

Marta Isabel de Matos Gonçalves

Estudos de adaptação e validação da Bateria Computorizada de Testes Neuropsicológicos de Cambridge para avaliação da doença de Alzheimer

Tese de doutoramento em Psicologia, especialidade em Neuropsicologia,
orientada pela Professora Doutora Maria Salomé Pinho e pelo Professor Doutor Mário Rodrigues Simões
e apresentada à Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra

Junho de 2016



UNIVERSIDADE DE COIMBRA



UNIVERSIDADE DE COIMBRA

FACULDADE DE PSICOLOGIA E DE CIÊNCIAS DA EDUCAÇÃO

**Estudos de Adaptação e Validação da Bateria Computorizada de
Testes Neuropsicológicos de Cambridge
para Avaliação da Doença de Alzheimer**

FICHA TÉCNICA

Autora:

Marta Isabel de Matos Gonçalves

Título do trabalho:

Estudos de adaptação e validação da Bateria Computorizada de Testes Neuropsicológicos de Cambridge para avaliação da doença de Alzheimer.

Tese de doutoramento em Psicologia, especialidade em Neuropsicologia, apresentada à Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.

Trabalho orientado por:

Professora Doutora Maria Salomé Pinho

Professor Doutor Mário Rodrigues Simões

Imagem da capa:

Neuron [fotografia]. (2015). Retirada de

<http://legionofleia.files.wordpress.com/2015/03/neuron.jpg>

Data:

Junho de 2016

Financiamento:

Projecto financiado pela Fundação para a Ciência e a Tecnologia (Bolsa Individual de Doutoramento, referência SFRH/BD/75281/2010) ao abrigo do Programa Operacional Potencial Humano do Quadro de Referência Estratégico Nacional (formação avançada) e participado pelo Fundo Social Europeu e fundos nacionais do Ministério da Ciência, Tecnologia e Ensino Superior.



ERRATA

DEDICATÓRIA

Ao meu avô João, que em mim semeou a noção do valor e gosto pelo estudo. À minha avó Leopoldina, que ainda me ampara em memórias e sonhos. Ao meu avô Francisco e à minha avó Joaquina, cuja sabedoria me vai sendo transmitida pelas histórias que meus pais, madrinha e tios me contam. E a todas as pessoas mais velhas com quem me cruzei e me deixaram um pouco delas. De algum modo, todos me conduziram a este trabalho que lhes dedico.

AGRADECIMENTOS

As páginas que se seguem condensam quase oito anos de trabalho árduo dos quais cerca de dois anos e meio foram dedicados à recolha da amostra. Foram muitas as vezes que senti transportar um menir às costas e aquelas em que pensei desistir. Chegar a esta meta não foi fácil e são muitos aqueles que me ajudaram nesta jornada e a quem gostaria de prestar o meu mais profundo agradecimento.

Aos Directores, Presidentes, Comissões de Ética e demais Profissionais que generosamente me concederam autorização para realizar este estudo nas Organizações que dirigem ou representam, nomeadamente, as Instituições José de Mello Residências e Serviços, Centro de Apoio Social de Oeiras do Instituto de Acção Social das Forças Armadas, Casa dos Professores de Carcavelos da Associação de Solidariedade Social dos Professores, Centro Psicogeriátrico Nossa Senhora de Fátima das Irmãs Hospitaleiras do Sagrado Coração de Jesus, Residência Assistida da Delegação da Costa do Estoril da Cruz Vermelha Portuguesa, Inválidos do Comércio Residências e Lares, Casa de Repouso da Enfermagem Portuguesa, Corpuscare Casa de Repouso, Casa de Saúde do Telhal do Instituto São João de Deus, Centro de Dia Oeiras São Julião e Centro Social Paroquial de Oeiras. Aos Médicos, Enfermeiros, Psicólogos, Técnicos Superiores de Serviço Social, Fisioterapeutas, Terapeutas Ocupacionais e demais Profissionais destas Instituições que me dedicaram parte do seu tempo a responderem a entrevistas e a um inventário para cuidadores formais, que me orientaram pela Instituição, que me cederam salas com condições adequadas para administração de testes e que tanto me ajudaram no recrutamento de participantes voluntários. A todos os familiares dos participantes que igualmente colaboraram neste estudo, dando consentimento informado, ajudando a transportar o seu familiar à Instituição ou recebendo-me em casa e respondendo ao inventário para cuidadores informais. E muito em especial, a todos os participantes mais velhos que pacientemente e amigavelmente colaboraram neste estudo, dispondo-se a realizar duas a três sessões de avaliação e ajudando-me na divulgação e recrutamento de mais participantes voluntários. Sem eles, este estudo teria sido impossível de ser realizado. A todos, muito, muito grata!

Aos meus Orientadores, Professora Doutora Maria Salomé Pinho e Professor Doutor Mário Rodrigues Simões, por me terem acompanhado durante toda esta jornada, suportando as minhas dúvidas existenciais, apoiando-me mesmo em férias, noites e fins-

de-semana, acreditando nas minhas capacidades e incentivando-me a continuar, sobretudo nos momentos mais difíceis em que tanto necessitei desse apoio. Infinitamente grata!

Aos meus pais, ao meu irmão e à minha cunhada, que tantas vezes ouviram as minhas lamentações e que de perto reconheceram as minhas angústias, grata por todo o apoio, encorajamento e paciência, por todo o afecto e tolerância perante os meus longos silêncios e ausências. Aos meus dois sobrinhos que nasceram neste tempo, seguramente, o que melhor me aconteceu nestes oito anos, grata pela alegria com que me enchem por dentro e que de tantas formas me dá força e coragem para seguir em frente.

A todos os meus amigos que considero família, grata pelo apoio incondicional, pela paciência, por suportarem as minhas ausências e por me retirarem de casa de maneira a não esquecer que existem outros mundos, outros temas, outros assuntos e outras vivências e por de tantas maneiras me fazerem soltar gargalhadas.

Aos meus colegas de doutoramento, por todas as partilhas e entre-ajuda, por todo o encorajamento e solidariedade.

A todos, bem hajam!

EPIGRAFE

«The known is finite, the unknown infinite; intellectually, we stand on an islet in the midst of an illimitable ocean of inexplicability. Our business in every generation is to reclaim a little more land, to add something to the extent and the solidity of our possessions.» (Huxley, n.d.).

Estudos de Adaptação e Validação da Bateria Computorizada de Testes Neuropsicológicos de Cambridge para Avaliação da Doença de Alzheimer

Marta Isabel de Matos Gonçalves

RESUMO

O aumento da esperança média de vida tem sido acompanhado por maior prevalência de demências, sendo a demência secundária à doença de Alzheimer (DA) o tipo mais frequente. As demências são o principal preditor de admissão em instituições geriátricas. Estas instituições necessitam por isso de proceder a numerosas avaliações neuropsicológicas. Em Portugal, a maioria dos testes neuropsicológicos são de papel-e-lápis, carece de normas representativas de grupos etários acima dos 80 anos e de normas para pessoas mais velhas institucionalizadas. Os testes computorizados apresentam muitas vantagens comparativamente aos testes de papel-e-lápis, como uniformização na administração, maior acordo entre examinadores e menor dispêndio de tempo na cotação e em materiais de resposta. A Bateria Computorizada de Testes Neuropsicológicos de Cambridge (*Cambridge Neuropsychological Test Automated Battery*, CANTAB) é das mais utilizadas no mundo, reunindo muitos estudos de validade preditiva e poder discriminante. Quatro testes da CANTAB têm sido especialmente recomendados para a avaliação da DA: os testes Processamento Rápido de Informação Visual (*Rapid Visual Information Processing*, RVP), Aprendizagem de Pares Associados (*Paired Associates Learning*, PAL), Tempo de Reacção (*Reaction Time*, RTI) e Memória de Trabalho Espacial (*Spatial Working Memory*, SWM). A presente dissertação envolveu três estudos. O primeiro teve por objectivo analisar a fiabilidade teste-reteste destas provas e a sua susceptibilidade aos efeitos da prática, em 34 pessoas mais velhas sem problemas neuropsiquiátricos e a residir em instituições geriátricas. O segundo estudo teve por objectivo examinar a influência de variáveis sócio-demográficas (idade, escolaridade, género, experiência com computadores, consumo de psicotrópicos compatível com a prescrição de medicamentos potencialmente inapropriados e tempo de institucionalização) no desempenho nestes testes e definir normas em função das variáveis com maior influência, em 128 pessoas

mais velhas sem problemas neuropsiquiátricos e a viver em instituições geriátricas. No terceiro estudo examinou-se a validade de constructo destes testes (por análise das correlações intra-teste e por análise factorial) e a validade concorrente (por comparação dos resultados nestes testes com os obtidos em testes de papel-e-lápis), em 137 pessoas mais velhas sem problemas neuropsiquiátricos e 37 pessoas com demência devido a DA em estágio ligeiro a moderado, vivendo em instituições geriátricas e na comunidade. Os participantes foram recrutados em lares, residências, centros de dia e clínicas geriátricas e todos deram consentimento informado antes da sua participação. Cada participante realizou duas sessões, a primeira para análise dos critérios de inclusão nos grupos (incluindo uma entrevista semi-estruturada e a administração de testes de papel-e-lápis) e a segunda (ocorrida o mais tardar uma semana depois) para administração dos testes da CANTAB e mais alguns testes de papel-e-lápis. Os participantes que colaboraram no estudo da fiabilidade teste-reteste realizaram ainda uma terceira sessão (quatro semanas após a segunda sessão) para readministração da CANTAB. Foram ainda entrevistados o médico de cada participante e/ou um cuidador (in)formal, para recolha de dados clínicos e resposta a um inventário de capacidade funcional. Os quatro testes da CANTAB apresentaram medidas com boa a adequada fiabilidade teste-reteste. Foram identificados pequenos efeitos da prática em três medidas de resultados (testes PAL e SWM), tendo estes sido eficazmente corrigidos por dois métodos de estimação de equações baseadas na regressão padronizada para correcção dos efeitos da prática. Diferentes variáveis sócio-demográficas tiveram impacto nas pontuações em diferentes testes e medidas de um mesmo teste. No geral, os testes da CANTAB apresentaram boa a adequada validade convergente e saturaram em factores distintos (i.e., atenção selectiva e velocidade de processamento [RVP], memória episódica visuo-espacial [PAL], velocidade psicomotora [RTI] e memória de trabalho visuo-espacial e planeamento [SWM]). As correlações entre estes testes e os testes de papel-e-lápis foram moderadas a fracas. Com esta dissertação espera-se impulsionar a realização de mais estudos com a CANTAB em Portugal, bem como a sua utilização na população mais velha institucionalizada.

Palavras-chave: CANTAB; normas portuguesas; fiabilidade teste-reteste; validade de constructo; validade concorrente; doença de Alzheimer; pessoas mais velhas.

Adaptation and Validity Studies of the Cambridge Neuropsychological Test Automated Battery for the Assessment of Alzheimer's Disease

Marta Isabel de Matos Gonçalves

ABSTRACT

The increasing mean life expectancy has been accompanied by greater prevalence of dementias, with the dementia due to Alzheimer's disease (AD) being the most frequent dementia type. The dementias are the main predictor of entry into geriatric institutions. Therefore, these institutions have to proceed to numerous neuropsychological assessments. In Portugal, most of the neuropsychological tests are pencil-and-paper tests, their norms are not representative of groups above 80 years of age and few tests have norms for institutionalized older persons. The computerized tests offer many advantages over the pencil-and-paper tests such as standardized administration, greater inter-examiners agreement, and less expenditure of time and of materials for registering responses. The Cambridge Neuropsychological Test Automated Battery (CANTAB) is one of the most commonly worldwide used batteries, with many predictive validity and discriminant power studies having been conducted with the CANTAB. Four CANTAB tests have been specially recommended for the assessment of AD: the Rapid Visual Information Processing (RVP), Paired Associates Learning (PAL), Reaction Time (RTI), and Spatial Working Memory (SWM) tests. The current dissertation involved three studies. The first study aimed to examine the test-retest reliability of these CANTAB tests and their susceptibility to practice effects, in 34 older persons without neuropsychiatric diagnoses and living in geriatric institutions. The second study aimed to examine the influence of socio-demographic variables (age, education, gender, computer experience, psychotropic drug use compatible with the prescription of potentially inappropriate medications, and institutionalization time) on the performance on these tests and to set norms considering the variables with greater influence on the test measures, in 128 older persons without neuropsychiatric diagnoses and living in geriatric institutions. In the third study, the construct validity of these tests was examined (through intra-test correlations and factor analysis) as well as the concurrent validity (by comparison of these tests' results with those obtained on the pencil-and-

paper tests) in 137 older persons without neuropsychiatric problems and in 37 persons with mild to moderate AD dementia, living in geriatric institutions and in the community. The participants were recruited in retirement homes, adult day-care centers, and geriatric clinics and all of them gave informed consent before participation. Each participant was enrolled in two sessions, with the first session having been held to analyze the inclusion criteria in the groups (through a semi-structured interview and the administration of pencil-and-paper tests) and with the second session having been conducted (at the latest one week after the first session) for administration of the CANTAB tests and additional pencil-and-paper tests. The participants who collaborated in the test-retest reliability study were enrolled in a third session (about four weeks after the second session) to perform again the CANTAB tests. Each participant's physician and/or an informal or formal caregiver were also interviewed to collect clinical data and to complete a functional capacity inventory. The four CANTAB tests revealed measures with good-to-adequate test-retest reliability. Small practice effects were identified for three outcome measures (PAL and SWM tests), with these practice effects having been effectively corrected by two estimated standardized regression-based methods for practice effect correction. Different socio-demographic variables influenced the scores in different tests and measures of a same test. Overall, the CANTAB tests presented good-to-adequate convergent validity and loaded onto distinct factors (i.e., selective attention and processing speed [RVP], visuospatial episodic memory [PAL], psychomotor speed [RTI], and visuospatial working memory and planning [SWM]). The correlations between these tests and the pencil-and-paper tests were moderate-to-weak. This dissertation is expected to promote more studies with the CANTAB in Portugal, as well as to favor its use in the institutionalized older population.

Keywords: CANTAB; Portuguese norms; test-retest reliability; construct validity; concurrent validity; Alzheimer's disease; older persons.

SUMÁRIO

Introdução	19
A demência expressa em números	19
Importância da avaliação neuropsicológica da doença de Alzheimer e outras demências.....	20
<i>Avaliação da doença de Alzheimer</i>	<i>21</i>
Testes neuropsicológicos para avaliação da doença de Alzheimer e outras demências.....	25
<i>Problemas dos testes neuropsicológicos utilizados em Portugal</i>	<i>25</i>
<i>Importância da validação dos testes neuropsicológicos para a população residente em instituições geriátricas</i>	<i>29</i>
<i>Testes de papel-e-lápis e testes computadorizados.....</i>	<i>31</i>
A Bateria Computorizada de Testes Neuropsicológicos de Cambridge.....	34
<i>Bateria Computorizada de Testes Neuropsicológicos de Cambridge para avaliação da doença de Alzheimer</i>	<i>38</i>
Objectivos e estrutura do projecto de investigação e da presente dissertação	47
Referências.....	51
Apêndices.....	65

Capítulo 1. Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes	81
Abstract.....	81
Introduction.....	82
Method.....	86
<i>Participants</i>	<i>86</i>
<i>Design</i>	<i>88</i>
<i>Apparatus and materials</i>	<i>89</i>
<i>Procedure</i>	<i>92</i>
<i>Statistical analysis.....</i>	<i>92</i>
Results.....	94

<i>Test-retest reliability</i>	94
<i>Changes in test performance and practice effects</i>	96
Discussion	99
References	105

Capítulo 2. Effects of socio-demographic variables on performance on the Cambridge neuropsychological automated tests for the assessment of dementia and Portuguese norms for older adults living in retirement homes..... 111

Abstract	111
Introduction.....	112
<i>The influence of socio-demographic variables on CANTAB test performance</i>	114
Method	117
<i>Participants and procedure</i>	117
<i>Apparatus and materials</i>	121
<i>Statistical analysis</i>	123
Results.....	125
<i>Analyses of the effects of the socio-demographic variables on the CANTAB measures</i>	125
<i>Comparative analysis of the effects of the socio-demographic variables on the CANTAB and pencil-and-paper test measures</i>	128
<i>Portuguese normative values</i>	130
Discussion	132
<i>Effect of age</i>	133
<i>Effect of education</i>	134
<i>Effect of gender</i>	135
<i>Effect of computer experience</i>	135
<i>Effect of psychotropic drug use</i>	136
<i>Effect of institutionalization time</i>	137
<i>Study limitations, strengths, and future directions</i>	138
References.....	141
Appendices.....	147

Capítulo 3. Construct and concurrent validity of the Cambridge neuropsychological automated tests for the assessment of dementia in Portuguese older adults.....	161
Abstract.....	161
Introduction.....	162
<i>Correlations between the CANTAB tests and pencil-and-paper tests</i>	<i>163</i>
<i>Correlations between CANTAB tests</i>	<i>166</i>
<i>Factor analytic studies.....</i>	<i>166</i>
Method.....	168
<i>Participants and procedure.....</i>	<i>168</i>
<i>Apparatus and materials.....</i>	<i>173</i>
<i>Statistical analysis.....</i>	<i>176</i>
Results.....	177
<i>Correlational analysis.....</i>	<i>177</i>
<i>Factorial analysis.....</i>	<i>189</i>
Discussion.....	192
<i>Concurrent validity of the CANTAB tests in the group without neuropsychiatric diagnoses.....</i>	<i>192</i>
<i>Construct validity of the CANTAB tests in the group without neuropsychiatric diagnoses.....</i>	<i>194</i>
<i>Comparative analysis of the concurrent and construct validity of the CANTAB tests on a sample of patients with mild to moderate Alzheimer’s disease dementia.....</i>	<i>196</i>
<i>Study limitations, strengths, and future directions.....</i>	<i>199</i>
References.....	202
Discussão global.....	209
Síntese e reflexão integrativa dos resultados	209
Considerações práticas acerca da utilização dos testes computadorizados com pessoas mais velhas.....	216
<i>Problemas relacionados com as plataformas de resposta.....</i>	<i>217</i>
<i>Problemas relacionados com as características das provas</i>	<i>219</i>
<i>Problemas relacionados com a base de dados normativa original.....</i>	<i>220</i>

Limitações e vantagens deste projecto de investigação.....	221
Direcções futuras	224
Conclusão.....	226
Referências.....	227

Lista integral de referências	233
--	------------

INTRODUÇÃO

A DEMÊNCIA EXPRESSA EM NÚMEROS

Em Portugal, à semelhança do verificado em muitos países desenvolvidos (United Nations, 2015), a esperança média de vida tem aumentado, tendo-se registado em 2014 mais de 2 milhões de residentes com 65 ou mais anos de idade (i.e., 20.3% do total de residentes em Portugal; Instituto Nacional de Estatística, 2015a). Este aumento na esperança média de vida tem, porém, sido acompanhado por maior prevalência de demências.

