

## QUALIDADES PSICOMÉTRICAS DE UM CONJUNTO DE 45 TESTES COGNITIVOS

Cristiano Mauro Assis Gomes★  
Oto Borges★★

### RESUMO

*O presente texto relata a análise das qualidades psicométricas de um conjunto de 45 testes cognitivos, avaliando a função dos mesmos como marcadores de fatores cognitivos. Os testes foram traduzidos e adaptados a partir de uma seleção entre os testes do Kit of Factor-Referenced Cognitive Tests do Educational Testing Service. Eles foram aplicados a uma amostra de 160 estudantes do terceiro ano de uma escola brasileira de ensino médio da rede federal de ensino. Foram empregados a análise fatorial exploratória, o método dos eixos principais e a análise paralela por permutação, com critérios rigorosos para a identificação de testes marcadores. Seis tipos de escalas foram identificados: escalas muito boas, escalas adequadas, escalas quase adequadas, escalas fracas, escalas com uma discrepância e escalas inadequadas. Alguns testes e escalas analisados apresentaram características de marcadores. Por outro lado, todos os testes e escalas podem ser considerados adequados, diante dos critérios psicométricos usuais.*

*Palavras-chave: Qualidades Psicométricas. Testes de Inteligência. Validação.*

## PSYCHOMETRIC PROPERTIES OF A SET OF 45 COGNITIVE TESTS

### ABSTRACT

*This paper reports the analysis of the test's psychometrical qualities from a set of 45 cognitive tests. The tests had been translated and adapted to Brazilian Portuguese from a selection of the tests of the Kit of Factor-Referenced Cognitive Tests of Educational Testing Service. The tests were applied in a sample of 160 students of the third year of one federal high school. Exploratory Factor Analysis, Principal Axis Factoring and Permutation Parallel Analysis were used in the study*

---

★ Psicólogo, Doutor em Educação. Professor adjunto da Universidade Federal de Minas Gerais. Endereço: Universidade Federal de Minas Gerais, Departamento de Psicologia – FAFICH. Av. Antônio Carlos, 6627 - Sala 4010 – Pampulha. CEP.: 31270-901 – Belo Horizonte, MG – Brasil.

*E-mail: cristianogomes@ufmg.br*

★★ Doutor em Física. Professor Associado na Universidade Federal de Minas Gerais. Coordena o Programa de Pós-graduação em Educação: Conhecimento e Inclusão Social da FAE/UFMG. (Apoio: CNPq).

*E-mail: otoborges@ufmg.br.*

*of reliability of the extracted factors. Markers tests were identified using rigorous criteria. Six scales were identified using the rigorous criteria of markers tests: very good scales, satisfactory scales, acceptable scales, weak scales, discrepant scales and inadequate scales. The results indicated that some tests and some of the generated scales could be regarded as marker tests of cognitive factors. All tests and scales can be considered adjusted by the usual psychometric criterion.*

*Keywords: Psychometrical Properties. Intelligence Tests. Validation.*

Dois aspectos fundamentais norteiam este artigo. O primeiro deles envolve a carência na psicologia brasileira de um amplo número de testes cognitivos, capazes de trazer ao pesquisador da psicologia, amplitude e verticalidade em seus estudos sobre a natureza da inteligência e as diversas implicações desses estudos sobre os modelos teóricos relacionados à arquitetura, o desenvolvimento cognitivo e o desempenho humano em diversos contextos humanos.

A literatura psicométrica sugere, por exemplo, que para se descobrir ou identificar um componente da arquitetura intelectual humana, deve-se usar pelo menos três testes teoricamente relacionados a esse componente e elaborados especificamente para mensurá-lo (CARROLL, 1993; GORSUCH, 1974). Identificar empiricamente qualquer construto da arquitetura cognitiva requer a disponibilidade de um vasto arsenal de testes que, teoricamente, estão envolvidos na mensuração do mesmo construto. No Brasil, o número de testes cognitivos para as pesquisas na área ainda é pequeno, e, a julgar pela quantidade de testes aprovados pelo CFP no ano de 2005, pode-se considerar esse número também incipiente para aplicações nos diversos setores da prática do psicólogo.

