

Características psicométricas da Bateria Padrão do Universal Nonverbal Intelligence Test (UNIT): um estudo preliminar¹

Bateria Padrão do UNIT

*Renata Ferrarez Fernandes Lopes
Ana Paula Gomes Moreira
Claudiane Aparecida Guimarães
Ederaldo José Lopes*

Resumo

Este trabalho teve como objetivo realizar a análise preliminar das características psicométricas da Bateria Padrão do Universal Nonverbal Intelligence Test (UNIT) numa amostra com 77 sujeitos. Os subtestes avaliaram as habilidades de atenção para detalhe, processamento seqüencial, mediação simbólica, formação de conceitos, integração viso-motora, pensamento abstrato, processamento holístico, capacidade de síntese, dentre outros. Como se trata de um teste não-verbal, a comunicação entre o aplicador e o avaliado foi feita a partir de gestos universais, e 100% dos participantes foram capazes de compreender as instruções. Uma análise de componentes principais indicou a presença de um componente (fator *g*) que explicou 90,01% da variância total do teste. O coeficiente de consistência interna (método das metades) é bastante satisfatório (0,99). Os dados obtidos, até o presente, mostram a viabilidade de aplicação deste teste, configurando-se numa alternativa para os testes intelectuais verbais conhecidos.

Palavras-chave: Universal Nonverbal Intelligence Test; Psicometria; Avaliação psicológica.

Psychometric characteristics of the Universal Nonverbal Intelligence Test (UNIT) Standard Battery: a preliminary study

Abstract

This work aimed to study psychometric characteristics of the Universal Nonverbal Intelligence Test (UNIT) - Standard Battery in a sample of 77 subjects. The UNIT subtests evaluated the attention abilities for detail, sequential processing, symbolic mediation, concepts formation, integration visor-motor, abstract thought, holistic processing, synthesis capacity, among others. As the UNIT is a non-verbal test, the communication between applicator and the subject was made starting from universal gestures, and 100% of the participants were able of comprehend instructions. An analysis of main components (PC) indicated the presence of 1 component (factor *g*) that explains 90,01% of the total variance. The Internal consistency coefficient (split-half) was very satisfactory (0,99). The results obtained, until the present moment, showed that the test has application viability, configuring itself as an option for the verbal well-known intellectual tests.

Key words: Universal Nonverbal Intelligence Test; Psychometrics; Psychological assessment.

¹ Esta pesquisa foi desenvolvida com o apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

Características Psicométricas de la Bateria Padrón del Universal Nonverbal Intelligence Test (UNIT): un Estudio Preliminar

Resumen

Este trabajo tuvo como objetivo realizar el análisis preliminar de las características psicométricas de la Bateria Padrón del *Universal Nonverbal Intelligence Test (UNIT)* en una muestra de 77 sujetos. Los sub-testes de esa batería evalúan las habilidades de atención para detalle, procesamiento secuencial, mediación simbólica, formación de conceptos, integración viso-motora, pensamiento abstracto, procesamiento holístico, capacidad de síntesis, entre otros. Como se trata de un test no-verbal, la comunicación entre el aplicador y el evaluado fue hecha a partir de gestos universales, y 100% de los participantes fueron capaces de entender las instrucciones. Un análisis de componentes principales (PC) ha indicado la presencia de 1 componente (factor g) que explica 90,01% de la variancia total, con $eigenvalue = 3,6$. El coeficiente de consistencia interna (método de las mitades) es bastante satisfactorio (0,99). Los datos obtenidos hasta este momento muestran la viabilidad de aplicación de este test no-verbal, de gran potencial para evaluar intelectualmente muestras brasileñas, configurándose en una alternativa para los testes intelectuales verbales conocidos.

Palabras-clave: Universal Nonverbal Intelligence Test; Psicometría; Evaluación psicológica.

Introdução

Aspectos Gerais

O UNIT (*Universal Nonverbal Intelligence Test*) é um instrumento que se baseia no conceito de inteligência como uma estrutura hierárquica formada por três níveis: a inteligência como uma habilidade geral (g) no topo, fatores como memória e raciocínio fluido no segundo nível e fatores mais particularmente definidos, como orientação espacial e flexibilidade de organização, no terceiro nível (Bracken & McCallum, 1998). De acordo com Bracken e McCallum (1998), os testes desenvolvidos a partir desse modelo, como é o caso do UNIT, permitem combinar uma cobertura abrangente das aptidões com flexibilidade de uso e têm sido cada vez mais utilizados nos países que o validaram tanto para fins teóricos como práticos (Anastasi, 1992, 1994).