A demência é uma condição clínica tipicamente de início tardio e caracterizada pelo desenvolvimento de vários défices cognitivos, a ponto de pôr em causa a autonomia da pessoa na realização de actividades de vida diária (AVDs) como a gestão financeira, a preparação de refeições ou mesmo a higiene pessoal (McKhann et al., 2011; Prince et al., 2013). A demência pode ter distintas patologias na origem, sendo por isso classificada em diferentes tipos (Santana, 2005a). Na população mais velha, os tipos mais frequentes são a demência decorrente da doença de Alzheimer (DA) e a demência vascular, nas formas pura ou mista (Jellinger, 2006; Prince et al., 2013). Estas e outras demências (e.g., demência secundária à doença de Parkinson [DP], demência com corpos de Lewy e demências fronto-temporais) são doenças neurodegenerativas crónicas de evolução lenta e progressiva (especialmente a demência devido a DA). Por conseguinte, elas comportam elevados custos a nível governamental e das famílias, os quais podem ser estimados em termos do número de consultas realizadas por vários técnicos especializados, fármacos administrados, tempo despendido pelo cuidador no apoio das AVDs e na supervisão da pessoa com demência (tempo em que o cuidador, no caso de ser trabalhador activo, fica impedido de exercer funções), serviços de apoio domiciliário e outros relacionados com a institucionalização (Wimo, Jönsson, Bond, Prince, & Winblad, 2013).

A nível mundial, a estimativa de pessoas a viver com demência em 2010 era de 35.6 milhões de indivíduos (com custos totais calculados na ordem dos 531 mil milhões de euros) sendo a expectativa a de quase duplicação deste número de pessoas a cada 20 anos (Prince et al., 2013; Wimo et al., 2013). Quanto à Europa, a prevalência média estimada para 2010 era de 6.9%, verificando-se um aumento da prevalência de

demência com o aumento do escalão etário, nomeadamente, de 1.6% no grupo dos 60 aos 64 anos a 43.1% no grupo dos 90 ou mais anos (Prince et al., 2013).

Em Portugal, a prevalência média de demência estimada para 2013 era de 5.9% (variando de 0.9% no grupo etário dos 60 aos 65 anos a 24.8% no grupo dos 85 ou mais anos), tendo os custos com a medicação sido calculados (para um número inferior de pacientes) na ordem dos 37 milhões de euros (Santana, Farinha, Freitas, Rodrigues, & Carvalho, 2015). Os dados nacionais relativos a 2013 revelaram ainda, para o escalão dos 65 ou mais anos de idade, 2001 óbitos por demência e 1670 por DA (respectivamente, 2.2% e 1.9% do total de óbitos assinalados para este grupo etário), quando em 2000 tinham sido reportados 616 óbitos por DA (i.e., menos de metade dos óbitos por DA registados em 2013) (Instituto Nacional de Estatística, 2001, 2015b).

Estes números são provavelmente subestimativas se considerarmos que muitas pessoas com demência ou com DA morrem por outras condições clínicas e que muitos casos encontram-se por diagnosticar (Santana et al., 2015). Não obstante, estas estatísticas permitem compreender que a prevalência de demência tem vindo a aumentar e que o cenário é preocupante. Por conseguinte, são urgentes todas as medidas que permitam uma avaliação precisa e um diagnóstico diferencial e precoce, pois a intervenção clínica é tanto mais útil quando mais cedo forem identificados os sintomas.

IMPORTÂNCIA DA AVALIAÇÃO NEUROPSICOLÓGICA DA DOENÇA DE ALZHEIMER E OUTRAS DEMÊNCIAS

Ao contrário de outras doenças cujo diagnóstico é possível apenas pelo exame físico do doente, o diagnóstico de demência requer o recurso a meios complementares de diagnóstico, outros informadores além do paciente (preferencialmente um cuidador próximo) e o parecer clínico de um médico neurologista e/ou psiquiatra.

Os meios complementares de diagnóstico mais comumente utilizados na avaliação da DA e outras demências são a avaliação neuropsicológica, os exames neurológicos e de imagem (e.g., electroencefalograma, tomografia axial computadorizada e ressonância magnética) e os exames laboratoriais (e.g., análises ao sangue, estudo genético e análise do líquido céfalo-raquidiano) (Guerreiro, 2005; Lezak, Howieson, Bigler, & Tranel, 2012; Sales, 2005; Santana, 2005b, 2005c; Santana, Baldeiras, & Oliveira, 2005; Santana & Oliveira, 2005; Santiago, 2005).

Os exames médicos e laboratoriais visam identificar possíveis causas da demência e detectar alterações histológicas, anatómicas e fisiológicas no sistema nervoso. Por sua vez, a avaliação neuropsicológica permite o estudo dos seguintes aspectos: (1) anamnese da pessoa examinada, sendo fundamental para identificar antecedentes familiares e constituir a história clínica do indivíduo em termos do modo de aparecimento e evolução dos sintomas; (2) estado psicológico, para identificação de possíveis sintomas comportamentais, do pensamento ou do humor e para rastreio de perturbações psicológicas que poderão prejudicar as funções cognitivas; (3) grau de capacidade funcional da pessoa para as AVDs; (4) nível de funcionamento pré-mórbido, para testar a hipótese das alterações cognitivas serem atribuíveis, por exemplo, a uma oligofrenia; e (5) funcionamento cognitivo da pessoa comparativamente ao seu funcionamento prévio ou aos processos de envelhecimento normal, para análise de possíveis alterações, declínios ou défices cognitivos, tipos de défices e gravidade dos mesmos, sendo os resultados interpretados de acordo com a relação que se conhece entre funções cognitivas e áreas cerebrais (Guerreiro, 2005; Lezak et al., 2012; Santana, 2005b; Santiago, 2005; Simões, 2012).

Avaliação da doença de Alzheimer

Considere-se o caso particular da DA. Até ao momento foram identificadas duas formas desta doença, designadas familiar e esporádica (Moreira & Oliveira, 2005; Santana, 2005a). A forma familiar inicia-se precocemente (i.e., antes dos 65 anos), é pouco frequente e manifesta-se em todos os portadores de mutações nos genes autossómicos dominantes da proteína precursora da beta-amilóide, da pré-senilina 1 ou da pré-senilina 2. O seu diagnóstico é definitivo e pode ser feito durante a vida da pessoa, por estudo genético. A forma esporádica apresenta início tardio (i.e., geralmente após os 65 anos), é a forma mais frequente e aquela a que geralmente se faz referência quando se considera a população mais velha. O diagnóstico definitivo da forma esporádica de DA só é actualmente possível *post mortem*, por estudo neuropatológico dos tecidos cerebrais; nomeadamente, pela identificação (sobretudo nas regiões límbica e temporoparietal do córtex cerebral) de placas senis (i.e., aglomerados de fragmentos de proteína beta-amilóide [β A] depositados no espaço extracelular) e tranças neurofibrilares (i.e., filamentos helicoidais de proteínas anómalas como a tau fosforilada, encontrados nos neurónios). O diagnóstico *in vivo* da forma esporádica de DA é ainda incerto e por isso é considerado um diagnóstico possível ou provável.

Pensa-se que vários factores possam contribuir para o desenvolvimento dos processos fisiopatológicos da forma esporádica de DA (e.g., a idade, factores de risco vascular, ser portador de um ou dois alelos $\epsilon 4$ no gene da apolipoproteína E; Jack et al., 2011; Moreira & Oliveira, 2005; Santana, 2005a) mas muitos destes mecanismos continuam por esclarecer. O modelo de progressão da DA que mais tem sido defendido, e no qual assentam os critérios de diagnóstico clínico (Albert et al., 2011; McKhann et al., 2011; Sperling et al., 2011), encontra-se exemplificado na Figura 1.

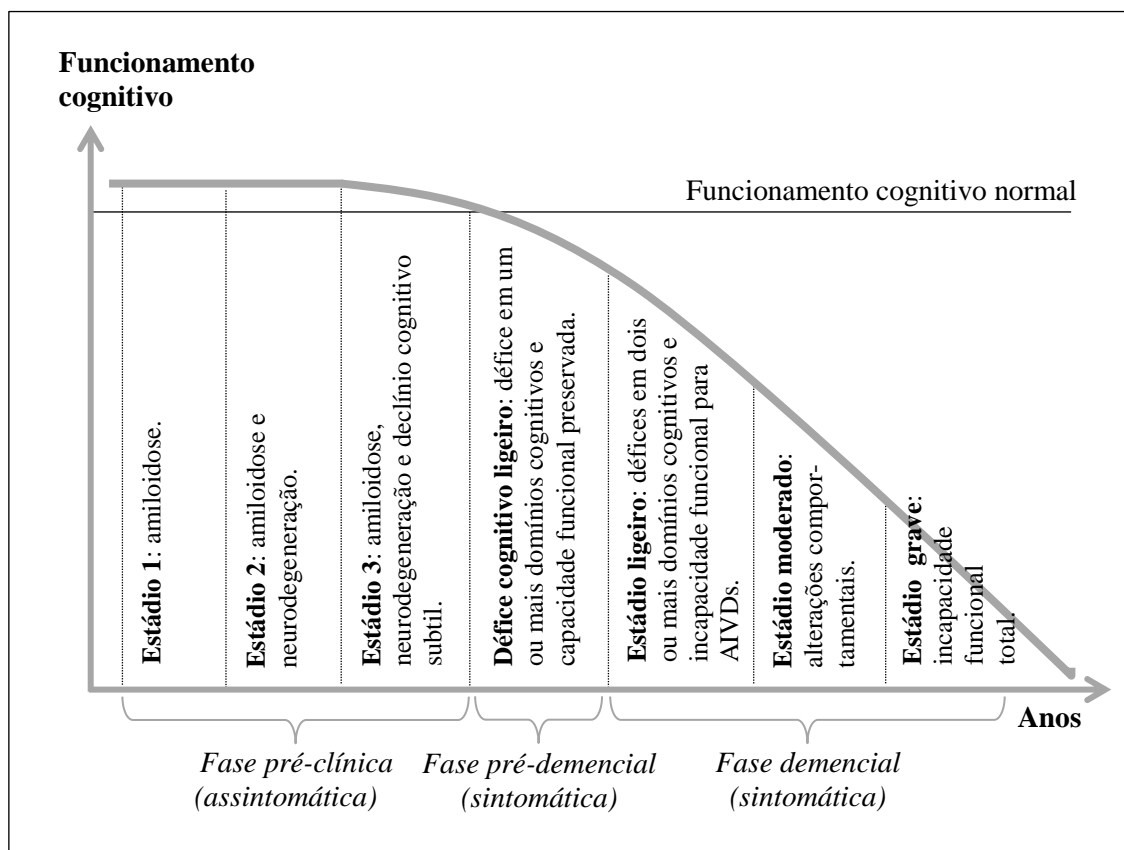


Figura 1. Modelo teórico de progressão da doença de Alzheimer. Legenda: AIVDs = actividades instrumentais de vida diária.

De acordo com este modelo, a DA inicia-se com a fase pré-clínica ou prodrómica ao longo da qual se desenvolvem processos fisiopatológicos e de neurodegeneração (especialmente na área do córtex entorrinal e do giro para-hipocampal; Khan et al., 2014), na ausência de sintomas clínicos (Sperling et al., 2011). O primeiro estágio desta fase caracteriza-se pela acumulação de βA . Este biomarcador é avaliado indirectamente por técnicas como a tomografia por emissão de positrões (PET) (uma técnica de neuroimagem requerendo a injeção no paciente de ligandos específicos) e a punção lombar (em que se infere a amiloidose cerebral pela presença de concentrações baixas

do peptídeo β A com 42 aminoácidos no líquido céfalo-raquidiano). No segundo estágio da fase pré-clínica, estas técnicas e outras (e.g., ressonância magnética funcional e estrutural) permitem inferir a ocorrência de processos de neurodegeneração (e.g., concentrações elevadas das proteínas tau e tau fosforilada no líquido céfalo-raquiano, neuroimagens sugestivas de disfunção ou perda neuronal, activação da glia e atrofia cortical nas regiões hipocampal e temporal média). Durante os dois primeiros estádios, ou mesmo antes, a avaliação neuropsicológica é importante para identificação da linha de base do funcionamento cognitivo da pessoa. Deste modo, no terceiro estágio da fase pré-clínica, a avaliação neuropsicológica poderá mais facilmente detectar declínios subtis comparativamente a essa linha de base, ainda que no terceiro estágio o funcionamento cognitivo continue a inserir-se nos parâmetros considerados normais para a idade e/ou escolaridade do indivíduo.

A fase seguinte da DA designa-se pré-demencial e corresponde ao défice cognitivo ligeiro (DCL) devido a DA (Albert et al., 2011). Em termos clínicos, esta é uma fase particularmente relevante, pois é quando surgem as primeiras manifestações clínicas da doença, como os défices em pelo menos um domínio cognitivo (e.g., memória episódica, capacidade visuo-espacial, linguagem, atenção e função executiva) na ausência de dificuldades quanto à capacidade funcional para as AVDs. Enquanto na fase pré-clínica os meios complementares de diagnóstico são utilizados somente para fins de investigação (pelo menos até à actualidade), na fase pré-demencial da DA estas técnicas são essenciais para fins clínicos. Assim, elas são importantes na identificação precoce dos sintomas, rastreio de causas possíveis e diagnóstico diferencial, de modo a que a intervenção clínica possa ser correctamente equacionada. No caso particular da avaliação neuropsicológica, ela é revelante na identificação dos défices cognitivos (geralmente deduzidos por resultados nos testes cognitivos inferiores à média normativa em mais de um desvio-padrão ou um desvio-padrão e meio), no rastreio de causas psicológicas e na análise do modo de início e tipo de sintomas e de défices. Esta informação é, portanto, de grande utilidade para o diagnóstico precoce e diferencial. Atendendo a que nem todos os casos com DCL progridem para demência e que muitos outros factores podem contribuir para baixos resultados num teste, é desejável que a classificação de DCL seja estabelecida com base em avaliações repetidas que demonstrem a consistência dos défices no tempo (Darby, Maruff, Collie, & McStephen, 2002; De Jager & Budge, 2005; Mitchell, Arnold, Dawson, Nestor, & Hodges, 2009; Summers & Saunders, 2012). A avaliação neuropsicológica é importante também neste

processo de monitorização. Actualmente, o diagnóstico de DCL devido a DA é considerado altamente provável quando os critérios para o seu diagnóstico são corroborados pelos resultados da avaliação neuropsicológica e pelos exames médicos e laboratoriais (Albert et al., 2011).

A fase demencial da DA caracteriza-se pela presença de défices em pelo menos dois domínios cognitivos (aumentando o número e a gravidade dos défices ao longo do tempo), incapacidade funcional para as AVDs (inicialmente ao nível das actividades mais complexas mas, mais tarde, também ao nível das actividades básicas) e sintomas psicológicos ou comportamentais (sobretudo a partir do estágio moderado) (McKhann et al., 2011). O diagnóstico de demência devido a DA é considerado provável quando os critérios são corroborados pelos resultados da avaliação neuropsicológica e por exames médicos e laboratoriais, não coexistindo outras causas ou doenças primárias com a demência. No caso de coexistirem outras causas ou a evolução ser atípica (e.g., início súbito), o diagnóstico de demência devido a DA considera-se possível. Na fase demencial, a DA torna-se mais notória e, caso ainda não tenha sido diagnosticada, é importante que o seja nos estádios ligeiro a moderado, para maior eficácia da intervenção clínica. A avaliação neuropsicológica assume-se como relevante na fase demencial da DA por motivos semelhantes aos indicados para a fase pré-demencial, sendo que na fase demencial se destaca o seu papel ao nível da monitorização dos défices cognitivos, análise da gravidade dos sintomas e avaliação da eficácia da intervenção clínica e/ou farmacológica.