O segundo aspecto que norteia as discussões deste artigo diz respeito ao ambicioso projeto da psicologia de elaborar testes capazes de identificar e mensurar componentes da inteligência de uma maneira precisa, de modo a mensurar apenas o componente visado, sem envolver significativamente outros componentes da inteligência. A elaboração de testes para a mensuração específica de um determinado componente da inteligência é uma atividade antiga relatada na literatura da psicologia (GORSUCH, 1974; GUSTAFSSON; 1992a, 1992b; MULLAIK; 1993, 1996). Esse tipo de teste construído especialmente para medir forte e particularmente um único fator é denominado de teste de referência ou teste marcador. O *Educational Testing Service* (ETS), é uma instituição conhecida pela pesquisa sobre a elaboração de testes educacionais como o SAT, teste americano de competências educacionais. O ETS também realiza muitas pesquisas no campo dos testes marcadores, tendo realizado três meta-análises sobre testes marcadores, e elaborado baterias de testes cognitivos para os pesquisadores da área. Seu último conjunto de testes marcadores foi gerado por meio de uma meta-análise de 23 fatores primários estudados até o final dos anos 1970 (EKSTROM *et al.*, 1976; EKSTROM; FRENCH; HARMAN, 1979).

Os testes marcadores representam avanços no campo dos estudos sobre a inteligência, pois indicam uma intenção deliberada de identificar com maior profundidade, e com instrumentos mais incisivos, a ampla diversidade de com-

ponentes que podem compor a arquitetura cognitiva humana (CARROLL; 1993, 1995, 1997). O desafio dos testes marcadores é bem maior do que o dos testes usuais, na medida em que eles devem indicar com maior propriedade os componentes da arquitetura intelectual.

O maior desafio dos testes cognitivos tem sido justamente poder ser considerado um bom marcador. Segundo Carroll (2003), boa parte dos testes de inteligência é explicada por fatores cognitivos em diferentes níveis da arquitetura cognitiva, de forma que tanto componentes gerais ou amplos, como componentes muito especializados, são mobilizados nos testes. Esse é um problema, já que os testes marcadores pretendem separar e diferenciar os componentes cognitivos e concentrar o seu foco em um componente específico.

Conforme o argumento de Carroll (2003), se boa parcela dos testes confeccionados ativa ou mobiliza vários níveis de processos cognitivos, pode-se considerar o empreendimento dos testes marcadores um desafio com uma trajetória ainda bastante complicada. No fundo, há uma dificuldade muito forte para se elaborar testes capazes de mensurar fortemente ou exclusivamente um determinado processo, pois os testes existentes são matizados por vários processos. Por exemplo, quando se tenta elaborar três testes específicos para verificar a capacidade das pessoas em rodar mentalmente uma figura tridimensional, constata-se que muitas vezes esses testes também mobilizam a capacidade de pensamento espacial amplo, de caráter não específico, assim como mobilizam o fator geral que não está relacionado especificamente a processos de pensamento espaciais. Essa múltipla influência de processos ativados na resolução dos testes é prejudicial para um teste marcador.

Além da dificuldade para se elaborar testes capazes de mensurar quase que exclusivamente apenas um componente da arquitetura intelectual, seja este uma habilidade muito especializada, uma habilidade abrangente ou uma habilidade geral, há de se considerar a falta de critérios para que um teste seja efetivamente considerado um teste marcador. A tradição psicométrica acordou alguns critérios objetivos para a definição das relações entre os testes e os componentes cognitivos identificados por meio de procedimentos da análise fatorial (CARROLL, 1995), o que indiretamente também contribui para a análise dos testes marcadores. Um desses critérios é que um teste somente pode ser considerado relacionado a um determinado componente cognitivo se possuir uma carga fatorial de pelo menos 0,30 no fator indicativo do componente cognitivo. Além disso, as cargas fatoriais desse mesmo teste em outros fatores (indicativos de outros componentes cognitivos) não devem ser superiores à carga fatorial no fator visado.

Uma estratégia bastante útil para saber se um determinado teste pode ser considerado marcador de um fator cognitivo é a análise da unidimensionalidade dos testes e a verificação da comunalidade de cada teste em uma escala unidimensional. Essa estratégia é interessante na medida em que a unidimensionalidade indica que os testes que fazem parte da escala medem uma única e mesma dimensão, de forma que é possível identificar a força com que cada um desses

testes é explicado pela própria dimensão encontrada (SPEARRIT, 1996). Se a comunalidade de um teste em uma escala unidimensional é alta isso significa que ele é um bom marcador daquela dimensão.

