

# Fidedignidade e validade de construto do Teste de Aprendizagem Auditivo-Verbal de Rey em idosos brasileiros

Reliability and construct validity of the Rey-Auditory Verbal Learning Test in Brazilian elders

Jonas Jardim de Paula<sup>1</sup>, Larissa Pacheco Cunha Melo<sup>1</sup>, Rodrigo Nicolato<sup>2</sup>, Edgar Nunes de Moraes<sup>3</sup>, Maria Aparecida Bicalho<sup>3</sup>, Amer Cavalheiro Hamdan<sup>4</sup>, Leandro Fernandes Malloy-Diniz<sup>2</sup>

<sup>1</sup>Laboratório de Investigações Neuropsicológicas do Instituto Nacional de Ciência e Tecnologia de Medicina Molecular (INCT-MM) da Universidade Federal de Minas Gerais (UFMG).

<sup>2</sup>Departamento de Saúde Mental da UFMG.

<sup>3</sup>Departamento de Clínica Médica da UFMG.

<sup>4</sup>Departamento de Psicologia da Universidade Federal do Paraná (UFPR).

Recebido: 7/6/2011 – Aceito: 31/10/2011

## Resumo

**Contexto:** O Teste de Aprendizagem Auditivo-Verbal de Rey (RAVLT) é amplamente utilizado para a avaliação da memória episódica. Suas propriedades psicométricas, porém, não foram bem analisadas no Brasil. **Objetivos:** Buscar evidências de fidedignidade e validade de construto do RAVLT e analisar a influência de idade, gênero, escolaridade e sintomas depressivos no desempenho. **Métodos:** Cento e vinte e seis idosos saudáveis realizaram o RAVLT, o Miniexame do Estado Mental (MEEM), o Desenho do Relógio (DR) e a Escala de Depressão Geriátrica. A fidedignidade foi avaliada pela análise de consistência interna e a validade de construto, pela estrutura fatorial e correlações com o MEEM e o DR. A influência da idade, escolaridade e sintomas depressivos foi estimada mediante regressão linear, enquanto diferenças de gênero foram avaliadas comparando o desempenho de homens e mulheres. **Resultados:** O teste apresenta alta consistência interna e estrutura bifatorial relacionada aos processos de armazenamento e evocação da memória episódica. O teste mostrou, em geral, correlações fracas com o MEEM e o DR. Apenas a idade e o gênero influenciaram o desempenho na tarefa. **Conclusão:** Nossos resultados indicam que a versão do RAVLT analisada apresenta bons indícios de fidedignidade e validade de construto, atestando sua aplicabilidade em contextos clínicos e de pesquisa para a população estudada.

Paula JJ, et al. / *Rev Psiq Clín.* 2012;39(1):19-23

**Palavras-chave:** Teste de Aprendizagem Auditivo-Verbal de Rey, RAVLT, validade, fidedignidade, memória episódica, aprendizagem verbal.

## Abstract

**Background:** The Rey Auditory-Verbal Learning Test (RAVLT) is widely used for the assessment of episodic memory. However, there are few studies in Brazil assessing its psychometric properties. **Objectives:** To search for evidence of reliability and construct validity of the RAVLT, and to assess the influence of age, schooling, gender, and depressive symptoms on test performance. **Methods:** One hundred twenty-six healthy older adults (aged 60 and over) performed the RAVLT, Mini-Mental State Exam (MMSE), Clock Drawing Test (CDT) and the Geriatric Depression Scale. Reliability was assessed by analysis of internal consistency, and construct validity by factor analysis and correlations with the MMSE and CDT. The influence of age, schooling and depressive symptoms was estimated by conducting linear regression analysis, and the role of gender by comparing the performance of males and females. **Results:** The RAVLT showed a high internal consistency, weak correlations with the MMSE and CDT, and a bifactorial structure, which is related to the processes of learning and episodic memory retrieval. Only age and gender affected test performance. **Discussion:** Our results provide evidence of reliability and construct validity in the tested RAVLT version, attesting its potential for clinical and research purposes for the Brazilian elderly population.

Paula JJ, et al. / *Rev Psiq Clín.* 2012;39(1):19-23

**Keywords:** Rey Auditory-Verbal Learning Test, RAVLT, validity, reliability, episodic memory, verbal learning.

## Introdução

A população idosa brasileira tem aumentado significativamente desde meados do século passado e deve superar a marca de 20 milhões de habitantes ainda nesta década<sup>1</sup>. Com o maior número de idosos e a maior expectativa de vida, crescem a prevalência de quadros de demências e seu estágio pródrômico, o comprometimento cognitivo leve, dado que o envelhecimento é um dos principais fatores de risco para o desenvolvimento de tais quadros. Um estudo populacional no contexto brasileiro<sup>2</sup> encontrou prevalência de 7,1%, sendo a doença de Alzheimer o tipo de demência mais comumente diagnosticado (cerca de 55% dos casos). A prevalência do comprometimento cognitivo leve ainda é debatida, oscilando entre 3% e 20% em pacientes idosos<sup>3</sup>.