Ao contrário da maioria dos exames médicos referidos, a avaliação neuropsicológica é um meio complementar de diagnóstico não invasivo, envolvendo métodos de avaliação como a entrevista, a observação comportamental e a utilização de instrumentos de avaliação (e.g., questionários, inventários e testes neuropsicológicos) (Simões, 2012). Comparativamente a alguns exames médicos, a avaliação neuropsicológica é também menos dispendiosa e não comporta riscos para a saúde. No entanto, ela não está isenta de problemas e erros de medida, à semelhança das outras técnicas. Assim, por exemplo, dois laboratórios podem chegar a resultados discrepantes para uma mesma colheita de líquido céfalo-raquidiano, ao utilizarem diferentes valores de referência (Albert et al., 2011). Do mesmo modo, os resultados de um indivíduo num teste neuropsicológico podem ser interpretados de modo discrepante por neuropsicólogos que utilizem diferentes pontos de corte. Um indivíduo também pode

obter resultados discrepantes em dois testes de memória episódica, tal como distintas técnicas de neuroimagem podem conduzir a conclusões divergentes.

Além destes problemas, também é frequente obter-se informação conflitante entre meios complementares de diagnóstico. Por exemplo, a PET e a punção lombar podem conduzir a resultados distintos quanto a possível amiloidose, tal como a PET, a punção lombar e a ressonância magnética estrutural podem indicar resultados discordantes quanto à presença de processos de neurodegeneração (McKhann et al., 2011). Por sua vez, a avaliação neuropsicológica pode identificar declínios ou défices cognitivos na ausência de alterações observáveis em técnicas de neuroimagem (Lezak et al., 2012). O diagnóstico clínico da DA e outras demências é tanto mais rigoroso quantas mais técnicas forem utilizadas. Os avanços na identificação dos biomarcadores da DA (Albert et al., 2011; Jack et al., 2011; McKhann et al., 2011; Sperling et al., 2011) ainda não permitiram solucionar o problema de tais biomarcadores serem comuns a outras demências e serem observados em pessoas que não chegam a desenvolver DCL ou demência. Por conseguinte, a avaliação neuropsicológica continua a ser um meio complementar de diagnóstico essencial na avaliação da DA e outras demências, não excluindo nem devendo ser excluída pelas restantes técnicas.

Atendendo à importância dos meios complementares de diagnóstico na avaliação da DA e outras demências, é fundamental investir no aperfeiçoamento destas técnicas. No caso particular da avaliação neuropsicológica, tal envolve o estudo e melhoria dos meios de avaliação e das propriedades psicométricas dos instrumentos utilizados, bem como a análise da adequação da aplicabilidade desses meios e instrumentos de avaliação a diferentes contextos e grupos populacionais.

TESTES NEUROPSICOLÓGICOS PARA AVALIAÇÃO DA DOENÇA DE ALZHEIMER E OUTRAS DEMÊNCIAS

Problemas dos testes neuropsicológicos utilizados em Portugal

A diversidade de testes neuropsicológicos existentes no mundo para avaliação da DA e outras demências é considerável (vidé Burns, Lawlor, & Craig, 2004; Lezak et al., 2012; Strauss, Sherman, & Spreen, 2006). Em Portugal também é possível identificar um aumento considerável do número de instrumentos disponíveis (ver Guerreiro, 2005; Maia, Correia, & Leite, 2009; Santana, 2005b; Simões, 2012; Simões & Santana, 2015).

Contudo, muitos destes instrumentos carecem de estudos sistemáticos de aferição e validação para a população portuguesa.

No caso dos instrumentos que dispõem de valores normativos para Portugal, é frequente apresentarem normas já datadas ou mesmo não publicadas (e.g., Bateria de Lisboa para Avaliação das Demências; Guerreiro, 1998), as normas terem sido definidas com base em resultados de pequenas amostras (e.g., Teste da Figura Complexa de Rey; Espírito-Santo et al., 2015) ou para grupos pouco representativos de idades acima dos 80 anos (e.g., Avaliação Cognitiva de Montreal; Freitas, Simões, Alves, & Santana, 2011), a amostra pertencer a uma área geográfica não representativa de todas as regiões de Portugal (e.g., Exame Breve do Estado Mental; Morgado, Rocha, Maruta, Guerreiro, & Martins, 2009) ou incluir somente indivíduos a viver na comunidade e os testes apresentarem normas considerando uma única variável demográfica (e.g., idade), quando outras variáveis sócio-demográficas poderão apresentar tanta ou maior influência no desempenho nos testes (e.g., Escala de Inteligência de Wechsler para Adultos – III edição [ou *Wechsler Adult Intelligence Scale-III*, WAIS-III] e Escala de Memória de Wechsler – III edição [ou *Wechsler Memory Scale-III*, WMS-III]; Wechsler, 2008a, 2008b, 2008c).

Além dos problemas relacionados com as normas e as amostras (Simões, 2012), vários testes neuropsicológicos carecem também de um conjunto aceitável de estudos quanto às suas propriedades psicométricas (i.e., fiabilidade, validade e poder discriminante das pontuações).

Os estudos de fiabilidade têm por objectivo analisar a confiabilidade do teste, sendo esta tanto maior quanto menor for o erro da medida. Fazem parte deste grupo de estudos os seguintes: a análise da consistência interna (i.e., as correlações entre itens e o total do teste), a fiabilidade inter-examinadores (i.e., o acordo entre avaliadores quanto aos resultados no teste e sua interpretação); a fiabilidade teste-reteste (i.e., a consistência dos resultados em avaliações repetidas ao longo do tempo) e que inclui a análise da susceptibilidade do teste aos efeitos da prática e aos efeitos de chão e de tecto (i.e., obter, respectivamente, o resultado mínimo ou máximo possível no teste); e a fiabilidade de formas paralelas (i.e., a correlação entre formas alternativas de um teste) (American Psychological Association, 2015; Anastasi, 1988; Kline, 1986).

Como mencionado anteriormente, a avaliação da DA e outras demências requer avaliações repetidas ao longo do tempo. Porém, a maioria dos testes neuropsicológicos utilizados em Portugal dispõe de uma forma única de teste ou não mais do que duas

formas paralelas. Quando se utilizam formas únicas, o risco de ocorrerem efeitos da prática é maior (Calamia, Markon, & Tranel, 2013). Um exemplo comum é o Exame Breve do Estado Mental (*Mini-Mental State Examination*, MMSE; Folstein, Folstein, & McHugh, 1975), um instrumento de rastreio cognitivo que, por ser tão simples, é facilmente aprendido. Quando utilizado em avaliações repetidas (e.g., de seis em seis meses) é frequente os indivíduos responderem mesmo antes de se terminar uma questão. Simultaneamente, é um instrumento com elevados efeitos de tecto. Tanto os efeitos da prática como os efeitos de tecto aumentam o risco de falsos negativos, tal como os efeitos de chão aumentam o risco de falsos positivos. Por conseguinte, é necessário conhecer a fiabilidade teste-reteste de cada instrumento (de preferência, analisando mais de dois momentos de avaliação) e a sua susceptibilidade a efeitos da prática, de chão e de tecto para que possam ser adoptados procedimentos que permitam uma avaliação mais objectiva (Benedict & Zgaljardic, 1998; Calamia et al., 2013; Collie, Maruff, Darby, & McStephen, 2003; Falleti, Maruff, Collie, & Darby, 2006; Lowe & Rabbitt, 1998). São exemplos desses procedimentos o estabelecimento de normas para avaliações repetidas, os métodos de correcção dos efeitos da prática ou o desenvolvimento de formas de teste mais extensas que incluam itens de baixa a elevada dificuldade e critérios de interrupção ou retrocesso, para ajustamento da dificuldade do teste à capacidade cognitiva da pessoa examinada (Attix et al., 2009; Bruggemans, van de Vijver, & Huysmans, 1997; Crawford & Garthwaite, 2007; Cysique et al., 2011; Duff, 2012; Iverson, 2001; McSweeny, Naugle, Chelune, & Lüders, 1993; Wechsler, 2008a).

Quanto aos estudos de validade, estes têm por objectivo analisar até que ponto um teste mede o que é suposto medir ou até que ponto existe evidência empírica ou modelos teóricos a sustentar a adequação das conclusões derivadas desse teste (American Psychological Association, 2015; Anastasi, 1988; Kline, 1986). Para que um teste seja considerado válido são necessários estudos de diferentes tipos de validade. Assim, por exemplo, os estudos de validade de constructo procuram analisar até que ponto um teste mede um determinado constructo e não outro, recorrendo com frequência à análise factorial. Este procedimento estatístico exige amostras de grande dimensão (e.g., Marôco, 2014a, 2014b), pelo que muitos instrumentos não dispõem ainda de estudos sobre a validade de constructo em amostras portuguesas compostas por indivíduos saudáveis ou por pessoas com condições clínicas como DA ou outras demências. Já os estudos de validade relativa a um critério examinam a correlação entre

o resultado no teste e um determinado critério, sendo exemplos os estudos de validade concorrente, validade preditiva e validade retrospectiva. A validade concorrente é a mais frequentemente avaliada, nomeadamente, pela análise da correlação do resultado no teste com os resultados em outros testes medindo o mesmo constructo e administrados no mesmo momento. Os estudos de validade preditiva e retrospectiva são mais raros pois requerem que o resultado no teste seja correlacionado com uma variável que irá ser medida algum tempo depois ou que foi medida algum tempo antes. No caso da DA e outras demências, este período de tempo pode envolver anos. Se, por exemplo, se pretender examinar a correlação do resultado num teste de deterioração cognitiva com o critério “classificação quanto a demência” num grupo de pessoas com DA prodrómica, o intervalo de espera até avaliar o critério poderá ser superior a quatro anos, o que é impraticável em projectos de investigação de curta duração. Todavia, os estudos de validade preditiva são de grande importância para a utilização futura de tais testes neuropsicológicos na avaliação das fases pré-clínica e pré-demencial da DA.

A análise do poder discriminante de um teste tem por objectivo verificar se este permite diferenciar o desempenho entre grupos (American Psychological Association, 2015). Este tipo de estudos é fundamental para o diagnóstico diferencial e precoce da DA e outras demências. Todavia, muitos dos testes neuropsicológicos utilizados em Portugal ainda não possuem estudos comparando o desempenho entre grupos saudáveis, com DA prodrómica, com DCL devido a DA, com demência devido a DA, com DCL devido a outras causas que não DA ou com outros tipos de demência (Simões, 2012). Por exemplo, apesar da WMS-III ser utilizada há vários anos, só muito recentemente passou a dispôr de um estudo de análise do seu poder discriminante para diferenciar os desempenhos de indivíduos saudáveis, com demência devido a DA e com demência vascular subcortical (C. Gonçalves et al., 2016), carecendo ainda de estudos analisando as outras condições clínicas citadas. O efeito da ausência deste tipo de estudos é desconhecer-se a sensibilidade e especificidade dos testes para tais condições clínicas, entendendo-se por sensibilidade a probabilidade do teste detectar uma condição (e.g., défice cognitivo) em alguém com essa condição e por especificidade a probabilidade do teste indicar ausência dessa condição em alguém em que a condição esteja ausente. Consequentemente, desconhecem-se também os pontos de corte desses testes (definidos em termos do resultado correspondendo à maior sensibilidade e especificidade possíveis) para distinguir os desempenhos, por exemplo, de um indivíduo com DCL do de outro com demência devido a DA.

Todos estes problemas indicam que há ainda muito trabalho a fazer em Portugal em termos de investigação em testes neuropsicológicos. Esse trabalho é necessário para maior objectividade da avaliação neuropsicológica da DA e outras demências, repercutindo-se em maior rigor na pesquisa destas condições clínicas em neurociências e na avaliação da eficácia dos tratamentos propostos.

Importância da validação dos testes neuropsicológicos para a população residente em instituições geriátricas

O facto da maioria dos testes neuropsicológicos apresentarem normas baseadas no desempenho de pessoas mais velhas portuguesas a residir na comunidade e com idades geralmente inferiores aos 80 anos gera questões quanto à adequação destas normas para residentes em instituições geriátricas.

Sabe-se que as pessoas com demência apresentam maior probabilidade de serem institucionalizadas e que a demência se associa a internamentos de longa duração (Nihtilä et al., 2007). Nos Estados Unidos, a deterioração cognitiva e a incapacidade funcional para as AVDs (duas das características da demência) são os maiores preditores de admissão em instituições geriátricas, sendo mais relevantes do que a idade (Gaugler, Duval, Anderson, & Kane, 2007). Naturalmente que nem todas as pessoas admitidas em instituições geriátricas (i.e., lares e residências séniores) apresentam demência, incapacidade funcional ou outros problemas de saúde. Frequentemente, estas instituições incluem instalações de vida independente dirigidas a pessoas mais velhas sem problemas neuropsiquiátricos. Para estas pessoas, a instituição é uma opção quando pretendem, por exemplo, viver sem preocupações domésticas, junto de um cônjuge dependente a necessitar de cuidados especializados ou mais perto de outros familiares ou amigos, de modo a evitar o isolamento ou garantir cuidados de apoio no futuro. Como estas pessoas tendem a residir nas instituições até ao final das suas vidas, é comum algumas desenvolverem demência durante a sua permanência, se ainda não a apresentavam no momento da admissão. Por conseguinte, as instituições geriátricas têm a necessidade imperiosa de proceder a avaliações neuropsicológicas repetidas, de modo a detectar precocemente declínios ou défices cognitivos e monitorizar a sua evolução, com vista à implementação de terapêuticas ou intervenções clínicas mais eficazes e ajustadas à gravidade e tipo de sintomas.

Perante a escassez de normas para pessoas mais velhas a viver em lares e residências geriátricas, estas instituições têm de fazer uso das normas existentes para

indivíduos da comunidade. Um dos problemas desta comparação reside na diferença de idades entre os que vivem em instituições geriátricas e os que compõem as amostras normativas de indivíduos da comunidade. Segundo os dados do Instituto Nacional de Estatística (2012a), do total de pessoas que em 2011 residiam em instituições de apoio social em Portugal, 79% tinham 65 ou mais anos de idade, sendo que a maioria dos residentes apresentavam idades acima dos 80 anos (ver Figura 2). Assim, o desempenho de pessoas com 80 ou mais anos de idade a residirem em instituições geriátricas é frequentemente comparado ao de um grupo de indivíduos da comunidade composto por um número pouco representativo ou mesmo sem pessoas com essas idades.

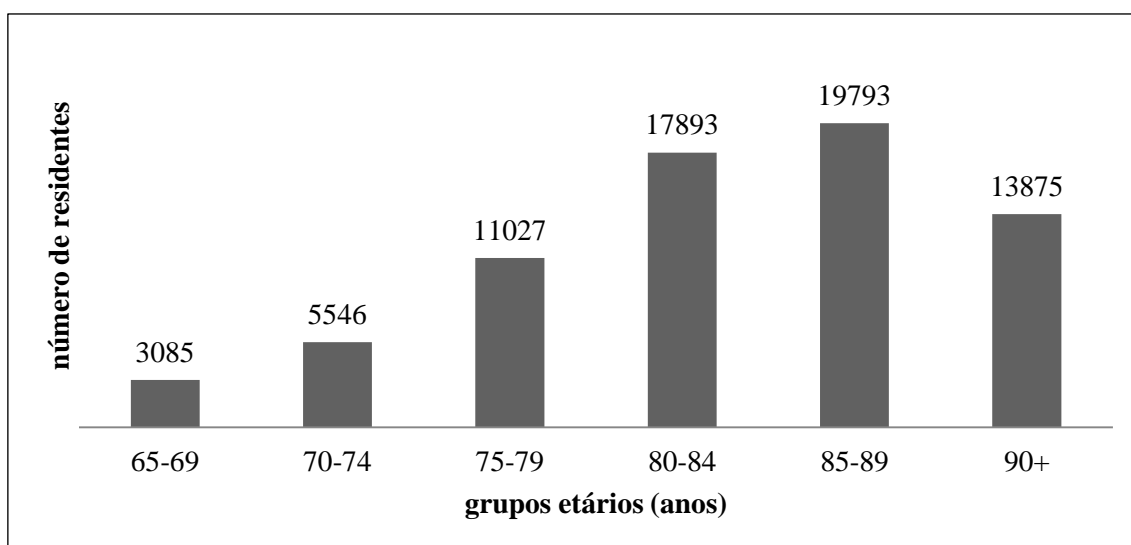


Figura 2. Número de pessoas mais velhas a residir em instituições de apoio social em 2011, em Portugal (Instituto Nacional de Estatística, 2012a).

Outro problema em aplicar normas para indivíduos da comunidade a residentes em instituições geriátricas é o facto destes grupos populacionais viverem em contextos diferentes. Assim como o desempenho num teste tende a variar de país para país e entre áreas geográficas distintas (e.g., rurais e urbanas), também é de esperar que difira entre pessoas a viver na comunidade e em instituições. Não sendo a vida numa instituição geriátrica necessariamente melhor ou pior do que a vida na comunidade, a cultura ambiental é diferente (Kahana, 2006). Nas instituições geriátricas a diversidade etária é menor, existem regras e rotinas a cumprir, o acesso a apoio técnico especializado é mais fácil e, em muitos casos, as pessoas dispõem de equipamentos e actividades recreativas. Atendendo a que as instalações requerem frequentemente a partilha de espaços (e.g., quartos, refeitórios, salas comuns), o convívio social também pode ser mais frequente. Por conseguinte, é possível que as diferenças entre a cultura ambiental das instituições e

da comunidade se reflectam em diferentes tipos de estimulação cognitiva, problemas quotidianos e estratégias de confronto utilizadas, com impacto a nível comportamental, emocional e cognitivo. Neste sentido, alguns estudos têm efectivamente revelado diferenças de desempenho nos testes entre pessoas mais velhas institucionalizadas e a residir na comunidade (Kahana, 2006; Plati, Covre, Lukasova, & Macedo, 2006; Rapp, Schnaider-Beerli, Sano, Silverman, & Haroutunian, 2005).