Este artigo analisa a qualidade psicométrica de um conjunto de 45 testes cognitivos. Esse conjunto foi inicialmente pensado como uma seleção dentre os 72 testes do *Kit of Factor-Referenced Cognitive Tests* do ETS (EKSTROM *et al.*, 1976; EKSTROM; FRENCH; HARMAN, 1979). Os resultados levantados neste artigo salientam as qualidades psicométricas dos testes e das escalas obtidas para a análise dos fatores visados, de forma a contribuir para um estudo mais apurado sobre o papel desses testes enquanto marcadores de fatores em uma amostra no Brasil.

Inicialmente são descritos os participantes da pesquisa, os testes selecionados e os fatores visados, assim como as estratégias de análise dos dados e considerações éticas. Os resultados são discutidos, em termos das qualidades psicométricas dos testes e das escalas geradas. Uma discussão final é realizada sobre as principais evidências encontradas e as contribuições das mesmas na análise dos testes marcadores e no estudo sobre os fatores cognitivos no Brasil.

## METODOLOGIA

### PARTICIPANTES

São considerados os resultados de 160 estudantes do terceiro ano de uma escola de ensino médio da rede federal de ensino. Tal escola seleciona estudantes através de um exame público com as vagas distribuídas por três faixas socioeconômicas, e por progressão direta a partir de uma escola de ensino fundamental da mesma instituição mantenedora. Os participantes pertenciam a oito turmas da terceira série, duas delas em 2004 e seis delas em 2005. Apesar da diversidade de estudantes nas classes socioeconômicas, os participantes fazem parte de uma amostra com desempenho escolar superior à média nacional. Enquanto a média brasileira alcançou um valor de 40,56% de acertos na prova objetiva do ENEM de 2001, por exemplo, os participantes desta pesquisa alcançaram uma pontuação média de 64,32% na mesma prova.

### INSTRUMENTOS

Foram selecionados, traduzidos e adaptados para os participantes da pesquisa 45 testes cognitivos do *Kit of Factor-Referenced Cognitive Tests*, produzido pelo *Educational Testing Service* (EKSTROM *et al.*, 1976). A tradução e adaptação foi realizada pelo primeiro dos autores deste trabalho. Na adaptação foi considerado tanto o conteúdo lingüístico do inglês para o português, como os aspectos socioculturais e suas implicações no entendimento do teste pelos participantes da pesquisa. Foi feito um piloto em 2004 para verificar se a tradução e adaptação das instruções e dos itens dos testes não ocasionavam problemas importantes.

A meta-análise de Ekstrom, French e Harman (1979) indica que os 45 testes selecionados para uso no presente trabalho mensuram 15 fatores cognitivos especializados, e fazem isso de forma que cada grupo de três testes se relaciona a um fator cognitivo específico. Os testes selecionados para o presente estudo estão relacionados a seguir, agrupados segundo o fator cognitivo a que se referem:

Flexibilidade de Fechamento (CF): Figuras Escondidas (CF1); Teste de Padrões Escondidos (CF2); Teste de Cópia (CF3).

Fechamento Verbal (CV): Palavras Misturadas (CV1); Palavras Escondidas (CV2); Palavras Incompletas (CV3).

Fluência Figural (FF): Teste de Ornamentação (FF1); Teste de Elaboração (FF2); Teste de Símbolos (FF3).

Fluência Ideativa (FI): Teste de Tópicos (FI1); Teste de Temas (FI2); Teste de Categorias de Objetos (FI3).

Fluência de Palavras (FW): Teste do Término das Palavras (FW1); Teste do Início das Palavras (FW2); Teste do Início e Término das Palavras (FW3).

Indução (I): Teste dos Arranjos de Letras (I1); Teste das Posições (I2); Classificação de Figuras (I3).

Memória Associativa (MA): Teste Numérico-Figural (MA1); Teste de Numérico-Objeto (MA2); Teste de Nomes e Sobre-Nomes (MA3).

Memória de Retenção (MS): Teste de Retenção Auditivo-Numérica (MS1); Teste de Retenção Numérico-Visual (MS2); Teste Auditivo de Retenção de Letras (MS3).

Memória Visual (MV): Teste da Memória de Formas (MV1); Memória de Construções (MV2); Memória de Mapas (MV3).

Facilidade Numérica (N): Teste de Adição (N1); Teste de Multiplicação e Subtração (N3); Correção de Adição e Subtração (N4).

Velocidade Perceptiva (P): Teste “Encontrando a Letra A” (P1); Teste de Comparação Numérica (P2); Teste de Figuras Idênticas (P3).

Raciocínio Geral (RG): Teste de Aptidão Aritmética (RG1); Teste de Aptidão Matemática (RG2); Teste de Operações Aritméticas Necessárias (RG3);

Raciocínio Lógico (RL): Teste de Silogismos Sem-Sentido (RL1); Representação de Relações (RL2); Teste de Inferências (RL3).