A avaliação cognitiva nos quadros demenciais ou de comprometimento cognitivo leve é fundamental para o diagnóstico clínico

e precoce. O uso de testes neuropsicológicos é uma das principais ferramentas para tal avaliação, pois eles permitem que o paciente seja comparado a um grupo ou referencial normativo apropriado, possibilitando o exame objetivo da memória. Tal postura é essencial, visto que muitas vezes pacientes apresentam dificuldades significativas em relatar suas dificuldades cognitivas, conforme alguns estudos sobre o tema na revisão de Dourado *et al.*<sup>4</sup>, e a percepção destas é influenciada por diferentes fatores como o próprio comprometimento cognitivo e sintomas de ansiedade<sup>5</sup>.

O Teste de Aprendizagem Auditivo-Verbal de Rey (RAVLT) foi proposto originalmente como instrumento neuropsicológico para avaliação dos processos de aprendizagem e memória<sup>6</sup>. Foi adaptado para a população brasileira por Malloy-Diniz *et al.*<sup>7</sup>, com uma segunda versão publicada posteriormente pelo mesmo autor<sup>8</sup>, com algumas modificações que favorecem sua aplicação na popu-

lação idosa brasileira como alteração da lista de palavras original por uma nova contendo dissílabos de alta frequência no português brasileiro. Em um estudo recente<sup>9</sup> foram avaliadas as relações entre idade, escolaridade e gênero e desempenho no teste, encontrando correlações significativas entre essas três variáveis e os processos de aprendizagem e memória avaliados.

Em estudos com populações de outros países, o RAVLT tende a apresentar boas propriedades psicométricas. A consistência interna é, em geral, acima de 0,9 e ele apresenta boas correlações com outros instrumentos para avaliação da memória episódica, como o *California Verbal Learning Test*<sup>10</sup>.

Fichman *et al.*<sup>11</sup> realizaram o primeiro estudo de validade de construto do RAVLT para a população idosa no contexto brasileiro. Utilizando correlações convergentes (instrumentos de exame cognitivo com correlação positiva tendem a mensurar, em algum nível, construtos semelhantes), os autores correlacionaram os componentes "A7", "REC", "Total" e um índice composto computado pelo total de palavras evocadas em A5 menos o total de palavras evocadas em A1 (em que se consideraria o total de palavras memorizadas, ponderando-se a memória de curto prazo), com um teste de aprendizagem de figuras que é independente de escolarização formal. Seus achados indicaram correlações moderadas ( $r = 0,528$ ,  $p < 0,01$ ) no componente de evocação e fraca ( $r = 0,197$ ,  $p < 0,01$ ) no de reconhecimento.

Contudo, uma questão metodológica importante ao utilizar correlações convergentes ou divergentes com outros instrumentos para a análise da validade de construto é que, embora significativas, a magnitude de efeito obtida por tais correlações tende a ser baixa. No trabalho de Fichman *et al.*<sup>11</sup>, a correlação moderada das etapas de evocação representa aproximadamente 27% de variância ( $r^2$ ), enquanto a tarefa de reconhecimento, apenas 4%. Assim, respectivamente, 73% e 96% da variância comum entre os dois instrumentos são explicados por outros fatores que não o construto latente a ser validado. Uma das possíveis explicações para a variância não explicada naquele estudo é que o RAVLT possui estímulos verbais, enquanto o outro instrumento analisado possui estímulos visuais. Ainda devem ser consideradas a intercorrelação da idade sobre tais testes (o total de variância explicada por tal variável pode diferir de instrumento para instrumento, mascarando as correlações) e demais variáveis como escolaridade, inteligência geral e sexo. Diante de tais vieses, a validade de construto realizada mediante correlações convergentes deve ser complementada por outras formas de validação, buscando evidências convergentes de tal propriedade.

Quando o desempenho em diversos testes neuropsicológicos é correlacionado, os resultados geralmente agrupam-se em um componente único, responsável pela maior parte da variância encontrada<sup>12</sup>. Tal metodologia visa à identificação de um traço latente, ou seja, um construto cognitivo relacionado às tarefas em questão que possa ser inferido com base na associação dos desempenhos em diferentes tarefas cognitivas. A identificação de traços latentes foi historicamente associada ao desenvolvimento da psicologia das diferenças individuais e aos estudos sobre inteligência e personalidade. O método mais comumente utilizado em tal investigação é a análise fatorial exploratória, técnica estatística multivariada que trabalha com o padrão de correlações ou covariância apresentado pelos itens de um determinado conjunto de dados, buscando padrões latentes de intercorrelações<sup>13</sup>. Tal metodologia é uma das mais indicadas na avaliação da validade de construto de testes psicológicos e neuropsicológicos, pois permite o teste de hipóteses acerca dos traços latentes avaliados pelos instrumentos. Em um estudo recente<sup>14</sup>, a estrutura latente de um protocolo de exame cognitivo composto por vários testes foi analisada em uma população clínica de idosos brasileiros, encontrando-se um único fator responsável por mais de 58% da variância total.