Deste modo, comparar o desempenho de pessoas mais velhas institucionalizadas com o de pessoas da comunidade e mais jovens pode comportar o risco de se cometerem erros graves de avaliação. Por isso, a realização de estudos normativos e de validação de testes neuropsicológicos para a população mais velha institucionalizada é uma necessidade urgente em Portugal.

Testes de papel-e-lápis e testes computadorizados

A maioria dos testes administrados à população mais velha portuguesa é do tipo de papel-e-lápis. No entanto, os testes computadorizados já existem há algumas décadas (Kline, 1986; Larrabee & Crook, 1996). Comparativamente aos testes de papel-e-lápis, os testes automatizados apresentam as seguintes vantagens: (1) a administração, o registo e a cotação das respostas são automáticos, garantindo procedimentos estandardizados que diminuem a probabilidade de se cometerem erros e de ocorrerem divergências entre avaliações conduzidas por um examinador ou entre examinadores; (2) os estímulos são apresentados de modo dinâmico, o que ajuda a manter a pessoa examinada atenta e motivada para a tarefa; (3) é possível registar vários aspectos do desempenho além dos acertos e erros, como a latência de resposta a qual é medida com uma precisão de milissegundos; (4) podem ser geradas múltiplas formas alternativas de teste de modo a reduzir os efeitos da prática, pelo que são testes mais adequados para avaliações repetidas; (5) é mais fácil ajustar a dificuldade do teste à capacidade da pessoa examinada, de modo a reduzir os efeitos de chão e de tecto; (6) os resultados estão disponíveis imediatamente após a administração das provas, o que representa uma enorme poupança de tempo; (7) é possível armazenar uma quantidade elevada de dados, sem consumo de papel nem necessidade de espaço para arquivos; (8) podem ser gravadas cópias dos resultados em vários sistemas de armazenamento e definidas palavras-chave para acesso aos programas, o que se traduz em medidas de maior segurança; (9) múltiplos testes e resultados são facilmente transportados num *tablet*, o que permite administrar as provas mesmo a pessoas acamadas e rapidamente comparar

o desempenho actual com resultados anteriores; e (10) os materiais de teste são menos susceptíveis a desgaste e podem ser mais facilmente substituídos ou actualizados sem que para isso seja necessário comprar todo o material (Cernich, Brennana, Barker, & Bleiberg, 2007; Collerton et al., 2007; Collie et al., 2003; Elwood, 2001; Gualtieri & Johnson, 2005; Kline, 1986; Mead & Drasgow, 1993; Wild, Howieson, Webbe, Seelye, & Kaye, 2008; Zygouris & Tsolaki, 2014).

Apesar destas vantagens, pouco se tem investido na adaptação e validação de testes computadorizados para a população mais velha portuguesa. É possível que este atraso, comparativamente a outros países, esteja relacionado com o custo elevado destes instrumentos. De facto, os testes computadorizados ainda hoje são caros, se bem que muito menos do que na época do seu aparecimento (Gualtieri & Johnson, 2005; Mead & Drasgow, 1993). Por outro lado, se considerarmos os custos que os testes de papel-e-lápis envolvem a longo prazo (e.g., folhas de resposta, grelhas de cotação, aquisição de conjuntos completos de material de teste para substituir os danificados ou para obtenção de novas versões), então talvez os testes automatizados sejam mais económicos.

Outro motivo que poderá justificar alguma resistência na administração de testes computadorizados a pessoas mais velhas é o receio de clínicos e investigadores de que a tecnologia seja mal aceite pelos mais velhos ou prejudique o seu desempenho devido à falta de experiência na utilização de computadores. Curiosamente, existem estudos empíricos indicando que as pessoas mais velhas toleram melhor os testes automatizados do que os testes de papel-e-lápis (Collerton et al., 2007). Do mesmo modo, tem-se registado em Portugal um aumento progressivo no número de utilizadores de computador pertencentes à faixa etária dos 65 aos 74 anos (ver Figura 3) (Instituto Nacional de Estatística, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012b, 2013, 2014, 2015c).

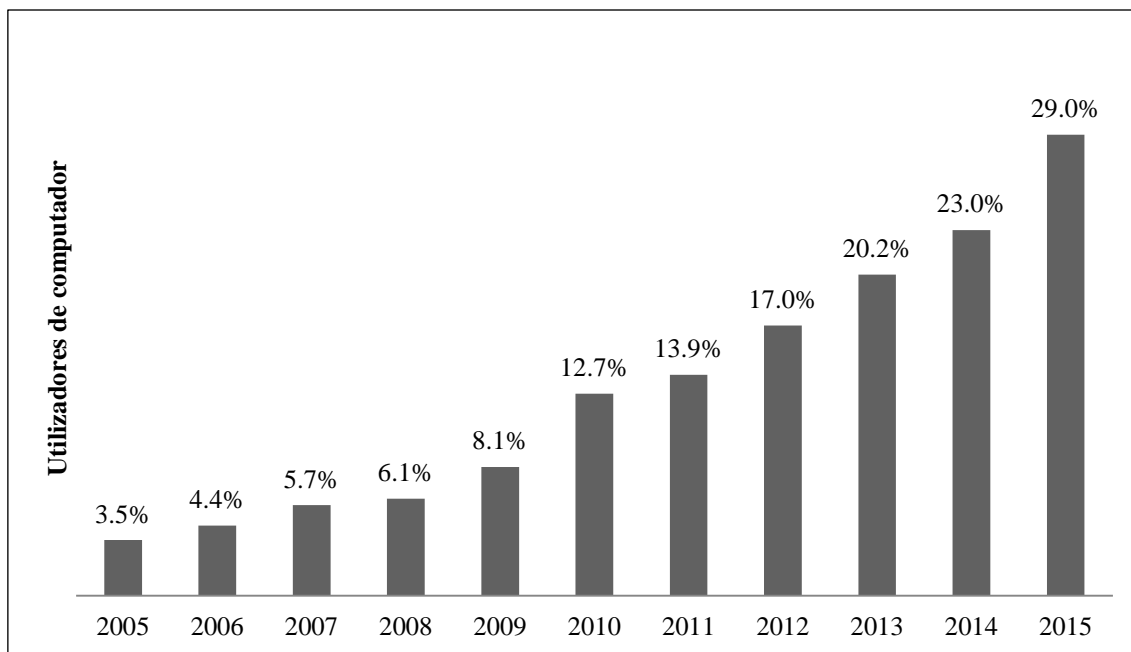


Figura 3. Percentagem de residentes em Portugal com 65 a 74 anos de idade que utilizam computador (Instituto Nacional de Estatística, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012b, 2013, 2014, 2015c).

É inquestionável que nos últimos tempos a tecnologia tem dominado a maioria dos sectores empresariais, de comunicação, lazer e atendimento ao público. Por muito que se queira evitá-la, é quase impossível fazê-lo. Os telemóveis e a Internet têm permitido estabelecer comunicações mais frequentes e rápidas. Muitos idosos encontram vantagens em aprender a dominar estas tecnologias, como forma de se manterem em contacto com familiares distantes ou mesmo pedir auxílio em qualquer lugar onde estejam. A tecnologia *touch screen* está hoje presente em balcões de atendimento ao público para obtenção de senhas, nos postos multibanco para a realização de operações e nos próprios telemóveis. Sendo uma tecnologia muito intuitiva (e menos exigente em termos de coordenação motora comparativamente à escrita), torna-se fácil de aprender pelos mais velhos. Deste modo, existem hoje poucos motivos para crer que uma pessoa mais velha se recuse a responder a um teste devido ao facto deste ser apresentado em suporte informático. Além disso, os jovens de hoje são os mais velhos de amanhã sendo que poucos jovens estão habituados à escrita manual. Todos estes aspectos sugerem que os testes de papel-e-lápis venham a ser substituídos pelos testes computadorizados num futuro próximo, pelo que se justifica o estudo deste tipo de instrumentos de avaliação a fim de poderem ser usados na prática clínica.

A BATERIA COMPUTORIZADA DE TESTES NEUROPSICOLÓGICOS DE CAMBRIDGE

A Bateria Computorizada de Testes Neuropsicológicos de Cambridge (*Cambridge Neuropsychological Test Automated Battery*, CANTAB) foi inicialmente desenvolvida na Universidade de Cambridge em 1987, com a finalidade de ser aplicada ao estudo de demências como as secundárias à DA e à DP (Morris, Evenden, Sahakian, & Robbins, 1987). Nessa altura, a CANTAB era composta por quatro testes mas, desde então, novas provas têm vindo a ser adicionadas à bateria. A CANTAB é comercializada pela Cambridge Cognition (<http://www.cambridgecognition.com>) desde 1994 e desde 2002 que esta empresa é responsável pelo contínuo aperfeiçoamento dos testes.

Uma das particularidades de muitos dos testes da CANTAB é o facto de exibirem estímulos não verbais e abstractos e requererem respostas motoras, de modo a reduzir a interferência de factores culturais e linguísticos e torná-los mais fáceis de adaptar e utilizar em outros países (Sahakian & Owen, 1992). Estes testes foram assim construídos com o objectivo de universalizar as práticas neuropsicológicas e permitir comparar os resultados entre estudos internacionais. A universalização de práticas permitiria ainda combater a dificuldade que ocorre ao compararem-se os resultados de estudos com diferentes baterias de avaliação da demência, nomeadamente, a de saber se determinados resultados são atribuíveis à demência ou às diferentes características e propriedades psicométricas dos testes aplicados.

Uma outra singularidade dos testes da CANTAB é que a maioria baseia-se em tarefas e modelos da neuropsicologia animal com utilidade comprovada na compreensão dos processos e bases neuronais da demência (Bartko, Vendrell, Saksida, & Bussey, 2010; Fray, Robbins, & Sahakian, 1996; Robbins et al., 1994; Sahakian & Owen, 1992; Taffe, Weed, Gutierrez, Davis, & Gold, 2002; Talpos, Winters, Dias, Saksida, & Bussey, 2009). Os modelos animais permitem compreender com maior precisão os componentes nos quais se dividem os processos relacionados com determinadas funções cognitivas. O facto de muitos dos testes da CANTAB se basearem nestes modelos facilita a comparação entre estudos com humanos e animais e permite uma interpretação mais detalhada dos processos em análise, apoiada em fundamentação teórica robusta.

Devido a estas características dos testes da CANTAB, eles têm sido muito usados na investigação académica e em ensaios clínicos farmacêuticos. Nos últimos anos, a sua aplicação também se tem estendido ao contexto clínico. Entretanto, muitos dos testes foram adaptados a diferentes populações, alguns possuindo versões mais simples para

crianças ou pessoas com demência e versões mais difíceis para adultos saudáveis (Cambridge Cognition, 2006a, 2006b).

Em 2010, a Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra adquiriu a versão 3.0 do *software* CANTABeclipse[®] (Cambridge Cognition, 2006a, 2006b), um sistema recomendado para a investigação e incluindo 22 testes. Estas provas também eram vendidas num outro sistema mais apropriado para ensaios clínicos e designado CANTABelect[®]. As diferenças entre o CANTABeclipse[®] e o CANTABelect[®] é que este último continha especificações no *software* no sentido de assegurar o cumprimento rigoroso das normas da *Food and Drug Administration*, ao conter mais operações de segurança para evitar utilizações indevidas e erros no registo dos sujeitos e sendo os dados recolhidos recebidos pela Cambridge Cognition para verificação de erros e omissões. Com as constantes inovações e aperfeiçoamentos tecnológicos, estes sistemas foram melhorados e apresentam hoje outras designações (<http://www.cambridgecognition.com>). Actualmente, a CANTAB reúne 25 testes (ver Tabela 1) sendo a totalidade ou alguns destes testes ainda vendidos em sistemas de *software* específicos para os contextos de investigação, ensaios clínicos farmacêuticos e cuidados de saúde. Alguns destes sistemas estão disponíveis em 12 línguas, incluindo o português, à semelhança do *software* CANTABeclipse[®] versão 3.0.

A CANTAB é presentemente utilizada em mais de 68 Universidades do mundo (e.g., Reino Unido, Alemanha, Suécia, França, Holanda, Suíça, Estados Unidos da América, Canadá, Japão, Hong-Kong, Singapura e Austrália) e dispõe de mais de 1600 artigos publicados em revistas com revisão por pares. Entre estes artigos encontram-se estudos que analisam os testes da CANTAB em mais de 25 tipos de condições clínicas (<http://www.cambridgecognition.com>). No âmbito da DA e outras demências, salientam-se os estudos com grupos com DA em fase pré-clínica, DCL devido a DA ou a outras condições clínicas, demência devido a DA, demência devido a DP, demência com corpos de Lewy e demência fronto-temporal (Blackwell et al., 2004; Chamberlain et al., 2011; Fowler, Saling, Conway, Semple, & Louis, 2002; Kiddle et al., 2015; Lee, Rahman, Hodges, Sahakian, & Graham, 2003; Olde Dubbelink et al., 2014; Sahgal, McKeith, Galloway, Tasker, & Steckler, 1995; Stefanova et al., 2015; Swainson et al., 2001).

Tabela 1. Testes da Bateria Computorizada de Testes Neuropsicológicos de Cambridge.

Tipo de teste	Teste	Referência original
Iniciação ao ecrã táctil	Rastreio Motor (<i>Motor Screening</i> , MOT)	(Morris et al., 1987)
	Grande/Pequeno Círculo (<i>Big/Little Circle</i> , BLC)	(Eagger, Levy, & Sahakian, 1991)
Memória visual	Aprendizagem de Pares Associados (<i>Paired Associates Learning</i> , PAL)	(Morris et al., 1987)
	Correspondência com o Padrão Diferido (<i>Delayed Matching to Sample</i> , DMS)	(Morris et al., 1987)
	Memória de Reconhecimento de Padrões (<i>Pattern Recognition Memory</i> , PRM)	(Sahakian et al., 1988)
	Memória de Reconhecimento Espacial (<i>Spatial Recognition Memory</i> , SRM)	(Sahakian et al., 1988)
Funções executivas, memória de trabalho e planeamento	Meias de Cambridge (<i>Stockings of Cambridge</i> , SOC)	(Morris et al., 1987)
	Amplitude Espacial (<i>Spatial Span</i> , SSP)	(Morris et al., 1988)
	Memória de Trabalho Espacial (<i>Spatial Working Memory</i> , SWM)	(Morris et al., 1988)
	Mudança de Agrupamento Intra/Extra Dimensional (<i>Intra-Extra Dimensional Set Shift</i> , IED)	(Downes et al., 1989)
	Tarefa de Mudança de Atenção (<i>Attention Switching Task</i> , AST) ^a	(Eagger et al., 1991)
Atenção	Meias de Cambridge com um Toque (<i>One Touch Stockings of Cambridge</i> , OTS)	(Sahakian et al., 2009)
	Procura Visual de Correspondência com o Padrão (<i>Match to Sample Visual Search</i> , MTS)	(Downes et al., 1989)
	Processamento Rápido de Informação Visual (<i>Rapid Visual Information Processing</i> , RVP)	(Sahakian, Jones, Levy, Gray, & Warburton, 1989)
	Tempo de Reacção (<i>Reaction Time</i> , RTI)	(Eagger et al., 1991)
	Tempo de Reacção de Escolha (<i>Choice Reaction Time</i> , CRT)	(Racinais, Gaoua, & Grantham, 2008)
	Tempo de Reacção Simples (<i>Simple Reaction Time</i> , SRT)	(Edgin et al., 2010)
Memória verbal	Teste Gradativo de Nomeação (<i>Graded Naming Test</i> , GNT)	(Swainson et al., 2001)
	Memória de Reconhecimento Verbal (<i>Verbal Recognition Memory</i> , VRM)	(Need et al., 2009)
Tomada de decisão e controlo da resposta	Inibição Afectiva (<i>Affective Go/No-go</i> , AGN)	(Murphy et al., 1999)
	Tarefa de Apostas de Cambridge (<i>Cambridge Gambling Task</i> , CGT)	(Rahman, Robbins, & Sahakian, 1999)
	Tarefa de Paragem do Sinal (<i>Stop Signal Task</i> , SST)	(Turner et al., 2003)
	Tarefa de Amostragem de Informação (<i>Information Sampling Task</i> , IST)	(Chamberlain et al., 2007)
Cognição social	Tarefa de Reconhecimento de Emoções (<i>Emotion Recognition Task</i> , ERT) ^a	(Scoriels et al., 2011)
Questionários	Escalas Analógicas Visuais (<i>Visual Analogue Scales</i> , VAS) ^a	(Turner, Clark, Dowson, Robbins, & Sahakian, 2004)

Nota. ^aTestes não incluídos no *software* CANTABeclipse versão 3.0. No site <http://www.cambridgecognition.com> encontra-se informação mais detalhada e demonstrações em vídeo das tarefas de cada teste.

Atendendo a que os testes da CANTAB não foram desenvolvidos em simultâneo, alguns encontram-se mais estudados do que outros. Os mais antigos apresentam uma vasta bibliografia apontando para a sua elevada sensibilidade e especificidade, poder discriminante e validade preditiva relativamente a determinadas condições clínicas. Com base nesses estudos, a Cambridge Cognition tem vindo a auxiliar os utilizadores na escolha dos testes, definindo baterias de avaliação para condições clínicas específicas.