O UNIT foi desenvolvido inicialmente com o objetivo de promover uma avaliação da inteligência geral em ampla extensão, de habilidades de memória e raciocínio (escalas primárias) e da capacidade de mediação simbólica e não-simbólica (escalas secundárias). Memória e raciocínio são construtos fundamentais da inteligência na medida em que envolvem habilidades associativas e cognitivas, respectivamente. Os subtestes de memória do UNIT foram desenvolvidos a fim de aferir o construto de forma mais ampla dos outros testes de inteligência que avaliam esta habilidade. Enquanto os outros testes

avaliam a memória de curto prazo através de tarefas mnemônicas de formato serial simples (por exemplo, memória para palavras, *span* de dígitos), bem como apresentam uma mediação verbal no que tange à aplicação dos itens e às respostas da criança, o UNIT toca em habilidades associativas mais amplas à medida que apresenta os estímulos de forma visual para o sujeito do teste (exigindo muitas vezes uma mediação simbólica envolvendo rotulação, organização e categorização) e exige respostas motoras (por exemplo, rearranjo físico de cartões de resposta), ampliando a sensibilidade de mensuração desta habilidade associativa. O raciocínio é considerado o “coração” da inteligência pois requerem do indivíduo habilidades de planejamento e avaliação baseados na dedução de relações. No UNIT, tarefas clássicas de raciocínio como desenho de cubos, matrizes progressivas, analogias e labirintos foram transformadas para permitirem aplicações não-verbais.

As escalas secundárias são subordinadas as primeiras e descrevem processos mentais subjacentes que facilitam a performance da criança nas questões não-verbais, a mediação simbólica (envolvendo representação, analogia e conceitualização) e mediação não-simbólica (envolvendo abstração e processamento não verbal). Em resumo, os quocientes simbólicos e não-simbólicos do UNIT advêm de um conjunto de processos mediacionais como aqueles obtidos nos

testes de inteligência tradicionais tais como as escalas Wechsler (Bracken & McCallum, 1998).

Quanto à organização, o UNIT é um teste completamente não-verbal, desenvolvido para ser aplicado a indivíduos na faixa etária de 5 a 17 anos, e as atividades a serem desenvolvidas são indicadas ao indivíduo através de gestos universais. Assim, as respostas dadas pelos participantes também são completamente não-verbais e são demonstradas através dos materiais que compõem o próprio teste, como cubos e fichas de plástico. O UNIT constitui-se de seis subtestes: Memória Simbólica, Desenho de Cubo, Memória Espacial, Raciocínio Analógico, Memória de Objeto e Labirintos. Esses subtestes podem ser agrupados em três baterias distintas: a Bateria Reduzida, composta pelos dois primeiros subtestes; a Bateria Padrão, composta pelos quatro primeiros subtestes e a Bateria Extensa, composta por todos os seis subtestes. A Bateria Reduzida deve ser utilizada quando o objetivo é realizar um mapeamento do funcionamento intelectual do indivíduo; a Bateria Padrão é apropriadamente utilizada para estabelecer decisões referentes a posicionamento educacional, e a Bateria Extensa, para uma avaliação mais profunda e estabelecimento de diagnósticos (Bracken & McCallum, 1998).

Quanto à interpretação do UNIT, convém lembrar que a testagem psicológica é muito mais do que a simples administração e correção de testes psicológicos; ela é, na verdade, um processo no qual múltiplos métodos e múltiplos tipos de informação são integrados de maneira significativa para a compreensão mais profunda de características individuais. Noronha e Vendramini (2003) afirmam que os testes de inteligência têm sido entendidos como ferramentas auxiliares na coleta de dados, e, portanto, devem ser utilizados junto com as demais informações qualitativas ou subjetivas organizadas pelo psicólogo. Eles auxiliam a compreensão do problema estudado, de forma a facilitar o processo de tomada de decisão ou inferência diagnóstica, aspecto fundamental no campo da Psicopedagogia. Nesse sentido, a interpretação do UNIT é conduzida através de múltiplos estágios, nos quais as informações são consideradas de forma sucessiva, desde as

informações mais confiáveis e globais (o que é chamado de Quociente de Inteligência Geral – FSIQ) até as informações mais específicas (comparações entre escalas, subtestes ou itens). Cada nível de informação é examinado de forma qualitativa e quantitativa. Além disso, a avaliação do UNIT permite uma análise ampla dos construtos que se propõe a medir, uma vez que além da estratégia normativa (interindividual) de interpretação, ele possibilita a realização de uma interpretação ipsativa (intra-individual) (Bracken & McCallum, 1998).