O RAVLT é um teste destinado à avaliação da memória episódica, com componentes relacionados à memória de curto prazo, à aprendizagem, à memória de longo prazo imediata e tardia e à memória de reconhecimento. Sua estrutura fatorial, contudo, é heterogênea e dependente da população estudada. Em grupos clínicos de pacientes

com epilepsia ou transtornos psiquiátricos<sup>15</sup>, o teste tende a apresentar um modelo bifatorial, com um componente ligado à memória de curto prazo e um ligado à memória de longo prazo. Na população geral, são geralmente encontrados dois ou três fatores: memória de curto prazo, memória de longo prazo e memória de reconhecimento<sup>16</sup>. O teste tem sido amplamente utilizado em contextos clínicos e de pesquisa em diversos países, sendo considerado uma medida válida e eficaz para a mensuração da memória episódica e sensível às deficiências de memória encontradas em vários quadros clínicos como a demência de Alzheimer<sup>17</sup>, comprometimento cognitivo leve<sup>17</sup> e depressão maior<sup>18</sup>, demonstrando boa validade de critério.

O presente estudo tem por objetivo avaliar a fidedignidade e a validade de construto do RAVLT. Serão investigadas as hipóteses: 1) *os itens de memória episódica do RAVLT apresentarão elevada consistência interna (alfa de Cronbach > 0,8)*; 2) *as diversas etapas do RAVLT apresentarão correlações fracas ou moderadas com o escore total do Miniexame do Estado Mental e o Desenho do Relógio*; 3) *sua estrutura fatorial apresentará modelo bi ou trifatorial, com componentes relacionados aos processos de armazenamento e evocação* e 4) *a idade, escolaridade, gênero e sintomas depressivos influenciarão significativamente o desempenho da tarefa*.

O trabalho é associado ao projeto Binômio Depressão e Demência no idoso aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais (334/06).

## Método

### Participantes

Foram avaliados 126 idosos saudáveis recrutados por meio de anúncios locais na região metropolitana de Belo Horizonte, 63 mulheres e 63 homens. A idade média da amostra foi de 71,74 (DP = 7,9; Min = 60 e Max = 89) anos e escolaridade média de 6,74 (DP = 2,3; Min = 4 e Max = 17) anos. A idade apresentou distribuição normal na amostra estudada, atestada pelo teste Kolmogorov-Smirnov ( $Z = 1,031$ ,  $p = 0,239$ ), enquanto a escolaridade se polarizou em quatro (35%) e oito anos (50%). Os critérios de inclusão para o grupo controle foram idade superior a 60 anos, escolarização formal mínima de quatro anos, resultado do Miniexame do Estado Mental (MEEM) acima da nota de corte proposta por escolaridade<sup>19</sup> (média de 27,96, DP = 1,57), nenhum comprometimento funcional atestado pelos inventários Katz *et al.*<sup>20</sup> e Lawton e Brody<sup>21</sup>, que avaliam atividades de vida diária básicas e instrumentais, pontuação inferior a 6 na Escala de Depressão Geriátrica, instrumento de rastreio para sintomas depressivos<sup>22</sup> (média de 3,07, DP = 2,01), histórico negativo de sintomas neuropsiquiátricos ou comprometimentos neurológicos e pontuação na Escala de Avaliação Clínica das Demências (CDR)<sup>23</sup> igual a 0. As avaliações foram realizadas pela equipe de neuropsicologia do Laboratório de Investigações Neuropsicológicas (LIN-UFMG). Os participantes preencheram o termo de consentimento livre e esclarecido referente ao projeto.

A tabela 1 apresenta o perfil sociodemográfico da amostra, assim como o resultado dos testes neuropsicológicos. De forma a permitir melhor compreensão das características da amostra estudada, a tabela foi dividida por idade (60-69, 70-79, 80-89 anos) e por escolaridade (menos que oito e maior que oito anos).

### Instrumentos de exame neuropsicológico

**RAVLT:** Utilizou-se a versão adaptada para a população idosa brasileira<sup>8</sup>. As 15 palavras do teste são lidas pausadamente ao probando, solicitando-se que ele as repita após a leitura, de forma independente à ordem em que foram ditas (A1). Repete-se o mesmo procedimento nas etapas A2, A3, A4 e A5, salientando-se que o sujeito deve sempre lembrar-se de todas as palavras, inclusive as ditas anteriormente. Aplica-se, então, uma segunda lista de palavras (B1), distratora, em que se solicita ao probando evocar apenas as palavras desta. Pedem-se, então, as palavras da primeira lista (A6), sem a exposição dela, em uma tarefa de evocação da memória episódica de curto prazo.

Vinte e cinco minutos após tal etapa, o sujeito deve novamente evocar as palavras da primeira lista (A7), de forma a avaliar a evocação tardia da memória episódica verbal. Por fim, é apresentada uma lista de reconhecimento (REC) contendo 50 palavras, as 15 palavras da primeira lista, as 15 da segunda e 20 outras palavras semântica ou foneticamente relacionadas com as palavras anteriores, devendo o sujeito julgar se a palavra era pertencente ou não à primeira lista. Os escores de cada etapa são computados de acordo com o número de palavras corretamente evocadas, à exceção de REC, em que se pontuam todos os acertos (palavras corretamente classificadas) e subtrai-se 35 (total de distratores). De forma a evitar resultados negativos nas análises posteriores, nesta foi realizada a subtração dos resultados negativos, gerando um escore que oscila entre 0 e 50 pontos. Calcula-se também o total ( $\Sigma A1-A5$ ) de acertos nas etapas de aprendizagem.

