

Juliana Alves Batista

**CONFIABILIDADE DOS TESTES DA BATERIA DE FUNÇÃO
COGNITIVA DO ESTUDO LONGITUDINAL DE SAÚDE DO ADULTO
– ELSA-BRASIL**

**Universidade Federal de Minas Gerais
Programa de Pós-graduação em Saúde do Adulto
Belo Horizonte – MG
2012**

Juliana Alves Batista

**CONFIABILIDADE DOS TESTES DA BATERIA DE FUNÇÃO
COGNITIVA DO ESTUDO LONGITUDINAL DE SAÚDE DO ADULTO
– ELSA -BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Saúde do Adulto da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Ciências Aplicadas à Saúde do Adulto (área de concentração em Clínica Médica).

Orientadora: Valéria Maria de Azeredo Passos

Co-orientadora: Luana Giatti Gonçalves

**Belo Horizonte – MG
2012**

Batista, Juliana Alves.
B333c Confiabilidade dos testes de bateria de função cognitiva do estudo longitudinal de saúde do adulto [manuscrito]. / Juliana Alves Batista. -- Belo Horizonte: 2012.
52f.
Orientador: Valéria Maria de Azevedo Passos.
Co-Orientador: Luana Giatti Gonçalves.
Área de concentração: Clínica Médica.
Dissertação (mestrado): Universidade Federal de Minas Gerais, Faculdade de Medicina.

1. Reprodutibilidade dos Testes. 2. Ciência Cognitiva. 3. Estudos Longitudinais. 4. Saúde do Adulto. 5. Dissertações Acadêmicas. I. Passos, Valéria Maria de Azevedo. II. Gonçalves, Luana Giatti. III. Universidade Federal de Minas Gerais, Faculdade de Medicina. IV. Título

NLM : BF 311

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Reitor

Clélio Campolina Diniz

Vice-Reitor

Rocksane de Carvalho Norton

Pró-Reitor de Pós-Graduação

Ricardo Santiago Gomez

Pró-Reitor de Pesquisa

Renato de Lima Santos

FACULDADE DE MEDICINA

Diretor

Professor Francisco José Penna

Chefe do departamento de Clínica Médica

Prof. Anelise Impelizeri Nogueira

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS À SAÚDE DO ADULTO

Coordenador

Professora Teresa Cristina de Abreu Ferrari

Sub-coordenador

Professora Valéria Maria Azeredo Passos

Colegiado

Professor Marcus Vinícius Melo de Andrade

Professora Tereza Cristina de Abreu Ferrari

Professor Luiz Gonzaga Vaz Coelho

Professor Francisco Eduardo Costa Cardoso

Professora Suely Meireles Rezende

Professora Valéria Maria de Azeredo Passos

Andréa de Lima Bastos (discente titular)

DECLARAÇÃO DE DEFESA



FACULDADE DE MEDICINA
CENTRO DE PÓS-GRADUAÇÃO
Av. Prof. Alfredo Balena 190 / sala 533
Belo Horizonte - MG - CEP 30.130-100
Fone: (031) 3409.9641 FAX: (31) 3409.9640



DECLARAÇÃO

A Comissão Examinadora, abaixo assinada, composta pelos Professores Doutores Valéria Maria Azeredo Passos, Luana Giatti Gonçalves, Enrico Antônio Colosimo e Rogério Gomes Beato, aprovou a defesa da dissertação intitulada: "CONFIABILIDADE DOS TESTES DA BATERIA DE FUNÇÃO COGNITIVA DO ESTUDO LONGITUDINAL DE SAÚDE DO ADULTO – ELSA –BRASIL", apresentada pela mestrandia JULIANA ALVES BATISTA, para obtenção do título de Mestre em Saúde do Adulto, pelo Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto – área de concentração em Ciências Clínicas, da Faculdade de Medicina da Universidade Federal de Minas Gerais, realizada em 20 de abril de 2012.


Profa. Valéria Maria Azeredo Passos
Orientadora


Profa. Luana Giatti Gonçalves
Coorientadora


Prof. Enrico Antônio Colosimo


Prof. Rogério Gomes Beato

CONFERE COM ORIGINAL
Centro de Pós-Graduação
Faculdade de Medicina - UFMG

Aos participantes do ELSA

AGRADECIMENTOS

À Suprema Personalidade de Deus, luz em cada momento de trabalho;

ao querido e muito amado Xande, que emprestou-me seus olhos da fantasia, sempre doce e acolhedor;

à minha mamãe, Déia, Syl, Flávia, Sally, Biel, Lulu e Pedrinho;

ao meu pai, referência de vida, que na distância, na lembrança, sempre esteve presente;

à professora Valéria, por ter me recebido como aluna e pela presteza e cuidado com que conduziu a orientação;

às professoras Luana e Sandhi pelos momentos de direcionamento acadêmico;

à toda equipe do Projeto ELSA, momentos incríveis de convívio e aprendizado;

aos participantes deste estudo que se disponibilizaram com tanto carinho;

ao Rodrigo, pela calma, paciência e disponibilidade;

à Mari, amiga do coração, pelas conversas, apoio, incentivo, compreensão e por todos os almoços que tornaram o dia-a-dia muito mais agradável;

à Ana, acadêmica deste estudo, pelo trabalho cuidadoso e companhia agradável;

aos queridos amigos Felipe, Flávia, Dani, Marina, Natália e Susu;

Muito obrigada!

“O valor de praticar com rigor, por algum tempo, uma *ciência rigorosa* não está propriamente em seus resultados: pois eles sempre serão uma gota ínfima, ante o mar das coisas dignas de saber. Mas isso produz um aumento de energia, de capacidade dedutiva, de tenacidade; prende-se a *alcançar um fim de modo pertinente*. Neste sentido é valioso, em vista de tudo o que se fará depois, ter sido homem de ciência.”

Friedrich Nietzsche
(*Humano, demasiado humano*, 1878)

SUMÁRIO

1 CONSIDERAÇÕES INICIAIS.....	9
2 OBJETIVOS	18
3 ARTIGO: CONFIABILIDADE DA BATERIA DE TESTES DE FUNÇÃO COGNITIVA DO ESTUDO LONGITUDINAL DE SAÚDE DO ADULTO – ELSA.....	19
RESUMO	19
3.1 INTRODUÇÃO	19
3.2 MÉTODOS	21
3.3 RESULTADOS	23
3.4 DISCUSSÃO	26
3.5 CONCLUSÃO	28
REFERÊNCIAS BIBLIOGRÁFICAS	28
4. CONSIDERAÇÕES FINAIS.....	31
5. ANEXOS	
Anexo I – Termo de Consentimento Livre e Esclarecido.....	32
Anexo II – Documento de aprovação do Comitê de Ética.....	36
Anexo III – Organograma de convite ao participante ELSA para estudo da confiabilidade de testes cognitivos.....	39
Anexo IV - Dicionário das variáveis para o Estudo da Confiabilidade da Bateria de Testes de Função Cognitiva do ELSA.....	40
Anexo V –E-mail de envio de Trabalho ao XVIII Congresso Brasileiro de Geriatria e Gerontologia.....	44
Anexo VI – Artigo publicado na Revista Dement Neuropsychol.....	46

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

LISTA DE ABREVIATURAS

DP: desvio padrão

et al: e outros

T0: Tempo Zero

T1: Tempo Um

LISTA DE SIGLAS

CEFET/MG: Centro Federal de Educação Tecnológica de Minas Gerais

CERAD: *Consortium to Establish a Registry for Alzheimer's Disease*

CERAD-K: Versão coreana do *Consortium to Establish a Registry for Alzheimer's Disease*

COEP: Comitê de Ética em Pesquisa

CONEP: Comissão Nacional de Ética em Pesquisa

DA: Doença de Alzheimer

DECIT: Departamento de Ciência e Tecnologia

ELSA: Estudo Longitudinal de Saúde do Adulto

FINEP: Financiadora de Estudos e Projetos

FIOCRUZ: Fundação Oswaldo Cruz

HC/UFMG: Hospital das Clínicas da UFMG

MCT: Ministério da Ciência e Tecnologia

MS: Ministério da Saúde

TCLE: Termo de Consentimento Livre e Esclarecido

UFBA: Universidade Federal da Bahia

UFMG: Universidade Federal de Minas Gerais

UFRGS: Universidade Federal do Rio Grande do Sul

USP: Universidade de São Paulo



1. CONSIDERAÇÕES INICIAIS

A cognição se refere ao ato ou processo de possuir ou adquirir um conhecimento por meio das redes neuronais interconectadas que formam complexos circuitos que originam as atividades mentais ligadas à percepção, ação planejada e pensamento.¹ Em psicologia, o termo cognição é utilizado para se referir aos processos de pensamento.²

As funções cognitivas se organizam como sistemas funcionais complexos ou rede de conexões, as quais dependem da ação conjunta de diversas regiões cerebrais conectadas entre si. O estudo dessas funções trata da relação entre a cognição, o comportamento e a atividade do sistema nervoso central em condições normais e patológicas e, além da avaliação clínica, pode envolver testes para averiguação das capacidades funcionais intelectuais. Estas aferições tentam captar e descrever fenômenos que são complexos e transformá-los em formatos padronizados para que possam ser analisados em estudos clínicos e epidemiológicos.^{3,4}

Os testes neuropsicológicos, aplicados isoladamente ou em bateria de testes, são parte da avaliação neuropsicológica, que inclui entrevistas, exame físico, avaliação de características emocionais e sociais do indivíduo.⁵ As baterias constituem-se em sequências de testes que avaliam comportamento e cognição. Podem ser padronizadas (compostas pelos mesmos testes) ou flexíveis (compostas por testes agrupados de acordo com a necessidade). São úteis no diagnóstico de patologias cerebrais, na elaboração de planos de reabilitação de pacientes, nas políticas de prevenção em saúde mental e contribuem no manejo clínico dos pacientes psiquiátricos. Além disso, facilitam o estabelecimento de perfil cognitivo basal antes, durante e depois de tratamentos, bem como colaboram para o diagnóstico diferencial em condições que envolvam prejuízo cognitivo.^{6,7}

¹Kandel ER, Schwartz JH.; Jessell TM. Princípios da Neurociência. 4ª ed., Barueri (SP): Manole, 2003.

²Kolb, B.; Whishaw, I.Q. Neurociência do comportamento. Barueri (SP): Manole, 2002.

³Machado FP, Avaliação neuropsicológica do Idoso. In: Tavares A. Compêndio de neuropsiquiatria geriátrica. 2ª ed. Rio de Janeiro: Guanabara Koogan; 2005. P. 157-173.

⁴Kristensen CH, Almeida RMM, Gomes WB. Desenvolvimento Histórico e Fundamentos Metodológicos da Neuropsicologia Cognitiva. *Psicol. Reflex. Crit.*, 2001;14 (2):259-274.

⁵Howieson DB, Lezak, MD. A avaliação neuropsicológica. In: Stuart CY, Robert EH. Compêndio de Neuropsiquiatria. Porto Alegre: Artes Médicas; 1996. P. 91-105

⁶Abreu ID, Florenza OV, Barros HL. Demência de Alzheimer: correlação entre memória e autonomia. *Revista de Psiquiatria. Clínica* 2005 Jun; 32(3): 131-136.

⁷Malloy-Diniz LF, Carvalho, AM. O exame neuropsicológico e suas contribuições à psiquiatria. *Psiquiatr. biol.* 2001 Jun; 9(2): 66-77.

A avaliação da cognição, além do uso de testes neuropsicológicos, pode ser feita por queixa subjetiva de déficit de seus componentes, por exemplo, a memória. O uso dos testes tende a minimizar a subjetividade da queixa, uma vez que há evidências na literatura que mostram que a percepção subjetiva não corresponde necessariamente ao comprometimento objetivo da função. A queixa subjetiva pode estar relacionada ao estado afetivo dos indivíduos e não necessariamente ao declínio cognitivo.⁸

Na aplicação de testes cognitivos existem fatores que podem interferir no seu resultado. Existem aqueles relacionados ao examinador, como a entonação ao dar as instruções, a habilidade em criar um ambiente amigável, porém profissional, a experiência com o teste, o respeito à técnica, as respostas neutras frente às perguntas do entrevistado, a repetição da pergunta e não sua interpretação; aqueles relacionados ao entrevistado, como o humor, disposição para realização do teste, e as variáveis a que o teste se destina a medir; e aqueles relacionados ao próprio teste, como a validade, incluindo as medidas de sensibilidade e especificidade e a precisão ou confiabilidade, que mostra o quanto a medida pode ser replicada mantendo as mesmas condições de aplicação, apesar de serem raras as possibilidades de replicações de estudos epidemiológicos sob as mesmas condições.⁹ Em pesquisas epidemiológicas raramente é obtido o nível de reprodutibilidade encontrado em investigações de laboratório, onde há maior possibilidade de manter condições idênticas de avaliação.¹⁰

Em estudos de confiabilidade de um instrumento para coleta de dados, como testes cognitivos, considera-se o teste preciso quando existe constância dos resultados obtidos nas replicações. Nestas replicações, chamadas de retestes, busca-se ao máximo manter as mesmas condições de aplicação, em que devem ser consideradas as variáveis que interferem no desempenho do teste, como exemplos: o ambiente, a privacidade, a luminosidade, o examinador. O intervalo de tempo entre o teste e o reteste é outro fator importante a considerar. Períodos longos são suscetíveis às mudanças, como por exemplo, alterações da capacidade cognitiva. Períodos curtos podem favorecer a lembrança das respostas que o indivíduo deu no primeiro teste e depois, simplesmente, repetiu as respostas recordadas no reteste.¹¹

⁸Charchat-Fichman H, Caramelli P, Sameshima K, Nitrini R. Declínio da capacidade cognitiva durante o envelhecimento. Rev. Bras. Psiquiatr 2005 Mar; 27(1): 79-82.

⁹ Medronho, RA, et al. Validade em estudos epidemiológicos. In: _____. Epidemiologia. São Paulo: Atheneu 2006. P. 199-212.

¹⁰ Pereira MG, Aferição dos eventos. In: _____. Epidemiologia; teoria e prática. Rio de Janeiro: Guanabara Koogan 1995. P. 358-376.

¹¹ Martins, GA. Sobre confiabilidade e validade. RGBN 2006 jan/abr; 8 (20):1-12.

A confiabilidade é fundamental para a validade interna de um estudo, em que as informações geradas permitem inferências sobre a população-alvo de onde a amostra foi retirada. Esta deve ser avaliada quando um instrumento específico é empregado e monitorada sempre que possível. A sistematização da coleta de dados, com passos definidos e ordenados de conduta, tende a melhorar a qualidade da informação e, conseqüentemente, a concordância.⁹

Em estudos epidemiológicos as conclusões devem estar baseadas em resultados válidos, livre de erros metodológicos na concepção, desenho ou implementação do estudo ou no processo de análise dos dados. A validade externa de um estudo se refere, portanto, às generalizações que podem advir dele para além da amostra estudada, considerando os métodos utilizados, a representatividade da amostra estudada e a natureza da população de onde a amostra foi retirada.⁹

Neste momento, se desenvolve na Universidade Federal de Minas Gerais (UFMG) e em outras cinco instituições de ensino superior e de pesquisa do Brasil – UFBA, UFRGS, USP e FIOCRUZ – o Estudo Longitudinal de Saúde do Adulto (ELSA), pesquisa multicêntrica e multidisciplinar, com financiamento da FINEP/MCT e do DECIT/MS, que tem por meta identificar determinantes da incidência de doenças crônicas, principalmente as cardiovasculares e o diabetes. A linha de base do estudo foi concluída em 18/12/2010, composta por 15.102 servidores públicos.