Assim, em 2010, a Cambridge Cognition indicava quatro testes para a avaliação dos estádios ligeiro a moderado da demência devido a DA, antecedidos da administração de um teste de sensibilização ao ecrã táctil. Estes testes eram o Rastreo Motor (*Motor Screening*, MOT), Processamento Rápido de Informação Visual (*Rapid Visual Information Processing*, RVP), Aprendizagem de Pares Associados (*Paired Associates Learning*, PAL), Tempo de Reacção (*Reaction Time*, RTI) e Memória de Trabalho Espacial (*Spatial Working Memory*, SWM) e a bateria intitulava-se CANTAB Alzheimer™. Nessa altura, também existia uma bateria para avaliação do DCL designada CANTAB MCI™ (*Mild Cognitive Impairment*, MCI) e que incluía os testes MOT, RVP, PAL, RTI e Correspondência com o Padrão Diferido (*Delayed Matching to Sample*, DMS).

Como alguns dos testes da CANTAB Alzheimer™ têm revelado poder discriminante para mais do que uma condição clínica ou para além da demência (e.g., o teste SWM também tem sido recomendado para avaliação da DP e da demência com corpos de Lewy; Fray et al., 1996; Sahgal et al., 1995), a CANTAB Alzheimer™ passou a designar-se CANTAB *Dementia*™, algum tempo depois. Recentemente, a Cambridge Cognition voltou a alterar o nome desta bateria para CANTAB *Alzheimer's*, propondo a administração dos testes MOT, RTI, PAL e SWM e, opcionalmente, os testes RVP e DMS para avaliação das fases pré-demencial e demencial da DA. Simultaneamente, a Cambridge Cognition propõe a bateria CANTAB *Prodromal Alzheimer's* para avaliação da fase pré-clínica da DA, sendo esta bateria composta pelos mesmos testes que a CANTAB *Alzheimer's* (i.e., o MOT, RVP, PAL, RTI, SWM e DMS) e pelo teste Memória de Reconhecimento de Padrões (*Pattern Recognition Memory*, PRM).

O projecto de investigação que deu origem ao presente trabalho foi elaborado em 2010. A bateria seleccionada para estudo foi a então intitulada CANTAB AlzheimerTM. Devido às reformulações que o seu nome sofreu e aos novos critérios de diagnóstico da DA (Albert et al., 2011; Jack et al., 2011; McKhann et al., 2011; Sperling et al., 2011), consideramos apropriado referir-nos a esta bateria como sendo indicada para a avaliação da DA e outras demências, embora a selecção dos testes tenha tido como critério as especificidades da DA.

Bateria Computorizada de Testes Neuropsicológicos de Cambridge para avaliação da doença de Alzheimer

Os quatro testes da CANTAB recomendados para a avaliação da DA permitem examinar as principais funções cognitivas que tendem a sofrer declínio até ao estágio moderado da demência devido a DA. Além das tarefas se basearem em paradigmas de estudo utilizados com animais, estes testes apresentam a vantagem de já terem sido estudados e validados em diversos grupos clínicos e em pesquisas utilizando técnicas de neuroimagem. Por conseguinte, existe muita informação acerca das funções cognitivas avaliadas por estas provas e acerca das bases neurológicas envolvidas no desempenho das tarefas (Chase, Clark, Sahakian, Bullmore, & Robbins, 2008; Coull, Frith, Frackowiak, & Grasby, 1996; de Rover et al., 2011; Gould et al., 2005; Manes et al., 2002; McGaughy, Dalley, Morrison, Everitt, & Robbins, 2002; Oosterman et al., 2008; Owen, Evans, & Petrides, 1996; Sexton et al., 2012). Assim, o teste RVP avalia atenção visual sustentada e desempenho continuado, o PAL avalia aprendizagem e memória episódica associativa visuo-espacial, o RTI avalia atenção visual sustentada, tempo de reacção e tempo de movimento e o SWM avalia memória de trabalho e planeamento (Cambridge Cognition, 2006a, 2006b).

O tempo médio de administração desta bateria é aceitável (cerca de uma hora) e o modo de resposta aos testes é simples, nomeadamente por mero toque num ecrã táctil (o qual se encontra assente num suporte estável) e/ou por utilização de um botão de um bloco com dois interruptores. Os testes PAL, RTI e SWM encontram-se organizados por ordem crescente de dificuldade, de modo a evitar efeitos de chão e de tecto. Além disso, os testes RVP, RTI e SWM apresentam fases de treino, os testes PAL, RTI e SWM permitem várias tentativas de resposta e os testes PAL e RTI dispõem de critérios

de interrupção. Estes aspectos contribuem para diminuir a ansiedade e a frustração das pessoas com dificuldades cognitivas e facilitam que as provas sejam bem toleradas, tanto por pessoas com demência como por pessoas saudáveis. O facto dos estímulos serem coloridos, apresentarem movimento e serem acompanhados por sons ou frases de *feedback* (e.g., indicação de transição para um estágio seguinte) confere a estes testes a aparência de jogos, o que contribui para a motivação da pessoa examinada. A administração destes testes requer a presença de um examinador, para transmitir pequenas instruções verbais e executar comandos (e.g., passar de um problema para outro). Não obstante, esta presença é não invasiva, visto o examinador ficar ao lado da pessoa examinada (ver Figura 4), com a vantagem de poder observar o comportamento do indivíduo e ajudá-lo, caso surja alguma dificuldade técnica.

Como já indicado, alguns testes da CANTAB dispõem de versões mais difíceis, para adultos saudáveis, e versões mais simples, para crianças ou pessoas com demência. Antes da administração do teste, o examinador deve seleccionar qual é a versão (designada no manual da CANTAB por *mode* ou modo de teste; Cambridge Cognition, 2006b) com que pretende correr essa prova. A Tabela 2 sistematiza as versões ou modos disponíveis no *software* CANTABeclipse[®] versão 3.0 para os testes MOT, RVP, PAL, RTI e SWM, indicando as diferenças entre eles (Cambridge Cognition, 2006b). A Cambridge Cognition recomenda que a CANTAB para avaliação da DA seja administrada na seguinte ordem de testes e seleccionando as seguintes versões ou modos de execução das provas, caso a bateria seja administrada a pessoas com demência: teste MOT, versão “alta visibilidade”; teste RVP, versão “3-5-7”; e testes PAL, RTI e SWM, versão “clínica”. No caso da bateria ser administrada a pessoas mais velhas sem demência, a versão recomendada para administração do teste RVP é a “clínica” (Sahakian et al., 1989).

Terminada a administração da bateria, os resultados são automaticamente gravados. O examinador poderá depois, a partir do *software*, aceder a relatórios sumários, folhas detalhadas especificando as respostas e as condições de administração de cada estímulo ou problema ou, em alternativa, obter um ficheiro de tipo *commas and values* (.csv) com as medidas de resultados que pretender analisar. Este ficheiro .csv poderá ser exportado e convertido em ficheiros do *software* Microsoft[®] Excel (Microsoft[®] Office, Microsoft[®] Corporation) ou do *Statistical Package for the Social*

Sciences Statistics (SPSS, IBM®), o que torna mais fácil a análise estatística dos dados ao não ser necessária a inserção manual dos mesmos. Na Tabela 2 é possível consultar todas as medidas de resultados que podem ser seleccionadas para cada teste, pelo *software* CANTABeclipse® versão 3.0 (Cambridge Cognition, 2006a, 2006b).

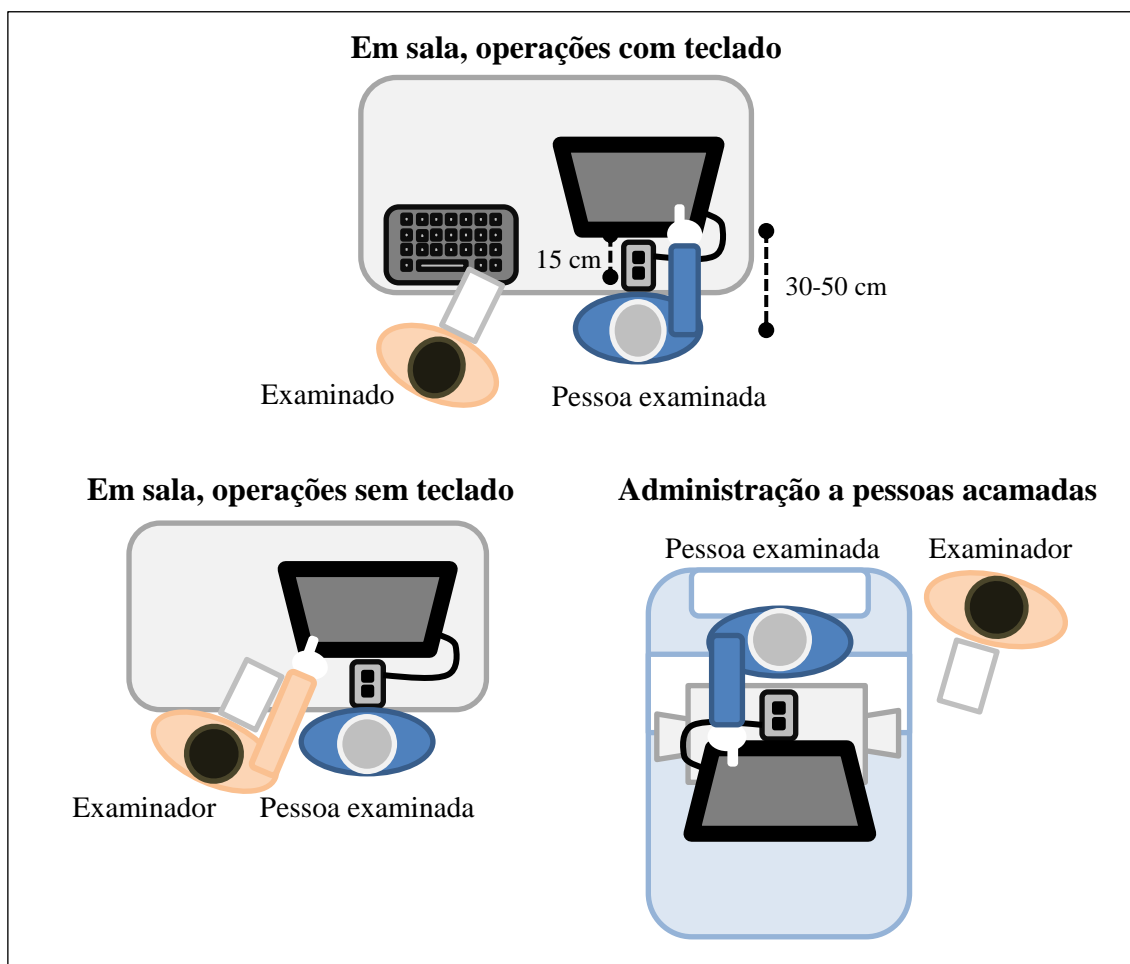


Figura 4. Formas de administração da Bateria Computorizada de Testes Neuropsicológicos de Cambridge para avaliação da doença de Alzheimer. O único material a que a pessoa examinada deve ter acesso é o ecrã táctil e o bloco de interruptores. O examinador pode optar por não usar teclado.

Tabela 2. Descrição dos testes da CANTAB para avaliação da doença de Alzheimer.

Teste	Descrição	Medidas de resultados
MOT		
Modo clínico	A tarefa consiste em tocar nas cruzes que vão aparecendo uma a uma e aleatoriamente no ecrã.	<ul style="list-style-type: none"> • Latência média (i.e., média de tempo desde que o aparecimento da cruz até se tocar com sucesso no ecrã)^a; • Média de erro (i.e., média da distância entre o centro da cruz e o local onde se toca no ecrã, para itens respondidos com sucesso; unidade de medida: <i>pixels</i>).
Modo alta visibilidade	Versão recomendada para pessoas mais velhas ou com menor acuidade visual. A tarefa é igual à anterior mas as cruzes aparecem a negro.	
RVP		
Modo clínico	Versão recomendada para adultos sem demência e adolescentes com mais de 14 anos. Os dígitos 2 a 9 são exibidos um a um e de modo aleatório no centro do ecrã, à velocidade de 100 dígitos por minuto. A tarefa consiste em detectar as sequências 2-4-6, 4-6-8 e 3-5-7 premindo um botão logo que surja o último número da sequência. O teste é composto por 7 blocos: os quatro primeiros são para treino (sendo dadas ajudas e <i>feedback</i> nos dois primeiros blocos) e os três últimos blocos são para avaliação (i.e., são contabilizados nas medidas de resultados).	<ul style="list-style-type: none"> • Total de sucessos (i.e., total de sequências-alvo detectadas dentro da janela de resposta de 1800 milissegundos); • Total de não respostas (i.e., total de sequências-alvo não detectadas dentro da janela de resposta); • Total de falsos alarmes (i.e., número de vezes em que o botão é premido fora da janela de resposta); • Total de rejeições correctas (i.e., número estímulos correctamente rejeitados por não serem a sequência-alvo); • Probabilidade de sucesso (i.e., ‘total de sucessos’/[‘total de sucessos’ + ‘total de não respostas’]); • Probabilidade de falso alarme (i.e., ‘total de falsos alarmes’/[‘total de falsos alarmes’ + ‘total de rejeições correctas’]); • A’ (i.e., índice de sensibilidade a detectar sequências-alvo, independentemente da tendência de resposta; a fórmula de cálculo^b inclui as medidas ‘probabilidade de sucesso’ e ‘probabilidade de falso alarme’); • B’’ (i.e., índice de tendência a responder independentemente da sequência-alvo estar presente; a fórmula de cálculo^b inclui as medidas ‘probabilidade de sucesso’ e ‘probabilidade de falso alarme’); • Latência média (i.e., média de tempo a detectar correctamente a sequência-alvo)^a.
Modo 1-2-3	Versão recomendada para crianças dos 4 aos 8 anos. A tarefa é igual à anterior mas a sequência a detectar é a 1-2-3.	
Modo 3-5-7	Versão recomendada para crianças dos 7 aos 14 anos e adultos com demência. A tarefa é igual à da versão ou modo clínico mas a sequência a detectar é a 3-5-7.	

(A Tabela 2 continua)

Tabela 2 (continuação 1). Descrição dos testes da CANTAB para avaliação da doença de Alzheimer.

Teste	Descrição	Medidas de resultados
PAL		
Modo clínico	Versão recomendada para quando se pretende realizar uma só avaliação. O sistema exhibe múltiplas formas de teste mas não garante que uma pessoa não responda a uma mesma forma, numa outra vez. A tarefa consiste em tocar na caixa em que se viu um determinado padrão aparecer. A ordem de apresentação dos padrões e o local onde estes surgem é aleatória. O teste inclui 8 estádios (dois de 1 padrão, dois de 2 padrões, dois de 3 padrões, um de 6 padrões e outro de 8 padrões). Para cada estádio, existem dez ensaios; se a pessoa falhar todos os ensaios de um estádio, o teste termina.	<ul style="list-style-type: none"> • Total de erros (i.e., número total de erros no teste; esta medida também pode ser calculada para cada estádio); • Total de erros ajustado (i.e., número total de erros com um ajustamento para cada estádio não completado), sendo o ajustamento calculado pela fórmula ('soma do número de padrões dos estádios não completados' – ['número de padrões dos estádios não completados'/'número de caixas dos estádios não completados']) × 'número de tentativas para cada estádio não completado'; esta medida também pode ser calculada para cada estádio, sendo o ajustamento aplicado a quem não tenha atingido esse estádio; • Total de ensaios (i.e., total de ensaios necessários até completar os estádios do teste; esta medida também pode ser calculada por estádios); • Total de ensaios ajustado (i.e., total de ensaios necessários até completar os estádios do teste com um ajustamento equivalente à soma de dez ensaios por cada estádio não completado; esta medida também pode ser calculada por estádios, sendo o ajustamento aplicado a quem não tenha atingido esse estádio); • Total de estádios completados; • Pontuação de memória no primeiro ensaio (i.e., total de padrões correctamente localizados no primeiro ensaio de cada estádio completado); • Total de estádios completados no primeiro ensaio; • Média de erros até ao sucesso (i.e., 'total de erros'/'total de estádios completados'); • Média de ensaios até ao sucesso (i.e., 'total de ensaios'/'total de estádios completados').
Modo paralelo	Versão recomendada para avaliações repetidas; o sistema garante que a pessoa não voltará a responder a uma mesma forma de teste. A tarefa e o critério de interrupção são os mesmos que os da versão ou modo clínico. O teste inclui 5 estádios (dois de 2 padrões, um de 3 padrões, um de 6 padrões e um de 8 padrões).	

(A Tabela 2 continua)

Tabela 2 (continuação 2). Descrição dos testes da CANTAB para avaliação da doença de Alzheimer.

