (Primi, 2003; Schelini, 2006). O primeiro nível da hierarquia indica habilidades bem especializadas e específicas, enquanto os níveis superiores indicam, progressivamente, habilidades cada vez mais genéricas e gerais (Bickley, Keith & Wolfle, 1995; Carroll, 1993; Cole & Randall, 2003; Johnson & Bouchard Jr., 2005a; Johnson & Bouchard Jr., 2005b; Johnson, Nijenhuis & Bouchard Jr., 2007; Marañón & Andrés-Pueyo, 2000; Primi, 2003; Schelini, 2006).

As habilidades cognitivas de níveis hierárquicos superiores da arquitetura intelectual têm interessado a vários pesquisadores por mostrarem bom poder preditivo em relação a diversos desempenhos humanos, como o

rendimento escolar e a performance profissional (Garcia & Noronha, 2007; McGrew, Keith, Flanagan & Vanderwood, 1997; Primi, Santos & Vendramini, 2002; Spearrit, 1996). Uma destas habilidades de alta-ordem é o Fator Geral (g), o construto mais estudado em termos de sua natureza, poder preditivo e conexões com outras variáveis cognitivas e socioculturais (Colom, Abad, Rebollo & Shih, 2005; Colom, Rebollo, Palácios, Juansepina & Kyllonen, 2004; Flores-Mendoza & Nascimento, 2001; Haier, White & Alkire, 2003; Helms-Lorenz, van de Vijver & Poortinga, 2003; Luo, Thompson & Detterman, 2003; Pasquali, 2002).

Apesar da importância das habilidades de alta-ordem (níveis hierárquicos que estão acima do primeiro nível), segundo a avaliação de Carrol (2003), naquela ocasião não existia nenhuma bateria cognitiva, em nível nacional e internacional, capaz de mensurar de forma equilibrada e integral essas habilidades. As baterias existentes mensuram de forma privilegiada algumas habilidades

cognitivas em detrimento de outras (Carroll, 2003; Roberts & cols, 2000; Tirre & Field, 2002; Watkins, Wilson, Kotz, Carbone & Babula, 2006). No Brasil, a carência é ainda maior (Wechsler & Schelini, 2006). Merecem destaques a validação do WJIII (Wechsler & Schelini, 2006) e a adaptação brasileira dos testes verbais do WJIII (Wechsler, Vendramini & Schelini, 2007), a bateria multidimensional para a inteligência infantil (Schelini & Wechsler, 2005, 2006) e o instrumento para a avaliação dinâmica da inteligência fluída (Primi, Cruz, Muniz & Petrini, 2006). Apesar destes esforços, faltam pequenas baterias de testes aprovados pelo CFP capazes de mensurar cada uma das habilidades de alta-ordem do campo psicométrico. Uma consulta à lista de testes aprovados no SATEPSI (www2.pol.org.Br/satepsi/sistema/admin.cfm?lista1=sim), atualizada até maio de 2008, revela que não há baterias aprovadas pelo CFP para medir a inteligência cristalizada (Gc); fluída (Gf), memória e aprendizagem (Gy), rapidez cognitiva (Gs) e outros fatores de alta ordem. Essa lacuna, no entanto, não se refere exclusivamente à psicometria, mas também a construtos de outras tradições que investigam a inteligência, como é o caso da psicologia do desenvolvimento e da teoria do processamento da informação (Adey, Csapó, Demetriou, Hautamäki & Shayer, 2007; Gardner, 2006; Sternberg, 2006).

Com relação à relevância da mensuração da Inteligência Fluida (Gf), alguns aspectos merecem atenção. O primeiro deles refere-se à própria definição da Inteligência Fluida (Gf) e sua relevância para o mundo contemporâneo (Primi, 2002; Primi & Almeida, 2002). Delimitada conceitualmente por Cattell nos anos de 1940, e identificada empiricamente por uma série de pesquisadores do campo psicométrico (Carroll, 1993; Cattell, 1971; Gustafsson, 1984; Horn, 1968; Spearrit, 1996), a Inteligência Fluida (Gf) é entendida como a capacidade das pessoas em aprender coisas novas que requerem relações lógicas e abstratas, assim como a capacidade em resolver problemas onde o conhecimento prévio é incipiente. Envolvido por um largo conjunto de atividades que demandam processos lógico-abstratos e situações onde o conhecimento prévio não é suficiente, o mundo atual parece requerer de forma significativa a Inteligência Fluida (Gf).

Um segundo aspecto que expõe a relevância da mensuração da Inteligência Fluida (Gf) diz respeito ao exposto de que no Brasil faltam

baterias específicas já aprovadas pelo CFP para identificar essa habilidade. Um outro aspecto envolve uma questão investigativa que trata da relação entre a Inteligência Fluida (Gf) e o Fator Geral (g) (Gustafsson, 1984; Undheim & Gustafsson, 1987). Segundo os estudos de Carroll (1993), a Inteligência Fluida (Gf) é a habilidade cognitiva de alta-ordem que mais se aproxima em termos correlacionais com o Fator Geral (g) de inteligência. Essa condição torna, por consequência, a mensuração da Inteligência Fluida um desafio, na medida em que boa parte dos testes marcadores desse construto carrega fortemente o Fator Geral (g), em detrimento de Gf (Carroll, 2003). Outro aspecto sobre a relevância da mensuração da Inteligência Fluida (Gf) refere-se às articulações entre esse construto teórico e o desenvolvimento cognitivo. Determinados estudos apontam para uma trajetória de desenvolvimento da Inteligência Fluida em função da idade e da maturação biológica (Almeida, Lemos, Guisande & Primi, 2008; McArdle, Hamagami, Meredith & Bradway, 2000; Salthouse, 2004). No trabalho de Almeida e seus colaboradores (2008) foram aplicadas provas cognitivas a uma amostra de 1835 estudantes das sétima, oitava e nona séries de Portugal. Os resultados indicam que o desempenho nos testes de habilidades cognitivas variou com a série e a idade, e segundo os autores a série tem maior impacto sobre a diferenciação cognitiva dos sujeitos pesquisados do que a idade.