O estudo da cognição em adultos e idosos é uma das linhas de investigação no ELSA-Brasil e tem por objetivos principais criar e manter coorte para estimar prevalência, incidência e história natural do comprometimento cognitivo. A vantagem do estudo longitudinal é permitir estabelecer comparações dos resultados dos testes da mesma pessoa em momentos diferentes. Isto é importante, pois existem controvérsias entre estudos transversais que indicam declínio da capacidade cognitiva durante o envelhecimento e longitudinais que revelam estabilidade ou melhora do desempenho cognitivo ao longo do tempo.¹²

¹²Salthouse TA. Selective review of cognitive aging. Journal of the International Neuropsychological Society 2010 Jun 4; 16(5):1-7.

Testes de função cognitiva no ELSA

Para estudar esta função no ELSA-Brasil estão sendo utilizados os seguintes testes: Memória da lista de palavras, testes de Fluência verbal semântica e fonêmica, Evocação da lista de palavras, Reconhecimento da lista de palavras e Teste de trilhas parte B. Com exceção do teste de Fluência verbal fonêmica com a letra F, todos fazem parte da bateria de testes neuropsicológicos do *Consortium to Establish a Registry for Alzheimer's Disease* (CERAD). Estes testes são amplamente utilizados em estudos com idosos. No ELSA, a coorte estudada é composta desde adultos jovens, pessoas acima de 35 anos de idade, até idosos com idade máxima de 74 anos.

O CERAD foi criado em 1986 por uma concessão do Instituto Nacional do Envelhecimento dos Estados Unidos para desenvolver medidas padronizadas, válidas, confiáveis e de fácil aplicação voltadas ao diagnóstico e acompanhamento de pacientes com doença de Alzheimer. A bateria do CERAD é composta de uma bateria de avaliação clínica e outra neuropsicológica.¹³ A bateria neuropsicológica tem sido usada, de forma integral ou em parte, em vários estudos em diferentes países, incluindo os EUA, Colômbia, Índia, China, Sudeste da Ásia, América Latina e Caribenhos e na África.^{13,14} Esta bateria inclui testes de memória (fixação, evocação e reconhecimento), linguagem (nomeação e fluência verbal categórica), praxia (cópia de desenhos geométricos), função executiva (Teste das trilhas), além de um teste geral (Mini-exame do Estado Mental).¹⁵

A confiabilidade teste-reteste, com intervalo de um mês entre as aplicações da bateria neuropsicológica do CERAD foi inicialmente medida por Morris e colaboradores (1989) em estudo realizado com indivíduos com idade igual ou superior a 50 anos, com e sem demência, e encontraram, de forma geral, medidas confiáveis nos escores dos testes. A maioria dos estudos tem por interesse a validação, com determinação de níveis de sensibilidade, especificidade e determinação de pontos de corte para discriminação para demência, mas poucos estudos que mediram a confiabilidade teste-reteste.

¹³ Morris JC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989 Sep;39:1159-1165.

¹⁴ Fillenbaum GG, et al. CERAD (Consortium to Establish a Registry for Alzheimer's Disease) The first 20 years. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 2008 Mar; 4(2):96-109.

¹⁵ Chandler MJ, Lacritz LH, Hynan LS, Barnard HD, Allen G, Deschner M, et al. A total score for the CERAD neuropsychological battery. *Neurology* 2005; 65(1):102-106.

A adaptação da bateria neuropsicológica do CERAD para uso no Brasil foi realizada no final da década de 90, em estudo conduzido por Bertolucci e colaboradores. Utilizou-se uma versão adaptada para o português em dois grupos, sendo um composto por 80 idosos saudáveis e outro de 21 idosos com provável Doença de Alzheimer. Os resultados indicaram que a adaptação foi adequada e fornecem um padrão para o desempenho esperado em relação às demências. Os autores consideraram os testes viáveis para a larga aplicação em nosso país.¹⁶

Os testes para avaliação da memória tentam captar esta complexa função do sistema nervoso, encarregada de reter informações ou percepções, para que possam ser posteriormente utilizadas. A recepção, o armazenamento e a liberação da informação são propriedades generalizadas da rede neurônica e relacionam-se a adaptação do indivíduo ao meio ambiente.

A memória de curto prazo armazena as tarefas a serem executadas dentro de um intervalo temporal. Também chamada memória de trabalho, é importante para realizar atividades cognitivas como compreender, raciocinar e resolver problemas e seria responsável por armazenar informação por um breve período, realizar operações e desempenhar, simultaneamente, outras tarefas.¹⁷ O teste de Memória da lista de palavras do CERAD avalia este tipo de memória e a curva de aprendizagem quando da repetição da lista¹⁶¹⁷.

O teste de Memória da lista de palavras do CERAD inclui lista de dez palavras não relacionadas, que o examinador mostra ao sujeito e imediatamente depois pede que o sujeito recorde em voz alta o maior número de palavras que ele puder. Este procedimento é realizado três vezes, para a avaliação da memória imediata. A pontuação é obtida pela soma das palavras lembradas nas três tentativas, com um escore máximo de trinta pontos.¹⁶

Outra parte do teste de memória, chamada Evocação, avalia a memória tardia e mede a capacidade de retenção das palavras ditas anteriormente após intervalo de tempo e realização de atividade distraidora. Com um ponto de corte igual a três para discriminação entre

¹⁶Consenza RM, Avaliação neuropsicológica do Idoso. In: Tavares A. Compêndio de neuropsiquiatria geriátrica. 1ª ed. Rio de Janeiro: Guanabara Koogan; 2005.P. 157-173.

pacientes com e sem demência, foi percebida sensibilidade entre 74,2% e 75% e especificidade entre 75% e 82,4%.^{16,18}

Também faz parte do teste de memória o reconhecimento das palavras apresentadas, quando misturadas a outras. Este teste mede a capacidade de resgate de informações com estímulo, por meio de uma lista que contém elementos da lista original mais uma série de distraidores. Estudo de validação em idosos, para um ponto de corte igual a oito, obteve sensibilidade de 75% e especificidade de 66%¹⁶ e para um ponto de corte igual a sete, sensibilidade de 76,2% e especificidade de 87,1%.¹⁹ A acurácia mostrou-se também elevada em estudo do Alzheimer em indivíduos com escolaridade média ou alta; para um ponto de corte igual a oito obteve-se valores de sensibilidade de 90,4% e especificidade de 74,5%.¹⁹

Para avaliação da linguagem foram utilizados os testes de Fluência verbal semântica e o teste de Fluência verbal fonêmica. Os testes de fluência verbal estão comumente inseridos em quase todas as baterias neuropsicológicas para estudo da cognição nos indivíduos com DA^{20,21,22}, doença de Parkinson²³, nas alterações cognitivas associadas ao envelhecimento normal ou a quadros demências^{24,25}, ou são utilizados isoladamente em estudos específicos de memória semântica, fonêmica, função executiva ou da linguagem.^{26,27} A fluência verbal é mensurada pela quantidade de palavras de uma mesma categoria lembradas em determinado período de tempo, geralmente, um minuto.²⁸

O teste de Fluência verbal mais empregado em estudos brasileiros é o de categoria semântica, geralmente utilizando a categoria animais. Este teste é recomendado pelo Departamento

¹⁸ Bertolucci PHF, Okamoto IH, Brucki SMD, Siviero MO, Neto JT, Ramos LR. Applicability of the CERAD neuropsychological battery to Brazilian elderly. *Arquivos de Neuropsiquiatria* 2001; 59(3-A):532-536

¹⁹ Nitrini R, et al. Brief cognitive battery in the diagnosis of mild Alzheimer's disease in subjects with medium and high levels of education. *Dementia and Neuropsychologia* 2007; 1:32-36

²⁰ Nitrini R, Caramelli P, Bottino CMC, Damasceno BP, Brucki SMD, Anghinah R. Diagnóstico de doença de Alzheimer no Brasil: avaliação cognitiva e funcional. *Arquivos de Neuropsiquiatria* 2005; 63(3-A):720-727.

²¹ Chen P, et al. Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. *Neurology* July 10, 2001 57:163-164

²² Clark LJ, et al. Longitudinal verbal fluency in normal aging, preclinical and prevalent Alzheimer disease. *Am J AlzheimersDisOtherDemen*. 2009 December; 24(6): 461-468.

²³ DOENÇA DE ALZHEIMER E DE PARKINSON: ESTUDO COMPARATIVO DA FLUÊNCIA VERBAL. Anais do XVI Encontro de Iniciação Científica e Encontro de Iniciação em Desenvolvimento Tecnológico e Inovação da PUC-Campinas 27 e 28 de setembro de 2011 ISSN 1982-0178

²⁴ Radanovic M, et al. Analysis of brief language tests in the detection of cognitive decline and dementia. *Dementia & Neuropsychologia* 2007;1:37-45

²⁵ Rodrigues AB, Yamashita ET, Chiappetta ALML. Teste de fluência verbal no adulto e no idoso: verificação da aprendizagem verbal. *Rev CEFAC* 2008 out-dez; 10(4):443-451.

²⁶ Brucki SMD, Rocha MSG. Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Brazilian Journal of Medical and Biological Research* 2004; 37: 1771-1777.

²⁷ Brucki SMD, Malheiros SMF, Okamoto IH, Bertolucci PHF. Dados normativos para o teste de fluência verbal categoria animais em nosso meio. *Arq Neuropsiquiat* 1997;55(1):56-61.

²⁸ Lezak, MD. *Neuropsychological Assessment*. 3rd ed. New York: Oxford University Press, 1995.

Científico de Neurologia Cognitiva e do Envelhecimento da Academia Brasileira de Neurologia por ter como vantagens a sua brevidade e disponibilidade de notas de corte definidas em função da escolaridade.²¹

Os testes de fluência podem ser utilizados para avaliar a capacidade de armazenamento da memória verbal (léxico e semântico), a velocidade psicomotora, a linguagem, em que pode ser avaliada a forma de expressão e como a pessoa utilizou o tempo disponível para a execução do teste, a habilidade de recuperar a informação guardada na memória e o processamento das funções executivas, com foco na concentração, na capacidade de organizar o pensamento e nas estratégias utilizadas para a busca de palavras.^{3,27,28}

O teste de Fluência verbal semântica, em estudo da adaptação da bateria do CERAD para uso no Brasil, apresentou equilíbrio entre sensibilidade e especificidade, 75% para ambos, além de ser considerado de fácil aplicação.¹⁶ Para este teste é importante utilizar níveis diferenciados de corte, considerando o efeito da escolaridade sobre o desempenho, uma vez que com o aumento da escolaridade aumenta-se o número de animais lembrados. Em relação a idade os estudos são controversos se existem diferenças de desempenho em função desta variável. Existem os que apontam não haver diferenças^{27,28} e os que pontuam declínio no número de palavras lembradas com o aumento da idade.^{29,30}

Estudos indicam as notas de corte para a população brasileira de idosos, ao redor de doze e treze para a categoria animais, em amostras de indivíduos com média de oito anos e mais e quatro a dezesseis anos de escolaridade, respectivamente.^{16,31} Ponto de corte para perda cognitiva igual a treze foi encontrado em estudo com a população brasileira com idade da amostra a partir de quinze anos em grupo de 336 indivíduos com média de escolaridade próxima a seis anos de estudo²⁸ e ponto de corte igual a 11 para amostra com média de escolaridade igual a 7,9 anos.²⁰

O teste de Fluência verbal fonêmica requer a geração de palavras iniciadas com dada letra em determinado tempo. Geralmente são utilizadas as letras F, A, S em intervalo de tempo de um

²⁹ Adam M. Brickman, et al. Category and letter verbal fluency across the adult lifespan: relationship to EEG theta power. *Archives of Clinical Neuropsychology* 20 (2005) 561–573

³⁰ Bolla KI, Gray S, Resnick SM, Galante R, Kawas C. Category and letter fluency in highly educated older adults. *The Clinical Neuropsychologist* 1998;12(3):330–338.

³¹ Nitirini R, Lefèvre BH, Mathias SC, Caramelli P, Carrilho PEM, Sauaia N, et al. Testes Neuropsicológicos de aplicação simples para o diagnóstico de demência. *Arquivos de Neuropsiquiatria* 1994; 52(4):457-465.

minuto. São especialmente úteis para detecção de disfunção executiva, especialmente à esquerda.³⁰ Pode ser uma abordagem valiosa para a detecção de alterações cognitivas, dada a sua estabilidade no processo de envelhecimento.³² Este teste, comparado ao de fluência verbal semântica, permite de maneira mais clara identificar as estratégias de pensamento para orientar a lembrança de palavras e é mais difícil para indivíduos que não desenvolvem suas próprias estratégias de recordação.²⁹

Estudos comparativos entre testes de Fluência semântica e fonêmica demonstram declínio mais rápido nos testes de Fluência semântica quando comparado aos de Fluência fonêmica tanto em indivíduos cognitivamente normais quanto naqueles com DA.^{23,33} As diferenças entre os diferentes testes de Fluência se devem, provavelmente, aos diferentes mecanismos fisiopatológicos envolvidos.

Para o teste de Fluência fonêmica obteve-se ponto de corte de 15, com sensibilidade 83,6% e especificidade 67,0%, observado em estudo de demência na DA em amostra de indivíduos com escolaridade média e alta.²⁰

O Teste de Trilhas, instrumento de avaliação das funções executivas, foi originalmente desenvolvido em 1944 para avaliação da rapidez visuo-motora.²⁹ O termo “função executiva” indica funções psíquicas fundamentais para a produção de respostas comportamentais adequadas aos diferentes contextos. Inclui formulação de objetivos e conceitos, motivação, planejamento, auto-regulação, insight, abstração, análise, manipulação de conhecimentos adquiridos e flexibilidade mental. O Teste de trilhas examina a velocidade de processamento, flexibilidade cognitiva, capacidade de manutenção do engajamento mental, o rastreamento visual, a destreza motora e a memória operacional.¹⁶

Este teste é aplicado nas formas A e B. Em ambas as formas o teste propriamente dito é precedido de um treinamento mais curto. Na forma A é apresentada uma folha com vinte e cinco círculos numerados, distribuídos ao acaso, que devem ser unidos em uma linha contínua. Na forma B, utilizada no ELSA, os círculos apresentam treze números e doze letras,

³²Steiner VAG *Dementia & Neuropsychologia* 2008 December;2(4):328-332 328 Phonemic verbal fluency and age Steiner VAG, et al. Phonemic verbal fluency and age: A preliminary study Veronique Agnes Guernet Steiner1, Leticia Lessa Mansur1,2,3, Sonia Maria D. Brucki3, Ricardo Nitri3

³³ Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis Julie D. Henry*, John R. Crawford, Louise H. Phillips

que devem ser unidos alternadamente (1-A, 2-B, etc.) e no menor tempo possível. Este exige maior capacidade de atenção e habilidade para fazer mudanças conceituais alternadas.³⁴Os erros são apontados e devem ser corrigidos pelo próprio testando. O teste pode ser encerrado após três erros ou cinco minutos.¹⁶

Como existem poucos estudos de confiabilidade, a maior parte realizada apenas em população idosa, neste trabalho será estudada a confiabilidade de testes cognitivos que avaliam as funções cognitivas, em voluntários adultos (35-74 anos) da coorte ELSA.