Finalmente, o UNIT oferece um valor adicional na predição de sucesso acadêmico e é utilizado como ferramenta de diagnóstico, sendo sensível às condições clínicas comuns e excepcionais (Bracken & McCallum, 1998; Fives & Flanagan, 2002).

A amostra de padronização: Estudos comparativos entre grupos

O estudo da comparação entre grupos geralmente leva em conta as diferenças entre os escores médios. Essa abordagem pode levar a vieses nos resultados, normalmente na direção daqueles grupos que obtêm um escore médio mais baixo num determinado teste. No entanto, não há nenhuma razão *a priori* para considerar que amostras que diferem em sexo, etnia, ou outras características apresentem o mesmo desempenho em testes de inteligência (Jensen, 1980).

O UNIT possui duas características que podem ser úteis na redução de vieses: o fato de ser um teste não-verbal e a extensa amostra utilizada na sua padronização. Essa amostra, constituída de 2.100 crianças e adolescentes, nas diferentes faixas etárias, foi escolhida levando-se em conta diferentes aspectos tais como gênero, etnia, região dos Estados Unidos, escolaridade etc. Dessa grande amostra, foi possível se fazer uma série de comparações entre subamostras, considerando os diferentes aspectos. Para ilustrar a seguir, algumas dessas comparações (todas extraídas de Bracken & McCallum, 1998) foram reduzidas tendo por base o sexo e a origem étnica.

Com relação à comparação entre grupos, por sexo, utilizando as três baterias do UNIT, pode-se dizer que eles não diferem em termos do coeficiente de inteligência geral. Entretanto, os quocientes de

raciocínio e não-simbólicos são levemente mais altos para os homens em relação às mulheres, enquanto os quocientes de memória e simbólicos são levemente mais altos para as mulheres, diferenças consistentes com os estudos de Halpern (1997).

Na comparação entre grupos, por origem étnica, a análise de uma subamostra de 352 examinandos americanos de origem africana mostrou que, na média, nas três baterias do UNIT, houve uma vantagem dos examinandos brancos, mas a diferença entre os escores dos indivíduos brancos e os de origem africana foi menor que 15 pontos, uma diferença freqüentemente encontrada na literatura (Suzuki & Valencia, 1997). Esses autores afirmam, todavia, que as diferenças encontradas entre os grupos podem ser devidas a fatores sócio-econômicos e não a fatores raciais. Por outro lado, a amostra de 49 americanos de origem asiática e de Ilhas do Pacífico mostrou um desempenho superior no UNIT em relação à amostra normativa de brancos americanos, resultado consistente com as conclusões de Suzuki e Valencia (1997). Já a amostra de hispânicos, comparada com uma amostra de não-hispânicos ($N = 194$), apresentou diferenças muito pequenas, variando de -0,14 a 3,06. Finalmente, num estudo comparativo entre 30 equatorianos e a amostra normativa americana mostrou diferenças no coeficiente de inteligência geral nos valores de 2,73, 5,33 e 3,26, respectivamente nas baterias abreviada, padrão e extensa.

Propriedades Psicométricas

Os índices de fidedignidade a partir dos escores padronizados do UNIT indicaram um alto nível de precisão na mensuração das habilidades avaliadas pelo teste. Uma análise dos índices de consistência interna, obtidos pelo método das metades com correção pela fórmula de Spearman-Brown (0,93 e 0,90 para a bateria padrão em amostras não-clínicas e clínicas, respectivamente), e estabilidade temporal obtida através de teste e reteste (0,88 para a bateria-padrão em amostras não-clínicas) assegura a eficiência do teste, podendo este ser usado em diagnósticos clínicos para excepcionidades específicas (distúrbios de aprendizagem, distúrbios da fala ou linguagem, retardo mental), distúrbios cognitivos, bem como para

a separação em grupos categorizados por sexo, raça e etnia (Bracken & McCallum, 1998).

Com relação à validade, estudos de validade de construto (Bracken & McCallum, 1998), realizados com a técnica de extração de componentes principais e análise fatorial (Gorsuch, 1983), com grupos de americanos de origem africana e hispânicos sustentam a interpretação de um fator de inteligência geral único, muito embora uma solução de dois fatores (memória e raciocínio) também seja razoável na interpretação dos resultados obtidos nesses estudos. Evidências de validade externa do UNIT também foram apresentadas. Comparativamente a outros instrumentos de medida da inteligência, o UNIT possui bons parâmetros correlacionais. Especificamente, o desempenho de sujeitos no UNIT foi estudado em relação ao desempenho de sujeitos em outros testes de inteligência, como o WISC-III (Wechsler, 1991) e o *Standard Progressive Matrices* (Raven, 1958). Por exemplo, com amostras de crianças com problemas de aprendizagem, as correlações com o WISC III foram, respectivamente, para as baterias abreviada, padrão e extensa do UNIT: 0,78, 0,84 e 0,83. Na mesma ordem de apresentação das baterias, as correlações do UNIT com as matrizes progressivas de Raven foram mais baixas: 0,50, 0,56 e 0,59.