Além dos pontos comentados, pode-se também destacar a investigação dos componentes específicos que compõem a Inteligência Fluida. Primi, Cruz, Nascimento e Petrini (2006) apresentam argumentos e evidências favoráveis a respeito da participação de dois componentes específicos para a explicação dos mecanismos da Inteligência Fluida: (1) o gerenciamento executivo e (2) a atenção seletiva e controle inibitório. O trabalho dos autores pauta-se em argumentos a favor da associação entre a Memória de Trabalho e a Inteligência Fluida. Por outro lado, apesar desses resultados é importante se considerar os argumentos de Gignac (2007), sobre os resultados de certos estudos internacionais de associação entre a Memória de Trabalho e a Inteligência, que segundo ele divulgam resultados equivocados, por inflarem as correlações entre esses dois construtos, em função do fator do procedimento de atenuação dos erros de estimativa realizados pela análise fatorial confirmatória

Por último, não se pode desconsiderar os argumentos de Johnson e Bouchard Jr. (2005a, 2005b) e de Johnson, Nijenhuis e Bouchard Jr. (2007), contrárias à existência da Inteligência Fluída. Os resultados dessas pesquisas apontam que os modelos de Cattell-Horn e de Carroll são inadequados, principalmente por considerarem a presença da Inteligência Fluída e da Inteligência Cristalizada. Em virtude da relevância do problema apontado, a falta de baterias elaboradas especificamente para a mensuração da Inteligência Fluída dificulta uma análise mais detalhada do tema em questão.

Este artigo pretende investigar as qualidades psicométricas de uma bateria brasileira elaborada para mensurar uma das habilidades cognitivas de alta-ordem: a Inteligência Fluída (Gf). A bateria analisada é o Conjunto de Testes de Inteligência Fluída, formado pelo Teste de Indução (I), Teste de Raciocínio Lógico (RL) e o Teste de Raciocínio Geral (RG), elaborados pelo primeiro autor deste artigo a partir de sua experiência com a tradução, adaptação e validação de 45 testes cognitivos do *Kit of Factor-Referenced Cognitive Tests* do *Educational Testing Service* (Ekstrom, French, Harman & Dirmen, 1976). Ela é uma dentre quatro baterias específicas desenhadas com o intuito de desenvolver uma bateria brasileira para mensurar fatores cognitivos de alta-ordem. As outras baterias específicas pretendem mensurar a Inteligência Cristalizada, a Habilidade Visuo-Espacial e a Habilidade de Memória.

Visando analisar as qualidades psicométricas do Conjunto de Testes de Inteligência Fluída serão investigadas a unidimensionalidade e a estrutura fatorial de cada um dos testes do Conjunto. Além disso, os testes do Conjunto serão avaliados em termos de sua validade convergente e validade discriminante. A influência da idade e da série não são investigadas, bem como não são investigados componentes específicos de cada teste da bateria. O foco do trabalho é a identificação de bons testes marcadores da inteligência fluída visando o desenvolvimento de uma bateria de fatores de alta ordem.

MÉTODO

Participantes

Fizeram parte do estudo 558 estudantes. Desse total, 384 (68,8%) eram alunos de uma escola da rede federal de ensino médio 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 destes 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.

Instrumentos

Conjunto de Testes de Inteligência Fluída

Teste de Indução (I):

Composto por 15 itens e um tempo limite de 14 minutos para sua execução. Cada item é composto por cinco grupos de quatro letras. Entre os cinco grupos há quatro grupos que apresentam um mesmo padrão, uma mesma regra de organização de suas letras. O respondente deve identificar o grupo que não apresenta esse padrão e marcá-lo com um x.

Teste de Raciocínio Lógico (RL):

O Teste de Raciocínio Lógico (RL) é composto por 30 itens e um tempo limite de 24 minutos para sua execução. Cada item é formado por uma conclusão proveniente de duas premissas lógicas que não têm nenhuma relação com o mundo. O objetivo do respondente é indicar se a conclusão presente no item é adequada ou inadequada em relação às premissas do item. O respondente deve marcar uma de duas opções dadas.

Teste de Raciocínio Geral (RG):

É composto por 15 itens e um tempo limite de 18 minutos para sua execução. Cada item é formado por um problema lógico-matemático, composto por um enunciado e um espaço para sua resolução. O respondente deve interpretar o enunciado, resolver o problema, e escolher uma das cinco opções de respostas do conjunto de múltiplas-escolhas.

Conjunto de Testes de Fluência (Gr)

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 o desenho em cada camiseta seja único.

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 idéias relacionadas a um tópico pré-determinado. Um exemplo de tópico é “Uma viagem de trem”. O escore é formado pelo número de idéias 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 pré-determinada. Um exemplo de categoria é “Objetos Vermelhos”. O escore é formado pelo número de objetos escritos relacionados com a categoria.

Coleta e Análise de Dados

Os testes foram aplicados conjuntamente, de forma coletiva, e sempre por psicólogos ou estudantes de psicologia devidamente treinados. Os testes foram aplicados em 31 turmas escolares, sendo duas sessões por turma, com tempo médio de 100 minutos por sessão. Além dos dois conjuntos de testes aqui relatados, foram aplicados 12 outros testes nas duas sessões. 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).

Quanto ao tratamento dos dados, para a identificação da estrutura fatorial de cada teste do Conjunto de Testes de Inteligência Fluida (Gf) foram utilizados os seguintes procedimentos fatoriais exploratórios de retenção de fatores: auto-valor-maior-do-que-um, *scree test*, análise paralela por permutação, 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, uma solução era identificada como a mais adequada. Para identificação da melhor solução foi empregada a análise fatorial confirmatória, usados os índices CFI e RMR e a comparação dos qui-

quadrados e graus de liberdade provenientes das diferentes soluções. Para análise da unidimensionalidade foi empregado o modelo Rasch, através do 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). Com relação à validade convergente e discriminante, foi gerada uma análise fatorial confirmatória do Conjunto de Testes de Inteligência Fluida (Gf), acrescentada de três testes elaborados especificamente para mensurar a Fluência (Gr), uma outra habilidade cognitiva de alta-ordem do campo psicométrico.