³⁴ Mota MMPE, Banhato EFC, Silva KCA, Cupertino APFB. Triagem cognitiva: comparações entre o mini-mental e o teste de trilhas. *Estudos de Psicologia* 2008; 25(3): 353-359.

2. OBJETIVOS

Objetivo geral:

- Avaliar a confiabilidade dos testes da bateria de função cognitiva do Estudo Longitudinal de Saúde do Adulto.

Objetivos específicos:

- Avaliar a confiabilidade dos seguintes testes: Memória da lista de palavras, Reconhecimento da lista de palavras, Evocação da lista de palavras, Fluência verbal semântica, Fluência verbal fonêmica, Teste de trilhas forma B.
- Avaliar a influência do sexo, da idade e da escolaridade na confiabilidade dos testes.

3. ARTIGO

Confiabilidade dos testes da bateria de função cognitiva do Estudo Longitudinal de Saúde do Adulto– ELSA-Brasil

Juliana Alves Batista¹, Luana GiattiGonçalves^{1,2}, Sandhi Maria Barreto^{1,2}, Valéria Maria de Azeredo Passos^{1,2}

¹ Programa de Pós-graduação em Ciências Aplicadas à Saúde do Adulto. Faculdade de Medicina da Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil.

²Faculdade de Medicina da Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil.

RESUMO

Introdução: A avaliação da função cognitiva é uma das linhas de investigação do Estudo Longitudinal de Saúde do Adulto – ELSA-Brasil. O objetivo deste estudo foi avaliar a confiabilidade dos testes cognitivos do ELSA-Brasil, em população adulta e idosa, fundamental para garantia de qualidade das informações. **Método:** Estudo de confiabilidade tipo teste-reteste em 160 participantes do ELSA-Brasil. **Resultados:** Foram encontrados os seguintes valores de correlação: teste de memória de palavras 0,56, reconhecimento 0,35, evocação 0,50, fluência verbal semântica 0,53, fluência verbal fonêmica 0,61 e teste de trilhas 0,91. **Conclusão:** Foi observada confiabilidade moderada para testes de cognição realizados em adultos e no intervalo de tempo de vinte dias. A melhora do desempenho nos testes em relação aos retestes sugere o efeito do aprendizado. O Teste de Trilhas pode ser útil para reaplicação em curto espaço de tempo.

Palavras-Chave: testes cognitivos, confiabilidade teste-reteste

3.1 INTRODUÇÃO

A cognição humana relaciona-se ao ato de adquirir conhecimento por meio de complexa interconexão das redes neuronais que formam os processos mentais ligados ao pensamento, percepção e ação planejada.¹ O estudo das funções cognitivas inclui a avaliação clínica e a neuropsicológica. Na avaliação neuropsicológica podem ser utilizados testes, aplicados isoladamente ou em baterias, para averiguação das capacidades funcionais e intelectuais. Estes testes tentam captar e descrever fenômenos complexos de forma padronizada para que possam ser analisados em estudos clínicos e epidemiológicos.

Em estudos epidemiológicos as conclusões devem estar baseadas em resultados válidos, livre de erros metodológicos na concepção, desenho ou implementação do estudo ou no processo de análise dos dados.

Desenhos de estudos longitudinais e transversais podem apresentar resultados contraditórios em relação ao desempenho nos testes cognitivos. Os estudos transversais, quase unanimemente, mostram declínio do desempenho cognitivo com aumento da idade, enquanto as comparações longitudinais algumas vezes revelam estabilidade ou melhora no desempenho cognitivo para a mesma faixa etária.²

A replicação dos testes em estudos longitudinais deve ser capaz de captar diferenças reais na cognição dos indivíduos, e não diferenças que possam ser atribuídas ao instrumento utilizado ou à forma de aplicação.

Uma bateria de testes cognitivos foi utilizada no Estudo Longitudinal de Saúde do Adulto (ELSA-Brasil) – coorte de 15.105 servidores de seis instituições de ensino e pesquisa no Brasil.³ Esta bateria utiliza parte dos testes neuropsicológicos do *Consortium to Establish a Registry for Alzheimer's Disease* (CERAD) que avaliam a memória, linguagem e função executiva⁵.

A bateria de testes cognitivos do CERAD, amplamente utilizada em estudos clínicos e epidemiológicos em diversas regiões do mundo, é descrita como tendo vantagens de detectar demência em fase inicial, possuir alta concordância entre examinadores, possibilidade de comparar resultados entre diferentes grupos e apresentar boa reprodutibilidade teste-reteste. Esta bateria foi adaptada para uso no Brasil, porém sua confiabilidade não foi testada.^{5,6,7,8}

Vários fatores podem interferir nos resultados dos testes cognitivos. Existem aqueles relacionados ao examinador, ao entrevistado, e aqueles relacionados ao próprio teste. Este deve ser avaliado em relação a dois requisitos essenciais: validade e confiabilidade.⁹ Muitos estudos descrevem medidas relacionadas à validade, tais como sensibilidade, especificidade, definição de pontos de corte e avaliação de desempenho. Estudo brasileiro com 80 idosos livres de doença neurológica ou psiquiátrica e 21 indivíduos com doença de Alzheimer apresentou pontos de corte igual a doze para o teste de fluência verbal, quatorze para o de memória de palavras, três para evocação e oito para reconhecimento. De modo geral, obteve-se sensibilidades e especificidades iguais ou maiores que 75%.⁷

A avaliação de desempenho de acordo com diversas variáveis mostrou que escolaridade mais alta relacionou-se ao melhor desempenho nos testes de fluência verbal^{10,11,12} e memória da lista de palavras.¹³ A idade mais baixa relacionou-se ao melhor desempenho em todos os testes do CERAD, à exceção do teste de reconhecimento¹³. Para o

teste de fluência semântica os resultados dos estudos são contraditórios em relação ao desempenho por grupo etário, alguns não encontraram diferenças entre as faixas de idade^{10,11} e outros que relatam diminuição com a idade no número de palavras lembradas^{12,13}. A fluência fonêmica mostrou-se estável no processo de envelhecimento¹¹. Em relação ao gênero os dados são contraditórios sobre o desempenho nos testes de fluência verbal.^{11,13,14}

Poucos estudos mediram a precisão ou confiabilidade teste-reteste, que mostra o quanto a medida pode ser replicada mantendo-se as mesmas condições de aplicação.^{8,14,15,16} Esses estudos mostram que os testes do CERAD são confiáveis em intervalo de tempo de um mês entre os testes. Nestes estudos não se investigou a influência da idade, sexo e escolaridade na confiabilidade.

Este estudo tem por objetivo avaliar a confiabilidade dos testes cognitivos aplicados no ELSA-Brasil, e investigar a variabilidade da mesma em função da idade, sexo e escolaridade.

3.2 MÉTODOS

O ELSA-Brasil é um estudo prospectivo multicêntrico desenvolvido em seis instituições, de ensino superior e pesquisa: Universidade Federal de Minas Gerais - UFMG, Universidade de São Paulo - USP, Fundação Oswaldo Cruz - FIOCRUZ-RJ, Universidade Federal do Espírito Santo - UFES, Universidade Federal da Bahia - UFBA e Universidade Federal do Rio Grande do Sul – UFRGS. Os principais objetivos desse estudo são investigar a incidência e a progressão do diabetes e das doenças cardiovasculares e examinar os fatores biológicos, comportamentais, ambientais, ocupacionais, psicológicos e sociais relacionados a essas doenças e às suas complicações, buscando compor modelo causal que contemple suas inter-relações. Foram considerados elegíveis para a pesquisa servidores, com idade entre 35 a 74 anos, incluindo funcionários ativos e aposentados, de cada uma das instituições participantes.³

Este estudo de confiabilidade tipo teste-reteste utilizou uma amostra de conveniência, composta por 160 participantes, de ambos os sexos, do centro de pesquisa ELSA – Minas Gerais. Para inclusão no estudo, os participantes deveriam ser servidores ativos ou aposentados da UFMG ou CEFET e consentir em participar por meio da assinatura do TCLE (Anexo I). Todos os 172 participantes realizaram a bateria de testes cognitivos duas vezes. Os testes aplicados foram: teste de palavras (composto pelos testes de memória de palavras, evocação e reconhecimento), de fluência verbal semântica categoria animais, de fluência verbal fonêmica com a letra F e teste de trilhas. Estes testes foram aplicados pelo mesmo

entrevistador, previamente treinado e certificado. Para aplicação buscou-se um ambiente tranquilo, com boa iluminação e baixos níveis de ruído ou outros estímulos distratores. Para garantir adesão ao estudo, foi dada opção de realização do reteste no local de trabalho da pessoa ou biblioteca próxima, desde que se mantivessem as condições adequadas para aplicação.

Foi denominado como “T0 – Tempo 0” a primeira aplicação da bateria de testes, realizada durante a coleta de dados do ELSA-Brasil. O reteste, realizado no intervalo entre 15 a 25 dias após o T0, foi denominado “T1 – Tempo 1”. Os mesmos testes foram aplicados nas duas ocasiões, porém em ordens distintas.

Os testes foram gravados e posteriormente revistos para conferência das informações. Os escores dos testes de fluência verbal foram definidos por duas supervisoras do ELSA-Brasil, previamente treinadas e certificadas, seguindo normas padronizadas.^{17,18}

Utilizou-se o Programa Epiinfo[®] 3.5.3¹⁹ para dupla entrada de dados e, para análise, o Programa STATA[®], versão 10.0²⁰.

Para análise, inicialmente foi feita a descrição das características da população de estudo. Procedeu-se a análise descritiva dos testes por meio da média e desvio padrão, e a amplitude de variação na primeira e segunda aplicação. Para verificar as diferenças das médias dos escores do teste e do reteste foi realizado teste de T pareado. A avaliação da homogeneidade da variância entre teste e reteste foi feita pelo teste de Bartlett. No teste de trilhas B, foi realizado teste de Mann-Whitney para verificar a diferença das medianas, uma vez que não atendeu ao pressuposto da homogeneidade da variância.

O Coeficiente de Correlação Intraclasse (ICC) foi utilizado como medida de estimativa de confiabilidade entre os testes e retestes. Este teste estima a fração da variabilidade total causada pela variação entre os indivíduos.²¹ Os valores do Coeficiente de Correlação Intraclasse foram classificados como pobres se iguais à zero, leves de 0,01 a 0,2; discretos de 0,21 a 0,4; moderados de 0,41 a 0,6; substanciais de 0,61 a 0,8 e quase perfeitos se valores entre 0,81 e 0,9.²²

Para avaliar o padrão dos dados foram obtidos os gráficos de dispersão. Para comparar nossos resultados com outros estudos foram estimados Coeficientes de Correlação de Pearson, que medem o grau com que um conjunto pareado de observações em um diagrama se aproxima de uma situação em que cada ponto se localiza exatamente na linha reta.²¹

O ELSA-Brasil foi aprovado pelos comitês de ética de cada uma das instituições envolvidas (Anexo II) e pelo Conselho Nacional de Ética em Pesquisa.

3.3 RESULTADOS

Dos 172 participantes, 12 foram excluídos por não terem respondido a todos os testes, assim foram analisados escores referentes a 160 indivíduos. As características demográficas dos participantes do estudo são apresentadas na Tabela1. Observa-se que a amostra é composta predominantemente por adultos com menos de 60 anos (idade média de 52,2 anos) com elevada escolaridade. O intervalo de tempo entre a aplicação do teste e o reteste variou de 14 a 27 dias, com média e mediana de aproximadamente 20 dias (DP=3,0). Destes, 88 (55%) responderam aos testes no ELSA e 71 (44,4%) no local de trabalho.

Tabela1. Características sócio-demográficas dos participantes do estudo de confiabilidade dos testes de função cognitiva do ELSA-Brasil.

Variável	Número	Percentual
Sexo		
Masculino	81	50,6
Feminino	79	49,4
Faixa etária		
35-59	121	75,6
60-74	39	24,4
Escolaridade		
Segundo grau incompleto	17	10,6
Segundo grau completo	46	28,8
Superior	97	60,6

Quando comparados os resultados de T0 e T1, observam-se maiores valores nos escores dos retestes de memória de palavras, evocação, de fluência verbal semântica e fonêmica com letra F, assim como redução do escore médio do tempo para realização do teste de trilhas B(Tabela 2).

Tabela 2. Amplitude, média e comparação dos testes e retestes da bateria de testes de função cognitiva do ELSA-Brasil. (n=160)

Medidas	Testes					
	Memória de Palavras	Evocação	Reconhecimento	Fluência animais	Fluência letra F	Teste de Trilhas
Amplitude teste	11-28	1-10	8-10	6-35	3-27	29-858
Amplitude reteste	12-30	2-10	7-10	10-34	2-26	31-526
Média (DP) teste	20,78(3,66)	6,89(1,85)	9,71(0,55)	19,06(5,00)	13,05(4,61)	124,51(102,92)
Média(DP) reteste	23,58(3,64)	7,86(1,76)	9,76(0,51)	20,21(5,11)	14,09(4,54)	97,56(58,85)
	p<0,0001	p<0,0001	p=0.42	p=0,0002	p<0,0001	p=0,009 ¹

¹A variância entre os escores do teste e do reteste não foi homogênea (teste de Bartlett p=0,91), foi utilizado teste Mann-Whitney para comparar as medianas. Mediana (p25-,p75) do teste: 90 (71-138). Mediana (p25-p75) do reteste: 81,5(64,5-111,5).

A medida da confiabilidade entre T0 e T1 é apresentada na Tabela 3. O ICC variou de 0,35 para o teste de reconhecimento a 0,91 para o teste de trilhas B. Todos os testes apresentaram correlação de Pearson positiva com valores estatisticamente significativos.

Tabela 3. Confiabilidade entre os testes e retestes da bateria ELSA (n=160).