Athnasiou (2000) realizou um estudo teórico comparativo entre cinco testes não-verbais de inteligência do ponto de vista de suas propriedades psicométricas. Os resultados desse estudo revelaram que o UNIT possui alta consistência interna (considerando todas as faixas etárias, o índice é 0,93); itens com gradientes de dificuldade apropriados, revelando sensibilidade para pequenas diferenças ou aumentos nas habilidades mensuradas; manual constituído de explicações adequadas, dentro dos padrões de validade; elevado índice de validade nos escores e itens específicos para diversos grupos de examinandos.

Além dos estudos citados acima, alguns outros autores têm aprofundado a análise das propriedades psicométricas do UNIT, considerando somente a amostra de padronização, ou acrescentando novas amostras. Por exemplo, Eikum e Baker (2001) realizaram um estudo para verificar a influência da

linguagem na avaliação da inteligência em diferentes culturas. Nesse estudo, crianças de origem norte-americana foram comparadas a crianças de origem norueguesa quanto aos escores obtidos nos subtestes do UNIT. As conclusões mostraram que a estruturação do teste permite maior redução da interferência de aspectos da linguagem nos resultados finais.

Caruso e Witkiewitz (2001) avaliaram as habilidades de memória e raciocínio, medidas no UNIT, através da análise de componente fidedigno (*Reliable Component Analysis – RCA*), uma alternativa à análise de componentes principais. O objetivo do estudo foi verificar, utilizando a amostra de padronização de 2.100 participantes, se havia uma vantagem substancial em utilizar a RCA, comparativamente à forma de análise dos escores apresentada no manual do UNIT (Bracken & McCallum, 1998). Os resultados com a RCA mostraram diferenças maiores entre os escores de memória e raciocínio, com aumento na fidedignidade entre os escores dessas habilidades (de 0,70 conforme o manual do UNIT para 0,88 utilizando RCA). De acordo com Caruso e Witkiewitz (2001), o aumento da fidedignidade dos escores aumenta a validade discriminante em relação ao escore total do teste. Um outro estudo dessa mesma natureza foi feito por Wilhoit e McCallum (2002). Utilizando estatística multivariada (Hair, Anderson, Tatham & Black, 2005), eles realizaram uma análise de perfis a partir dos escores da amostra de padronização do UNIT. Os resultados sustentaram a estrutura fatorial subjacente ao UNIT e permitiram avaliar diferenças mais profundas entre os vários *clusters* obtidos, o que permite um uso diagnóstico mais confiável do UNIT para objetivos educacionais, por exemplo.

Considerando amostras com algum tipo de dificuldade ou deficiência, Maller (2000) estudou o funcionamento diferencial dos itens do UNIT em quatro subtestes, considerando grupos com e sem deficiência auditiva: memória simbólica, memória espacial, raciocínio analógico e memória para objeto. Utilizando os procedimentos de detecção de funcionamento diferencial de itens, incluindo a teoria de resposta ao item, ela não encontrou nenhum item nesses subtestes que mostrasse funcionamento diferencial, indicando que a probabilidade de uma resposta correta parece não

ser afetada pela qualidade auditiva. Maller e French (2004) avaliaram a invariância da estrutura fatorial do UNIT através da análise da amostra de padronização e uma amostra de 102 participantes com deficiência auditiva. Utilizando procedimentos estatísticos (por exemplo, o índice de *goodness-of-fit*) e análises de *follow-up*, eles avaliaram dois modelos teóricos, um primário, com dois fatores (memória e raciocínio) e o outro secundário, também com dois fatores (simbólico e não-simbólico). As formas gerais dos dois modelos se mostraram invariantes através das amostras estudadas, sustentando os dois modelos teóricos. Num estudo voltado para problemas emocionais, Pendley, Myers e Brown (2004) avaliaram o desempenho de 29 crianças com transtorno de atenção e hiperatividade. Os resultados mostraram que as crianças com esse transtorno têm, em média, áreas específicas de *déficits* ligados aos processos executivos (*working memory*, planejamento) que contribuem para escores de quociente intelectual (QI) mais baixos. Entretanto, asseguram os autores, o UNIT precisa ser interpretado com precaução e não pode ser utilizado, neste caso, como ferramenta para se decidir se uma criança tem ou não o transtorno. Finalmente, um estudo de Borghese e Gronau (2005) abordou a questão da proficiência limitada em língua inglesa. Os autores compararam as diferenças dos escores médios e as correlações entre o UNIT e o WISC-III em estudantes americanos de origem mexicana. Os resultados revelaram correlações maiores que as esperadas entre as escalas completas do UNIT e do WISC-III e os QIs verbais. Os dados sustentam as validades convergente e discriminante do UNIT para a população alvo. Além disso, os QIs da escala completa do UNIT foram significativamente mais altos em relação ao WISC-III.