RESULTADOS E DISCUSSÃO

O uso de diferentes procedimentos de retenção de fatores da análise fatorial exploratória indicou um número variado de soluções fatoriais para os três testes do conjunto. Um exemplo é o Teste de Indução (I). Para esse teste, a análise paralela apontou para uma estrutura fatorial de três fatores de nível um e um fator geral (g). Por sua vez, o critério do autovalor-maior-do-que-um e o critério da máxima verossimilhança apontaram para uma solução de quatro fatores de nível um e um fator geral (g). O *scree test*, por outro lado, ampliou o número de fatores, e indicou uma estrutura fatorial de cinco fatores de primeiro nível e um fator geral (g). Para poder comparar as diferentes soluções geradas em cada teste e escolher a mais adequada, cada solução fatorial exploratória obtida foi modelada para ser analisada através da análise fatorial confirmatória.

A solução mais adequada de cada um dos três testes, obtida pelo método confirmatório, apresentou bom grau de ajuste aos dados, com um CFI muito próximo ou superior a 0,9 e um RMR inferior a 0,08 (Teste de Raciocínio Lógico, CFI=0,89, RMR=0,01, $\chi^2=567,27$, gl=363; Teste de Raciocínio Geral, CFI=0,99, RMR= 0,01, $\chi^2=97,74$, gl=72; Teste de Indução, CFI=0,98, RMR=0,01, $\chi^2=111,17$, gl=73). A descrição e a discussão da estrutura fatorial dos testes do Conjunto de Testes de Inteligência Fluida (Gf) será feita, pois, tendo como referência a melhor solução obtida pelo método confirmatório.

Em relação à estrutura fatorial dos testes, nenhum deles mensura apenas um único fator. Cada um dos testes apresenta um conjunto de fatores de primeiro nível, assim como um fator geral. Os fatores de nível um se relacionam a um número

limitado de itens e dizem respeito a processos específicos de cada teste, enquanto o fator geral interage diretamente com todos os itens do teste. A Figura 01 apresenta a estrutura complexa

encontrada em todos os testes, sem corresponder diretamente ao número de itens e de fatores de nível um encontrados especificamente em cada teste.

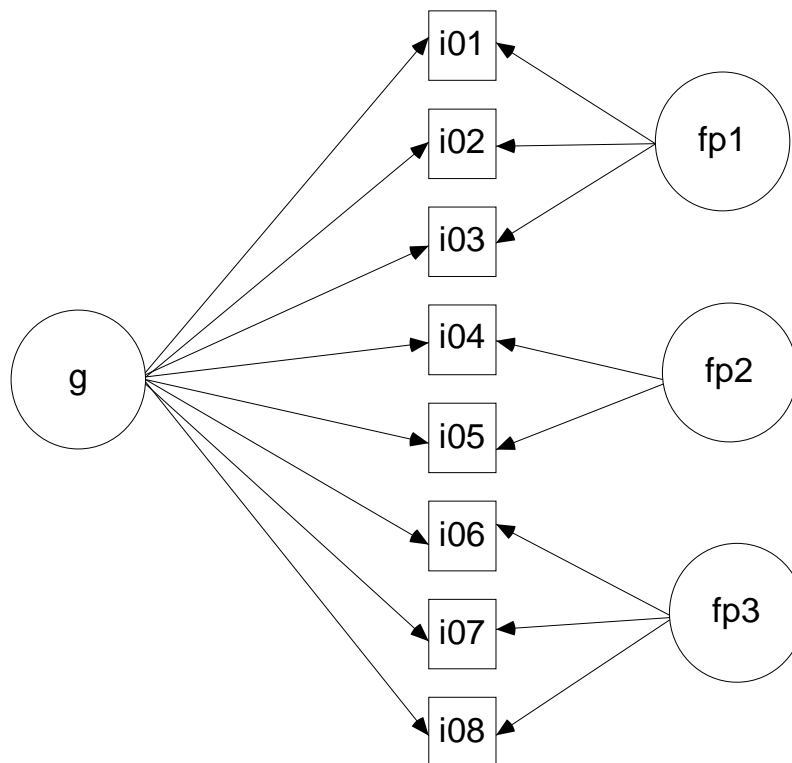


Figura 01. Estrutura Fatorial comum encontrada em todos os testes do Conjunto de Testes.
 Legenda: g = fator geral; fp1 = fator primário 1; fp2 = fator primário 2; fp3 = fator primário 3; i = item.

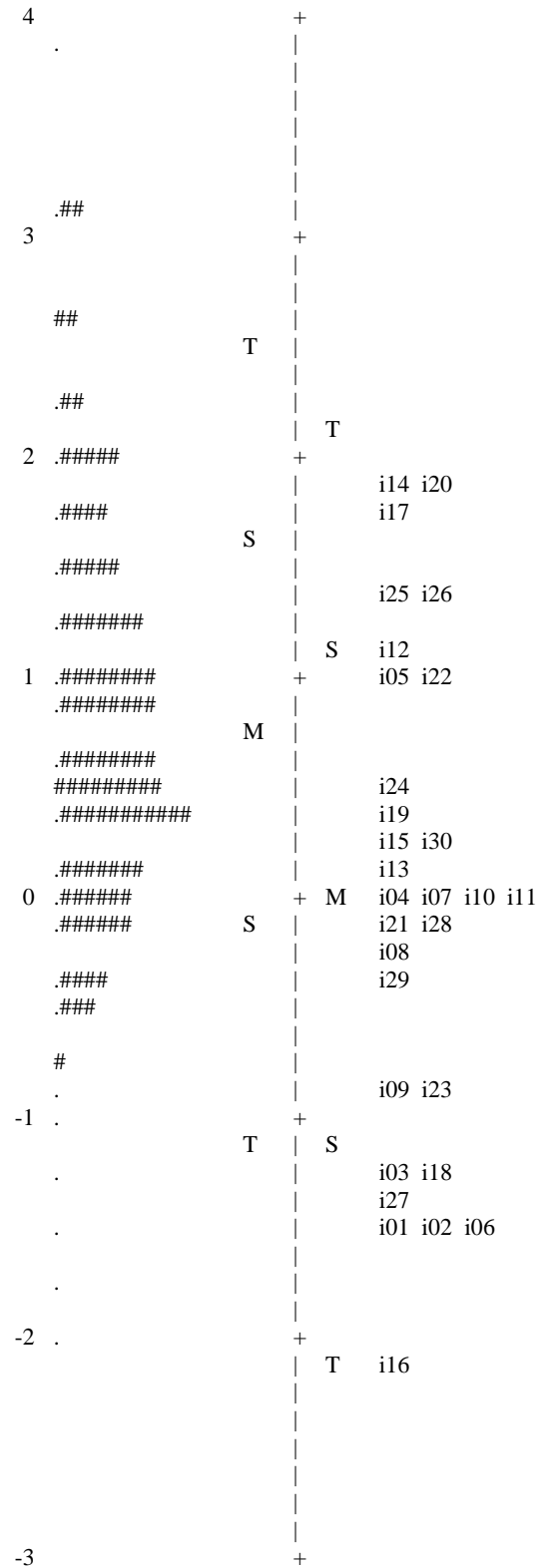
Com respeito à unidimensionalidade, todos os testes podem ser considerados unidimensionais. No entanto, alguns aparentemente apresentam uma unidimensionalidade mais robusta do que outros. Analisando cada teste em questão, o Teste de Raciocínio Lógico (RL) pode ser considerado unidimensional, mas alguns itens necessitam de novas análises. Uma evidência favorável à unidimensionalidade do teste refere-se ao amplo conjunto de itens que é carregado pelo fator geral com uma carga igual ou superior a 0,3. Por outro lado, essa condição não é presente em todos os itens e alguns apresentam uma carga inferior a 0,10 ou uma carga negativa em relação ao fator geral (itens oito, 14, 17, 20, 24, 25 e 26) (Tabela 01). Outra evidência favorável à unidimensionalidade é a preponderância do fator geral sobre os fatores de nível um na explicação da variância comum dos itens do teste. O fator geral explica 34,49% da variância comum do teste e, apesar de não explicar