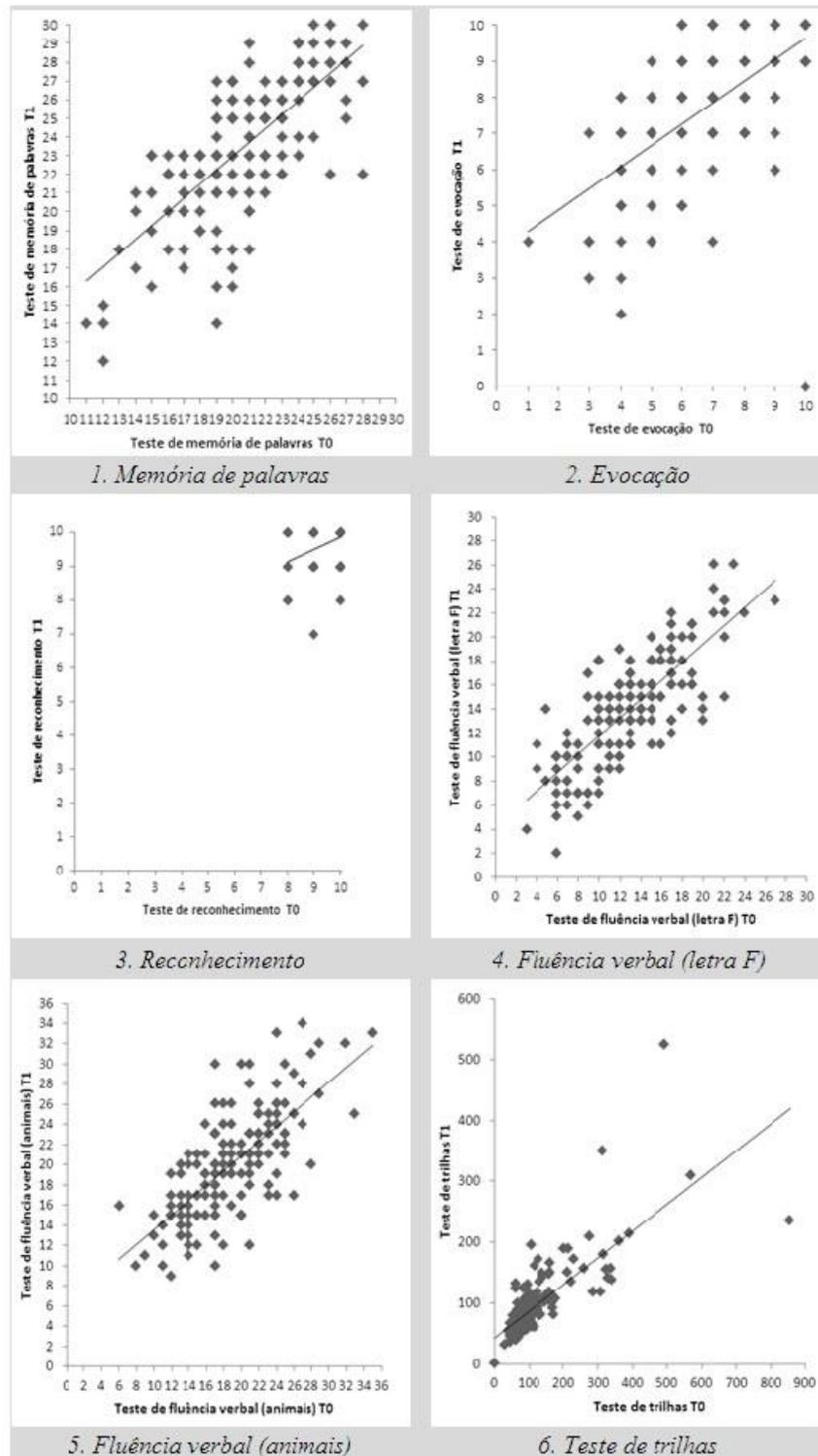
Teste	ICC (IC95%)	Pearson(valor de p)
Memória de palavras	0,56 (0,33-0,79) ¹	0,74 (p<0,0001)
Evocação	0,50 (0,17-0,83) ¹	0,68 (p<0,0001)
Reconhecimento	0,35 (0,00-0,94) ¹	0,40 (p<0,0001)
Fluência verbal (animais)	0,53 (0,33-0,74) ¹	0,72 (p<0,0001)
Fluência verbal (letra F)	0,61(0,41-0,81) ¹	0,77 (p<0,0071)
Teste de trilhas B	0,91 (0,87-0,95) ¹	0,76 (p<0,0001)

¹valor de p < 0,0001

A Figura 1 apresenta os gráficos de dispersão correspondentes aos coeficientes de correlação de Pearson dos valores encontrados nos testes de cognição. É possível notar pelo

início e inclinação da reta, diferente de 45^0 , que para os testes de memória e fluência os valores de T1 foram maiores que T0 e o inverso para o teste de trilhas. Observa-se no gráfico do teste de evocação a maior dispersão dos valores. É notável no teste de reconhecimento os escores próximos a dez, limite máximo do teste.

Figura 1. Gráficos de dispersão dos Testes de memória, fluência e trilhas.



A estabilidade entre os valores encontrados em T0 e em T1 foi aferida de acordo com as variáveis sexo, idade, escolaridade (Tabela 4). Verifica-se a variabilidade dos escores de todos os testes segundo sexo, faixa etária e escolaridade. Alguns testes como memória de palavras, de fluência verbal semântica e fonêmica não apresentaram ICC estatisticamente significativa.

Tabela 4. Coeficiente de Correlação Intraclasse por sexo, idade, escolaridade e local dos testes da bateria de função cognitiva do ELSA-Brasil.

Variável	Memória de palavras	Evocação	Reconhecimento	Fluência animais	Fluência letra F	Teste de trilhas
Sexo						
Feminino	0,63 (0,39-0,86) ¹	0,52 (0,16-0,88) ¹	0,61 (0,00-1,22) ¹	0,54 (0,30-0,78) ¹	0,53 (0,28-0,78) ¹	0,88 (0,80-0,96) ¹
Masculino	0,45 (0,16-0,74) ¹	0,50 (0,15-0,85) ¹	0,30 (0,00-0,89) ²	0,53(0,29-0,77) ¹	0,64 (0,42-0,86) ¹	0,97 (0,96-0,99) ¹
Faixa etária						
35-59	0,56 (0,31-0,80) ¹	0,49 (0,15-0,84) ¹	0,26 (0,00-0,78) ²	0,57 (0,36-0,78) ¹	0,61 (0,41-0,82) ¹	0,90 (0,85-0,95) ¹
60-74	0,56 (0,24-0,88) ²	0,47 (0,07-0,86) ²	0,50 (0,00-1,19) ²	0,23 (0,00-0,62) ³	0,78 (0,58-0,97) ¹	0,76 (0,28-1,24) ³
Escolaridade						
2º grau incompl.	0,27 (0,00-0,88) ³	0,07 (0,00-0,67) ³	0,64 (0,00-1,36) ²	0,07 (0,00-0,89) ³	0,93 (0,79-1,06) ³	0,72 (0,24-1,20) ³
2º grau completo	0,71 (0,47-0,94) ¹	0,56 (0,18-0,94) ¹	0,30(0,00-0,92) ²	0,41 (0,07-0,75) ²	0,49 (0,17-0,81) ²	0,89 (0,77-1,01) ²
Superior	0,49 (0,22-0,76) ¹	0,48 (0,11-0,84) ¹	0,32 (0,00-0,89) ²	0,45 (0,22-0,69) ¹	0,57 (0,33-0,80) ¹	0,86 (0,78-0,94) ¹

¹valor de p < 0,0001; ²valor de p < 0,05; ³valor de p >0,05

3.4 DISCUSSÃO

Poucas publicações avaliam a confiabilidade teste-reteste dos testes utilizados neste estudo^{8,14,15,16}, sendo este o primeiro estudo no Brasil. Na adaptação para o Brasil dos testes da bateria do CERAD considerou a confiabilidade teste-reteste de um estudo estadunidense realizado em amostra de 610 indivíduos, com ou sem doença de Alzheimer e idade média de 68 anos.⁸

Conhecer as medidas de confiabilidade pode ser de utilidade clínica ou em estudos epidemiológicos, principalmente em população de países e línguas distintas da qual os testes foram desenvolvidos.

Neste estudo, optou-se por intervalo de tempo também utilizado em estudo da variabilidade intrapessoal do desempenho cognitivo os quais responderam aos testes cognitivos em três ocasiões com intervalos de tempo dentro do período de duas semanas²³ e em estudos da confiabilidade da bateria do CERAD, onde o reteste ocorreu após um mês.^{8,15}

Longos períodos de tempo entre as medições poderiam favorecer o aparecimento de mudanças reais nas funções cognitivas, comprometendo o estudo da confiabilidade dos testes; enquanto períodos mais curtos, podem ser contaminados pelo efeito de aprendizado.⁹

Em cada aplicação, os testes foram dispostos em diferentes ordens, com objetivo de diminuir o aprendizado.^{23,24} A ordem dos testes foi disposta de tal forma que houvesse sempre um teste distraindo entre o teste de memória de palavras e o teste de evocação.⁴ Apesar destes cuidados, em todos os testes pode-se observar a melhora do desempenho à medida que os eles foram reaplicados. Para este efeito, também observado em outros estudos^{8,15,25} com intervalo de um dia a um mês entre os testes, infere-se que seja atribuído ao aprendizado.

Os testes de memória de palavras, evocação e fluência verbal semântica revelaram confiabilidade moderada, confiabilidade discreta para o teste de reconhecimento, substancial para o teste de fluência verbal fonêmica e quase perfeita para o teste de trilhas B, quando realizados em período de cerca de 20 dias.

Confiabilidade moderada para o teste de memória de palavras também foi encontrada em estudo realizado na Coreia e outro nos Estados Unidos da América, com amostra de 20 e 278 pessoas, respectivamente, com idade maior que 50 anos e intervalo de um mês entre os testes. Estes mesmos estudos obtiveram confiabilidade substancial para o teste de evocação (0,64 e 0,65).^{8,15}

O teste de Reconhecimento apresentou os menores valores de confiabilidade e um intervalo de confiança tão abrangente que não foi possível interpretar o valor do ICC. Baixo valor de confiabilidade, similar ao encontrado para este teste também foi observado no estudo estadunidense⁸ (0,36), porém, no estudo coreano observou-se valor igual a 0,74.¹⁵ A pouca confiabilidade no teste de reconhecimento pode ser explicada pelo efeito “teto”, em que a amostra perfaz valores nos testes próximos ao máximo, o que diminui a faixa de variabilidade dos escores. Este efeito também foi observado em estudo com amostra de 278 indivíduos sem doença de Alzheimer e intervalo de tempo de um mês entre os testes.⁸

Estudos coreano¹⁵ e estadunidense⁸ e inglês¹⁴, este último com amostra de 90 indivíduos com idade superior a 40 anos e elevada escolaridade que responderam aos testes com intervalo de tempo entre uma e oito semanas, obtiveram confiabilidades substanciais para os testes de fluência verbal semântica e fonêmica. Esta classificação, baseada nos valores encontrados pelo Coeficiente de Correlação de Pearson, é compatível àquela observada neste estudo, considerando o mesmo teste estatístico.

O teste de trilhas, dada sua confiabilidade quase perfeita, pode ser útil para aplicação em espaço de tempo de vinte dias entre o teste e o reteste. Confiabilidade elevada para este

teste foi encontrada em estudo alemão, com amostra de 55 indivíduos, média de idade de 46 anos e média de 10 anos de escolaridade.¹⁶

A confiabilidade medida de forma estratificada por sexo, idade e escolaridade variou de moderada a substancial para o teste de memória de palavras e teste de fluência verbal fonêmica. Manteve-se moderada para o teste de evocação e fluência verbal semântica e quase perfeita para o teste de trilhas. Para o teste de reconhecimento a confiabilidade variou de discreta a substancial, porém com intervalos de confiança amplos e sobrepostos, o que indica que a probabilidade de encontrar estes valores ao acaso é grande. Principalmente para o grupo de pessoas com segundo grau incompleto obteve-se valores que não são estatisticamente significativos, provavelmente devido ao reduzido número de pessoas, sendo necessários estudos adicionais com amostra maior.

As medidas de qualidade relativas ao ambiente da entrevista (ambiente calmo, reservado, com baixos níveis de estímulos distraidores) podem ter contribuído para não haver diferenças entre os valores de confiabilidade estratificados por local de realização dos testes.

Infere-se que o aprendizado tenha influenciado os resultados obtidos na confiabilidade dos testes. Evidência deste efeito é a diferença entre médias dos testes de memória e fluência, com exceção do teste de reconhecimento e diferença das medianas do teste de trilhas. No teste de reconhecimento a pequena variabilidade dos escores, todos muito próximos a dez, valor máximo possível para o teste, pode ter contribuído para que as diferenças entre as médias não tenham sido significativas.

3.5 CONCLUSÃO

Neste estudo, foi observada confiabilidade moderada para testes de cognição realizados em adultos e no intervalo de tempo de vinte dias. A melhora do desempenho em todos os testes em relação aos retestes sugere o efeito do aprendizado. Recomenda-se o reteste dos testes de memória da lista de palavras, reconhecimento, evocação e dos testes de fluência fonêmica e semântica com intervalo de tempo superior ao utilizado neste estudo. O teste de trilhas B, que apresentou confiabilidade quase perfeita, pode ser útil para reaplicação em curto espaço de tempo.

REFERENCIAS BIBLIOGRÁFICAS

1. Kandel ER, Schwartz JH.; Jessell TM. Princípios da Neurociência. 4 ed., Barueri (SP): Manole, 2003.
2. Salthouse, TA. Selective review of cognitive aging. Journal of the International Neuropsychological Society 2010; 16: 754-760.

3. Aquino EM, Barreto SM, Bensenor IM, et al. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): Objectives and Design. *Am J Epidemiol.* 2012; 175(4):315-24
4. Lezak, MD. *Neuropsychological assessment* 3ed. New York, NY: Oxford University Press, 1995.
5. Abreu, ID, Florenza OV, Barros HL. Demência de Alzheimer: correlação entre memória e autonomia. *Revpsiquiatr clín.* 2005; 32(3): 131-136.
6. Fillenbaum GG, Belle G, Morris JC, et al. CERAD (Consortium to Establish a Registry for Alzheimer's Disease): The first 20 years. *Alzheimer's & Dementia: The journal of the Alzheimer's Association.* Mar 2008; 4(2): 96-109
7. Bertolucci PHF, Okamoto IH, Neto JT, Ramos LR, Brucki SMD. Desempenho da população brasileira na bateria neuropsicológica do Consortium to Establish a Registry for Alzheimer's Disease (CERAD). *Revista de Psiquiatria Clínica* 1998 março/abril; 25(2).
8. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989 Sep;39:1159-1165.
9. Martins GA. Sobre Confiabilidade e Validade. *RGBN* jan/abr 2006; 8(20):1-12.
10. Brucki SMD, Malheiros SMF, Okamoto IH, Bertolucci PHF. Dados normativos para o teste de fluência verbal categoria animais em nosso meio. *ArqNeuropsiquiatr* 1997; 55(1):56-61.
11. Brucki SMD, Rocha MSG. Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Brazilian Journal of Medical and Biological Research* 2004; 37: 1771–1777.
12. Silva TBL, Yassuda MS, Guimarães VV, Florindo AA. Fluência Verbal e Variáveis Sociodemográficas no Processo de Envelhecimento: Um Estudo Epidemiológico. *Psicologia: Reflexão e Crítica* 2011; 24 (4): 739-746.
13. Ribeiro PCC, Oliveira BHD, Cupertino APFB, et al. Desempenho de Idosos na Bateria Cognitiva CERAD: Relações com variáveis sociodemográficas e saúde percebida. *PsicolReflexCrít*, Apr 2010, 23(1):102-109
14. Harrison JE, Buxton P, Husain M, Wise R. Short test of semantic and phonological fluency: Normal performance, validity and test-retest reliability. *British Journal of Clinical Psychology* 2000; 39:181-191.
15. Lee JH, Lee KU, Lee DY, et al. Development of the Korean Version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K): Clinical and Neuropsychological Assessment Batteries. *Journal of Gerontology* 2002; 57B(1): 47-53.
16. Wagner S, Helmreich I, Dahmen N, Lieb K, Tadić A. Reliability of Three Alternate Forms of the Trail Making Tests A and B. *ArchClinNeuropsychol* 2011; 26 (4): 314-321.
17. Passos VMA, Giatti L, Barreto SM, Figueiredo RC, Caramelli P, Bonseñor I, et al. Verbal fluency tests reliability in a Brazilian multicentric study, ELSA-Brasil. *ArqNeuropsiquiatr* 2011;69(5):814-816
18. Caramelli P, Carthery-Goulart MT, Porto CS, Charchat-Fichman Helenice, Nitrini R. Category fluency as a screening test for Alzheimer disease in illiterate and literate patients. *Alzheimer Dis Assoc Disord* 2007;21:65-67.
19. Dean AG, Arner TG, Sunki GG, et al. *Epi Info™*, a database and statistics program for public health professionals. Centers for Disease Control and Prevention, Atlanta, Georgia, USA, 2007.
20. StataCorp. *Stata Statistical Software: Release 10.* College Station, Texas: 2007 StataCorp LP.