Em resumo, todos esses estudos apontam para a viabilidade do emprego do UNIT em diversas circunstâncias, desde a avaliação da inteligência *per se* até o seu emprego para fins diagnósticos.

Tendo por base a importância de se utilizar um teste não-verbal de inteligência e a escassez desse tipo de instrumento em nosso meio, o objetivo deste trabalho foi apresentar dados preliminares sobre a aplicação, padronização e características psicométricas do UNIT

no contexto brasileiro, inicialmente com estudantes de escolas públicas da cidade de Uberlândia, MG.

Método

Participantes

Participaram deste estudo 77 crianças com idade variando de 6 anos e 11 meses a 11 anos e 10 meses, idade média de 7 anos e 9 meses, 37 do sexo masculino e 40 do sexo feminino todas estudantes do ensino fundamental de uma escola estadual da cidade de Uberlândia.

Instrumento

Utilizou-se um kit de aplicação do UNIT contendo: 1 dos livros-estímulo (livro 1) contendo os 4 subtestes avaliados na bateria-padrão: memória simbólica e espacial, desenho do cubo e raciocínio analógico, cronômetro e formulário de registro da resposta do sujeito .

Procedimento

Inicialmente, solicitou-se aos pais dos alunos autorização por escrito para a participação na pesquisa. Foi enviada a cada pai uma carta contendo as explicações sobre o UNIT e os objetivos da aplicação da bateria-padrão, assim como o termo de consentimento e autorização para publicação dos dados. As crianças, cujos pais forneceram o consentimento, foram testadas individualmente em sessões de cerca de 40 minutos. A aplicação do instrumento foi precedida pelo estabelecimento de um breve *rapport* pela pesquisadora informando os objetivos do teste, assim como o significado das pantomimas e gestos utilizados durante a sessão. A testagem só foi iniciada após a compreensão e a cooperação dos examinandos terem sido asseguradas.

A partir do início da aplicação do teste, as explicações sobre os procedimentos de administração se limitaram a pantomimas e a demonstrações não-verbais, conforme indicado no manual e demais materiais que compõem o UNIT.

Antes de iniciar cada subteste, o examinador posicionou o livro de estímulos a 45 cm sobre uma

mesa retangular, em frente ao examinando. Neste espaço de trabalho foram colocados os materiais de respostas do itens dos subtestes de memória simbólica, desenho de cubo e memória espacial à medida que cada subteste era aplicado. No subteste raciocínio analógico, os examinandos apontaram as respostas no próprio livro de estímulos. Para os examinandos destros, o examinador se posicionou sempre à direita do examinando e para os sinistros, o examinador se posicionou sempre à esquerda do examinando.

Resultados e Discussão

Com relação à aplicação do teste, é necessário considerar que os escores obtidos em um teste psicológico só serão representativos do construto que ele supõe medir se a administração do teste for realizada uniformemente pelos examinadores. Assim, o fato de as instruções do UNIT serem apresentadas de forma não-verbal o diferencia dos outros testes não-verbais de inteligência, tornando-se um desafio para o aplicador acostumado com a comunicação verbal das instruções durante a avaliação intelectual.

No processo de padronização de um teste não-verbal com as peculiaridades de aplicação do UNIT, é preciso enfatizar a precisão na administração das instruções, pois isso garante que os resultados da avaliação não se deva a uma tentativa não-padronizada de comunicação não-verbal para auxiliar na compreensão da demanda da tarefa exigida pelo subteste.

Desta forma, a aplicação dos subtestes da bateria padrão revelou que 100% dos participantes foram capazes de compreender as instruções não-verbais, dadas a partir dos itens de demonstração e de amostra; 95% acertaram os itens de checagem de compreensão da tarefa na primeira tentativa e 100% dos participantes compreenderam os sinais universais utilizados durante a aplicação. Conclui-se, portanto, que os resultados obtidos são confiáveis da perspectiva da padronização do teste quanto à sua aplicação.

Com relação à análise dos resultados, os escores médios obtidos pelos sujeitos estão apresentados na Tabela I.