pele menos a metade da variância comum, é o fator explicativo predominante. Por sua vez, se os itens com carga inferior a 0,2 em relação ao fator geral são retirados da análise, o fator geral passa a explicar 56,96% da comunalidade do teste, um forte acréscimo de 22,47%. Uma terceira evidência favorável à unidimensionalidade do teste refere-se à escala unidimensional gerada pelo modelo Rasch. Ela mostrou-se confiável em relação aos itens do teste (0,99), com um índice *infit* médio foi de 0,99 e desvio-padrão de 0,11. Apenas o item 20 não esteve entre os valores de 0,7 a 1,3 do *infit*, não se adequando à escala. Apesar da presença de uma boa distribuição quanto à dificuldade dos itens, essa condição não é plenamente verificada. Todos os itens difíceis da escala não carregam adequadamente o fator geral, de modo que, se são excluídos da análise, há uma considerável redução da amplitude da distribuição dos itens na escala unidimensional. (Figura 02).

Tabela 1. Estrutura Fatorial do Teste de Raciocínio Lógico

	g	fp1	fp2	fp3	fp4	fp5	fp6	fp7
i01	0,44	0,44	0,00	0,00	0,00	0,00	0,00	0,00
i02	0,49	0,50	0,00	0,00	0,00	0,00	0,00	0,00
i03	0,47	0,36	0,00	0,00	0,00	0,00	0,00	0,20
i04	0,19	0,00	0,00	0,00	0,00	0,00	0,30	0,00
i05	0,26	0,00	0,00	0,00	0,76	0,00	0,00	0,00
i06	0,35	0,17	0,00	0,00	0,00	0,00	0,21	0,00
i07	0,27	0,00	0,00	0,00	0,00	0,00	0,35	0,00
i08	0,10	0,00	-0,34	0,00	0,00	0,00	0,00	0,00
i09	0,49	0,00	0,00	0,08	0,00	0,00	0,00	0,24
i10	0,22	0,00	0,00	0,00	0,00	0,13	0,00	0,45
i11	0,35	0,16	0,00	0,11	0,00	0,00	0,00	0,00
i12	0,23	0,00	0,00	0,00	0,50	0,00	0,00	0,00
i13	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,56
i14	-0,07	0,00	0,51	0,00	0,00	0,00	0,00	0,00
i15	0,47	0,00	0,00	0,00	0,00	0,56	0,00	0,00
i16	0,41	0,19	0,00	0,22	0,00	0,00	0,00	0,00
i17	-0,15	0,00	0,63	0,35	0,00	0,00	0,00	0,00
i18	0,36	0,09	0,00	0,00	-0,18	0,00	0,25	0,00
i19	0,16	0,00	0,00	0,00	0,42	0,00	0,00	0,00
i20	-0,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00
i21	0,19	0,00	0,00	0,31	0,00	0,00	0,00	0,00
i22	0,15	0,00	0,33	0,00	0,00	0,40	0,00	0,00
i23	0,50	0,00	0,00	0,16	0,00	0,00	0,00	0,00
i24	0,04	0,00	0,00	0,00	0,00	0,00	0,65	0,23
i25	-0,15	0,00	0,53	0,00	0,00	0,00	0,00	0,00
i26	0,04	0,00	0,54	0,00	0,00	0,00	0,00	-0,20
i27	0,41	0,00	0,00	0,19	0,00	0,00	0,00	0,00
i28	0,33	0,00	0,00	0,00	0,00	0,13	0,23	0,00
i29	0,44	0,00	0,00	0,29	0,00	0,00	0,00	0,00
i30	0,43	0,00	0,00	0,00	0,00	0,40	0,00	0,00
%comun.	34,49	7,64	16,65	4,82	11,80	7,63	8,92	8,06

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



Legenda: # = quatro pessoas.

Figura 2. Distribuição dos Itens e Pessoas do Teste de Raciocínio Lógico pelo Modelo Rasch.