**Miniexame do Estado Mental:** instrumento de rastreio comumente utilizado para a identificação de comprometimento cognitivo e diagnóstico de demências<sup>19</sup>. Contém perguntas que avaliam a orientação temporal, orientação espacial, cálculos, memória de curto prazo, praxias motoras, nomeação, linguagem e praxia visuoconstrutiva. A pontuação total varia de 0 a 30, sendo os pontos de corte 16, 22 e 26 utilizados para a identificação de comprometimento cognitivo significativo em indivíduos analfabetos, escolarizados e com curso superior, respectivamente.

**Teste do Desenho do Relógio:** utilizado na avaliação das habilidades visuoconstrutivas e funções executivas, é um instrumento de aplicação rápida e simples, sensível a comprometimento cognitivo. No presente estudo, utilizou-se a pontuação variando de 0 (pior) a 5 (melhor)<sup>24</sup>.

## Análise estatística

A análise de fidedignidade consiste em avaliar a estabilidade ou precisão da medida de um teste. Um dos métodos para a análise de tal propriedade é a avaliação da consistência interna, realizada neste estudo por meio do cálculo do alfa de Cronbach (Pasquali, 2006) com os componentes A1, A2, A3, A4, A5, B1, A6, A7 e REC do RAVLT. A validade de construto do RAVLT foi realizada mediante suas correlações com os outros dois testes neuropsicológicos utilizados (MEEM e DR) e de uma análise fatorial exploratória, realizada pelo método de fatoração por eixos principais e rotação oblíqua (*direct oblimin*). Tal configuração foi selecionada tendo em vista o foco teórico do presente estudo e da provável relação de dependência entre as variáveis do RAVLT<sup>13</sup>. Os testes de adequação amostral de Kaiser-Meyer-Olkin

(KMO) e de Esfericidade de Bartlett foram calculados para atestar a viabilidade da análise na amostra selecionada. A proporção de observações e variáveis (dez para um) recomendada por Hair *et al.*<sup>13</sup> foi satisfeita. Foram, então, avaliadas as relações entre as diferentes etapas do RAVLT e a idade, escolaridade, sintomas depressivos e gênero. Para os três primeiros fatores, optou-se por modelos de regressão linear *stepwise*, tendo os componentes do RAVLT como variáveis dependentes e as demais variáveis como independentes. Para o gênero, o desempenho de homens e mulheres foi comparado por meio do teste *t* de Student com cálculo da magnitude de efeito por meio do *d* de Cohen.

## Resultados

Os nove componentes do RAVLT apresentaram alto índice de intercorrelações, gerando um alfa de Cronbach de 0,831 e indicando boa consistência interna. As correlações entre os diferentes componentes do teste com o MEEM e o DR oscilaram entre não significativas ( $r = 0,08$ ,  $p = 4,05$ ) a fracas ( $r = 0,448$ ,  $r^2 = 0,20$ ,  $p < 0,001$ ). A correlação encontrada entre o MEEM e o DR ( $r = 0,713$ ,  $r^2 = 0,50$ ,  $p < 0,001$ ) pode ser considerada alta. A tabela 2 mostra as correlações entre os testes.

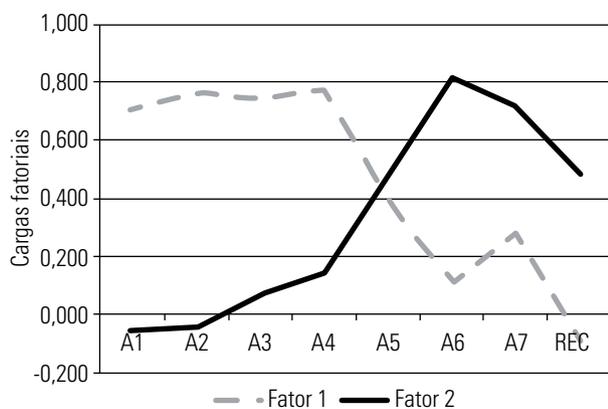
O teste KMO (0,856) e o teste de esfericidade de Bartlett ( $\chi^2 = 458,8$ ,  $p < 0,001$ ) indicam que a amostra é adequada ao procedimento de análise fatorial exploratória. Com base na análise dos autovalores e da interpretação do *screeplot*, uma solução de dois fatores foi considerada a mais adequada (autovalores: 4,411 e 1,139) para a amostra, responsável por 59% da variância total. A análise da matriz de padrões indica que o primeiro fator teve cargas predominantemente relacionadas com os itens A1 (0,704), A2 (0,763), A3 (0,742) e A4 (0,776), enquanto o segundo fator relacionou-se com os itens A6 (0,816), A7 (0,714) e REC (0,485). O componente A5 exibiu cargas significativas em ambos os fatores (0,393 no primeiro e 0,478 no segundo). Os fatores apresentaram, ainda, uma forte correlação ( $r = 0,752$ ,  $r^2 = 0,56$ ). Tais resultados estão expostos na figura 1.