Teste	Descrição	Medidas de resultados
RTI		
Modo clínico	Versão recomendada para quando se pretende realizar uma só avaliação. O sistema exhibe múltiplas formas mas não garante que a pessoa não responda a uma mesma forma, numa outra vez. O teste inclui 5 estádios, sendo os três primeiros de treino (estádio 1, condição 1 círculo: tocar o mais depressa possível no círculo, logo após uma bola amarela aparecer; estágio 2, condição 5 círculos: tocar no círculo onde a bola aparecer; estágio 3, condição 1 círculo: premir um botão e libertá-lo logo após a bola aparecer; estágio 4, condição 1 círculo: libertar o botão quando a bola aparecer e tocar no círculo com a mesma mão; estágio 5, condição 5 círculos: libertar o botão quando a bola aparecer e tocar no círculo onde ela apareceu). Se o critério de 5 respostas correctas em 6 ensaios não for cumprido (em qualquer estágio excepto o primeiro) o teste termina.	<ul style="list-style-type: none"> • Tempo de reacção simples (i.e., velocidade média com que a pessoa liberta o botão logo após a bola amarela aparecer, para os ensaios respondidos com sucesso na condição 1 círculo)^a; • Tempo de reacção escolha-cinco (i.e., o mesmo que o descrito para a medida anterior mas na condição 5 círculos)^a; • Tempo de movimento simples (i.e., tempo médio dispendido entre libertar o botão e tocar no círculo, para os ensaios respondidos com sucesso na condição 1 círculo)^a; • Tempo de movimento escolha-cinco (i.e., o mesmo que o descrito para a medida anterior mas na condição 5 círculos)^a; • Pontuação de exactidão simples (i.e., número total de ensaios respondidos com sucesso na condição 1 círculo); • Pontuação de exactidão escolha-cinco (i.e., o mesmo que o descrito para a medida anterior mas na condição 5 círculos); • Pontuação de erros simples (i.e., número total de ensaios em que a pessoa cometeu erros de um determinado tipo [‘localização incorrecta’, se tocou no círculo errado; ‘impreciso’, se não tocou no círculo com precisão; ‘demasiado cedo’, se tocou no círculo antes de se abrir a janela de resposta; ‘demasiado tarde’, se tocou no círculo após a janela de resposta] na condição 1 círculo); • Pontuação de erros escolha-cinco (i.e., o mesmo que o descrito para a medida anterior mas na condição 5 círculos).
Modo paralelo	Versão recomendada para avaliações repetidas; o sistema garante que a pessoa não voltará a responder a uma mesma forma de teste. O critério de interrupção é o mesmo que o da versão clínica. O teste inclui 4 estádios, sendo o primeiro e o terceiro para treino (estádios 1 e 2, condição 1 círculo: libertar o botão quando a bola amarela aparecer e tocar no círculo com a mesma mão; estádios 3 e 4, condição 5 círculos: libertar o botão quando a bola aparecer e tocar no círculo onde ela apareceu).	
Modo criança	A tarefa é igual à da versão paralela. Se o critério de 4 respostas correctas em 5 ensaios não for atingido (em qualquer estágio excepto o primeiro) o teste termina.	

(A Tabela 2 continua)

Tabela 2 (continuação 3). Descrição dos testes da CANTAB para avaliação da doença de Alzheimer.

Teste	Descrição	Medidas de resultados
SWM		
Modo clínico	A tarefa consiste em procurar fichas azuis e colocá-las dentro de uma casa até a preencher. Cada ficha é escondida, à vez, numa caixa diferente. Logo que a pessoa examinada encontre uma ficha dentro de uma caixa, ele(a) não deve lá voltar pois a ficha seguinte será escondida noutra caixa. O teste inclui quatro problemas de 3 caixas (fase de treino), quatro problemas de 4 caixas, quatro problemas de 6 caixas e quatro problemas de 8 caixas (fase de avaliação). A pessoa examinada é quem decide a ordem com que pesquisará as fichas nas caixas.	<ul style="list-style-type: none"> • Erros entre pesquisas (i.e., número de vezes em que se toca numa caixa onde já foi encontrada uma ficha)^c; • Erros dentro de uma pesquisa (i.e., número de vezes em que se toca numa caixa que já tinha mostrado estar vazia, numa mesma pesquisa)^c; • Erros duplos (i.e., número de vezes em que se comete um erro que pode ser categorizado como ‘entre pesquisas’ e ‘dentro de uma pesquisa’)^c; • Total de erros (i.e., ‘erros entre pesquisas’ + ‘erros dentro de uma pesquisa’ – ‘erros duplos’)^c; • Estratégia (i.e., número de vezes em que se começa uma pesquisa tocando numa caixa diferente; resultados mais altos indicam menor uso da estratégia de seguir sempre uma mesma sequência de caixas; medida calculada para o total de problemas de 6 e 8 caixas); • Média de tempo até à primeira resposta (i.e., tempo entre a apresentação do problema e o primeiro toque numa caixa)^{a,c}; • Média de tempo de preparação para procurar fichas (i.e., tempo entre toques para procurar fichas, nomeadamente, entre a apresentação do problema e tocar numa caixa e entre colocar uma ficha na casa e tocar na caixa seguinte)^{a,c}; • Média de tempo até à última resposta (i.e., tempo desde a apresentação do problema até tocar na caixa contendo a última ficha)^{a,c}.

Nota. A descrição aplica-se aos testes do *software* CANTABeclipse versão 3.0 (Cambridge Cognition, 2006a, 2006b). ^aUnidade de medida: milissegundos. ^bVer fórmulas em Roiser et al. (2005). ^cMedida calculada para o total de problemas de 4, 6 e 8 caixas ou para cada um destes três níveis de problemas. Legenda: CANTAB = *Cambridge Neuropsychological Test Automated Battery* (Bateria Computorizada de Testes Neuropsicológicos de Cambridge); MOT = *Motor Screening* (Rastreo Motor); PAL = *Paired Associates Learning* (Aprendizagem de Pares Associados); RTI = *Reaction Time* (Tempo de Reacção); RVP = *Rapid Visual Information Processing* (Processamento Rápido de Informação Visual); SWM = *Spatial Working Memory* (Memória de Trabalho Espacial).

Em termos de poder discriminante, sensibilidade, especificidade e valor preditivo dos testes RVP, PAL, RTI e SWM, são de salientar os estudos indicando várias medidas destes testes como tendo elevada sensibilidade a alterações cognitivas induzidas por substâncias psicoactivas, alterações essas observadas mesmo em períodos curtos de tempo, em pessoas mais velhas saudáveis, com DA ou com depressão major (e.g., RVP total de sucessos, RVP A', RVP latência média, PAL total de erros, PAL total de erros ajustado, PAL total de ensaios ajustado, PAL pontuação de memória no primeiro ensaio, RTI tempo de reacção simples, RTI tempo de reacção escolha-cinco, RTI tempo de movimento simples, RTI tempo de movimento escolha-cinco, SWM erros entre pesquisas, SWM erros dentro de uma pesquisa, SWM total de erros e SWM estratégia) (Herrera-Guzmán et al., 2009; Herrera-Guzmán et al., 2010; Jones, Sahakian, Levy, Warburton, & Gray, 1992; Lesk, Honey, & de Jager, 2009; M. Riekkinen, Laakso, Jäkälä, & Riekkinen, 1999; P. Riekkinen & Riekkinen, 1999; Sahakian & Coull, 1993; Sahakian et al., 1989).

Igualmente notáveis são os estudos indicando o poder das pontuações nestes testes em distinguirem os desempenhos entre grupos saudáveis e com demência devido a DA (Chamberlain et al., 2011; Fowler, Saling, Conway, Semple, & Louis, 1995; Fowler et al., 2002; Jones et al., 1992; Junkkila, Oja, Laine, & Karrasch, 2012; Lee et al., 2003; Sahakian et al., 1990; Sahakian et al., 1989; Sahgal et al., 1992; Sahgal et al., 1995; Saunders & Summers, 2010; Swainson et al., 2001), entre grupos saudáveis e com DCL (Alichniewicz, Brunner, Klünemann, & Greenlee, 2012; Alladi, Arnold, Mitchell, Nestor, & Hodges, 2006; Collie, Maruff, & Currie, 2002; Facal, Rodríguez, Juncos-Rabadán, Caamaño, & Sueiro, 2009; Fowler et al., 2002; Juncos-Rabadán, Facal, Pereiro, & Lojo-Seoane, 2014; Juncos-Rabadán, Pereiro, Facal, Reboredo, & Lojo-Seoane, 2014; Junkkila et al., 2012; Levy-Gigi, Kelemen, Gluck, & Kéri, 2011; Saunders & Summers, 2010, 2011), entre grupos com DCL amnésico e DCL não amnésico (Juncos-Rabadán, Facal, et al., 2014; Klekociuk & Summers, 2013; Saunders & Summers, 2010, 2011), entre grupos com DCL amnésico com múltiplos domínios afectados e DCL amnésico único (Juncos-Rabadán, Facal, et al., 2014), entre grupos com DCL e com demência devido a DA (Junkkila et al., 2012; Saunders & Summers, 2010), entre grupos saudáveis e com demência devido a DA e a causas vasculares (Kim et al., 2009), entre grupos saudáveis e com lesão frontal (Owen, Downes, Sahakian,

Polkey, & Robbins, 1990); entre grupos saudáveis e com atrofia de múltiplos sistemas com predominância estriatonigral (Robbins et al., 1992), entre grupos saudáveis e com depressão major (Beats, Sahakian, & Levy, 1996; Braw, Aviram, Bloch, & Levkovitz, 2011; Chamberlain et al., 2011; Sexton et al., 2012), entre grupos com demência devido a DA e com depressão major (Chamberlain et al., 2011), entre grupos com demência devido a DA e com demência com corpos de Lewy (Sahgal et al., 1995) e entre grupos com demência devido a DA e demência fronto-temporal nas variantes frontal e semântica (Lee et al., 2003).

Alguns estudos de sensibilidade e especificidade têm indicado que o teste PAL, quando isolado ou em associação com outros testes como o DMS, o Teste Gradativo de Nomeação (*Graded Naming Test*, GNT) (Cambridge Cognition, 2006a, 2006b) ou o Exame Cognitivo de Addenbrooke Revisto (*Addenbrooke's Cognitive Examination Revised*, ACE-R) (Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006), apresenta 80 a 100% de sensibilidade e 55 a 100% de especificidade relativamente à demência por DA (Ahmed, Mitchell, Arnold, Nestor, & Hodges, 2008; Blackwell et al., 2004; Fowler, Saling, Conway, Semple, & Louis, 1997; Fowler et al., 2002; Mitchell et al., 2009) e 48% de sensibilidade e 96% de especificidade relativamente ao DCL devido a DA (De Jager, Blackwell, Budge, & Sahakian, 2005).

Vários estudos de validade preditiva indicam que o teste PAL permite distinguir, em pessoas com DCL, o desempenho das que 12 a 32 meses depois vieram a desenvolver demência devido a DA das que não desenvolveram demência (Ahmed et al., 2008; Blackwell et al., 2004; Mitchell et al., 2009). Mais recentemente (Summers & Saunders, 2012), as medidas RVP A' e PAL total de erros ajustado foram indicadas como sendo os melhores preditores dos indivíduos com DCL que 20 meses depois desenvolveram demência devido a DA, ao distinguirem o seu desempenho inicial dos de indivíduos que se mantiveram com a classificação de DCL ou que recuperaram as funções cognitivas. Segundo este estudo, a medida PAL total de erros ajustado no estágio de 6 padrões permite igualmente distinguir os desempenhos iniciais dos indivíduos com DCL que se mantiveram com essa classificação dos que recuperaram as funções cognitivas. Neste sentido, um outro estudo revelou que um conjunto que incluía as medidas RVP A', RVP latência média, SWM total de erros e SWM estratégia era um bom preditor de indivíduos que mantinham a classificação de DCL, num intervalo de 11

meses (Klekociuk, Summers, Vickers, & Summers, 2014). Um estudo de validade preditiva com indivíduos saudáveis assinalou também o teste PAL como permitindo distinguir o desempenho inicial daqueles que dois anos depois desenvolveram declínio mnésico ou DCL amnésico dos que permanecem sem declínio cognitivo (Collie et al., 2001).

Todos estes aspectos sugerem que estes testes da CANTAB são uma alternativa vantajosa em relação às baterias mais tradicionais, de papel-e-lápis, para avaliação da DA, justificando-se a sua adaptação e validação para Portugal.

OBJECTIVOS E ESTRUTURA DO PROJECTO DE INVESTIGAÇÃO E DA PRESENTE DISSERTAÇÃO

Este projecto de investigação tem como tema a adaptação e validação para Portugal dos testes da CANTAB para avaliação da DA. Esta bateria foi seleccionada de modo a responder aos problemas da falta de testes computadorizados para avaliação destas condições clínicas em Portugal e, em particular, de testes com comprovada sensibilidade, poder discriminante e valor preditivo para a DA. Como população alvo, seleccionaram-se pessoas mais velhas a residirem em instituições geriátricas, devido à necessidade urgente de normas para este grupo populacional e às muitas vantagens que os testes da CANTAB têm a oferecer a estas instituições, se considerarmos o número elevado de avaliações a que têm de proceder e os muitos dados que necessitam armazenar.

Assim, os principais objectivos deste projecto de investigação foram a adaptação das instruções verbais dos testes MOT, RVP, PAL, RTI e SWM para a língua e cultura da população mais velha portuguesa (atendendo a que só as instruções exibidas pelo *software* CANTABeclipse® versão 3.0 estavam traduzidas para português), a análise das propriedades psicométricas dos testes RVP, PAL, RTI e SWM, o estabelecimento das normas destes testes para pessoas mais velhas portuguesas a residir em instituições geriátricas e o alargamento do conhecimento relativo à validade e utilidade clínica das pontuações nestas provas. Mais especificamente, pretendeu-se, na fase inicial deste projecto de investigação, examinar os seguintes aspectos:

- (1) A fiabilidade teste-reteste dos resultados no RVP, PAL, RTI e SWM (considerando um intervalo de quatro semanas) e a susceptibilidade destes

testes a efeitos da prática, de chão e de tecto, numa amostra de pelo menos 30 pessoas mais velhas portuguesas sem problemas neuropsiquiátricos e a residir em instituições geriátricas;

- (2) A influência de variáveis sócio-demográficas nos resultados nos testes, em particular a idade, género, escolaridade, familiaridade com o uso de computadores, tempo de institucionalização e consumo de medicação psicotrópica potencialmente inapropriada (um problema muito frequente na população mais velha portuguesa; Periquito et al., 2014), procedendo-se a esta análise numa amostra normativa de pelo menos 120 pessoas mais velhas portuguesas a residir em instituições geriátricas, para depois se definirem normas de teste em função das variáveis sócio-demográficas com maior influência nos desempenhos;
- (3) A validade de constructo dos quatro testes da CANTAB, examinando as relações entre as suas medidas (i.e., as intra-correlações e inter-correlações) e os factores em que essas medidas mais saturassem, com base nos resultados de uma amostra de pessoas mais velhas portuguesas sem problemas neuropsiquiátricos ($n = 120$), com demência ligeira devido a DA ($n = 40$) e com depressão major com défice objectivo na memória episódica ($n = 40$);
- (4) A validade concorrente destes testes da CANTAB, examinando as suas relações com os resultados em testes neuropsicológicos de papel-e-lápis já adaptados para a população portuguesa, ou seja, o Teste da Figura Complexa de Rey (Espírito-Santo et al., 2015; Rey, 2002), os subtestes Cenas de Família, Memória Lógica, Pares de Palavras e Localização Espacial da WMS-III (Wechsler, 2008b, 2008c), o subteste Memória de Dígitos da WAIS-III (Wechsler, 2008a) e o teste de fluência verbal semântica incluído no ACE-R (Firmino, Simões, Pinho, Cerejeira, & Martins, 2008; Mioshi et al., 2006), procedendo-se a esta análise na mesma amostra referida no ponto anterior.

Para a consecução destes objectivos, o projecto de investigação foi dividido em três fases. A primeira fase correspondeu ao estudo piloto. Nesta fase procurou-se testar a compreensão das instruções dos testes da CANTAB numa amostra composta por 10

participantes portugueses com 70 a 93 anos de idade, de ambos os géneros e com 4 a 25 anos de escolaridade, a residirem em instituições geriátricas. Os anexos A a E apresentam as instruções verbais finais para cada um dos testes (MOT, RVP, PAL, RTI e SWM, respectivamente). Com este estudo piloto pretendeu-se ainda ajustar a ordem de administração das provas incluídas no protocolo (i.e., os testes de papel-e-lápis administrados numa primeira sessão para verificação dos critérios de inclusão em cada um dos grupos e os testes de papel-e-lápis e da CANTAB administrados na sessão de teste), de modo a serem respeitados os intervalos de tempo entre a apresentação dos estímulos e a evocação e reconhecimento diferidos dos mesmos. Os testes e a ordem com que foram administrados serão referidos com maior detalhe ao longo deste trabalho.

A segunda fase deste projecto de investigação correspondeu à administração do protocolo de avaliação para constituição das amostras. Esta fase teve uma duração de quase dois anos e meio. Dois dos testes administrados para verificação dos critérios de inclusão nos grupos não dispunham de valores normativos para a população mais velha portuguesa, por serem itens de dois instrumentos de avaliação, nomeadamente, o teste de Fluência Verbal Semântica (item do ACE-R; Firmino et al., 2008; Mioshi et al., 2006) e a tarefa de Nomeação de Objectos e Dedos da Escala de Avaliação da Doença de Alzheimer (*Alzheimer's Disease Assessment Scale*, ADAS; Guerreiro, Fonseca, Barreto, & Garcia, 2008; Mohs, Rosen, & Davis, 1983). Por conseguinte, a par do presente estudo foi necessário constituir uma amostra normativa de pessoas mais velhas institucionalizadas, de modo a calcular os valores de referência para estes itens. Os critérios de inclusão desta amostra foram a obtenção de resultados não sugestivos de deterioração cognitiva no MMSE e total do ACE-R, ausência de sintomas depressivos graves avaliados pela Escala de Depressão Geriátrica de 30 itens (*Geriatric Depression Scale*, GDS-30; Barreto, Leuschner, Santos, & Sobral, 2008; Yesavage et al., 1982) e uma história clínica indicando ausência de problemas neuropsiquiátricos. O Anexo F apresenta as características demográficas dessa amostra e os valores normativos calculados por idade e escolaridade para os referidos itens.