21. Szklo M, Nieto FJ. Quality assurance and control. In: _____. Epidemiology: beyond the basics. 2^a ed. Sudbury, MA: Jones and Bartlett Publishers; 2007. P. 297-347.
22. Menezes PR, Scazufca M. Estudo da Confiabilidade da Versão em Português da “Escala de Avaliação da Incapacitação Psiquiátrica (WHO/DAS)”. Revista ABP-APAL 1993; 15(2): 65-67.
23. Salthouse TA, Nesselroade JR, Berish DE. Short-Term variability in cognitive performance and calibration of longitudinal change. Journal of Gerontology 2006; 61B(3): 144,151.
24. Salthouse TA. Implications of within-person variability in cognitive and neuropsychological functioning for the interpretation of change. Neuropsychology 2007; 21(4): 401-411
25. Rodrigues AB, Yamashita ET, Chiappetta ALML. Teste de fluência verbal no adulto e no idoso: verificação da aprendizagem verbal. Ver CEFAC, out-dez 2008; 10(4): 443-451.

4. CONSIDERAÇÕES FINAIS

O ELSA, por tratar-se de uma grande pesquisa multicêntrica e multidisciplinar, em que são coletados dados acerca de diversos aspectos da vida dos participantes tem possibilidades de estudar de forma abrangente a cognição humana. Dado o aspecto longitudinal do estudo, alterações cognitivas poderão ser identificadas e mensuradas por testes aplicados aos mesmos participantes ao longo do tempo, o que permite comparações diretas, diferentes das inferências advindas de comparações indiretas de estudos transversais. A maioria dos estudos dos déficits cognitivos, com aplicação de testes, é realizada em população idosa. Pouco se sabe da aplicabilidade dos testes e desempenho da população adulta em relação às funções cognitivas.

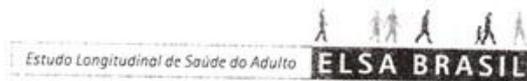
A preocupação com a garantia e controle de qualidade em todas as etapas de organização, planejamento e execução deste estudo de grande porte visa produção de dados fidedignos para que possam ser adequadamente analisados. Este cuidado reflete a compromisso ético e técnico do estudo para com seus participantes e para com a população em geral que se beneficiará com as inferências que podem advir dos achados de pesquisa.

Estudos de confiabilidade dos instrumentos utilizados fazem parte da fase inicial de estudos longitudinais. Foi publicado recentemente o estudo da concordância dos escores dos testes de Fluência verbal semântica e fonêmica aferidos por supervisores dos diferentes centros de pesquisa do ELSA, onde encontrou-se elevado grau de concordância entre cada centro e o padrão de referência. Estudos como este, realizados na fase inicial de estudos de longo prazo, permitem maior segurança no avanço dos estudos da função cognitiva.¹

5. ANEXOS

ANEXO I

Termo de Consentimento Livre e Esclarecido



UFMG Universidade Federal de Minas Gerais
Hospital Borges da Costa
Av. Alfredo Balena, 110 - Funcionários
30130-100 - Belo Horizonte MG
elsabg@ufmg.br
+55 31 3409-9140

Termo de Consentimento Livre e Esclarecido

Apresentação do estudo:

O Estudo Longitudinal de Saúde do Adulto – ELSA Brasil – é uma pesquisa sobre doenças crônicas que acometem a população adulta, principalmente as doenças cardiovasculares e o diabetes. É um estudo pioneiro no Brasil por ser realizado em várias cidades e por acompanhar as pessoas estudadas por um longo período de tempo. Graças a pesquisas semelhantes desenvolvidas em outros países, hoje se sabe, por exemplo, da importância de cuidados à pressão arterial e à dieta para a prevenção dessas doenças.

Objetivos do estudo:

O ELSA Brasil investigará fatores que podem levar ao desenvolvimento dessas doenças, ou ao seu agravamento, visando sugerir medidas mais eficazes de prevenção ou tratamento. Os fatores investigados incluem aspectos relacionados aos hábitos de vida, família, trabalho, lazer e saúde em geral, inclusive fatores genéticos.

Instituições envolvidas no estudo:

O ELSA Brasil envolverá 15.000 funcionários de instituições públicas de ensino e pesquisa localizadas em seis estados brasileiros (BA, ES, MG, RJ, RS e SP)¹. É coordenado por representantes de cada Centro de Investigação, do Ministério da Saúde e do Ministério da Ciência e Tecnologia, tendo sido aprovado pelos Comitês de Ética em Pesquisa dos seis centros. Em Belo Horizonte, o estudo está sob a responsabilidade da Universidade Federal de Minas Gerais.

Participação no estudo:

O/A Sr./a é convidado/a a participar do ELSA Brasil, que envolve o acompanhamento dos participantes por pelo menos sete anos, com a realização de entrevistas, de exames e medidas que ocorrerão em várias etapas.

¹ Fundação Oswaldo Cruz (Fiocruz), Universidade de São Paulo (USP), Universidade Federal da Bahia (UFBA), Universidade Federal de Minas Gerais (UFMG), Universidade Federal do Espírito Santo (UFES) e Universidade Federal do Rio Grande do Sul (UFRGS).

Inicialmente, o/a Sr./a fará a primeira parte da entrevista preferencialmente em sua unidade de trabalho e será agendado/a para comparecer ao Centro de Investigação ELSA (CI-MG), situado na Av. Alfredo Balena, nº 190, Santa Efigênia, em um único momento, com duração de cerca de cinco horas, ou seja, uma manhã inteira. No CI-MG, o/a Sr./a fará a segunda parte da entrevista, realizará algumas medidas (peso, altura, circunferência de cintura, quadril e pescoço e pressão arterial), exame de urina de 12 horas noturnas, ultrassom do abdome e carótidas, ecocardiograma, eletrocardiograma, fotografia do fundo de olho e exames especializados de fisiologia cardiovascular (Variabilidade da Frequência Cardíaca e Velocidade da Onda do Pulso). Realizará também exames de sangue², para os quais serão feitas duas coletas: a primeira quando chegar, em jejum de 12 horas, e a segunda, após duas horas da ingestão de uma bebida doce padrão (exceto os diabéticos que receberão um lanche específico em substituição). O total de sangue coletado será aproximadamente de 65 ml, e não traz inconveniências para adultos. Apenas um leve desconforto pode ocorrer associado à picada da agulha. Algumas vezes pode haver sensação momentânea de tontura ou pequena reação local, mas esses efeitos são passageiros e não oferecem riscos. A maioria desses exames já faz parte da rotina médica e nenhum deles emite radiação.

Caso necessário, será solicitada sua liberação para participar da pesquisa em horário de trabalho.

A coleta de sangue segue rotinas padronizadas e será realizada, assim como os demais procedimentos, por pessoal capacitado e treinado para este fim, supervisionados por profissional qualificado que poderá orientá-lo/a no caso de dúvida, ou alguma outra eventualidade.

Após esta primeira etapa do estudo, o/a Sr./a. será periodicamente contatado/a por telefone, correspondência ou e-mail para acompanhar as modificações no seu estado de saúde e para obtenção de informações adicionais. Estão previstas novas visitas ao CI-MG a cada três anos. Por isso, é muito importante informar seu novo endereço e telefone em caso de mudança.

Para poder monitorar melhor sua situação de saúde, é essencial obter detalhes clínicos em registros de saúde. Assim, necessitamos obter informações da UFMG e de outras instituições do sistema de saúde, a respeito da ocorrência de hospitalizações, licenças médicas, eventos de saúde, aposentadoria, ou afastamento de qualquer natureza. Para isso é imprescindível que nos autorize por escrito o acesso às mesmas ao final deste documento. Infelizmente, sem essa autorização, não será possível sua participação no estudo, pois dela depende a confirmação de eventos clínicos.

Armazenamento de material biológico:

Serão armazenadas amostras de sangue, urina e ácido desoxirribonucléico (DNA) por um período de cinco anos, sem identificação nominal, de forma segura e em locais especialmente preparados para a conservação das mesmas. Assim como em outras pesquisas no país e no mundo, essas amostras são fundamentais para futuras análises que possam ampliar o conhecimento sobre as doenças em estudo, contribuindo para o avanço da ciência.

² Hemograma completo, exames diagnósticos para diabetes (glicose e insulina em jejum e pós-ingestão e teste de tolerância à glicose), creatinina, dosagem de lipídios, hormônios associados ao diabetes ou à doença cardiovascular e provas de atividade inflamatória.

Análises adicionais, de caráter genético ou não, que não foram incluídas nos objetivos definidos no protocolo original da pesquisa, somente serão realizadas mediante a apresentação de projetos de pesquisa específicos, aprovados pelo Comitê Diretivo e pelos Comitês de Ética em Pesquisa de cada uma das instituições envolvidas, incluindo a assinatura de novos Termos de Consentimento Livre e Esclarecido.

Seus direitos como participante:

Sua participação é inteiramente voluntária, sendo fundamental que ocorra em todas as etapas do estudo. Entretanto, se quiser, poderá deixar de responder a qualquer pergunta durante a entrevista, recusar-se a fazer qualquer exame, solicitar a substituição do/a entrevistador/a, ou deixar de participar da pesquisa a qualquer momento.

Não será feito qualquer pagamento pela sua participação e todos os procedimentos realizados serão inteiramente gratuitos. Os participantes poderão ter acesso aos resultados das análises realizadas no estudo por meio de publicações científicas e do website oficial da pesquisa (www.elsa.org.br).

Os exames e medidas realizados no estudo não têm por objetivo fazer o diagnóstico médico de qualquer doença. Entretanto, como eles podem contribuir para o/a Sr/a. conhecer melhor sua saúde, os resultados destes exames e medidas lhe serão entregues e o/a Sr/a. será orientado a procurar as unidades da rede SUS ou outro serviço de saúde de sua preferência, quando eles indicarem alguma alteração em relação aos padrões considerados normais. Se durante a sua permanência no CI-MG forem identificados problemas que requeiram atenção de urgência/emergência, o/a Sr/a. será atendido/a no Hospital das Clínicas da UFMG.

Todas as informações obtidas do/a Sr/a. serão confidenciais, identificadas por um número e sem menção ao seu nome. Elas serão utilizadas exclusivamente para fins de análise científica e serão guardadas com segurança - somente terão acesso a elas os pesquisadores envolvidos no projeto. Com a finalidade exclusiva de controle de qualidade, sua entrevista será gravada e poderá ser revista pela supervisão do projeto. A gravação será destruída posteriormente. Como nos demais aspectos do projeto, serão adotados procedimentos para garantir a confidencialidade das informações gravadas. Em nenhuma hipótese será permitido o acesso a informações individualizadas a qualquer pessoa, incluindo empregadores, superiores hierárquicos e seguradoras.

Uma cópia deste Termo de Consentimento lhe será entregue. Se houver perguntas ou necessidade de mais informações sobre o estudo, ou qualquer intercorrência, o/a Sr/a. pode procurar a coordenadora do ELSA Brasil em Minas Gerais, Professora Sandhi Maria Barreto, Faculdade de Medicina da UFMG, no endereço: Av Alfredo Balena, 190, Campus Saúde; telefones (31) 3409-9938 ou (31) 3409-9140. O Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais pode ser contatado pelo seguinte telefone: (31) 3409-4502.

Sua assinatura a seguir significa que o/a Sr/a. leu e compreendeu todas as informações e concorda em participar da pesquisa ELSA Brasil.



Estudo Longitudinal de Saúde do Adulto

ELSA BRASIL



UFMG

UFMG Universidade Federal de Minas Gerais
Hospital Borges da Costa
Av. Alfredo Balena, 110 - Funcionários
30130-100 - Belo Horizonte - MG
elsamg@ufmg.br
+55 31 3409-9140

Termo de Consentimento Livre e Esclarecido

Nome do participante _____

Documento de Identidade _____

Data de nascimento _____

Endereço _____

Telefone para contato _____

Declaro que compreendi as informações apresentadas neste documento e dei meu consentimento para participação no estudo.

Autorizo os pesquisadores do Estudo Longitudinal de Saúde do Adulto – ELSA Brasil, a obter informações sobre a ocorrência de hospitalizações, licenças médicas, eventos de saúde, aposentadoria, ou afastamento de qualquer natureza em registros de saúde junto ao setor de recursos humanos da UFMG e a outras instituições de saúde públicas ou privadas, conforme indicar a situação específica.

No caso de hospitalização, autorizo, adicionalmente, que o/a representante do ELSA, devidamente credenciado/a, copie dados constantes na papeleta de internação, bem como resultados de exames realizados durante minha internação.

As informações obtidas somente poderão ser utilizadas para fins estatísticos e deverão ser mantidas sob proteção, codificadas e sem minha identificação nominal.

Assinatura _____

Declaro concordar que amostras de sangue sejam armazenadas para análises futuras sobre as doenças crônicas em estudo. Sim Não

Assinatura _____

Local _____ Data ____ / ____ / ____

Nome do/a entrevistador/a: _____

Código do/a entrevistador/a no CI-MG _____

Assinatura _____

ANEXO II**Documento de aprovação do Comitê de Ética**

Universidade Federal de Minas Gerais
Comitê de Ética em Pesquisa da UFMG - COEP

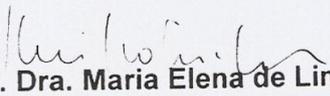
Parecer nº. ETIC 186/06

Interesse: Prof. (a) Sandhi Maria Barreto
Depto. De Medicina Preventiva e Social
Faculdade de Medicina -UFMG

DECISÃO

O Comitê de Ética em Pesquisa da UFMG – COEP, aprovou no dia 28 de junho de 2006 o projeto de pesquisa intitulado “**ELSA - Estudo longitudinal da saúde do adulto.**” bem como o Termo de Consentimento Livre e Esclarecido do referido projeto.