O Teste de Raciocínio Geral (RG) foi um teste que apresentou características pronunciadas de unidimensionalidade. Excetuando-se o item 13,

todos os outros apresentaram uma carga igual ou superior a 0,3 em relação ao fator geral (Tabela 02). Corroborando essa evidência, 63,36% da variância

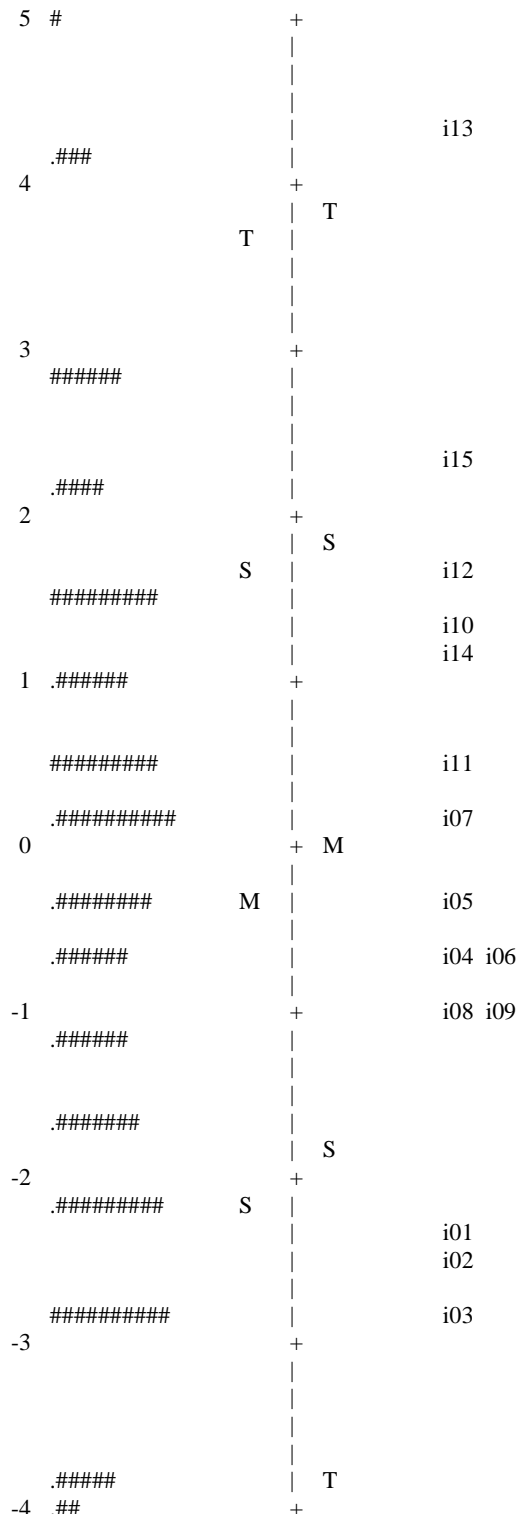
comum são explicados pelo fator geral. Outra evidência favorável de unidimensionalidade refere-se à escala gerada pelo modelo Rasch, com uma confiabilidade de 0,99 para os itens. O índice *infit* médio dos itens em relação à escala foi de 0,96 e

desvio-padrão de 0,17 e apenas o item quatro não esteve entre os valores de 0,7 a 1,3, não se adequando à escala. Além das condições apontadas, o teste mostrou boa distribuição quanto à dificuldade dos itens (Figura 03).

Tabela 2. Estrutura Fatorial do Teste de Raciocínio Geral

	g	fp1	fp2	fp3	fp4
i01	0,39	0,00	0,00	0,00	0,16
i02	0,44	0,00	0,00	0,00	0,25
i03	0,43	0,00	0,00	0,00	0,69
i04	0,40	0,00	0,24	0,00	0,00
i05	0,53	0,00	0,00	0,16	0,00
i06	0,71	0,00	-0,04	-0,01	0,00
i07	0,60	0,00	0,00	0,10	0,00
i08	0,71	0,00	0,22	0,00	0,00
i09	0,67	0,00	0,37	0,00	0,00
i10	0,57	0,00	0,00	0,29	0,00
i11	0,63	0,32	0,00	0,34	0,00
i12	0,55	0,31	0,00	0,40	0,00
i13	0,22	0,39	0,00	0,00	0,00
i14	0,47	0,66	0,00	0,00	0,00
i15	0,38	0,66	0,00	0,00	0,00
% comun.	63,36	18,43	3,66	6,00	8,55

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



Legenda: # = quatro pessoas.

Figura 3. Distribuição dos Itens e Pessoas do Teste de Raciocínio geral pelo Modelo Rasch.

Com relação ao Teste de Indução (I), excetuando-se os itens 12 e 13, todos os outros apresentaram uma carga igual ou superior a 0,3 em relação ao fator geral (Tabela 03). Corroborando

essa evidência, quase metade da variância comum é explicada pelo fator geral (45,47%). A escala gerada pelo modelo Rasch apresentou uma confiabilidade de 0,99 para os itens. O índice *infít*

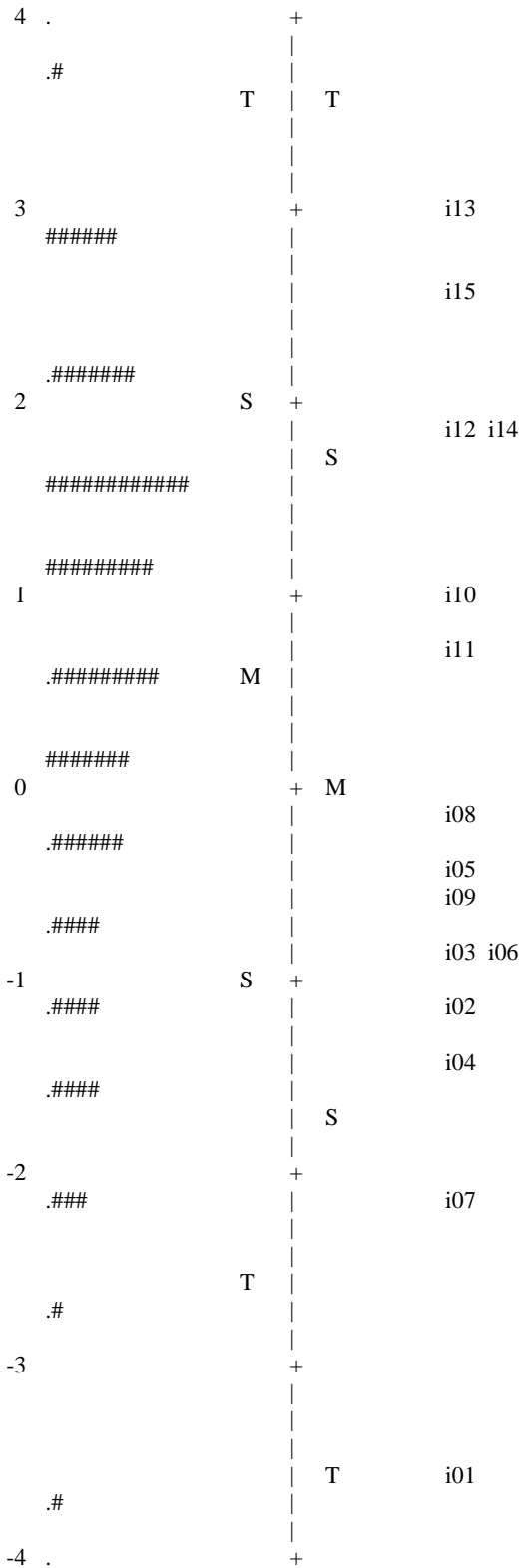
médio dos itens em relação à escala foi de 0,98 e desvio-padrão de 0,14 e todos os itens estiveram entre os valores de 0,7 a 1,3, se adequando à escala.