Para o estudo da fiabilidade teste-reteste referido no objectivo (1) constituiu-se uma amostra de 34 participantes. O estudo da influência das variáveis sócio-demográficas mencionado no objectivo (2) foi conduzido com uma amostra de 128

participantes. No caso dos grupos clínicos foi necessário alterar os objectivos iniciais para se conseguir uma dimensão amostral satisfatória. Assim, para o grupo com demência ligeira devido a DA, incluíram-se também pacientes em estágio moderado e foi ainda necessário procurar participantes em consultas externas de psicogeriatría (i.e., participantes a residir na comunidade). Com estas alterações nos objectivos iniciais constituiu-se um grupo DA com 37 pessoas. No caso do grupo com depressão major com défice objectivo de memória episódica, apesar de se terem procurado participantes entre os utentes de consultas externas de psicogeriatría, não foi possível obter um número significativo, pelo que este grupo deixou de fazer parte dos objectivos. Tendo sido incluídos participantes a residir na comunidade no grupo DA, procuraram-se ainda participantes sem problemas neuropsiquiátricos que também residissem na comunidade. Deste modo, os estudos de validade referidos nos objectivos (3) e (4) foram conduzidos com uma amostra de 137 participantes sem problemas neuropsiquiátricos e um grupo com 37 pessoas com demência ligeira a moderada devido a DA, a residir em instituições geriátricas e na comunidade.

A terceira fase deste projecto de investigação correspondeu à realização dos estudos referidos nos objectivos (1) a (4) e à redacção de artigos para publicação. Neste contexto, na presente dissertação optou-se pelo formato de compilação de artigos. Assim, o primeiro capítulo apresenta o estudo da fiabilidade teste-reteste (Gonçalves, Pinho, & Simões, 2016a) indicado no objectivo (1). O segundo capítulo diz respeito ao estudo da influência das variáveis sócio-demográficas e ao estabelecimento de normas dos quatro testes CANTAB (i.e., RVP, PAL, RTI e SWM) para a população mais velha portuguesa a residir em instituições geriátricas (Gonçalves, Pinho, & Simões, 2016b) concretizando o objectivo (2). Por último, o terceiro capítulo apresenta os estudos de validade de constructo e concorrente (Gonçalves, Pinho, & Simões, 2016c) indicados nos objectivos (3) e (4). Após estes capítulos, seguir-se-á uma discussão geral dos resultados, uma reflexão acerca da experiência na utilização dos testes CANTAB com pessoas mais velhas portuguesas e implicações práticas relativamente a trabalhos futuros.

REFERÊNCIAS

- Ahmed, S., Mitchell, J., Arnold, R., Nestor, P. J., & Hodges, J. R. (2008). Predicting rapid clinical progression in amnesic mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, 25(2), 170-177. doi: 10.1159/000113014
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., ... Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270-279. doi: 10.1016/j.jalz.2011.03.008
- Alichniewicz, K. K., Brunner, F., Klünemann, H. H., & Greenlee, M. W. (2012). Structural and functional neural correlates of visuospatial information processing in normal aging and amnesic mild cognitive impairment. *Neurobiology of Aging*, 33(12), 2782-2797. doi: 10.1016/j.neurobiolaging.2012.02.010
- Alladi, S., Arnold, R., Mitchell, J., Nestor, P. J., & Hodges, J. R. (2006). Mild cognitive impairment: Applicability of research criteria in a memory clinic and characterization of cognitive profile. *Psychological Medicine*, 36(4), 507-515. doi: 10.1017/s0033291705006744
- American Psychological Association. (2015). *APA dictionary of psychology* (2nd ed.). Washington, DC: Author.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan Publishing Company.
- Attix, D. K., Story, T. J., Chelune, G. J., Ball, J. D., Stutts, M. L., Hart, R. P., & Barth, J. T. (2009). The prediction of change: Normative neuropsychological trajectories. *The Clinical Neuropsychologist*, 23(1), 21-38. doi: 10.1080/13854040801945078
- Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de Depressão Geriátrica (GDS). In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* (2ª ed., pp. 69-72). Lisboa, Portugal: Novartis.
- Bartko, S. J., Vendrell, I., Saksida, L. M., & Bussey, T. J. (2010). A computer-automated touchscreen paired-associates learning (PAL) task for mice: Impairments following administration of scopolamine or dicyclomine and improvements following donepezil. *Psychopharmacology*, 214(2), 537-548. doi: 10.1007/s00213-010-2050-1
- Beats, B. C., Sahakian, B. J., & Levy, R. (1996). Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. *Psychological Medicine*, 26(3), 591. doi: 10.1017/s0033291700035662
- Benedict, R. H. B., & Zgaljardic, D. J. (1998). Practice effects during repeated administrations of memory tests with and without alternate forms. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 20(3), 339-352. doi: 10.1076/jcen.20.3.339.822
- Blackwell, A. D., Sahakian, B. J., Vesey, R., Semple, J. M., Robbins, T. W., & Hodges, J. R. (2004). Detecting dementia: Novel neuropsychological markers of preclinical Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 17(1-2), 42-48. doi: 10.1159/000074081
- Braw, Y., Aviram, S., Bloch, Y., & Levkovitz, Y. (2011). The effect of age on frontal lobe related cognitive functions of unmedicated depressed patients. *Journal of Affective Disorders*, 129(1-3), 342-347. doi: 10.1016/j.jad.2010.07.032

- Bruggemans, E. F., van de Vijver, F. J. R., & Huysmans, H. A. (1997). Assessment of cognitive deterioration in individual patients following cardiac surgery: Correcting for measurement error and practice effects. *Journal of Clinical and Experimental Neuropsychology*, 19(4), 543-559. doi: 10.1080/01688639708403743
- Burns, A., Lawlor, B., & Craig, S. (2004). *Assessment scales in old age psychiatry* (2nd ed.). London, UK: Martin Dunitz of the Taylor & Francis Group.
- Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological measures: Meta-analyses of test-retest correlations. *The Clinical Neuropsychologist*, 27(7), 1077-1105. doi: 10.1080/13854046.2013.809795
- Cambridge Cognition. (2006a). *CANTABeclipse™ version 3: Software user guide*. Cambridge, UK: Author.
- Cambridge Cognition. (2006b). *CANTABeclipse™ version 3: Test administration guide*. Cambridge, UK: Author.
- Cernich, A. N., Brennana, D. M., Barker, L. M., & Bleiberg, J. (2007). Sources of error in computerized neuropsychological assessment. *Archives of Clinical Neuropsychology*, 22(Suppl. 1), S39-S48. doi: 10.1016/j.acn.2006.10.004
- Chamberlain, S. R., Blackwell, A. D., Nathan, P. J., Hammond, G., Robbins, T. W., Hodges, J. R., ... Sahakian, B. J. (2011). Differential cognitive deterioration in dementia: A two year longitudinal study. *Journal of Alzheimer's Disease*, 24(1), 125-136. doi: 10.3233/jad-2010-100450
- Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2007). A neuropsychological comparison of obsessive-compulsive disorder and trichotillomania. *Neuropsychologia*, 45(4), 654-662. doi: 10.1016/j.neuropsychologia.2006.07.016
- Chase, H. W., Clark, L., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2008). Dissociable roles of prefrontal subregions in self-ordered working memory performance. *Neuropsychologia*, 46(11), 2650-2661. doi: 10.1016/j.neuropsychologia.2008.04.021
- Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, 55(10), 1630-1635. doi: 10.1111/j.1532-5415.2007.01379.x
- Collie, A., Maruff, P., & Currie, J. (2002). Behavioral characterization of mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 24(6), 720-733. doi: 10.1076/j.cen.24.6.720.8397
- Collie, A., Maruff, P., Darby, D. G., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *Journal of the International Neuropsychological Society*, 9(3), 419-428. doi: 10.1017/s1355617703930074
- Collie, A., Maruff, P., Shafiq-Antonacci, R., Smith, M., Hallup, M., Schofield, P. R., ... Currie, J. (2001). Memory decline in healthy older people: Implications for identifying mild cognitive impairment. *Neurology*, 56(11), 1533-1538. doi: 10.1212/wnl.56.11.1533

- Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia*, *34*(11), 1085-1095. doi: 10.1016/0028-3932(96)00029-2
- Crawford, J. R., & Garthwaite, P. H. (2007). Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology*, *21*(5), 611-620. doi: 10.1037/0894-4105.21.5.611
- Cysique, L. A., Franklin, D., Abramson, I., Ellis, R. J., Letendre, S., Collier, A., ... Heaton, R. K. (2011). Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. *Journal of Clinical and Experimental Neuropsychology*, *33*(5), 505-522. doi: 10.1080/13803395.2010.535504
- Darby, D., Maruff, P., Collie, A., & McStephen, M. (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*, *59*(7), 1042-1046. doi: 10.1212/wnl.59.7.1042
- De Jager, C., Blackwell, A. D., Budge, M. M., & Sahakian, B. J. (2005). Predicting cognitive decline in healthy older adults. *The American Journal of Geriatric Psychiatry*, *13*(8), 735-740. doi: 10.1097/00019442-200508000-00014
- De Jager, C., & Budge, M. (2005). Stability and predictability of the classification of mild cognitive impairment as assessed by episodic memory test performance over time. *Neurocase*, *11*(1), 72-79. doi: 10.1080/13554790490896820
- de Rover, M., Pironti, V. A., McCabe, J. A., Acosta-Cabronero, J., Arana, F. S., Morein-Zamir, S., ... Sahakian, B. J. (2011). Hippocampal dysfunction in patients with mild cognitive impairment: A functional neuroimaging study of a visuospatial paired associates learning task. *Neuropsychologia*, *49*(7), 2060-2070. doi: 10.1016/j.neuropsychologia.2011.03.037
- Downes, J. J., Roberts, A. C., Sahakian, B. J., Evenden, J. L., Morris, R. G., & Robbins, T. W. (1989). Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: Evidence for a specific attentional dysfunction. *Neuropsychologia*, *27*(11-12), 1329-1343. doi: 10.1016/0028-3932(89)90128-0
- Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology*, *27*(3), 248-261. doi: 10.1093/arclin/acr120
- Eagger, S. A., Levy, R., & Sahakian, B. J. (1991). Tacrine in Alzheimer's disease. *The Lancet*, *337*(8748), 989-992. doi: 10.1016/0140-6736(91)92656-m
- Edgin, J. O., Mason, G. M., Allman, M. J., Capone, G. T., DeLeon, I., Maslen, C., ... Nadel, L. (2010). Development and validation of the Arizona Cognitive Test Battery for Down syndrome. *Journal of Neurodevelopmental Disorders*, *2*(3), 149-164. doi: 10.1007/s11689-010-9054-3
- Elwood, R. W. (2001). MicroCog: Assessment of cognitive functioning. *Neuropsychology Review*, *11*(2), 89-100. doi: 10.1023/a:1016671201211
- Espírito-Santo, H., Lemos, L., Ventura, L., Moitinho, S., Pinto, A. L., Rodrigues, F., ... Daniel, F. (2015). Teste da Figura Complexa de Rey-Osterrieth-A. In M. R. Simões & I. Santana (Eds.), *Escalas e testes na demência* (3ª ed., pp. 100-107). Porto Salvo, Portugal: Novartis.
- Facal, D., Rodríguez, N., Juncos-Rabadán, O., Caamaño, J. M., & Sueiro, J. (2009). Use of the Cambridge Neuropsychological Test Automated Battery for the diagnosis of mild cognitive impairment. A pilot study in a Spanish sample. *Revista*

- Española de Geriátria y Gerontología*, 44(2), 79-84. doi: 10.1016/j.regg.2008.07.004
- Falleti, M. G., Maruff, P., Collie, A., & Darby, D. G. (2006). Practice effects associated with the repeated assessment of cognitive function using the CogState Battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, 28(7), 1095-1112. doi: 10.1080/13803390500205718
- Firmino, H., Simões, M. R., Pinho, M. S., Cerejeira, L., & Martins, C. (2008). *O Exame Cognitivo de Addenbrooke - Revisto*. Coimbra, Portugal: Serviço de Avaliação Psicológica da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State. *Journal of Psychiatric Research*, 12(3), 189-198. doi: 10.1016/0022-3956(75)90026-6
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. *Applied Neuropsychology*, 2(2), 72-78. doi: 10.1207/s15324826an0202_4
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1997). Computerized neuropsychological tests in the early detection of dementia: Prospective findings. *Journal of the International Neuropsychological Society*, 3(2), 139-146.
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (2002). Paired associate performance in the early detection of DAT. *Journal of the International Neuropsychological Society*, 8(1), 58-71. doi: 10.1017.S1355617701020069
- Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry*, 11, 329-336. doi: 10.1002/(sici)1099-1166(199604)11:4%3c329::aid-gps453%3e3.3.co;2-y
- Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2011). Montreal Cognitive Assessment (MoCA): Normative study for the Portuguese population. *Journal of Clinical and Experimental Neuropsychology*, 33(9), 989-996. doi: 10.1080/13803395.2011.589374
- Gaugler, J. E., Duval, S., Anderson, K. A., & Kane, R. L. (2007). Predicting nursing home admission in the U.S.: A meta-analysis. *BMC Geriatrics*, 7(13), 1-14. doi: 10.1186/1471-2318-7-13
- Gonçalves, C., Pinho, M. S., Cruz, V., Gens, H., Oliveira, F., Pais, J., ... Santos, J. M. (2016). Portuguese version of Wechsler Memory Scale-3rd edition's utility with demented elderly adults. *Applied Neuropsychology: Adult*, 1-14. doi: 10.1080/23279095.2015.1135440
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016a). Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, 23(4), 251-263. doi: 10.1080/23279095.2015.1053889
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016b). Effects of socio-demographic variables on performance on the Cambridge neuropsychological automated tests for the assessment of dementia and Portuguese norms for older adults living in retirement homes. *The Clinical Neuropsychologist*, 30(2), 284-317. doi: 10.1080/13854046.2016.1156745

- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016c). *Construct and concurrent validity of the Cambridge neuropsychological automated tests for assessment of dementia in Portuguese older adults*. Manuscript submitted for publication.
- Gould, R. L., Brown, R. G., Owen, A. M., Bullmore, E. T., Williams, S. C., & Howard, R. J. (2005). Functional neuroanatomy of successful Paired Associate Learning in Alzheimer's disease. *American Journal of Psychiatry*, *162*(11), 2049-2060. doi: 10.1176/appi.ajp.162.11.2049
- Gualtieri, C. T., & Johnson, L. G. (2005). Neurocognitive testing supports a broader concept of mild cognitive impairment. *American Journal of Alzheimer's Disease and Other Dementias*, *20*(6), 359-366. doi: 10.1177/153331750502000607
- Guerreiro, M. (1998). *Contributo da neuropsicologia para o estudo das demências*. (Tese de doutoramento não publicada), Faculdade de Medicina de Lisboa, Lisboa, Portugal.
- Guerreiro, M. (2005). Avaliação neuropsicológica das demências degenerativas. In A. Castro-Caldas & A. Mendonça (Eds.), *A doença de Alzheimer e outras demências em Portugal* (pp. 83-109). Lisboa, Portugal: Lidel, edições técnicas, lda.
- Guerreiro, M., Fonseca, S., Barreto, J., & Garcia, C. (2008). Escala de Avaliação da Doença de Alzheimer (ADAS). In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* (2ª ed., pp. 41-68). Lisboa, Portugal: Novartis.
- Herrera-Guzmán, I., Gudayol-Ferré, E., Herrera-Guzmán, D., Guàrdia-Olmos, J., Hinojosa-Calvo, E., & Herrera-Abarca, J. E. (2009). Effects of selective serotonin reuptake and dual serotonergic–noradrenergic reuptake treatments on memory and mental processing speed in patients with major depressive disorder. *Journal of Psychiatric Research*, *43*(9), 855-863. doi: 10.1016/j.jpsychires.2008.10.015
- Herrera-Guzmán, I., Herrera-Abarca, J. E., Gudayol-Ferré, E., Herrera-Guzmán, D., Gómez-Carbajal, L., Peña-Olvira, M., ... Joan, G.-O. (2010). Effects of selective serotonin reuptake and dual serotonergic–noradrenergic reuptake treatments on attention and executive functions in patients with major depressive disorder. *Psychiatry Research*, *177*(3), 323-329. doi: 10.1016/j.psychres.2010.03.006
- Instituto Nacional de Estatística. (2001). *Óbitos pela doença de Alzheimer 2000*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=70753&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2005). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2005*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=73455&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2006). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2006*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=74055&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2007). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2007*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de

- http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=5582190&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2008). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2008*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.anacom.pt/streaming/inq_familias_ineumic2008.pdf?contentId=726858&field=ATTACHED_FILE.
- Instituto Nacional de Estatística. (2009). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2009*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=56910276&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2010). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2010*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=83386604&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2011). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2011*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=107940220&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2012a). *Censos 2011 resultados definitivos - Portugal*. Lisboa, Portugal: Autor. Recuperado de http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=ine_censos_publicacao_det&contexto=pu&PUBLICACOESpub_boui=73212469&PUBLICACOESmodo=2&selTab=tab1&pcensos=61969554.
- Instituto Nacional de Estatística. (2012b). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2012*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=133548146&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2013). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2013*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=206557735&att_display=n&att_download=y.
- Instituto Nacional de Estatística. (2014). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2014*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=211422735&DESTAQUEStema=00&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2015a). *Estatísticas demográficas 2014*. Lisboa, Portugal: Autor. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=139659&PUBLICACOESmodo=2.
- Instituto Nacional de Estatística. (2015b). *Causas de morte 2013*. Lisboa, Portugal: Autor. Recuperado de

http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOEspub_boui=229905695&PUBLICACOEStema=00&PUBLICACOESmodo=2.