O relatório final ou parcial deverá ser encaminhado ao COEP um ano após o início do projeto.

pi 
Profa. Dra. Maria Elena de Lima Perez Garcia
Presidente do COEP/UFMG

Processo 0186/06

Enviado ao COEP em 25/05/2006
 Enviado ao parecerista em 12/06/2006
 Devolvido à COEP em 28/06/2006

Título: Estudo Longitudinal da Saúde do Adulto

Área: Ciências da Saúde

Pesquisadora: Sandhi Maria Barreto

Curriculum Vitae: Lattes

Apoio Institucional: Constam do processo documento assinado pelo Diretor da Faculdade de Medicina, e da chefe do Departamento de Medicina Preventiva e Social. Não consta parecer substanciado da Câmara Departamental.

Informações adicionais:

A pesquisadora informa que o projeto já foi aprovado pelo Comitê de Ética do Hospital das Clínicas da Universidade de São Paulo, fato que pode ser comprovado em documento anexado ao processo em questão. Informa, ainda, que o projeto já foi encaminhado ao CONEP e conta com financiamento de um convênio com a FINEP. Segundo a pesquisadora, o documento relativo ao convênio será enviado ao COEP/UFMG, assim que chegue em suas mãos.

Data prevista para o início do projeto: março de 2007

Resumo

Trata-se de um projeto multicêntrico envolvendo, além da UFMG, a Universidade Federal da Bahia, a Fundação Oswaldo Cruz, a Universidade de São Paulo, Universidade do Estado do Rio de Janeiro e a Universidade Federal do Rio Grande do Sul.

O objetivo do estudo é investigar aspectos epidemiológicos, clínicos e moleculares de doenças crônicas não transmissíveis, principalmente as cardiovasculares e o diabetes *mellitus* tipo 2.

Os sujeitos da pesquisa serão funcionários não terceirizados das seis instituições suoracitadas que integram a pesquisa. Serão recrutados 15 000 funcionários nos seis centros participantes, sendo 3000 na UFMG, entendendo-se que esse conjunto de funcionários “representa a diversidade étnica, econômica e social do país” (p.18). Por ser um estudo longitudinal, os participantes serão acompanhados por pelo menos sete anos, ao longo dos quais serão realizados novos exames e novas entrevistas.

Inicialmente os sujeitos serão convidados voluntariamente para participar do projeto. Se houver um número superior ao estipulado (no caso da UFMG, superior a 3000) far-se-á uma seleção para propiciar uma distribuição proporcional por faixa etária, sexo e categoria ocupacional.

A coleta de dados será feita pelos seguintes procedimentos: a) aplicação de questionário, envolvendo vários aspectos sócio-demográficos, b) medidas antropométricas, c) medidas de pressão arterial, d) índice tornozelo braquial, e) coleta de sangue, f) coleta de urina, g) eletrocardiograma, h) ultrassonografia, i) ecocardiograma e k) velocidade de onda de pulso.

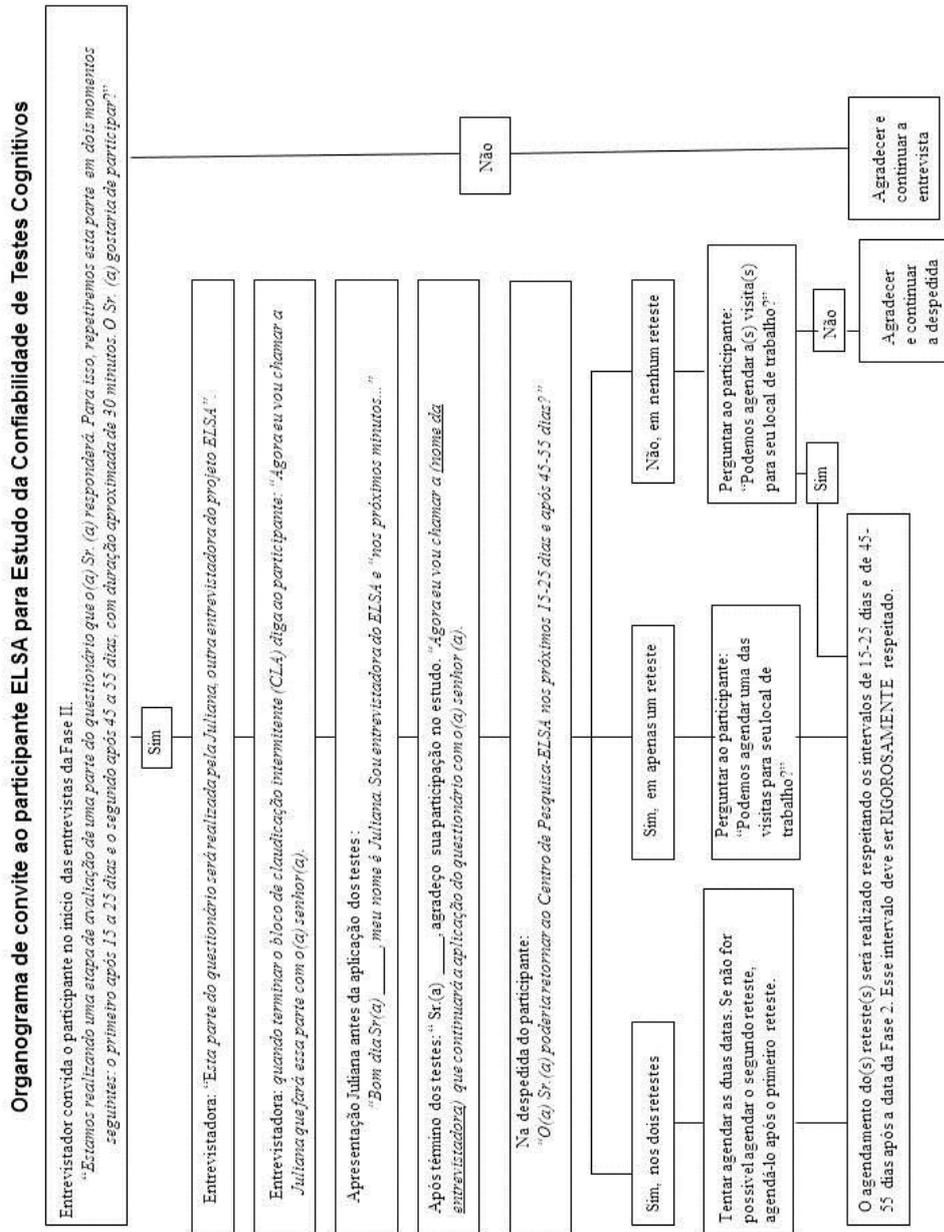
O Termo de Consentimento Livre e Esclarecido está bastante detalhado e formulado em linguagem acessível, fornecendo informações relativas aos objetivos do estudo e a todos os procedimentos (entrevistas, exames e acompanhamento) aos quais os participantes serão submetidos. Esclarece, também, como serão feitos os exames e as coletas de material biológico e por quem serão feitos. E ainda, informa sobre o que será feito com o material biológico coletado.

SJM somos pela aprovação do projeto.

Belo Horizonte, 28 de junho de 2006

ANEXO III

Organograma de convite ao participante ELSA para estudo da confiabilidade de testes cognitivos



ANEXO IV

Dicionário das variáveis para o Estudo da Confiabilidade da Bateria de Testes de Função Cognitiva do ELSA

Resumo das variáveis

I – Identificação

I.1 – ID

I.2 – Livro

I.3 – Sex

I.4 – Age

I.5 – Escolaridade

I.5.1 Esc

I.5.2 ser1 | 1 |

I.5.3 ser2 | 2 |

I.5.4 ser3 | 3 |

I.5.5 ser4 | 4 |

I.5.6 ser5 | 5 |

I.5.7 ser6 | 6 |

I.5.8 ser7 | 7 |

I.5.9 gcom1 | 8 |

I.5.10 ginc2 | 9 |

I.5.11 gcom2 | 10 |

I.5.12 uniinc | 11 |

I.5.13 unicom | 12 |

I.5.14 unipos | 13 |

I.5.15 Segundo Grau incompleto | 1 |

I.5.16 Segundo Grau completo | 2 |

I.5.17 Superior | 3 |

II. Codificação dos testes de Função Cognitiva

As variáveis seguintes correspondem à codificação dos testes. A legenda T0, T1 e T2 no início da variável identifica a mesma questão do teste, respondida pelo mesmo participante, em momentos distintos.

Legenda das variáveis:

T0- Tempo zero. Corresponde a bateria de testes cognitivos que o participante responde no dia da Fase II do ELSA.

T1 – Tempo Um. Corresponde a bateria de testes cognitivos que o participante responde na primeira repetição dos testes.

Detalhamento das variáveis

I – Identificação

I.1 – ID

Número de sete dígitos que identifica o participante do ELSA. Para os participantes de Minas Gerais este número está compreendido entre 3000000 e 3900000.

| _ _ _ _ _ _ _ |

I.2 – Livro

Número na capa que identifica o encadernado dos retestes. Compreendido entre 0 e 12.

| _ _ _ |

I.3 – Sex

Sexo do participante

Feminino | 0 |

Masculino | 1 |

I.4 – Age

Idade do participante em anos completos na data do preenchimento do Questionário de Fase I do ELSA. Compreendida entre 35 e 74 anos.

| _ _ _ |

I.5 – Escolaridade

I.5.1Esc Anos de escolaridade do participante

| _ _ _ |

I.5.2 Cursou com aprovação até a primeira série **ser1** | 1 |

I.5.3 Cursou com aprovação até a segunda série **ser2** | 2 |

I.5.4 Cursou com aprovação até a terceira série **ser3** | 3 |

I.5.5 Cursou com aprovação até a quarta série **ser4** | 4 |

I.5.6 Cursou com aprovação até a quinta série **ser5** | 5 |

I.5.7 Cursou com aprovação até a sexta série **ser6** | 6 |

I.5.8 Cursou com aprovação até a sétima série **ser7** | 7 |

I.5.9 Primeiro grau completo **gcom1** | 8 |

I.5.10 Segundo grau incompleto **ginc2** | 9 |

I.5.11 Segundo grau completo **gcom2** | 10 |

I.5.12 Universitário incompleto **uniinc** | 11 |

I.5.13 Universitário completo **unicom** | 12 |

I.5.14 Pós-graduação **unipos** | 13 |

I.5.15 Segundo Grau incompleto | 1 |

I.5.16 Segundo Grau completo|2 |

I.5.17 Superior | 3 |

II. Codificação dos testes de Função Cognitiva T0

III. Codificação dos testes de Função Cognitiva T1

II.1 Data/ III.1 Data

Data da realização da Bateria de Testes Cognitivos em dia, mês e ano com 4 dígitos.

|__ __| |__ __||__ __ __ __|

II.2. FCO01A/ III.2. FCO01A

Questão; “Gostaria que o(a) senhor(a) lesse algumas palavras, e ao final o(a) senhor(a) repetisse para mim. Podemos continuar?”

Sim | 0 |

Não | 1 |

II.3 MR1/ II.3 MR1

Escore da primeira tentativa do Teste da Lista de Palavras, compreendido entre 0 e 10.

|__ __|

II.4 MR2/ III.4 MR2

Escore da segunda tentativa do Teste da Lista de Palavras, compreendido entre 0 e 10.

|__ __|

II.5 MR3/ III.5 MR3

Escore da terceira tentativa do Teste da Lista de Palavras, compreendido entre 0 e 10.

|__ __|

II.6 Palavras/ III.6 Palavras

Soma dos escores MR1+MR2+MR3, compreendido entre 0 e 30.

|__ __|

II.7 MT/ III.7 MT

Escore do Teste de evocação tardia da lista de palavras, compreendido entre 0 e 10.

|__ __|

II.8 Rec1/ III.8 Rec1

Escore 1 do Teste de Reconhecimento.Número de palavras corretamente reconhecidas, compreendido entre 0 e 20.

|__ __|

II.9Rec2/ III.9Rec2

Escore 2 do Teste de Reconhecimento.Número de palavras equivocadamente reconhecidas, compreendido entre 0 e 20.

|__ __|

II.10 FluA/ III.10 FluA

Fluência verbal semântica,: número de animais lembrados.

|__ __|

II.11 FluF/ III.11 FluF/ IV.11 FluF

Fluência verbal fonêmica: número de palavras iniciadas com letra F lembradas.

|__ __|

II.12 Tri1/ III.12 Tri1/ IV.12 Tri1

Conseguir terminar o Teste de Trilhas parte I.

| 0 | Sim

| 1 | Não

| 2 | Não quis fazer o teste

II.13 Tri2/ III.13 Tri2/ IV.13 Tri2

Conseguir terminar o teste de trilhas parte II

| 0 | Sim

| 1 | Não

| 2 | Não quis fazer o teste

II.14 Tritime/ III.14 Tritime/ IV.14 Tritime

Tempo do exame: |__ __|hora|__ __| minutos |__ __|segundos

II.15 Seg/ III.15 Seg/ IV.15 Seg

Tempo do exame: |__ __|segundos

ANEXO V

E-mail de envio de Trabalho ao XVIII Congresso Brasileiro de Geriatria e Gerontologia

Assunto:	Fwd: XVIII Congresso Brasileiro de Geriatria e Gerontologia – Envio de Trabalho
De:	Valéria Passos (passos.v@gmail.com)
Para:	julyalba@yahoo.com.br; luana.giatti@gmail.com; sandhi.barreto@gmail.com;
Data:	Domingo, 11 de Março de 2012 17:19

----- Mensagem encaminhada -----

De: <trabalhos@ccmeventos.com.br>

Data: 11 de março de 2012 17:18

Assunto: XVIII Congresso Brasileiro de Geriatria e Gerontologia - Envio de Trabalho

Para: passos.v@gmail.com

Prezado(a) Sr(a). VALÉRIA MARIA DE AZEREDO PASSOS (76788)

Seu trabalho foi recebido com sucesso. Ele será encaminhado à Comissão Organizadora para análise e seleção. Você poderá acompanhar o status da avaliação de seu trabalho através da área restrita do site, informando seu login e senha.

Caso a Apresentação do trabalho seja delegada a um de seus co-autores, por favor, passe esta informação por email para a Secretaria do Evento.