A análise de regressão linear *stepwise*, contando com os componentes do RAVLT como variáveis dependentes e idade, escolaridade (em anos) e sintomas depressivos apresentados (GDS-15) como variáveis independentes, gerou modelos em que apenas a idade foi preditora de desempenho no RAVLT (significante em todas as etapas), enquanto a escolaridade e os sintomas depressivos não mostraram significância. A tabela 2 sintetiza tais resultados. Mulheres obtiveram desempenho superior ao dos homens em A1 ( $p < 0,001$ ,  $d = 0,66$ ), A3 ( $p = 0,031$ ,  $d = 0,41$ ), A5 ( $p = 0,015$ ,  $d = 0,48$ ) A6 ( $p = 0,032$ ,  $d = 0,49$ ) e Total ( $p = 0,008$ ,  $d = 0,51$ ).

**Tabela 1.** Estatística descritiva dos participantes e desempenho nos testes neuropsicológicos

	60-69 anos (n = 51)		70-79 anos (n = 43)		80-89 anos (n = 32)	
	< 8 anos n = 26 F = 12	> 8 anos n = 25 F = 15	< 8 anos n = 8 F = 5	> 8 anos n = 35 F = 18	< 8 anos n = 24 F = 14	> 8 anos n = 8 F = 3
	Média (DP)	Média (DP)	Média (DP)	Média (DP)	Média (DP)	Média (DP)
MEEM	29,1 (1,1)	29,1 (1,0)	27,7 (0,5)	27,4 (0,9)	26,0 (1,3)	26,5 (1,4)
DR	4,7 (0,8)	4,8 (0,6)	4,3 (0,8)	4,0 (0,5)	3,4 (0,6)	3,3 (0,5)
GDS-15	2,0 (1,8)	1,5 (1,7)	3,3 (1,6)	3,8 (1,1)	5,2 (1,2)	4,4 (1,8)
A1	5,1 (1,1)	5,3 (0,9)	5,2 (2,6)	5,0 (0,7)	3,9 (1,7)	3,5 (0,5)
A2	7,4 (1,8)	7,8 (1,0)	6,7 (2,1)	6,7 (0,9)	5,8 (1,7)	5,5 (0,8)
A3	8,7 (2,0)	8,8 (1,6)	7,8 (2,1)	7,5 (1,4)	7,1 (1,7)	6,9 (2,2)
A4	9,9 (2,0)	10,2 (1,4)	8,7 (2,6)	9,3 (1,6)	8,0 (2,1)	8,5 (1,7)
A5	11,1 (2,0)	11,1 (1,8)	8,8 (1,3)	10,7 (1,2)	9,9 (2,5)	9,4 (2,1)
B1	4,7 (1,1)	4,5 (1,3)	4,0 (0,6)	4,0 (0,7)	2,9 (1,5)	3,5 (1,2)
A6	9,8 (2,6)	9,3 (2,3)	9,0 (1,3)	8,7 (1,1)	8,4 (1,9)	8,0 (1,8)
A7	9,1 (2,8)	9,0 (2,0)	7,7 (2,4)	8,4 (1,2)	7,0 (2,0)	7,4 (2,0)
REC	31,7 (5,8)	28,9 (9,0)	29,7 (2,8)	28,7 (4,7)	26,9 (8,2)	25,8 (8,5)
RAVLT - Total	42,2 (7,2)	43,2 (4,9)	37,2 (9,2)	39,3 (4,1)	34,6 (8,2)	33,8 (5,6)

MEEM: Miniexame do Estado Mental; DR: Desenho do Relógio; GDS-15: Escala de Depressão Geriátrica; RAVLT: Teste de Aprendizagem Auditivo-Verbal de Rey; DP: desvio-padrão; F: participantes do sexo feminino.



**Figura 1.** Cargas fatoriais dos componentes do RAVLT.

**Tabela 2.** Análise de regressão linear e correlações do RAVLT com o MEEM e o DR

RAVLT	$\beta$	Modelo	Sig.	R <sup>2</sup> Ajustado	Cor. MEEM	Cor. DR
A1	-0,417	Idade	$p < 0,001$	0,17	0,244 ( $p = 0,010$ )	0,370 ( $p < 0,001$ )
A2	-0,491	Idade	$p < 0,001$	0,23	0,370 ( $p < 0,001$ )	0,377 ( $p < 0,001$ )
A3	-0,396	Idade	$p < 0,001$	0,15	0,372 ( $p < 0,001$ )	0,377 ( $p < 0,001$ )
A4	-0,441	Idade	$p < 0,001$	0,29	0,349 ( $p < 0,001$ )	0,390 ( $p < 0,001$ )
A5	-0,269	Idade	$p = 0,005$	0,07	0,193 ( $p < 0,001$ )	0,319 ( $p < 0,001$ )
B1	-0,448	Idade	$p < 0,001$	0,19	0,260 ( $p < 0,006$ )	0,335 ( $p < 0,001$ )
A6	-0,241	Idade	$p < 0,001$	0,05	0,153 ( $p < 0,110$ )	0,345 ( $p < 0,001$ )
A7	-0,338	Idade	$p < 0,001$	0,10	0,289 ( $p = 0,002$ )	0,344 ( $p < 0,001$ )
REC	-0,300	Idade	$p = 0,001$	0,08	0,080 ( $p = 0,405$ )	0,180 ( $p = 0,059$ )
Total	-0,493	Idade	$p < 0,001$	0,24	0,381 ( $p < 0,001$ )	0,448 ( $p < 0,001$ )

RAVLT: Teste de Aprendizagem Auditivo-Verbal de Rey; Sig.: significância; Cor.: correlação de Pearson; MEEM: Minixame do Estado Mental; DR: Desenho do Relógio.