Além das condições apontadas, o teste mostrou boa distribuição quanto à dificuldade dos itens (Figura 04).

Tabela 3. Estrutura Fatorial do Teste de Indução

	g	fp1	fp2	fp3	fp4	fp5
i01	0,40	0,00	0,06	0,00	0,00	0,59
i02	0,34	0,00	0,54	0,00	0,00	0,00
i03	0,41	0,31	0,00	0,00	0,00	0,00
i04	0,41	0,31	0,00	0,00	0,00	0,18
i05	0,51	0,49	0,00	0,00	0,00	0,00
i06	0,54	0,44	0,00	0,00	0,00	0,00
i07	0,39	0,00	0,36	0,00	0,00	0,25
i08	0,42	0,42	0,00	0,00	0,00	0,00
i09	0,57	0,45	0,00	0,00	0,00	0,00
i10	0,47	0,33	-0,24	0,00	0,00	0,00
i11	0,51	-0,01	0,00	0,00	0,00	0,00
i12	0,21	0,00	0,00	0,55	0,00	0,00
i13	0,14	0,00	0,00	0,72	0,00	0,00
i14	0,30	0,00	0,00	0,00	0,29	0,00
i15	0,31	0,00	0,00	0,00	0,33	0,00
%comun.	45,47	19,99	8,60	14,57	3,40	7,96

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



Legenda: # = sete pessoas.

Figura 4. Distribuição dos Itens e Pessoas do Teste de Indução pelo Modelo Rasch.

É relevante ressaltar que a estratégia utilizada neste trabalho foi a de valorizar a análise do fator geral encontrado em cada teste, na medida em que se defende a idéia de que testes bem

explicados (em sua variância comum) pelo fator geral tendem a apresentar evidências favoráveis de unidimensionalidade, característica importante quando se pretende que diferentes instrumentos de

medida sejam capazes de mensurar, com uma menor quantidade de erro, fatores de mais alta ordem.

Há abordagens que podem ser consideradas complementares à utilizada neste artigo e que têm procurado analisar detidamente os componentes específicos de testes relacionados com a Inteligência Fluida. Primi, Cruz, Nascimento e Petrini (2006) elaboraram um teste de analogias visando mensurar dois componentes bastante particulares, o gerenciamento executivo, e a atenção seletiva e controle inibitório. Segundo os autores, dois processos fundamentais definidores da Inteligência Fluida são esses componentes específicos, baseando-se em evidências que supõem haver uma forte associação entre a Memória de trabalho, seus componentes, e a Inteligência Fluida.

Apesar do aspecto promissor da busca por testes mais refinados e com itens bem específicos, é necessário melhor compreender esses componentes. Gignac (2007) contesta evidências que favorecem uma maior associação entre a Inteligência e a Memória de Trabalho. Analisando dois estudos que, respectivamente, relacionam diretamente g à Memória de Trabalho e g à Inteligência Fluida, ele

apresenta evidências de que esses resultados se devem ao fator de desatenuação das análises fatoriais confirmatórias. Segundo o autor, a desatenuação infla as cargas fatoriais dos fatores de menor ordem em relação ao fator de ordem superior, quando os fatores de menor ordem apresentam uma baixa confiabilidade. Assim, segundo seu argumento, várias cargas fatoriais que apontam uma relação de igualdade entre g e a Memória de Trabalho podem apresentar o viés apontado por ele.

Buscando analisar a validade convergente e a validade discriminante do Conjunto de Testes de Inteligência Fluida (Gf), foi feita uma análise fatorial confirmatória que abrange os três testes do Conjunto e três testes elaborados especificamente para mensurar a Habilidade de Recuperação Abrangente, também denominada de Fluência (Gr) (Figura 05). A análise fatorial confirmatória foi obtida a partir do escore total dos participantes em cada um dos seis testes mencionados e a solução fatorial gerada obteve um CFI de 1,00 e um RMSEA de 0,00, implicando em uma adequação perfeita aos dados.

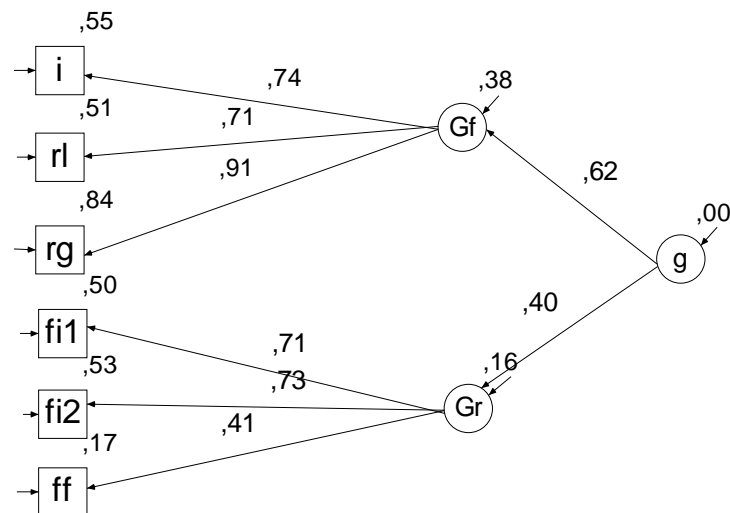


Figura 5. Solução Fatorial dos Seis Testes

Há uma clara distinção entre os testes relacionados com a Inteligência Fluida (Gf) e os testes relacionados com a Fluência (Gr). Nenhum teste de Fluência se relaciona com a Inteligência Fluida (Gf) e vice-versa. Com relação à validade convergente, todos os três testes do Conjunto de Testes de Inteligência Fluida convergem para a mensuração do construto visado e, além disso, possuem uma carga fatorial expressiva em relação a

ele. Essas condições apontam para uma validade convergente e divergente dos testes.