Compreensão Verbal (V): Teste de Vocabulário Ampliado (V3); Teste de Vocabulário Avançado I (V4); Teste de Vocabulário Avançado II (V5).

Visualização (VZ): Teste de Formas (VZ1); Teste de Dobraduras de Papel (VZ2); Teste de Desenvolvimento de Superfícies (VZ3).

## ANÁLISE DOS DADOS

Os testes foram analisados separadamente em 15 grupos de três, de acordo com o fator mensurado. Dessa maneira, foram obtidas 15 matrizes de correlação e 15 matrizes fatoriais. Para análise da unidimensionalidade de cada grupo de testes foi realizada a análise fatorial exploratória, com a utilização do Método dos Eixos Principais, com análise paralela por permutação. A confiabilidade dos dados foi verificada através da consistência interna entre os itens de cada teste, assim como da consistência interna da escala gerada em cada grupo de três testes, ambas medidas pelo *alpha* de Cronbach. Na medida em que todos os testes de fluência não possuem itens *a priori*, pois dependem da produção do respondente, o *alpha* de Cronbach descrito nesses testes refere-se à consistência interna entre o resultado total da primeira parte e o resultado total da segunda parte de cada um desses testes.

Neste estudo somente será considerado como marcador um teste cuja comunalidade for igual ou maior do que 50% da sua variância. Tal critério é mais exigente do que o critério usual para que um teste seja considerado indicador de um fator cognitivo, que exige apenas que o teste tenha uma carga mínima de 0,30 no fator. Essa carga, em análises unidimensionais, indica que 9% da variância do teste é explicada pelo fator cognitivo, o que consideramos um critério fraco, na medida em que os testes marcadores são aqueles instrumentos especificamente criados para mensurarem fortemente um determinado fator.

## CONSIDERAÇÕES ÉTICAS DA PESQUISA

Os 45 testes cognitivos foram aplicados coletivamente em seis ou sete baterias. Antes de realizarem qualquer tipo de teste, todos os participantes leram e assinaram um termo formal de adesão à pesquisa. Menores de idade deveriam trazer a assinatura de um dos responsáveis. O tempo máximo de cada bateria foi de 70 minutos, com um tempo extra de 30 minutos para organização e gestão da sala, totalizando 100 minutos por bateria. Os estudantes que perdiam um teste, um conjunto de testes ou uma ou mais baterias, marcavam um momento posterior com o aplicador (psicólogo) para que pudessem fazer os testes perdidos. Os testes foram aplicados em horários e dias letivos normais da escola dos participantes da pesquisa. No entanto, o aplicador combinava previamente com as turmas a seqüência da aplicação, local, dia e hora. Os dias e horários foram previamente combinados com os professores e os departamentos da escola, de forma a não prejudicar o andamento do percurso escolar dos estudantes.

## RESULTADOS

A análise fatorial dos 15 grupos de três testes indicou a presença de um único fator confiável em cada grupo de testes analisado, evidenciando a presença de uma escala unidimensional em cada grupo. Isso significa que os três testes de cada grupo, quando analisados entre si, estão relacionados e mensuram apenas um fator cognitivo.

Além de verificar se cada grupo dos três testes efetivamente se relacionava a apenas um fator cognitivo, também foi analisado o equilíbrio da participação de cada um dos três testes na mensuração do fator cognitivo. A análise da consistência interna das 15 escalas unidimensionais geradas indicou que alguns grupos de testes são bem consistentes entre si, enquanto outros grupos de testes possuem uma consistência bem menor, abaixo do desejado, inclusive. Essa evidência é percebida por meio do *alpha* de Cronbach dessas escalas. Os resultados variaram de 0,53 a 0,86. A escala gerada pelos testes N1, N3 e N4 apresentou o *alpha* de Cronbach mais alto (0,86), demonstrando uma boa consistência interna entre esses três testes. Por sua vez, a escala gerada pelos testes CV1, CV2 e CV3 apresentou o menor *alpha* de Cronbach (0,53), de modo que esses testes não demonstraram ser bem consistentes entre si. No caso de escalas bem consistentes, pode-se dizer que há um equilíbrio na participação de cada um dos testes na identificação do fator cognitivo, aspecto importante para uma boa mensuração. Diferentemente das escalas bem consistentes, as escalas menos consistentes, como no caso da escala dos testes CV1, CV2 e CV3, demonstraram um desequilíbrio na participação dos testes, de modo a indicar que possivelmente um ou dois testes possam ter tido influência maior na mensuração do fator, em detrimento do terceiro teste.