## Discussão

O presente estudo investigou algumas propriedades psicométricas do RAVLT, um dos testes de aprendizagem mais utilizados no contexto brasileiro. O teste da primeira hipótese proposta, acerca de sua consistência interna, indica que os diferentes itens do RAVLT apresentam um padrão consistente de intercorrelações, uma evidência de fidedignidade. Embora haja uma evidência robusta de tal propriedade, não foram encontrados outros estudos brasileiros avaliando a fidedignidade do RAVLT por outras metodologias, como a teste-reteste.

As hipóteses acerca da validade de construto do RAVLT apresentam evidências convergentes de tal propriedade para a população estudada. O teste apresentou correlações baixas, mas em geral significativas com dois instrumentos clássicos para avaliação cognitiva de pacientes idosos, o MEEM e o DR. Tal achado sugere que, embora partilhem variância, o RAVLT mensura construtos distintos daqueles avaliados pelos dois outros testes, um indicativo de validade de construto. Em concordância, os dois instrumentos de rastreamento apresentaram uma correlação forte entre si, evidenciando a distinção entre seus construtos e os do RAVLT.

A estrutura fatorial do teste na população estudada está de acordo com aquela encontrada na literatura<sup>15,16</sup>. Em idosos saudáveis, com o perfil sociodemográfico dos participantes deste estudo, uma estrutura de dois fatores foi suficiente para explicar mais da metade da variância apresentada no teste. Uma análise pormenorizada das cargas fatoriais de cada item do RAVLT indica que há uma mudança quanto ao processo cognitivo envolvido na tarefa: enquanto os itens A1, A2, A3 e A4 sugerem um processo relacionado ao armazenamento das informações apresentadas, os itens A6, A7 e, em menor grau, REC sugerem a busca de um conteúdo já armazenado no sistema mnésico. O componente REC apresentou carga fatorial inferior aos dois componentes do Fator 2. Conjecturamos que em uma população maior e mais heterogênea tal tarefa forme, por si só, um fator isolado, como apontado por outros autores<sup>16</sup>, considerando que tal componente difere da estrutura dos demais componentes do RAVLT.

Com base nessas informações, denominamos os dois fatores encontrados como, respectivamente, armazenamento e evocação. Tal solução é corroborada pelas cargas fatoriais de A5, último item da fase de aprendizagem do RAVLT, relacionado aos dois fatores encontrados, como uma transição entre processos. A estrutura fatorial encontrada é respaldada tanto pelos modelos neuropsicológicos clássicos de memória episódica<sup>25</sup> quanto pelos mais recentes<sup>26,27</sup>, correlações anatomoclínicas<sup>28</sup> e estudos com populações clínicas<sup>16,17,26</sup>. A alta correlação apresentada entre os dois fatores ainda encontra respaldo na relativa interdependência do sistema de memória episódica. Huijbers *et al.*<sup>26</sup> demonstraram em um estudo de neuroimagem funcional que os processos relativamente independentes de armazenamento e evocação partilham sobreposição no que tange à ativação da formação hipocampal, estrutura que atuaria como um elemento de síntese e transição entre as informações relacionadas ao processo de memorização, concepção partilhada por outros autores.

A influência de variáveis sociodemográficas no RAVLT corrobora alguns achados prévios, como a influência significativa da idade e gênero<sup>10</sup> no desempenho dos sujeitos (as normas populacionais brasileiras são divididas pela faixa etária e sexo). A idade, único preditor significativo nos modelos de regressão linear, explicou entre 5% e 29% da variância encontrada nos diferentes componentes do RAVLT, enquanto diferenças de gênero se manifestaram nos componentes A1, A3, A5, A6 e no total de palavras armazenadas, com feminino superior ao masculino. As magnitudes se aproximaram a meio desvio-padrão. A relevância clínica de tais diferenças é significativa, visto que em alguns quadros cínicos (como o comprometimento cognitivo leve), a interpretação quantitativa do desempenho neuropsicológico tende a ser fundamental para o diagnóstico<sup>3</sup>.