Atenciosamente,
Secretaria do Evento

Título:

CONFIABILIDADE DOS TESTES DA BATERIA DE FUNÇÃO COGNITIVA DO ESTUDO LONGITUDINAL DE SAÚDE DO ADULTO – ELSA-BRASIL

Inscrição Responsável:

VALÉRIA MARIA DE AZEREDO PASSOS

Modalidade:

Pôster

Eixo Principal:

Demências e outros distúrbios cognitivos e neurológicos, *Delirium*

Texto (resumo):

Introdução: A avaliação da função cognitiva é uma das linhas de investigação do Estudo Longitudinal de Saúde do Adulto – ELSA-Brasil, coorte de caráter multicêntrico desenvolvido em seis instituições de ensino superior e pesquisa do país, que tem por objetivos verificar a incidência de déficit cognitivo e demências em população adulta (34-75 anos). Objetivo: Avaliar a confiabilidade dos testes cognitivos usados no ELSA-Brasil. Método: Estudo de confiabilidade tipo teste-reteste aplicados a 172 participantes por entrevistador treinado. O intervalo entre o teste e reteste variou de 15 a 25 dias. O Coeficiente de Correlação Intraclassa (ICC) foi utilizado como estimativa da confiabilidade. Resultados: A amostra foi composta predominantemente por adultos (75,6%), idade média de 52 anos, com alta escolaridade e distribuição por sexo homogênea. Foram encontrados os seguintes valores de ICC: Teste de Memória de Palavras (0,58), Reconhecimento (0,35), Evocação (0,50), Fluência animais (0,57), Fluência letra F (0,58) e Teste de trilhas B (0,91). Discussão: Este é o primeiro estudo brasileiro a aferir a confiabilidade desses testes utilizados no ELSA-Brasil. Os valores encontrados foram similares àqueles de outros estudos. A confiabilidade de todos os testes foi classificada como moderada, com exceção do Teste de Reconhecimento, que não permitiu análise devido à amplitude do intervalo de confiança e do Teste de Trilhas B, cuja confiabilidade foi classificada como quase perfeita. Foi observado melhor desempenho nos retestes em relação aos testes, sugerindo aprendizado. Conclusão: Houve confiabilidade moderada para testes de cognição realizados em adultos e no intervalo de tempo de vinte dias. A melhora do desempenho em todos os testes em relação aos retestes sugere o efeito do aprendizado. Recomenda-se o reteste dos testes de memória e fluência com intervalo de tempo superior ao utilizado neste estudo. O Teste de Trilhas, que apresentou confiabilidade de quase perfeita, pode ser útil para reaplicação em curto espaço de tempo.

Apresentador:

Valéria Maria de Azeredo Passos

Autores:

- > Juliana Alves Batista¹ - Batista, J A - UNIVERSIDADE FEDERAL DE MINAS GERAIS
- > Luana Giatti - Giatti, L - UNIVERSIDADE FEDERAL DE MINAS GERAIS
- > Sandhi Maria Barreto - Barreto, S M - UNIVERSIDADE FEDERAL DE MINAS GERAIS
- > Valéria Maria de Azeredo Passos - Passos, V M A - UNIVERSIDADE FEDERAL DE MINAS GERAIS

ANEXO VI

Reliability of Cognitive Tests of ELSA-Brasil, the Brazilian Longitudinal Study of Adult Health

Juliana Alves Batista¹, Luana Giatti², Sandhi Maria Barreto³,
Ana Roscoe Papini Galery⁴, Valéria Maria de Azeredo Passos⁵

ABSTRACT. Cognitive function evaluation entails the use of neuropsychological tests, applied exclusively or in sequence. The results of these tests may be influenced by factors related to the environment, the interviewer or the interviewee. **Objectives:** We examined the test-retest reliability of some tests of the Brazilian version from the *Consortium to Establish a Registry for Alzheimer's disease*. **Methods:** The ELSA-Brasil is a multicentre study of civil servants (35-74 years of age) from public institutions across six Brazilian States. The same tests were applied, in different order of appearance, by the same trained and certified interviewer, with an approximate 20-day interval, to 160 adults (51% men, mean age 52 years). The Intraclass Correlation Coefficient (ICC) was used to assess the reliability of the measures; and a dispersion graph was used to examine the patterns of agreement between them. **Results:** We observed higher retest scores in all tests as well as a shorter test completion time for the Trail Making Test B. ICC values for each test were as following: Word List Learning Test (0.56), Word Recall (0.50), Word Recognition (0.35), Phonemic Verbal Fluency Test (VFT, 0.61), Semantic VFT (0.53) and Trail B (0.91). The Bland-Altman plot showed better correlation of executive function (VFT and Trail B) than of memory tests. **Conclusions:** Better performance in retest may reflect a learning effect, and suggest that retest should be repeated using alternate forms or after longer periods. In this sample of adults with high schooling level, reliability was only moderate for memory tests whereas the measurement of executive function proved more reliable.

Key words: cognitive assessment, reliability, cohort studies.

CONFIABILIDADE DE TESTES COGNITIVOS DO ELSA-BRASIL, ESTUDO LONGITUDINAL DE SAÚDE DO ADULTO

RESUMO. A avaliação da cognição utiliza testes neuropsicológicos, aplicados isoladamente ou em sequência. Os resultados destes testes podem sofrer influências relacionadas ao ambiente, ao entrevistador ou ao entrevistado. **Objetivos:** Investigar a confiabilidade teste-reteste de alguns testes da versão brasileira do *Consortium to Establish a Registry for Alzheimer's disease*. **Métodos:** O ELSA-Brasil é uma coorte multicêntrica de servidores (35-74 anos) de instituições públicas de seis estados. Os mesmos testes foram aplicados, em ordens distintas, pelo mesmo entrevistador treinado e certificado, com intervalo aproximado de 20 dias, entre 160 participantes (51% homens, média de idade de 52 anos). O Coeficiente de Correlação Intraclassa (CCI) foi usado na verificação da confiabilidade das medidas, e um gráfico de dispersão evidenciou padrões de concordância entre teste e reteste. **Resultados:** Foram observados maiores escores em todos os retestes, assim como menor tempo para completar o Teste de Trilhas B (Trilhas B). Os valores de CCI para cada teste foram: memória imediata de palavras (0,56), evocação (0,50), reconhecimento (0,35), Teste de Fluência Verbal fonêmica (TFV, 0,61), TFV semântica (0,53) e Trilhas B (0,91). O gráfico de Bland-Altman mostrou melhor correlação dos testes de função executiva (TFV e Trilhas B) que dos testes de memória. **Conclusões:** O melhor desempenho nos retestes parece refletir efeito de aprendizado e sugere que o reteste seja aplicado após períodos mais longos ou com formas alternativas dos testes. Nesta população com predominância de adultos jovens e de alta escolaridade, a confiabilidade foi moderada para testes de memória e maior para testes de função executiva.

Palavras-chave: avaliação cognitiva, confiabilidade, estudos de coortes.

¹Enfermeira, Mestre em Ciências da Saúde pelo Programa de Pós-graduação em Ciências da Saúde, Universidade Federal de Minas Gerais, Belo Horizonte MG, Brasil. ²Médica, Doutora em Epidemiologia, Professora Adjunta da Escola de Nutrição, Universidade Federal de Ouro Preto, Ouro Preto MG, Brasil. ³Médica, Doutora em Epidemiologia, Professora Titular do Departamento de Medicina Preventiva e Social, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte MG, Brasil. Coordenadora do ELSA-Brasil. ⁴Estudante de Iniciação Científica, Faculdade de Enfermagem, Universidade Federal de Minas Gerais, Belo Horizonte MG, Brasil. ⁵Médica, Especialista em Geriatria, Doutora em Medicina, Professora Associada do Departamento de Clínica Médica, Faculdade de Medicina, Universidade Federal de Minas Gerais. Vice-Coordenadora do Programa de Pós Graduação em Ciências da Saúde, Universidade Federal de Minas Gerais, Belo Horizonte MG, Brasil.

Valéria M.A. Passos. Centro de Investigação ELSA-MG / Hospital Borges da Costa – Av. Alfredo Balena, 110 – 30130-100 Belo Horizonte MG – Brasil. E-mail: vpastos@medicina.ufmg.br; passos.v@gmail.com

Disclosure: The authors report no conflicts of interest.

Received May 29, 2013 Accepted in final form August 15, 2013.

INTRODUCTION

Human cognition refers to the acquirement of knowledge by means of a complex interaction of the neural networks, which form the mental processes connected to thinking, perception, memory and pre-meditated action. The study of cognitive functions includes clinical and neuropsychological evaluations. Neuropsychological evaluation includes the use of tests, applied exclusively or in sequence, to assess functional and intellectual abilities. These tests attempt to capture and describe complex phenomena in a standardized manner, so they can be analysed in clinical and epidemiologic studies.¹

In scientific studies, evidence must be based on valid results, with no methodological errors in the conception, design and implementation of the study, or in the process of data analysis. The application of cognitive tests may be influenced by many factors, which can interfere in their results. There are factors that can be minimized by controlling the quality of the study. Training and certification can prevent sources of variability associated with the examiner, such as: intonation while giving the instructions, the ability to create a professional and friendly environment, experience with the test, following the correct technique, giving neutral answers to patients' questions, repeating the questions and not interpreting them etc. The test orientations should be able to prevent or attenuate some aspects related to the patient, such as fatigue, sleep deprivation, mood and readiness to take the test. Nonetheless, there are sources of variability associated with the test itself. These may be assessed by the test's validity and reliability. The latter indicates the extent to which the test can obtain the same results when reapplied, maintaining the same original conditions.²

In studies that examine the reliability of cognitive assessments, the test is considered precise when the results obtained upon its reapplication are consistent with the results from the first application. In the retests, one strives to maintain the same application conditions, considering the variables that interfere with performance, such as: environment, privacy, luminosity, the examiner's and the participant's situation. The time gap between the test and the retest is also an important factor to be considered. Long periods are associated with changes, such as alterations in cognitive capacity. Short periods, on the other hand, increase the probability of the learning effect, whereby participants remember their answers from the first test and simply repeat them in the retest.³

A battery of cognitive tests was used in the Lon-

gitudinal Study of Adult Health (ELSA-Brasil), which involves a cohort of 15,105 public civil servants. The object of the study is to investigate the incidence and progression of non-communicable chronic diseases, and to examine the biological, behavioural, environmental, occupational, psychological and social factors associated with these diseases and their complications, in an attempt to build a causal model which reflects their inter-relations.⁴ This battery of tests employs some of the neuropsychological tests from the Consortium to Establish a Registry for Alzheimer's disease (CERAD).⁵ The CERAD cognitive test battery, widely used in clinical and epidemiological studies, is described as having many advantages, such as: detecting dementia in its initial phase, allowing comparison of results from different groups, and offering good test-retest reproducibility and substantial interrater reliabilities.⁶ This battery was adapted and validated for use in Brazil in 1998, but its reliability has not yet been tested in this country. However, evaluation of performance on cognition tests show that younger age and higher schooling levels are associated with better performance.⁷⁻⁹

The reliability of a test is highly influenced by the characteristics of the population that takes it. Therefore, the objective of this study was to assess the reliability, by test and retest, of those cognitive tests applied in the ELSA-Brasil population and, furthermore, to investigate their reliability according to age, sex and schooling.

METHODS

Participants. The ELSA-Brasil is an ongoing multicentre study of volunteer adults (35 to 74 years of age) from public teaching and research institutions across six states in Brazil: Bahia, Espírito Santo, Minas Gerais, Rio de Janeiro, Rio Grande do Sul and São Paulo. This test-retest reliability study was performed on a convenience sample formed by 160 participants, selected according to pre-established quotas for sex and age groups (35-44, 45-64 and 65+ years), from one ELSA research centre in Minas Gerais. ELSA-Brasil was approved by the Research Ethics Committee of each of the institutions, including the Research Ethics Committee of the Federal University of Minas Gerais (COEP UFMG), and all participants signed an informed consent form (ETIC 186/06).⁴

Instruments and procedures. The battery of cognitive tests was applied twice, with an interval of 14-27 days (mean = 20 ± 3 days). The memory tests (immediate recall, evocation and recognition) compriseded a list of ten un-

related words printed in large letters on cards, with the words shown every 2 seconds and presented in a different order on each of the three learning trials, with immediate recall. After a 5 minutes' delay, retention and recollection were tested by a free recall and by the recognition of ten previous words that were intermixed with ten distractor words. Verbal fluency tests (VFT) consisted of asking participants to say in one minute as many words as possible related to a specific category of animals (semantic test) or beginning with the letter F (phonemic test). The Trail Making Test B (Trail B), part a, was used to train for Trail B, part b, with the time taken to complete the task computed only for part b. The participant was instructed to draw lines connecting letters and numbers in an order that alternates between increasing numeric value and alphabetic order (1,A, 2,B, 3,C, etc.). The participant had to draw as quickly as possible, without lifting the pencil tip from the page. Supervisors were instructed to point out the errors. The test score was the total time to complete the condition, including the time necessary to correct errors.⁵

The same tests were applied, albeit in a different order, between 22/02/2010 and 03/12/2010 by the same previously-trained and certified interviewer, in a quiet environment, with good lighting and low levels of noise or other distracting stimulations. The order of the tests was arranged in such a manner that there was always a diverting test, category/phonemic or phonemic/category VFT, between the word memory test and the recall and recognition tests. The Trail B was always the first or the last test to be performed. The tests were recorded and later revised. VFT scores were defined by previously-trained and certified supervisors from the ELSA-Brasil research centres. A high level of agreement was observed between each of the six centres and the reference standard.¹⁰

Statistical analysis. The Epiinfo® 3.5.3 Program, 10 was used for the double data entry, and the STATA® Program, 12 for the statistical analysis.

Descriptive analysis of the tests and retests was generated by means of the average and the range of variation in first and second application. As homogeneity was found only for Trail B data (Bartlett Test <0.05), the Mann Whitney test was used to compare the average time between test and retest.

The Intraclass Correlation Coefficient (ICC) was used as the main measure for estimating reliability, since this test assesses the total variability caused by differences between individuals. The ICC reliability test was done according to the characteristics of the participants: sex,

age range (35-59 and 60-74 years-old) and educational level (uncompleted high school, completed high school, University).

Reliability, according to ICC values, was classified as poor when equal to zero; slight – from 0.01 to 0.2; fair – from 0.21 to 0.4; moderate – from 0.41 to 0.6; substantial – from 0.61 to 0.8; almost perfect – from 0.81 to 0.9.¹³

In order to compare our results with other studies, Pearson Correlation Coefficients were also estimated for memory tests and VFT. The Spearman Coefficient was estimated to compare the Trail B test. The Pearson coefficient measures the degree to which a paired group of observations in a diagram approaches a situation where each point is located precisely over the straight line, which means the absence of difference between two observations. Dispersion graphics were used to evaluate the pattern and distribution of scores.

RESULTS

The study sample had the same sex and age distribution as the ELSA cohort and comprised 81 (50.6%) men and 79 (49.4%) women, 121 (75.6%) adults (35-59 years old) and 39 (24.4%) elderly (60-74 years old). A higher schooling level (10.6% had uncompleted high school, 28.8% completed high school and 60.6% had a University degree) than the participants of the cohort was observed.⁴

In addition, higher retest scores on the word memory, recall, semantic and phonemic VFT tests and a shorter retest time to perform the Trail B (Table 1), were also observed. The ICC varied from 0.35, for the recognition test, to 0.91, for the Trail B, which means that the capacity of the different tests to discriminate between individuals ranged from between moderate and almost perfect, respectively. All the tests presented a positive correlation, with statistically significant values, revealing that the retest scores tended to increase linearly in relation to the test scores (Table 2).