As evidências da análise sobre o equilíbrio ou desequilíbrio encontrado nas 15 escalas unidimensionais podem ser corroboradas com a análise da comunalidade de cada teste. Como a comunalidade é o percentual da variância de cada teste explicado pelos fatores extraídos de uma análise fatorial, a ocorrência de um único fator em cada grupo dos três testes fez da comunalidade um índice do quanto cada teste seria explicado pelo fator único encontrado. Os resultados da comunalidade de cada teste dos grupos de testes indicam que os grupos com comunalidades mais semelhantes em seus testes proporcionaram uma escala unidimensional mais consistente internamente. Essa tendência pode ser constatada por meio do seguinte exemplo: o grupo de testes CV1, CV2 e CV3 foi relatado como apresentando o menor *alpha* na escala unidimensional gerada (0,53). Concomitantemente, cada teste contribuiu muito diferentemente para o fator identificado. O teste CV1 possuiu 30% (0,30) da sua variância explicada pelo fator obtido. Já o teste CV2 apresentou 64% da sua variância explicada pelo fator obtido, e para o teste CV3 apenas 8% da variância foi explicada pelo fator. Os resultados indicaram uma grande discrepância no papel de cada um desses testes na mensuração do fator cognitivo, corroborando com os dados do *alpha* de Cronbach sobre a consistência interna da escala unidimensional gerada. Por sua vez, o grupo de testes N1, N3 e N4 foi o que apresentou a escala unidimensional com maior consistência interna (0,86). Ao mesmo tempo, também se constata que esses testes contribuíram de forma muito mais equilibrada para a identificação do fator cognitivo obtido, com os valores respectivos de 56%, 85% e 65% da sua variância explicados pelo fator primário obtido. Dessa maneira, pôde-se perceber que a consistência interna das escalas unidimensionais geradas foi proporcional ao equilíbrio da influência de cada teste na obtenção do fator único obtido em cada grupo de testes.

Além de observar a presença de unidimensionalidade em cada grupo de testes e a consistência interna das escalas unidimensionais geradas, também foi analisada a presença da consistência interna entre os itens de cada teste. Os *alpha* de Cronbach variaram de 0,60 a 0,99, com exceção do teste de fluência FI3, com um *alpha* de 0,54. O resultado do teste FI3 deve-se substancialmente à presença de apenas dois itens no teste, a saber, o resultado total de produções do respondente na parte um e o resultado total de produções do respondente na parte dois do teste. Como se sabe, quanto maior o número de itens de uma escala analisada, maior o escore do *alpha* de Cronbach e maior a consistência interna. Nesse sentido, pode-se afirmar que todos os testes apresentaram, pelo menos, uma razoável consistência interna entre seus itens, demonstrando que as respostas dos participantes da pesquisa foram confiáveis em todos os testes.

Discutindo os testes, em termos de seu papel como marcador de fatores cognitivos, definiu-se como critério marcador aquele teste capaz de ser explicado pelo fator cognitivo em pelo menos 50% de sua variância. Na medida em que a outra metade da variância, ou menos, é devida a condições específicas do próprio teste e de erros de medida, pelo menos a metade do teste deve dizer respeito ao fator cognitivo visado.

Tabela 1 – Propriedades Psicométricas dos Testes Marcadores dos Fatores RL, RG, I, V e CV

testes	fator	$\alpha$ do fator	$\alpha$ teste	carga no fator	comunalidade
rl1	rl	0,64	0,68	0,57	0,32
rl2	rl	0,64	0,80	0,66	0,44
rl3	rl	0,64	0,60	0,60	0,36
rg1	rg	0,77	0,80	0,89	0,80
rg2	rg	0,77	0,76	0,75	0,57
rg3	rg	0,77	0,75	0,56	0,31
i1	i	0,61	0,79	0,55	0,30
i2	i	0,61	0,80	0,66	0,43
i3	i	0,61	0,96	0,56	0,31
v3	v	0,79	0,80	0,86	0,73
v4	v	0,79	0,66	0,71	0,50
v5	v	0,79	0,79	0,69	0,48
cv1	cv	0,53	0,82	0,54	0,30
cv2	cv	0,53	0,77	0,80	0,64
cv3	cv	0,53	0,94	0,28	0,08

No entanto, os testes não devem ser classificados apenas isoladamente, na medida em que a comunalidade gerada em cada teste está vinculada evidentemente ao agrupamento dos três testes e ao cálculo da sua correlação. Dessa maneira, não somente os testes foram analisados como marcadores, mas também as 15 escalas geradas. A análise das escalas unidimensionais foi feita por meio da avaliação da sua consistência interna e do equilíbrio e da força das comunalidades de todos os três testes da escala.