Apesar da validade convergente e divergente expressa, há alguns desafios. Dois deles são apontados por Lohman (2001) e por Beauducel, Brocke e Liepmann, (2001) e merecem destaque. Segundo Lohman (2001), há evidências de que itens ordenados randomicamente tendem a medir mais pobremente Gf, na medida em que são

reduzidas as probabilidades de que o respondente possa aprender as regras envolvidas nos testes. Lohman (2001) indica que a disposição dos itens de um teste pode possibilitar de forma mais saliente que o respondente aprenda as regras envolvidas entre os itens, salientando uma característica fundamental da Inteligência Fluida (Gf), a capacidade de aprender coisas novas. Testes desse tipo tendem a ser melhor marcadores de Gf. Por sua vez, Beauducel e colaboradores (2001) enfocam outro aspecto problemático da mensuração de Gf: sua articulação com a modalidade figural. Para evitar esse viés, os autores elaboraram testes de Inteligência Fluida (Gf) e Inteligência Cristalizada (Gc), de maneira equilibrada nas modalidades verbal, figural e numérica, obtendo evidências de validade dos construtos sem o viés de associação entre Gf e a modalidade figural, e Gc e a modalidade verbal.

Em função dos argumentos, pode-se apontar a necessidade futura de incorporação de novas modalidades no Conjunto de Testes. Apesar do Conjunto possuir um teste da modalidade numérica e dois da modalidade verbal, variando processos dedutivos e indutivos, pode-se ampliar o número de modalidades. Concomitantemente, uma necessidade futura aponta para que os testes do Conjunto sejam formados por seqüências de itens capazes de oportunizar a aprendizagem das regras envolvidas nos testes, no momento em que o respondente faz os próprios testes sem a interferência do aplicador. Em função desses aspectos, espera-se melhorar a validade convergente e a validade divergente, assim como ampliar a característica dos testes como marcadores de Gf.

CONSIDERAÇÕES FINAIS

A carência de baterias específicas para cada uma das habilidades de alta-ordem prejudica a pesquisa na área dos estudos sobre a inteligência. No caso da Inteligência Fluida (Gf), a ausência de baterias específicas dificulta a identificação empírica e investigação desse construto, assim como impede que os estudos nacionais possam incorporar questões relevantes traçadas pela literatura internacional.

A literatura internacional psicométrica tem apontado e levado à tona algumas questões importantes relacionadas à Inteligência Fluida (Gf). Uma delas diz respeito à sua associação direta com o g, desafio apontado por Undheim e Gustafsson

(1987). Segundo as evidências dos autores, g se confunde com a Inteligência Fluida, de modo que esse último seria o melhor definidor do primeiro. Outra questão interessante diz respeito à dificuldade da psicométrica em elaborar testes capazes de mensurar separadamente os dois construtos descritos (Carroll, 2003). Segundo Carroll (2003), a falta de bons testes marcadores para a Inteligência Fluida, sem a forte influência de g, tem dificultado uma investigação mais efetiva sobre a real natureza da Inteligência Fluida, de modo que uma agenda de pesquisa sobre esse construto envolve a elaboração de melhores instrumentos de medida.

Considerando a necessidade de elaboração de baterias específicas voltadas para a mensuração de determinados construtos psicológicos, as mesmas devem possuir testes unidimensionais, com itens bem distribuídos quanto ao seu grau de dificuldade, assim como apresentar validade convergente e discriminante. Procurou-se, ao longo deste artigo, avaliar o Conjunto de Testes de Inteligência Fluida analisando esses aspectos.

De uma maneira geral, o Conjunto de Testes de Inteligência Fluida é válido e seus testes são unidimensionais. Apesar dessa condição geral favorável, o Teste de Raciocínio Geral e o Teste de Indução apresentaram uma característica mais pronunciada de unidimensionalidade do que o Teste de Raciocínio Lógico. Houve neste último teste a presença de muitos fatores de primeiro nível, o que contribuiu para uma menor explicação do fator geral sobre a comunalidade do teste. Por sua vez, se a eliminação dos itens deste teste com carga fatorial menor do que 0,2 em relação ao fator geral gerou um acréscimo substancial na capacidade deste fator em explicar a comunalidade do teste, essa estratégia não foi suficiente, pois os itens mais difíceis da escala unidimensional foram eliminados, diminuindo a amplitude da distribuição dos itens. Novos estudos são, pois, necessários, para investigar o comportamento dos itens do Teste de Raciocínio Lógico. Apesar da necessidade de revisões e ajustes possíveis, as evidências apontam para a possibilidade de que o Conjunto de Testes de Inteligência Fluida possa ser investigado em nível nacional, de modo a poder se tornar futuramente uma ferramenta para quem se propõe a estudar a Inteligência Fluida em amostras brasileiras.

REFERÊNCIAS

Adey, P., Csapó, B., Demetriou, A., Hautamäki, J., & Shayer, M. (2007). Can we be intelligent