Figure 1 depicts the dispersion graphs corresponding to Pearson coefficient values for the cognitive tests. The inclination of the line deviating from 45° shows the memory tests and VFT retest scores were higher than the test scores, while the opposite occurred with the Trail B. The recall test graph shows a higher dispersion of values. In the recognition tests, the presence of scores close to ten, the maximum limit in the test, is notable.

No influence of sex, age or schooling on reliability was found when variability of all test scores were analysed according to these variables, using stratified ICC values and their confidence intervals (Table 3).

Table 1. Score distribution of cognitive tests and retests among 160 participants of ELSA-Brasil.

Measures	Tests					
	Word memory	Recall	Recognition	VFT* (animals)	VFT (letter F)	TRAIL B (seconds)
Range – test	11-28	1-10	8-10	6-35	3-27	29-858
Range – retest	12-30	2-10	7-10	10-34	2-26	31-526
Average – test	21	7	10	19	13	90.0
Average – retest	23.5	8	10	20	14	81.5**
Difference	p<0.001	p<0.001	p=0.42	p<0.001	p<0.001	p=0.009

*VFT: Verbal Fluency Tests; **Trail B mean execution time and Mann-Whitney test.

DISCUSSION

The knowledge of reliability measures can be of clinical use and are very important for epidemiological studies, especially those with populations from countries with different schooling, culture and language than those where the tests were developed. Few publications have evaluated the reliability of the tests used in the present study.^{5,14-16} Although we used the adaption of the CERAD protocol proposed for Brazil,¹⁷ it is important to know the reliability of cognitive tests for the ELSA-Bra-

sil population, as well as for any study which includes a study population other than the population originally investigated. In the CERAD study, the reliability was assessed among 610 individuals with or without Alzheimer's disease, with an average age of 68 years.⁵

In epidemiological studies, one rarely obtains the reproducibility level found in laboratory investigations, where it is easier to maintain identical evaluation conditions. In the present study, the word memory, recall and semantic VFT test revealed moderate reliability; the phonemic VFT, substantial reliability; and the Trail B almost perfect reliability. The higher reliability of the category VFT and Trail B suggests that they are more precise and less influenced by time of reapplication, since processing speed is common to both tests and is less affected by the test-retest effect.

In this study, it was decided to measure reliability using the ICC. The Pearson coefficient was used only for the sake of comparison with other studies that employed it, since the Pearson coefficient allows assessment of the correlation between variables, but not the difference between the evaluations. Our results are similar to the findings of studies carried out with other populations. Moderate reliability for the word memory test was found in a study from Korea¹⁵ and another study from

Table 2. Test-retest reliability of cognitive tests performed in 160 participants of ELSA-Brasil.

Test	ICC* (95% CI**)	Pearson coefficient***
Word memory	0.56 (0.33-0.79)	0.74
Recall	0.50 (0.17-0.83)	0.68
Recognition	0.35 (0.00-0.94)	0.40
VFT****(animals)	0.53 (0.33-0.74)	0.72
VFT (letter F)	0.61(0.41-0.81)	0.77
Trail making test B*****	0.91 (0.87-0.95)	0.76

*ICC: Intraclass Correlation Coefficient; **CI: Confidence Interval; ***p<0.0001; ****VFT: Verbal Fluency Tests; *****Spearman coefficient.

Table 3. Intraclass Correlation Coefficient of cognitive function tests in ELSA-Brasil, by sex, age and schooling.

Variable	Word memory	Recall	Recognition	VFT*(animals)	VFT (letter F)	Trail B	
Sex	Female	0.63 (0.39-0.86)	0.52 (0.16-0.88)	0.61 (0.00-1.22)	0.54 (0.30-0.78)	0.53 (0.28-0.78)	0.88 (0.80-0.96)
	Male	0.45 (0.16-0.74)	0.50 (0.15-0.85)	0.30 (0.00-0.89)	0.53 (0.29-0.77)	0.64 (0.42-0.86)	0.97 (0.96-0.99)
Age group (years)	35-59	0.56 (0.31-0.80)	0.49 (0.15-0.84)	0.26 (0.01-0.78)	0.57 (0.36-0.78)	0.61 (0.41-0.82)	0.90 (0.85-0.95)
	60-74	0.56 (0.24-0.88)	0.47 (0.07-0.86)	0.50 (0.00-1.19)	0.23 (0.00-0.62)	0.78 (0.58-0.97)	0.76 (0.28-1.24)
Schooling	Uncompleted High School	0.27 (0.00-0.88)	0.07 (0.00-0.67)	0.64 (0.01-1.36)	0.07 (0.00-0.89)	0.93 (0.70-1.06)	0.72 (0.24-1.20)
	Completed High School	0.71 (0.47-0.94)	0.56 (0.18-0.94)	0.30 (0.00-0.92)	0.41 (0.07-0.75)	0.49 (0.17-0.81)	0.89 (0.77-1.00)
	University	0.49 (0.22-0.76)	0.48 (0.11-0.84)	0.32 (0.00-0.89)	0.45 (0.22-0.69)	0.57 (0.33-0.80)	0.72 (0.24-1.20)

*VFT: Verbal Fluency Tests; **Trail B: Trail Making Test B.

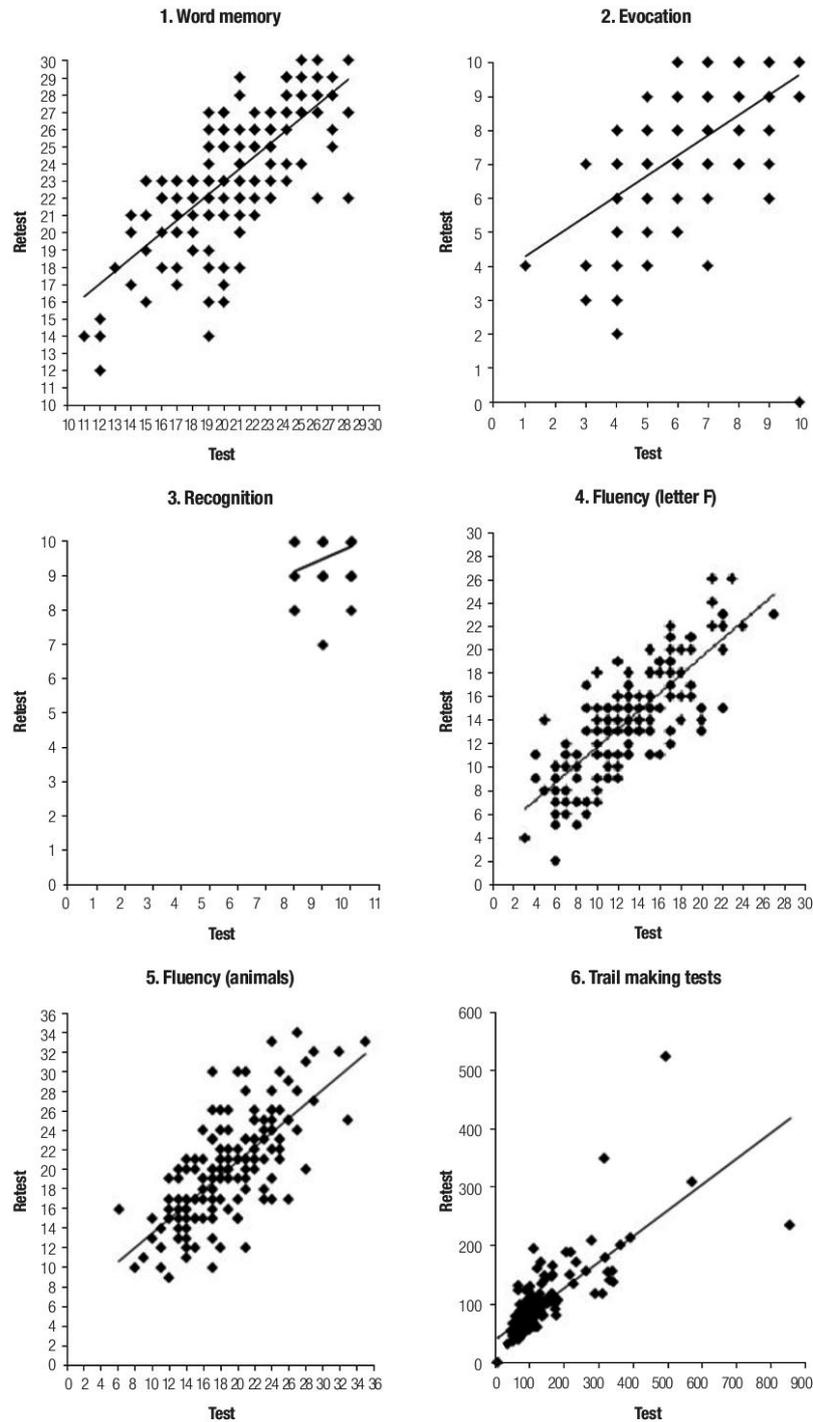


Figure 1. Test and retest dispersion graphs for battery of cognitive function tests among 160 participants of ELSA-Brasil.

the United States of America,⁵ which used samples of 20 and 278 people, respectively, aged under 50 years, and a one-month interval between test and retest. These same studies revealed substantial reliability for the recall test (Pearson coefficient=0.64). Lower reliability for the recognition test was also observed in the American study (Pearson=0.36), but not in the Korean investigation (Pearson=0.74).¹⁵ In our study, lower reliability for the recognition test may be explained by the ceiling effect, where the test values achieved by the sample are close to the maximum, reducing the variability between scores.

The better performance in the retests strongly suggests a learning effect, as observed in other studies.¹⁸ In the present study, we chose an interval of time similar to that adopted in other investigations, which ranged from two to four weeks. Longer periods between the tests could increase the probability of real changes in cognitive function, compromising the test reliability of the investigation; whilst shorter periods are more easily contaminated by the learning effect. As in other studies, in an attempt to avoid the learning effect, the retests were arranged in a different order. Despite these precautions, the influence of the learning effect may have contributed to decreasing the reliability of the tests.

Considering the Trail B showed almost perfect reliability, it may be useful when a short reapplication interval is necessary. High reliability for the Trail B test was also found in a German study, using a sample of 55 individuals, with a mean age of 46 years and, on average, 10 years of schooling.¹⁶

The studied tests presented the advantage of maintaining reliability regardless of sex, age and schooling. The test's capacity to be precise even when applied to

different people should not, however, lead to the conclusion that these variables have no effect when assessing the validity of these tests. There is evidence that these variables interfere with the capacity to distinguish between cognitive levels.^{9,19-20}

One limitation of this study is that it was conducted in one of the six ELSA research centers, as it was decided to reduce the variability of using different interviewers.

In conclusion, we observed moderate reliability for cognitive tests applied in adults, after a short interval averaging twenty days. The slight improvement in performance across all the retests, compared to the initial tests, suggests a learning effect. To avoid this effect, the ELSA-Brasil cognitive evaluation should use alternate equivalent versions of the test during study waves, estimated to be every three to four years, in order to reduce the influence of learning on prospective comparisons of cognitive tests in this Brazilian adult population.

Funding. This work was supported by a grant from the Ministry of Health and Ministry of Science and Technology (FINEP- Financiadora de Estudos e Projetos) for ELSA (n° 01 06 0278.00 MG). Prof. Barreto has a grant from the National Research Council of Brazil (CNPq, n° 01 06 0278.00), Prof. Passos has a grant from the State of Minas Gerais Agency for Research and Technology (FAPEMIG, n° 17767) and Prof. Giatti has a fellowship from the Coordination for the Improvement of Higher Education Personnel (CAPES).

Acknowledgments. The authors thank the ELSA-Brasil participants and the research team involved in the baseline study for their contribution to this study.

REFERENCES

1. Lezak MD, Howieson DB, Loring DW. Neuropsychological assessment. 4 ed. New York, NY: Oxford University Press; 2004.
2. Szklo M, Nieto FJ. Quality Assurance and Control. In: Sklo M, Nieto FJ (eds). Epidemiology beyond the basics. 2nd edition. Sudbury, MA, USA: Jones and Bartlett Publishers; 2007:297-348
3. Huley SB, Martin JN, Cummings SR. Planning the measurements: accuracy and precision In: Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB (eds). Designing clinical research. 3rd edition, Philadelphia, PA, USA. Lippincott Williams and Wilkins; 2006:55-67.
4. Aquino EM, Barreto SM, Benseñor IM, et al. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design. Am J Epidemiol 2012; 175:315-324.
5. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989; 39:1159-1165.
6. Fillenbaum GG, Belle G, Morris JC, et al. CERAD (Consortium to Establish a Registry for Alzheimer's disease): The first 20 years. Alzheimer's Dement 2008;4:96-109.
7. Brucki SMD, Malheiros SMF, Okamoto IH, Bertolucci PHF. Dados normativos para o teste de fluência verbal categoria animais em nosso meio. Arq Neuropsiquiatr 1997;55:56-61.
8. Brucki SMD, Rocha MSG. Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. Braz J Med Biol Res 2004;37:1771-1777.
9. Foss MP, Vale FAC, Special JG. Influência da escolaridade na avaliação neuropsicológica de idosos. Arq Neuropsiquiatr 2005;63:119-126.
10. Passos VMA, Giatti L, Barreto SM, et al. Verbal fluency tests reliability in a Brazilian multicentric study, ELSA-Brasil. Arq Neuropsiquiatr 2011; 69:814-816.
11. Dean AG, Armer TG, Sunki GG, et al. Epi Info™, a database and statistics program for public health professionals. Centres for Disease Control and Prevention, Atlanta, Georgia, USA; 2007.
12. Stata Statistical Software: Release 10. College Station, Texas: Stata Corporation; 2007.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

14. Harrison JE, Buxton P, Husain M, Wise R. Short test of semantic and phonological fluency: Normal performance, validity and test-retest reliability. *Brit J Clin Psychol* 2000;39:181-191.
15. Lee JH, Lee KU, Lee DY, et al. Development of the Korean Version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K): Clinical and Neuropsychological Assessment Batteries. *J Gerontol* 2002;57:47-53.
16. Wagner S, Helmreich I, Dahmen N, Lieb K, Tadic A. Reliability of Three Alternate Forms of the Trail Making Tests A and B. *Arch Clin Neuropsychol* 2011;26:314-321.
17. Bertolucci PHF, Okamoto IH, Neto JT, Ramos LR, Brucki SMD. Desempenho da população brasileira na bateria neuropsicológica do Consortium to Establish a Registry for Alzheimer's disease (CERAD). *Rev Psiquiatr Clin* 1998;25:80-83.
18. Salthouse TA. Selective review of cognitive aging. *J Int Neuropsychol Soc* 2010;16:754-760.
19. Charchat-Fichman H, Caramelli P, Sameshima K, Nitrini R. Declínio da capacidade cognitiva durante o envelhecimento. *Rev Bras Psiquiatr* 2005;27:79-82.
20. Christofolletti G, Olliani MM, Stella F, Gobbi S, Gobbi LTB. The influence of schooling on cognitive screening test in the elderly. *Dement Neuropsychol* 2007;1:46-51.