Tabela 1. Médias e desvios-padrão dos subtestes do UNIT da amostra brasileira e americana nas diferenças em pontos

Quocientes	Amostra Brasileira (N=77)		Amostra Americana (N= 1.159)		Diferença entre as amostras
	Média	Desvio Padrão	Média	Desvio Padrão	
Quociente de Memória	77,64	11,99	100,82	14,49	23,18
Quociente de Raciocínio	76,99	10,85	100,63	13,66	23,64
Quociente Simbólico	76,96	12,20	101,00	14,09	24,04
Quociente não simbólico	79,40	11,04	100,41	13,87	21,01
FSIQ	75,20	11,45	100,83	13,79	25,63

A despeito do fato de que é necessário estabelecer uma amostra normativa brasileira, com a qual os dados obtidos pelos sujeitos brasileiros devem ser comparados, foram feitas comparações baseadas nas tabelas americanas, uma vez que o estabelecimento da amostra normativa brasileira requererá mais sujeitos do que os atuais 77. Comparados com a amostra normativa americana, os participantes apresentam escores rebaixados com diferenças em torno de 20 pontos para mais. Isto pode ser explicado, em parte, pelo tamanho da amostra brasileira, mas há outras explicações plausíveis: se se levar em conta um grupo de sujeitos culturalmente mais próximos dos brasileiros (no caso, os equatorianos), pode-se verificar uma semelhança maior em termos de

desempenho no UNIT (Tabela 2). Se os escores médios da amostra brasileira por idade forem levados em consideração, pode-se observar que a amostra de 9-11 anos obteve um desempenho mais parecido com a amostra equatoriana de 30 sujeitos (idade média de 9,4 anos), nos dados apresentados por Bracken e McCallum (1998). Mesmo assim, há um rebaixamento dos escores da amostra brasileira, o que implica dizer, em termos de comparações transculturais, que há variáveis das quais não se tem controle, tais como nível sócio-econômico e escolaridade.

Considerando que a interpretação dos dados do UNIT é conduzida através de múltiplos estágios que vão da fonte de dados mais global e confiável que é o quociente de inteligência geral (FSIQ - dado é

Tabela 2. Médias e desvios-padrão dos subtestes do UNIT da amostra brasileira e equatoriana nas diferenças em pontos

Quocientes	Amostra Brasileira (N=32)		Amostra Equatoriana (N=30)		Diferença entre as amostras
	Média de idade: 10 anos e 7 meses		Média de idade: 9 anos e 4 meses		
	Média	Desvio Padrão	Média	Desvio Padrão	
Quociente de Memória	93,53	14,61	98,00	13,16	4,47
Quociente de Raciocínio	89,40	12,47	98,33	14,66	8,93
Quociente Simbólico	90,09	12,54	93,78	14,40	3,69
Quociente não simbólico	92,75	13,67	102,2	14,06	9,45
FISIQ	90,31	13,25	98,07	14,00	7,76

composto a partir das duas escalas do teste) até suas fontes específicas (quocientes de raciocínio, memória, simbólico e não-simbólico), realizou-se uma análise das discrepâncias entre os quocientes que compõem as escalas para verificar a confiabilidade do FSIQ .

A interpretação das escalas primária e secundária permite os seguintes padrões de discrepância possíveis. Quociente de memória (QM) < Quociente de raciocínio (QR); Quociente de memória (QM) > Quociente de raciocínio (QR), Quociente simbólico (QS) < Quociente não-simbólico (QNS) e Quociente simbólico (QS) > Quociente não-simbólico (QNS). Além disso, pode não haver discrepância significativa entre os quocientes (NS). Estes cálculos são feitos a partir dos escores padrões dos quocientes anteriores em função das idades. A distribuição da frequência desta amostra, em relação a estes padrões, pode ser observada na Tabela 3.

(validade e fidedignidade), optando-se, neste momento, por utilizar a análise de componentes principais (CP), que, a exemplo da análise fatorial (AF), fornece meios de agrupar um conjunto de variáveis observadas, expressas por meio de um conjunto reduzido de variáveis hipotéticas (Gorsuch, 1983; Pasquali, 2000).