Analisando cada uma das escalas, em seu papel de marcadora dos fatores primários, a escala do fator de Raciocínio Geral (RG) pode ser entendida como adequada. Apesar de um dos seus testes não apresentar comunalidade maior do que 0,50, dois dos seus testes apresentaram esse critério. Além disso, a escala apresentou-se de forma consistente, com um *alpha* de Cronbach de 0,77. A escala de Compreensão Verbal (V) também se apresentou adequada, com dois dos seus testes com uma comunalidade maior do que 0,50 e o terceiro teste com uma comunalidade bastante próxima a este valor. A consistência interna da escala apresentou-se adequada, com um *alpha* de Cronbach de 0,79. A escala de Visualização (VZ) também se mostrou adequada, com dois testes com comunalidade superior a 0,50 e um *alpha* de Cronbach de 0,76, o mesmo acontecendo com a escala de Flexibilidade de Fechamento (CF), com um *alpha* de Cronbach de 0,74. A escala de Facilidade Numérica mostrou-se ótima, com todos os testes com comunalidade acima de 0,50 e um *alpha* de Cronbach de 0,86, o mesmo acontecendo com a escala de Memória Associativa (MA), com um *alpha* de Cronbach de 0,81. A escala de Memória de Curto-Termo pode ser considerada quase adequada com um teste superior a 0,50 e um teste próximo de 0,50, e um *alpha* de Cronbach de 0,72. O mesmo acontece com a escala de Fluência de Palavras (FW), pois um de seus testes apresenta comunalidade maior do que 0,50 e outro teste apresenta uma comunalidade de 0,44, e um *alpha* de Cronbach de 0,68.

Tabela 2 – Propriedades Psicométricas dos Testes Referentes aos Fatores VZ, CF, P, N e MS

testes	fator	$\alpha$ do fator	$\alpha$ teste	carga no fator	comunalidade
vz1	vz	0,76	0,99	0,63	0,40
vz2	vz	0,76	0,80	0,76	0,58
vz3	vz	0,76	0,94	0,76	0,58
cf1	cf	0,74	0,87	0,61	0,38
cf2	cf	0,74	0,73	0,71	0,50
cf3	cf	0,74	0,93	0,76	0,58
p1	p	0,57	0,75	0,33	0,11
p2	p	0,57	0,93	0,67	0,45
p3	p	0,57	0,95	0,68	0,46
n1	n	0,86	0,90	0,75	0,56
n3	n	0,86	0,94	0,92	0,85
n4	n	0,86	0,96	0,81	0,65
ms1	ms	0,72	0,74	0,69	0,47
ms2	ms	0,72	0,75	0,73	0,54
ms3	ms	0,72	0,69	0,61	0,37

As escalas menos adequadas podem ser descritas da seguinte maneira. A escala de Indução (I) e a escala de Raciocínio Lógico (RL) não apresentaram grandes discrepâncias entre as comunalidades de seus testes, e têm uma consistência interna razoável. No entanto, nenhum dos testes apresentou uma comunalidade acima de 0,50, de forma que a maior parcela da variância desses testes deveu-se a aspectos específicos dos próprios testes ou a erros de medida. Apesar de a escala de Fluência Figural (FF) apresentar dois testes com comunalidade

bem superior a 0,50 e um *alpha* de Cronbach de 0,72, um dos seus testes, o FF3, apresentou uma comunalidade bem baixa, de apenas 0,24, salientando uma forte discrepância entre esse teste e os outros dois. A escala de Fluência Ideacional (FI) pode ser entendida como bastante fraca, pois um dos seus testes, o FI1, pode ser compreendido quase como o próprio fator, na medida em que 94% da sua variância é explicada pelo fator obtido, enquanto os outros dois testes são explicados em 10% e 22% de sua variância, respectivamente. A escala de Memória Visual (MV) também foi insuficiente, pois o teste MV2 concentrou uma comunalidade de 0,80, enquanto o teste MV1 de apenas 0,27 e o teste MV3 minguados 0,15. A escala de Velocidade Perceptiva (P) também não foi adequada, pois dois de seus testes apresentaram uma comunalidade próxima a 0,50, enquanto o terceiro teste apresentou uma comunalidade de apenas 0,11, com um *alpha* de Cronbach de 0,61. A escala de Fechamento Verbal (CV) também seguiu a mesma tendência, com o teste CV2 com uma comunalidade de 0,64, o teste CV1 uma comunalidade de 0,30 e o teste CV3 com uma comunalidade de apenas 0,08, com um *alpha* de Cronbach de apenas 0,53.