- about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review*, 23. Acessado em 25/01/2008. Disponível em <<http://doi:10.1016/j.edurev.2007.05.001>>.
- Almeida, L. S., Lemos, G., Guisande, M. A., & Primi, R. (2008). Inteligência, escolarização e idade: normas por idade ou série escolar?. *Avaliação Psicológica*, 7, 117-125.
- Beauducel, A. (2001). Problems with parallel analysis in data sets with oblique simple structure. *Methods of Psychological Research Online*, 6 (2). Acessado em 25/01/2007. Disponível em <<http://www.mpr-online.de>>.
- Beauducel, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence: facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences*, 30, 977-994.
- 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.
- Carroll, J. B. (1993). *Human cognitive abilities: a survey of factor analytic studies*. New York: 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). Elsevier Science/Pergamon Press.
- Cattell, R. B. (1971). *Abilities: Their structure growth and action*. Boston: Houghton Mifflin.
- 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.
- Colom, R., Abad, F. J., Rebollo, I., & Shih, P. C. (2005). Memory span and general intelligence: a latent-variable approach. *Intelligence*, 33, 623-642.
- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, 32, 277-296.
- Ekstrom, R. B., French, J. W., Harman, H. H. & Dirmen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service. *Estudos de Psicologia (Natal)*, 11 (3), 323-332.
- Flores-Mendoza, C., & Nascimento, E. (2001). Inteligência: o construto melhor investigado em psicologia. *Boletim de Psicologia*, 51 (114), 37-64.
- 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.
- Garcia Jr., A. J., & Noronha, A. P. P. (2007). Inteligência emocional e provas de raciocínio: um estudo correlacional. *Psicologia. Reflexão e Crítica*, 20, 480-489.
- Gardner, H. (2006). On failing to grasp the core of MI theory: a response to Visser et al. *Intelligence*, 34, 503-505.
- Gignac, G. E. (2007). Working memory and fluid intelligence are both identical to g?! Reanalyses and critical evaluation. *Psychology Science*, 49 (3), 187-207.
- 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.
- Haier, R. J., White, N. S., & Alkire, M. T. (2003). Individual differences in general intelligence correlate with brain function during nonreasoning tasks. *Intelligence*, 31, 429-441.
- Helms-Lorenz, M., van de Vijver, F. J. R., & Poortinga, Y. H. (2003). Cross-cultural differences in cognitive performance and Spearman's hypothesis: g or c? *Intelligence*, 31, 9-29.
- 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.
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75 (3), 242-259.
- 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 Buiten's (1963) battery of 46 tests of mental ability. *Intelligence*, 35, 69-81.
- Linacre J.M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16 (2), 878.
- Lohman, D. F. (2001). Fluid intelligence, inductive reasoning, and working memory: Where the theory of Multiple Intelligences falls short. Em N. Colangelo, & S. Assouline (Orgs.), *Anais, Talent Development IV: Proceedings from the 1998 Henry B. & Jocelyn Wallace National Research Symposium on talent development* (pp. 219-228). Scottsdale, AZ: Gifted Psychology Press.
- Luo, D., Thomson, L. A., & Detterman, D. K. (2003). The causal factor underlying the correlation between psychometric g and scholastic performance. *Intelligence*, 31, 67-83.
- 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.
- McArdle, J. J., Hamagami, F., Meredith, W. & Bradway, K. P. (2000). Modeling the dynamic hypotheses of Gf-Gc theory using longitudinal life-span data. *Learning and Individual Differences*, 12, 53-79.
- McGrew, K.S., Keith, T.Z., Flanagan, D.P., & Vanderwood, M. (1997). Beyond g: The impact of Gf-Gc specific cognitive ability research on the future use and interpretation of intelligence tests in the schools. *School Psychology Review*, 26, 189-201.
- Pasquali, L. (2002). Inteligência: um conceito equívoco. Em: R. Primi. (Org.), *Temas em avaliação psicológica*. (pp. 56-60). Campinas, SP: IBAP.
- Primi, R. (2002). Inteligência fluida: definição fatorial, cognitiva e neuropsicológica. *Cadernos de Psicologia e Educação - Paidéia*, 12 (2/3), 57-77.
- Primi, R. (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica*, 2 (1), 67-77.
- Primi, R., Cruz, M. B. Z., Muniz, M., & Petrini, M. C. (2006). Validade de construto de um instrumento informatizado de avaliação dinâmica da inteligência fluida. *Psico (PUCRS)*, 37, 109-122.
- Primi, R., & Almeida, L. S. (2002). Inteligência geral ou fluida: desenvolvimentos recentes na sua concepção. *Sobredotação*, 3 (2), 127-144.
- Primi, R., Cruz, M. B. Z., Nascimento, M. M., & Petrini, M. C. (2006). Validade de construto de um instrumento informatizado de avaliação dinâmica da Inteligência Fluida, *PSICO*, 37 (2), 109-122.
- Primi, R., Santos, A. A. A., & Vendramini, C. M. (2002). Habilidades básicas e desempenho acadêmicos em universitários ingressantes. *Estudos de Psicologia (Natal)*, 7 (1), 47-55.
- Roberts, R. D., Goff, G. N., Anjoul, F., Kyllonen, P. C., Pallier, G., & Stankov, L. (2000). The Armed Services Vocational Aptitude Battery (ASVAB) little more than acculturated learning (Gc)!? *Learning and Individual Differences*, 12, 81-103.
- Salthouse, T. A. (2004). Localizing age-related individual differences in a hierarchical structure. *Intelligence*, 32, 541-561.
- 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. (2006). Estudo da estrutura fatorial da Bateria Multidimensional de Inteligência Infantil. *Estudos de Psicologia (Campinas)*, 23, 105-112.
- Schelini, P. W., & Wechsler, S. M. (2005). Bateria multidimensional de inteligência infantil: desenvolvimento de instrumento. *Psico-USF*, 10 (2), 129-139.
- Spearrit, D. (1996). Carroll's model of cognitive abilities: educational implications. *International Journal of Educational Research*, 25 (2), 107-198.
- Sternberg, R. J. (2006). The rainbow project: enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence*, 34, 321-350.
- Tirre, W. C., & Field, K. A. (2002). Structural models of abilities measured by the Ball Aptitude Battery. *Educational and Psychological Measurement*, 62, 830-856.
- 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.
- Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (2006). Factor structure of the Wechsler Intelligence Scale for Children-Fourth Edition among refereed students. *Educational and Psychological Measurement*, 66, 975-983.

- Wechsler, S. M., & Schelini, P. W. (2006). Bateria de habilidades cognitivas Woodcock-Johnson III: validade de construto. *Psicologia: Teoria e Pesquisa*, 22, 287-295.
- Wechsler, S. M., Vendramini, C. M., & Schelini, P. W. (2007). Adaptação brasileira dos testes verbais da bateria de habilidades cognitivas

Woodcock-Johnson III. *Interamerican Journal of Psychology*, 41, 45-55.

Recebido em Junho de 2008
Reformulado em Fevereiro de 2009
Aceito em Março de 2009

SOBRE OS AUTORES:

Cristiano Mauro Assis Gomes: Psicólogo, Doutor em Educação, é 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. Apoio: Fapemig.

Oto Borges: Físico, Doutor em Física, é 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. Apoio: CNPq.