A análise de componentes principais com rotação oblíqua (OBLIMIN), realizada a partir dos quocientes de memória, raciocínio, simbólico e não-simbólico (escores padronizados), indicou a presença de apenas um componente principal (Quociente de inteligência geral - fator g- FSIQ) que explica 90,01% variância total. Este resultado se refere à validade de construto (o quanto se pode dizer que um teste mede um construto ou um traço, neste caso o fator g). Comparativamente ao estudo original de validação

Tabela 3. Distribuição da frequência absoluta desta amostra em relação às possíveis discrepâncias entre os quocientes das escalas primárias ou secundárias

<i>Discrepâncias</i>	<i>Frequência absoluta observada</i>
NS (raciocínio/memória)	56
NS (simbólico/não-simbólico)	61
QM>QR	3
QM<QR	4
QS>QNS	10
QS<QNS	11

Legenda: NS = Não-significativo; QM = Quociente De Memória; QR = Quociente de Raciocínio; QS = Quociente Simbólico; QNS = Quociente Não-Simbólico.

O teste qui-quadrado, a partir dos dados da Tabela 3, apontou um valor igual a 148,9 ($p < 0,001$), ou seja, há diferença na frequência observada em relação à frequência esperada de participantes da amostra para as categorias discrepâncias não-significativas entre quocientes da amostra (frequências absolutas maiores) em relação às discrepâncias significativas dos mesmos quocientes (frequências absolutas menores), indicando que os FSIQs da amostra são confiáveis.

Em seguida, procedeu-se ao estudo preliminar e exploratório das características psicométricas

do UNIT, que também usou o método de análise de componentes com rotação oblíqua (Bracken & McCallum, 1998), a análise dos dados mostrou um *eigenvalue* = 3,60 ($n=77$), enquanto no estudo original o valor do *eigenvalue* foi 2,33 ($n=1.996$). Pode-se dizer que os dados da análise de componentes apontam de forma bastante contundente (considerando o pequeno tamanho da amostra) para o estabelecimento de um critério de validade de construto bastante aceitável.

Para a análise da fidedignidade da bateria-padrão, utilizou-se o coeficiente de consistência interna

calculado pelo método das metades com correção pela fórmula de Spearman-Brown, a exemplo do que foi feito por Bracken e McCallum (1998). O resultado obtido foi 0,99. Este índice é um requisito fundamental do construto psicológico medido, pois diz respeito à uniformidade e coerência dos itens que compõem o teste (Anastasi & Urbina, 2000). No estudo original (Bracken & McCallum, 1998), o coeficiente geral obtido a partir da aplicação da bateria-padrão foi 0,89 e para a bateria completa, 0,88. Tanto os coeficientes obtidos no estudo original quanto o obtido com os participantes deste estudo, pode-se verificar uma consistência interna muito elevada, considerando os instrumentos psicológicos.

Em conjunto, os dados apresentados com a amostra estudada até o presente mostram uma boa perspectiva na adaptação, padronização e características psicométricas do UNIT. A boa compreensão dos sujeitos em faixas etárias mais baixas e mais altas, assim como a aplicação do teste tal qual construído e aplicado em sujeitos americanos reflete a capacidade que o instrumento possui de minimizar ou neutralizar as diferenças de ordem cultural (Fives & Flanagan, 2002), as quais podem dificultar o relacionamento entre examinador e examinando.

Tendo atendido de forma satisfatória à padronização das normas de aplicação, é necessário prosseguir os estudos no sentido de estabelecer um outro aspecto importante da padronização do teste, qual seja, a construção de normas brasileiras que sirvam de marco de comparação para as diversas populações cuja inteligência se constitui em variável importante de se conhecer. Isto vale, principalmente, no campo educacional em que propostas pedagógicas devem levar em conta o tipo de aprendizagem que se quer estabelecer e para quem. O valor diagnóstico e prognóstico do UNIT é de fundamental importância na avaliação qualitativa do tipo de habilidade que se quer avaliar e/ou desenvolver.

Finalmente, os índices psicométricos obtidos numa amostra ainda insuficiente para estudos fatoriais mais confiáveis já revelam o potencial da bateria como um todo. Tanto a validade do construto quanto à consistência interna do teste asseguram a viabilidade de utilização deste instrumento no contexto escolar.

Nesse sentido, a avaliação intelectual permitida pelo UNIT, quantitativa e qualitativa, dá importante referencial para psicólogos e educadores proporem atividades psicopedagógicas que assegurem, de fato, o desenvolvimento das habilidades cognitivas necessárias para a vida presente e futura do jovem estudante.