Ao contrário de estudos anteriores<sup>9,11</sup>, não foi encontrada influência significativa da escolaridade nos escores do teste. Tal fato pode ter como explicação a homogeneidade dos participantes avaliados em relação à escolaridade, além da distribuição desproporcional dos sujeitos em quatro e oito anos de escolaridade, o que reflete uma limitação do presente estudo. O estudo de tal variável em populações com escolaridade predominantemente baixa (0-3 anos) e alta (> 14 anos) torna-se fundamental para o bom uso do teste, de forma a investigar o desempenho em diferentes níveis educacionais, tendo em vista a importância disso para fins diagnósticos. Embora o escore da GDS não tenha mostrado influência nas diversas etapas do RAVLT, a baixa pontuação na escala não indica sintomas depressivos de grande intensidade, sendo inclusive um dos critérios de inclusão para o presente estudo. Em um grupo clínico com sintomas depressivos mais evidentes, tal fator poderia influenciar de forma negativa o desempenho no teste, como constatado por alguns estudos que utilizaram o RAVLT em pacientes com depressão maior<sup>18</sup>.

A análise das propriedades psicométricas do RAVLT conduzida neste estudo fornece evidências significativas de fidedignidade (por meio da análise da consistência interna), validade de construto (mediante o uso de correlações divergentes e análise da estrutura fatorial), além de investigar a influência de variáveis sociodemográficas, idade, escolaridade e gênero e da variável clínica sintomas depressivos. Contudo, o presente estudo apresenta as limitações metodológicas apontadas pelo trabalho de Delis *et al.*<sup>29</sup>. O RAVLT é

um instrumento marcado por processos de aprendizagem, em que cada etapa do teste depende fundamentalmente da etapa anterior, o que tende a aumentar a variância partilhada por cada componente. Tais autores argumentam que a avaliação da validade de construto por meio de correlações com outros testes neuropsicológicos e a análise da estrutura fatorial é excessivamente enviesada pelas características da amostra utilizada. Em um delineamento experimental, eles demonstraram que o perfil clínico da memória episódica de pacientes com demência de Alzheimer e doença de Huntington (marcados pela dissociação de desempenho entre memória de evocação imediata e memória de evocação tardia) não encontrava respaldo na análise de correlações ou estrutura fatorial nessa população. Contudo, o comentário de Larrabee<sup>30</sup> sobre tais evidências sugere que a análise fatorial pode ser uma ferramenta útil para a validade de construto, desde que conduzida em amostras de tamanho e homogeneidade adequados e, preferencialmente, em associação com outras medidas cognitivas análogas (por exemplo, usar as evocações tardias do RAVLT em conjunto com as da Figura Complexa de Rey e evocação das praxias da bateria CERAD).

Comparar a estrutura fatorial específica de cada grupo clínico também pode trazer evidências importantes sobre os padrões de variância apresentados pela amostra, permitindo inferências quanto à especificidade dos construtos avaliados e aplicabilidade clínica. Larrabee<sup>30</sup> conclui que o uso dessas metodologias, por si só, não exaure as investigações sobre validade de construto de um instrumento, mas, em conjunto com estudos clínicos de populações bem caracterizadas e estudos de caso específicos, torna-se fundamental para a análise.

O presente estudo apresenta evidências de validade e fidedignidade no que tange às propriedades psicométricas do RAVLT. A favor disso, encontram-se a convergência multimetodológica de tais evidências, a homogeneidade da amostra (baseada em uma série de critérios objetivos de inclusão) e o tamanho amostral adequado para as análises. Portanto, associado à primeira análise da validade de construto do teste<sup>11</sup>, fornece evidências convergentes de fidedignidade e validade de construto. Contudo, os resultados aqui encontrados não podem ser generalizados para toda a população idosa, sendo necessários estudos mais aprofundados sobre a influência de fatores como a escolaridade e sintomas depressivos em amostras mais heterogêneas, de forma a aumentar a validade do RAVLT para o exame clínico da população idosa. Nessa perspectiva, destacamos que o volume crescente de estudos para obtenção de parâmetros normativos e de aplicabilidade de testes neuropsicológicos para diferentes segmentos da população brasileira<sup>7,8,11,14,19</sup> deve ser acompanhado de estudos sobre as propriedades psicométricas desses mesmos instrumentos. Tal iniciativa propiciará o refinamento do uso de tais instrumentos tanto para a prática clínica quanto para a de pesquisa.

## Conclusão

O RAVLT apresentou boa consistência interna, correlações fracas com instrumentos globais de exame neuropsicológico e estrutura fatorial composta por dois componentes, armazenamento e evocação. A análise das variáveis sociodemográficas sugere influência significativa da idade em todos os componentes do teste e do gênero em A1, A3, A5, A6 e Total. Os resultados encontrados evidenciam a fidedignidade e a validade de construto do teste, respaldando seu potencial para uso clínico.