Tabela 3 – Propriedades Psicométricas dos Testes Referentes aos Fatores MV, MA, FW, FI e FF

testes	fator	$\alpha$ do fator	$\alpha$ teste	carga no fator	comunalidade
mv1	mv	0,61	0,86	0,52	0,27
mv2	mv	0,61	0,81	0,89	0,80
mv3	mv	0,61	0,70	0,39	0,15
ma1	ma	0,81	0,92	0,82	0,68
ma2	ma	0,81	0,88	0,76	0,57
ma3	ma	0,81	0,82	0,72	0,51
fw1	fw	0,68	0,61	0,74	0,54
fw2	fw	0,68	0,67	0,66	0,44
fw3	fw	0,68	0,61	0,54	0,29
fi1	fi	0,56	0,85	0,97	0,94
fi2	fi	0,56	0,84	0,31	0,10
fi3	fi	0,56	0,54	0,46	0,22
ff1	ff	0,76	0,88	0,89	0,80
ff2	ff	0,76	0,88	0,80	0,63
ff3	ff	0,76	0,86	0,48	0,24

Tendo analisado os testes e as escalas como marcadores dos fatores cognitivos, é importante salientar que essa análise é rigorosa, e muito superior ao critério tradicional da psicometria. Conforme salientado na metodologia, a literatura internacional tem considerado que um teste se relaciona a um fator cognitivo desde que o mínimo de 9% da variância do teste seja explicado pelo fator. Analisando por esse critério, todos os testes e as escalas podem ser considerados adequados, no que se refere às qualidades psicométricas.

## DISCUSSÃO

Algumas evidências podem ser enumeradas e discutidas. Primeiramente, os testes cognitivos apresentaram boas qualidades psicométricas, o que traz a possibilidade de uso no Brasil de um amplo conjunto de testes marcadores para identificação dos fatores cognitivos. Evidentemente, há o reconhecimento de que os participantes desta pesquisa representam uma parcela bastante específica da população brasileira, e não podem ser compreendidos como representantes fiéis dos estudantes do ensino médio. No entanto, os resultados destacados nesta pesquisa indicam evidências importantes para o estudo dos processos cognitivos e suas implicações, e indicam caminhos para estudos maiores, com amostras representativas da população de estudantes do ensino médio.

Os resultados indicaram diferenças entre os testes, de modo que se pôde criar categorias diferentes para as escalas obtidas. De modo a classificar as diferenças encontradas nos testes, seis grupos foram elaborados e nomeados da seguinte maneira: escalas muito boas, escalas adequadas, escalas quase adequadas, escalas fracas, escalas com uma discrepância e escalas inadequadas.

As escalas muito boas foram denominadas dessa forma por possuírem boa consistência interna e porque todos os seus testes demonstraram ser marcadores. Esse foi o caso das escalas de Facilidade Numérica (N) e Memória Associativa (MA).

Foram chamadas de escalas adequadas aquelas escalas com dois testes considerados marcadores e uma consistência interna adequada. Esse foi o caso das escalas de Raciocínio Geral (RG), Compreensão Verbal (V), Visualização (VZ) e Flexibilidade de Fechamento (CF).

As escalas quase adequadas demarcaram aquelas escalas que apresentaram apenas um teste como marcador e um segundo teste que chegava mais próximo ao requisito de marcador, apresentando adequada consistência interna. Esse foi o caso da escala de Memória de Curto-Termo (MS) e da escala de Fluência de Palavras (FW).

As escalas fracas envolveram as escalas que se mostraram consistentes internamente, mas com todos os seus testes com poder de indicação apenas razoável em relação ao fator cognitivo visado, com comunalidades menores do que 0,50. Esse foi o caso da escala de Raciocínio Lógico (RL) e da escala de Indução (I).

As escalas com uma discrepância foram aquelas que apresentaram dois testes como marcadores ou bastante próximos do critério, e o terceiro teste se apresentando bem insatisfatório. Esse foi o caso da escala de Fluência Figural (FF) e da escala de Velocidade Perceptiva (P).