Referências

- Anastasi, A. (1992). Are there unifying trends in the psychologies of the 1990s? In M.E. Donnelly (Ed.), *Reinterpreting the legacy of William James* (pp. 29-48). Washington, DC: American Psychological Association.
- Anastasi, A. (1994). Aptitude testing. Em *Encyclopedia of Human Behavior* (pp. 211-221). San Diego, CA: Academic Press.
- Anastasi, A., & Urbina, S. (2000). *Testagem psicológica* (7a. ed.). Trad. Maria Adriana Veríssimo Veronese. Porto Alegre, RS: Artes Médicas Sul.
- Athanasίου, M. S. (2000). Current nonverbal assessment instruments: a comparison of psychometric integrity and test fairness. *Journal of Psychoeducational Assessment*, 18, 211-229.
- Bracken, B. A., & McCallum, R. S. (1998). *Universal Nonverbal Intelligence Test: Examiner's Manual*. Itasca, IL: Riverside Publishing Company.
- Borghese, P. & Gronau, R.C. (2005). Convergent and discriminant validity of the Universal Nonverbal Intelligence Test with limited English proficient Mexican-American elementary students. *Journal of Psychoeducational Assessment*, 23, 128-139.
- Caruso, J.C. & Witkiewitz, K. (2001). Memory and reasoning abilities assessed by the Universal Nonverbal Intelligence Test: A reliable component analysis (RCA) study. *Educational and Psychological Measurement*, 61, 5-22.
- Eikum, F.S., & Baker, K. D. (2001). *A Cross Cultural Study of the Universal Non-verbal Intelligence Test* [On line]. Disponível: <https://www.swpsych.org/programs.heitml?ref=262&year=2001>.
- Fives, C. J., & Flanagan, R. (2002). A review of the Universal Nonverbal Intelligence Test (UNIT): An Advance for Evaluating Youngsters with Diverse Needs. *School Psychology International*, 23, 425-448.

- Gorsuch, R. L.(1983). *Factor analysis* (2nd. ed.). Hillsdale, NJ: Erlbaum.
- Hair, J.F.Jr., Anderson, R.E., Tatham, R.L., & Black, W.C. (2005). *Análise multivariada de dados* (5^a Ed.). Trad. Adonai Schlup Sant'Anna e Anselmo Chaves Neto. Porto Alegre, RS: Bookman.
- Halpern, D.F.(1997). Sex differences in intelligence: Implications for education. *American Psychologist*, 52, 1091-1102.
- Jensen, A. (1980). *Bias in mental testing*. New York, NY: The Free Press.
- Maller, S.J. (2000). Item invariance in four subtests of the Universal Nonverbal Intelligence Test (UNIT) across groups of deaf and hearing children. *Journal of Psychoeducational Assessment*, 18, 240-254.
- Maller, S.J. & French, B.F. (2004). Universal Nonverbal Intelligence Test factor invariance across deaf and standardization samples. *Educational and Psychological Measurement*, 64, 647-660.
- Noronha, A. P. P., & Vendramini, C. M. M.(2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16, 177-182.
- Pasquali, L.(2000). *Análise fatorial para pesquisadores*. Brasília, DF: UnB Editora.
- Pendley, J.D., Myers, C.L. & Brown, R.D. (2004). The Universal Nonverbal Intelligence Test with children with attention-deficit hiperactivity disorder. *Journal of Psychoeducational Assessment*, 22, 124-135.
- Raven, J. C.(1958). *Standard Progressive Matrices*. Cambridge, MA: Cambridge University Press.
- Suzuki, L.A., & Valencia, R. R.(1997). Race-ethnicity and measured intelligence: Educational implications. *American Psychologist*, 52, 1103-1114.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-3rd*. Edition. San Antonio, TX: The Psychological Corporation.
- Wilhoit, B.E. & McCallum, R.S. (2002). Profile analysis of the Universal Nonverbal Intelligence Test standardization sample. *School Psychology Review*, 31, 263-281.
- Recebido em: 01/06/2006
Revisado em: 14/11/2006
Aprovado em: 18/12/2006

Sobre os autores

Renata Ferrarez Fernandes Lopes (rfernandeslopes@fapsi.ufu.br) é professora Adjunta do Instituto de Psicologia, Universidade Federal de Uberlândia, MG.

Ederaldo José Lopes (ederaldol@umuarama.ufu.br) é professor Associado do Instituto de Psicologia, Universidade Federal de Uberlândia, MG.

Ana Paula Gomes Moreira (anapaulamoreira05@yahoo.com.br) é graduanda em Psicologia e Bolsista de Iniciação Científica da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, Processo SHA 639/04) do Instituto de Psicologia, Universidade Federal de Uberlândia, MG.

Claudiane Aparecida Guimarães (claudiane@com4.com.br) é graduanda em Psicologia e Bolsista de Iniciação Científica da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, Processo SHA 639/04) do Instituto de Psicologia, Universidade Federal de Uberlândia, MG.

Endereço para correspondência

Renata Ferrarez Fernandes Lopes
IP/UFU - Avenida Pará, 1720 Bloco 2C- Campus Umuarama
38405320, Uberlândia, MG