## Referências

- Lima-Costa MF, Veras R. Saúde pública e envelhecimento. *Cad Saude Publica*. 2003;19(3):700-1.
- Herrera E Jr, Caramelli P, Silveira AS, Nitrini R. Epidemiologic survey of dementia in a community-dwelling Brazilian population. *Alzheimer Dis Assoc Disord*. 2002;16:103-8.
- Busse A, Biskopf J, Riedel-Heller SG, Angermeyer MC. Mild cognitive impairment: prevalence and incidence according to different diagnostic criteria: results of the Leipzig Longitudinal Study of the Aged (Leila75+). *Br J Psychiatry*. 2003;182:449-54.
- Dourado M, Laks J, Leibing A, Engelhardt E. Consciência da doença na demência. *Rev Psiq Clin*. 2000;33(6):313-21.
- Paulo DLV, Yassuda MS. Memória de idosos e sua relação com a escolaridade, desempenho cognitivo e sintomas de depressão e ansiedade. *Rev Psiq Clin*. 2010;37(1):40-3.
- Rey A. *Lexamen clinique en psychologie*. Paris: Press Universitaire de France; 1958.
- Malloy-Diniz LF, Da Cruz MF, Torres V, Cosenza R. O Teste de Aprendizagem Auditivo-Verbal de Rey: normas para uma população brasileira. *Rev Bras Neurol*. 2000;36(3):79-83.
- Malloy-Diniz LF, Lasmar VA, Gazinelli LS, Fuentes D, Salgado JV. The Rey Auditory-Verbal Learning Test: applicability for the Brazilian elderly population. *Rev Bras Psiquiatr*. 2007;29(4):324-9.
- Salgado J, Abrantes SSC, Freitas D, Souza W, Oliveira J, Moreira L, et al. Applicability of the Rey Auditory-Verbal Learning Test to an adult sample in Brazil. *Rev Bras Psiquiatr*. 2011 [online]. ahead of print, p. 0-0. Epub Mar 18. Available from: <http://www.scielo.br/scielo.php>.
- Stallings G, Boake C, Sherer M. Comparison of the California Verbal Learning Test and the Rey Auditory-Verbal Learning Test in head-injured patients. *J Clin Exp Neuropsychol*. 1995;17:706-12.
- Fichman HC, Dias LBT, Fernandes CS, Lourenço R, Caramelli P, Nitrini R. Normative data and construct validity of the Rey Auditory Verbal Learning Test in a Brazilian Elderly population. *Psychol Neurosci*. 2010;3(1):79-84.
- Carroll JB. *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge: Cambridge University Press; 1993.
- Hair J, Anderson RE, Tatham RL, Black WC. *Análise multivariada de dados*. Porto Alegre: Bookman; 2005.
- De Paula JJ, Schlottfeldt CG, Santos LM, Cotta M, Bicalho MA, Moraes EN, et al. Psychometric properties of a brief neuropsychological protocol for use in geriatric populations. *Rev Psiq Clin*. 2010;37(6):246-50.
- Mueller H, Hasse-Sander R, Horn C, Helmstaedter C, Elger CE. Rey-Auditory Verbal Learning Test: structure of a modified German version. *J Clin Psychol*. 1997;53:663-71.
- Strauss E, Sherman EMS, Spreen O. *A Compendium of Neuropsychological tests*. 3. ed. New York: Oxford University Press; 2006.
- Balthazar MLF, Yasuda CL, Cendes F, Damasceno BP. Learning, retrieval, and recognition are compromised in aMCI and mild AD: are distinct episodic memory processes mediated by the same anatomical structures? *J Int Neuropsychol Soc*. 2010;16:205-9.
- Leposavić I, Leposavić L, Gavrilović P. Depression vs. dementia: a comparative analysis of neuropsychological functions. *Psihologija*. 2010;43(2).
- Brucki SM, Nitrini R, Caramelli P, Bertolucci PH, Ivan H, Okamoto IH. Sugestões para o Uso do Minixame do Estado Mental no Brasil. *Arq Neuropsiquiatr*. 2003;61(3-B):777-81.
- Katz S, Downs TD, Cash HR, Grotz RC. Progress in the development of the index of ADL. *Gerontologist*. 1970;10:20-30.
- Lawton MP, Brody EM. Assessment of older people: self-monitoring and instrumental activities of daily living. *Gerontologist*. 1969;9:179-86.
- Paradela EMP, Lourenco RA, Veras RP. Validação da Escala de Depressão Geriátrica em um ambulatório geral. *Rev Saude Publica*. 2005;39(6):918-23.
- Montaño MBMM, Ramos LR. Validade da versão em português da Clinical Dementia Rating. *Rev Saude Publica*. 2005;39(6):912-7.
- Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry*. 2000;15(6):548-61.
- Cohen NJ, Eichenbaum H. *Memory, amnesia and the hippocampal system*. Cambridge, Massachusetts: MIT Press; 1993.
- Huijbers W, Cyriel MAP, Cabeza R, Sander MD. The hippocampus is coupled with the default networking memory retrieval but not during memory encoding. *PLoS One*. 2010;6(4):1-10.
- Henke K. A model for memory systems based on processing modes rather than consciousness. *Nat Rev Neurosci*. 2010;11:523-32.
- Dickerson BC, Eichenbaum H. The episodic memory system: neuro-circuitry and disorders. *Neuropsychopharmacology*. 2010;35:86-104.
- Delis DC, Jacobson M, Bondi, MW, Hamilton JM, Salmon DP. The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: lessons from memory assessment. *J Int Neuropsychol Soc*. 2003;9:936-46.
- Larrabee GJ. Lessons on measuring construct validity: a commentary on Delis, Jacobson, Bondi, Hamilton, and Salmon. *J Int Neuropsychol Soc*. 2003;9:947-53.