As escalas inadequadas foram aquelas escalas onde um dos testes apresentava forte ou fortíssima relação com o fator obtido, assim como os outros testes não demonstravam importância na identificação do fator. Esse foi o caso da esca-

la de Fluência Ideacional (FI), a escala de Fechamento Verbal (CV) e a escala de Memória Visual (MV). Devido ao caráter bastante discrepante, essas escalas são aquelas que também apresentaram a menor consistência interna.

A análise dos testes marcadores é bastante rigorosa. Por que utilizar um critério tão rigoroso? Avaliar as qualidades psicométricas de testes como marcadores não significa avaliar pura e simplesmente as qualidades psicométricas de um teste cognitivo. Testes marcadores são aqueles elaborados especialmente para mensurar especificamente uma dimensão cognitiva. Através dessa condição conceitual pode-se entender o motivo por tamanho rigor na análise desses testes, enquanto marcadores, é claro. Por isso mesmo, é fundamental declarar que todas as escalas geradas nesta pesquisa são adequadas para mensurar os fatores primários visados. Nenhuma das escalas apresentou menos de dois testes que não fossem explicadas em pelo menos 9% de sua variância comum pelo fator obtido, conforme o critério internacional de carga de 0,3 no fator. Assim, pois, as evidências desta pesquisa indicam a adequação básica dos testes para mensurar fatores cognitivos.

Novos estudos são necessários, utilizando-se uma amostra mais ampla numericamente e em termos de características populacionais, no intuito de estabelecer de forma mais definitiva as qualidades psicométricas dos testes adaptados e traduzidos, enquanto marcadores de fatores cognitivos.

Com relação à análise dos testes marcadores este estudo traz uma inovação ao considerar adequados somente aqueles testes marcadores com pelo menos de 50% de sua variância comum atribuída ao fator visado, e ao considerar como escalas marcadoras adequadas aquelas com consistência interna adequada e com pelo menos dois testes com comunalidade superior a 50% e um terceiro teste razoavelmente próximo dessa condição. A psicologia brasileira tem carecido substancialmente de testes capazes de nortear a análise mais ampla, abrangente, e, ao mesmo tempo, verticalizada de processos cognitivos. Este estudo contribui para diminuir essa lacuna.

## REFERÊNCIAS

CARROLL, J. B. *Human cognitive abilities: a survey of factor-analytic studies*. New York: Cambridge University Press, 1993.

\_\_\_\_\_. On methodology in the study of cognitive abilities. *Multivariate Behavioral Research*, [S.l.], v. 30, n. 3, p. 429-452, 1995.

\_\_\_\_\_. Psychometrics, intelligence, and public perception. *Intelligence*, [V.l.], v. 24, n.1, p. 25-52, 1997.

\_\_\_\_\_. The higher-stratum structure of cognitive abilities: current evidence supports g and about ten broad factors. In: NYBORG, H. (Ed.). *The scientific study of general intelligence: tribute to Arthur R. Jensen*. Elsevier Science/Pergamon Press, 2003. p. 1-20.

CONSELHO FEDERAL DE PSICOLOGIA. *Lista dos Testes Aprovados*, 2005. Disponível em: <<http://pol.org.br>>. Acesso em: 2 dez. 2005.

EKSTROM, R. B.; FRENCH, J. W.; HARMAN, H. H. Cognitive factors: their identification and replication. *Multivariate Behavioral Research Monographs*, [S.l.], v. 79, n. 2, p. 1-85, 1979.

EKSTROM, R. B. *et al. Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service, 1976.

GORSUCH, R. L. *Factor analysis*. Philadelphia: W. B. Saunders Company, 1974.

GUSTAFSSON, J-E. The relevance of factor analysis for the study of group differences. *Multivariate Behavioral Research*, [S.l.], v. 27, n. 2, p. 239-247, 1992a.

\_\_\_\_\_. The “Spearman hypothesis” is false. *Multivariate Behavioral Research*, [S.l.], v. 27, n. 2, p. 265-267, 1992b.

MULAİK, S. A. Objectivity and multivariate statistics. *Multivariate Behavioral Research*, [S.l.], v. 28, n. 2, p. 171-203, 1993.

\_\_\_\_\_. Factor analysis is not just a model in pure mathematics. *Multivariate Behavioral Research*, [S.l.], v. 31, n. 4, p. 655-661, 1996.

SPEARRIT, D. Carroll’s model of cognitive abilities: educational implications. *International Journal of Educational Research*, [S.l.], v. 25, n. 2, p. 107-198, 1996.

Recebido em: março de 2007

Aceito em: novembro de 2007