- Instituto Nacional de Estatística. (2015c). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2015*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUE_Sdest_boui=224732374&DESTAQUESmodo=2.
- Iverson, G. L. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology*, 16(2), 183-191. doi: 10.1093/arclin/16.2.183
- Jack, C. R., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., ... Phelps, C. H. (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 257-262. doi: 10.1016/j.jalz.2011.03.004
- Jellinger, K. A. (2006). Clinicopathological analysis of dementia disorders in the elderly - An update. *Journal of Alzheimer's Disease*, 9(Suppl. 3), 61-70.
- Jones, G. M. M., Sahakian, B. J., Levy, R., Warburton, D. M., & Gray, J. A. (1992). Effects of acute subcutaneous nicotine on attention, information processing and short-term memory in Alzheimer's disease. *Psychopharmacology*, 108(4), 485-494. doi: 10.1007/bf02247426
- Juncos-Rabadán, O., Facal, D., Pereiro, A. X., & Lojo-Seoane, C. (2014). Visual memory profiling with CANTAB in mild cognitive impairment (MCI) subtypes. *International Journal of Geriatric Psychiatry*, 29(10), 1040-1048. doi: 10.1002/gps.4095
- Juncos-Rabadán, O., Pereiro, A. X., Facal, D., Reboredo, A., & Lojo-Seoane, C. (2014). Do the Cambridge Neuropsychological Test Automated Battery episodic memory measures discriminate amnesic mild cognitive impairment? *International Journal of Geriatric Psychiatry*, 29(6), 602-609. doi: 10.1002/gps.4042
- Junkkila, J., Oja, S., Laine, M., & Karrasch, M. (2012). Applicability of the CANTAB-PAL computerized memory test in identifying amnesic mild cognitive impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 34(2), 83-89. doi: 10.1159/000342116
- Kahana, E. (2006). Institutionalization. In R. Schulz (Ed.), *The encyclopedia of aging* (4th ed., Vol. I, pp. 595-600). New York, NY: Springer Publishing Company, Inc.
- Khan, U. A., Liu, L., Provenzano, F. A., Berman, D. E., Profaci, C. P., Sloan, R., ... Small, S. A. (2014). Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nature Neuroscience*, 17(2), 304-311. doi: 10.1038/nn.3606
- Kiddle, S. J., Steves, C. J., Mehta, M., Simmons, A., Xu, X., Newhouse, S., ... Dobson, R. J. B. (2015). Plasma protein biomarkers of Alzheimer's disease endophenotypes in asymptomatic older twins: early cognitive decline and regional brain volumes. *Translational Psychiatry*, 5(6), 584-589. doi: 10.1038/tp.2015.78
- Kim, C., Lee, J. Y., Ha, T. H., Choi, J. H., Yu, E.-M., Lee, S. J., ... Park, M. Y. (2009). The usefulness of the Cambridge Neuropsychological Test Automated Battery (CANTAB) for assessing cognitive functions in the elderly: A pilot study. *Journal of the Korean Geriatrics Society*, 13(2), 69-78. doi: 10.4235/jkgs.2009.13.2.69

- Klekociuk, S. Z., Summers, J. J., Vickers, J. C., & Summers, M. J. (2014). Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. *European Journal of Neurology*, *21*(10), 1330-1336. doi: 10.1111/ene.12488
- Klekociuk, S. Z., & Summers, M. J. (2013). The self-fulfilling prophecy of episodic memory impairment in mild cognitive impairment: Do episodic memory deficits identified at classification remain evident when later examined with different memory tests? *Neurology Research International*, *2013*, 1-6. doi: 10.1155/2013/437013
- Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. London, UK: Methuen & Co. Ltd.
- Larrabee, G. J., & Crook, T. H. (1996). Computers and memory. In L. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed., pp. 102-117). New York, NY: Oxford University Press.
- Lee, A. C. H., Rahman, S., Hodges, J. R., Sahakian, B. J., & Graham, K. S. (2003). Associative and recognition memory for novel objects in dementia: Implications for diagnosis. *European Journal of Neuroscience*, *18*(6), 1660-1670. doi: 10.1046/j.1460-9568.2003.02883.x
- Lesk, V. E., Honey, T. E., & de Jager, C. A. (2009). The effect of recent consumption of caffeine-containing foodstuffs on neuropsychological tests in the elderly. *Dementia and Geriatric Cognitive Disorders*, *27*(4), 322-328. doi: 10.1159/000207445
- Levy-Gigi, E., Kelemen, O., Gluck, M. A., & Kéri, S. (2011). Impaired context reversal learning, but not cue reversal learning, in patients with amnesic mild cognitive impairment. *Neuropsychologia*, *49*(12), 3320-3326. doi: 10.1016/j.neuropsychologia.2011.08.005
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford, NY: Oxford University Press.
- Lowe, C., & Rabbitt, P. (1998). Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. *Neuropsychologia*, *36*(9), 915-923. doi: 10.1016/S0028-3932(98)00036-0
- Maia, L., Correia, C., & Leite, R. (2009). *Avaliação e intervenção neuropsicológica: Estudos de casos e instrumentos*. Lisboa, Portugal: Lidel - edições técnicas, lda.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, *125*(3), 624-639. doi: 10.1093/brain/awf049
- Marôco, J. (2014a). *Análise de equações estruturais: Fundamentos teóricos, software e aplicações* (2ª ed.). Pêro Pinheiro, Portugal: Report Number - Análise e Gestão de Informação, Lda.
- Marôco, J. (2014b). *Análise estatística com o SPSS statistics* (6ª ed.). Pêro Pinheiro, Portugal: Report Number - Análise e Gestão de Informação, Lda.
- McGaughy, J., Dalley, J. W., Morrison, C. H., Everitt, B. J., & Robbins, T. W. (2002). Selective behavioral and neurochemical effects of cholinergic lesions produced by intrabasis infusions of 192 IgG-saporin on attentional performance in a five-choice serial reaction time task. *The Journal of Neuroscience*, *22*(5), 1905-1913.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., ... Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's

- Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263-269. doi: 10.1016/j.jalz.2011.03.005
- McSweeney, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). "T scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *Clinical Neuropsychologist*, 7(3), 300-312. doi: 10.1080/13854049308401901
- Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114(3), 449-458. doi: 10.1037/0033-2909.114.3.449
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21(11), 1078-1085. doi: 10.1002/gps.1610
- Mitchell, J., Arnold, R., Dawson, K., Nestor, P. J., & Hodges, J. R. (2009). Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm. *Journal of Neurology*, 256(9), 1500-1509. doi: 10.1007/s00415-009-5152-0
- Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin*, 19(3), 448-450.
- Moreira, P. I., & Oliveira, C. R. (2005). Fisiopatologia da doença de Alzheimer e de outras demências. In A. Castro-Caldas & A. Mendonça (Eds.), *A doença de Alzheimer e outras demências em Portugal* (pp. 41-60). Lisboa, Portugal: Lidel, edições técnicas, lda.
- Morgado, J., Rocha, C. S., Maruta, C., Guerreiro, M., & Martins, I. P. (2009). Novos valores normativos do Mini-Mental State Examination. *Sinapse*, 2(9), 10-16.
- Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 51(6), 757-766. doi: 10.1136/jnnp.51.6.757
- Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1987). Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer-type and Parkinson's disease. In S. M. Stahl, S. D. Iversen & E. C. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21-36). Oxford, UK: Oxford University Press.
- Murphy, F. C., Sahakian, B. J., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., & Paykel, E. S. (1999). Emotional bias and inhibitory control processes in mania and depression. *Psychological Medicine*, 29(6), 1307-1321. doi: 10.1017/s0033291799001233
- Need, A. C., Attix, D. K., McEvoy, J. M., Cirulli, E. T., Linney, K. L., Hunt, P., ... Goldstein, D. B. (2009). A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Human Molecular Genetics*, 18(23), 4650-4661. doi: 10.1093/hmg/ddp413
- Nihtilä, E. K., Martikainen, P. T., Koskinen, S. V. P., Reunanen, A. R., Noro, A. M., & Häkkinen, U. T. (2007). Chronic conditions and the risk of long-term institutionalization among older people. *The European Journal of Public Health*, 18(1), 77-84. doi: 10.1093/eurpub/ckm025

- Olde Dubbelink, K. T. E., Hillebrand, A., Twisk, J. W. R., Deijen, J. B., Stoffers, D., Schmand, B. A., ... Berendse, H. W. (2014). Predicting dementia in Parkinson disease by combining neurophysiologic and cognitive markers. *Neurology*, *82*(3), 263-270. doi: 10.1212/wnl.0000000000000034
- Oosterman, J. M., Vogels, R. L. C., van Harten, B., Gouw, A. A., Scheltens, P., Poggesi, A., ... Scherder, E. J. A. (2008). The role of white matter hyperintensities and medial temporal lobe atrophy in age-related executive dysfunctioning. *Brain and Cognition*, *68*(2), 128-133. doi: 10.1016/j.bandc.2008.03.006
- Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, *28*(10), 1021-1034. doi: 10.1016/0028-3932(90)90137-d
- Owen, A. M., Evans, A. C., & Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: A positron emission tomography study. *Cerebral Cortex*, *6*(1), 31-38. doi: 10.1093/cercor/6.1.31
- Periquito, C. M. N., Silva, P. C., Oliveira, P., Carneiro, C., Fernandes, A. I., & Costa, F. A. (2014). Revisão da medicação em idosos institucionalizados: Aplicação dos critérios STOPP e START. *Revista Portuguesa de Farmacoterapia*, *6*, 211-220.
- Plati, M. C. F., Covre, P., Lukasova, K., & Macedo, E. C. (2006). Depressive symptoms and cognitive performance of the elderly: Relationship between institutionalization and activity programs. *Revista Brasileira de Psiquiatria*, *28*(2), 118-121. doi: 10.1590/s1516-44462006000200008
- Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia*, *9*(1), 63-75. doi: 10.1016/j.jalz.2012.11.007
- Racinais, S., Gaoua, N., & Grantham, J. (2008). Hyperthermia impairs short-term memory and peripheral motor drive transmission. *The Journal of Physiology*, *586*(19), 4751-4762. doi: 10.1113/jphysiol.2008.157420
- Rahman, S., Robbins, T. W., & Sahakian, B. J. (1999). Comparative cognitive neuropsychological studies of frontal lobe function: Implications for therapeutic strategies in frontal variant frontotemporal dementia. *Dementia and Geriatric Cognitive Disorders*, *10*(Suppl. 1), 15-28. doi: 10.1159/000051207
- Rapp, M. A., Schnaider-Beeri, M., Sano, M., Silverman, J. M., & Haroutunian, V. (2005). Cross-domain variability of cognitive performance in very old nursing home residents and community dwellers: Relationship to functional status. *Gerontology*, *51*(3), 206-212. doi: 10.1159/000083995
- Rey, A. (2002). *Teste de Cópia de Figuras Complexas* (2ª ed.). Lisboa, Portugal: Cegoc-Tea.
- Riekkinen, M., Laakso, M. P., Jäkälä, P., & Riekkinen, P. (1999). Clonidine impairs sustained attention and memory in Alzheimer's disease. *Neuroscience*, *92*(3), 975-982. doi: 10.1016/s0306-4522(99)00070-6
- Riekkinen, P., & Riekkinen, M. (1999). THA improves word priming and clonidine enhances fluency and working memory in Alzheimer's disease. *Neuropsychopharmacology*, *20*(4), 357-364. doi: 10.1016/s0893-133x(98)00093-1
- Robbins, T. W., James, M., Lange, K. W., Owen, A. M., Quinn, N. P., & Marsden, C. D. (1992). Cognitive performance in multiple system atrophy. *Brain*, *115*(1), 271-291. doi: 10.1093/brain/115.1.271

- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): A factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitive Disorders*, 5(5), 266-281. doi: 10.1159/000106735
- Roiser, J. P., McLean, A., Ogilvie, A. D., Blackwell, A. D., Bamber, D. J., Goodyer, I., ... Sahakian, B. J. (2005). The subjective and cognitive effects of acute phenylalanine and tyrosine depletion in patients recovered from depression. *Neuropsychopharmacology*, 30, 775-785. doi: 10.1038/sj.npp.1300659
- Sahakian, B. J., & Coull, J. T. (1993). Tetrahydroaminoacridine (THA) in Alzheimer's disease: An assessment of attentional and mnemonic function using CANTAB. *Acta Neurologica Scandinavica*, 88(Suppl. 149), 29-35. doi: 10.1111/j.1600-0404.1993.tb04251.x
- Sahakian, B. J., Downes, J. J., Egger, S., Everden, J. L., Levy, R., Philpot, M. P., ... Robbins, T. W. (1990). Sparing of attentional relative to mnemonic function in a subgroup of patients with dementia of the Alzheimer type. *Neuropsychologia*, 28(11), 1197-1213. doi: 10.1016/0028-3932(90)90055-s
- Sahakian, B. J., Elliott, R., Low, N., Mehta, M., Clark, R. T., & Pozniak, A. L. (2009). Neuropsychological deficits in tests of executive function in asymptomatic and symptomatic HIV-1 seropositive men. *Psychological Medicine*, 25(6), 1233-1246. doi: 10.1017/s0033291700033201
- Sahakian, B. J., Jones, G., Levy, R., Gray, J., & Warburton, D. (1989). The effects of nicotine on attention, information processing, and short-term memory in patients with dementia of the Alzheimer type. *The British Journal of Psychiatry*, 154(6), 797-800. doi: 10.1192/bjp.154.6.797
- Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain*, 111(3), 695-718. doi: 10.1093/brain/111.3.695
- Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, 85(7), 399-402. doi: 10.1177/014107689208500711
- Sahgal, A., Lloyd, S., Wray, C. J., Galloway, P. H., Robbins, T. W., Sahakian, B. J., ... Edwardson, J. A. (1992). Does visuospatial memory in senile dementia of the Alzheimer type depend on the severity of the disorder? *International Journal of Geriatric Psychiatry*, 7(6), 427-436. doi: 10.1002/gps.930070608
- Sahgal, A., McKeith, I. G., Galloway, P. H., Tasker, N., & Steckler, T. (1995). Do differences in visuospatial ability between senile dementias of the Alzheimer and Lewy body types reflect differences solely in mnemonic function? *Journal of Clinical and Experimental Neuropsychology*, 17(1), 35-43. doi: 10.1080/13803399508406579
- Sales, F. (2005). Electroencefalograma. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 38-39). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I. (2005a). A doença de Alzheimer e outras demências: Diagnóstico diferencial. In A. Castro-Caldas & A. Mendonça (Eds.), *A doença de Alzheimer e outras demências em Portugal* (pp. 61-82). Lisboa, Portugal: Lidel, edições técnicas, lda.

- Santana, I. (2005b). Avaliação neuropsicológica. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 23-29). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I. (2005c). Estudos de imagem. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 40-42). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I., Baldeiras, I., & Oliveira, C. (2005). Exames laboratoriais. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 35-37). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I., Farinha, F., Freitas, S., Rodrigues, V., & Carvalho, A. (2015). Epidemiologia da demência e da doença de Alzheimer em Portugal: Estimativas da prevalência e dos encargos financeiros com a medicação. *Acta Médica Portuguesa*, 28(2), 182-188.
- Santana, I., & Oliveira, C. (2005). Estudos genéticos. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 43-47). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santiago, B. (2005). Avaliação psicológica. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 30-34). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Saunders, N. L. J., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 32(4), 350-357. doi: 10.1080/13803390903042379
- Saunders, N. L. J., & Summers, M. J. (2011). Longitudinal deficits to attention, executive, and working memory in subtypes of mild cognitive impairment. *Neuropsychology*, 25(2), 237-248. doi: 10.1037/a0021134
- Scoriels, L., Barnett, J. H., Murray, G. K., Cherukuru, S., Fielding, M., Cheng, F., ... Jones, P. B. (2011). Effects of modafinil on emotional processing in first episode psychosis. *Biological Psychiatry*, 69(5), 457-464. doi: 10.1016/j.biopsych.2010.09.043
- Sexton, C. E., McDermott, L., Kalu, U. G., Herrmann, L. L., Bradley, K. M., Allan, C. L., ... Ebmeier, K. P. (2012). Exploring the pattern and neural correlates of neuropsychological impairment in late-life depression. *Psychological Medicine*, 42(06), 1195-1202. doi: 10.1017/s0033291711002352
- Simões, M. R. (2012). Instrumentos de avaliação psicológica de pessoas idosas: Investigação e estudos de validação em Portugal. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 34(1), 9-33.
- Simões, M. R., & Santana, I. (2015). *Escalas e testes na demência* (3ª ed.). Porto Salvo, Portugal: Novartis.
- Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., ... Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 280-292. doi: 10.1016/j.jalz.2011.03.003
- Stefanova, E., Ziropadja, L., Stojkovic, T., Stankovic, I., Tomic, A., Jecmenica-Lukic, M., ... Kostic, V. (2015). Mild cognitive impairment in early Parkinson's disease using the movement disorder society task force criteria: Cross-sectional study in Hoehn and Yahr stage 1. *Dementia and Geriatric Cognitive Disorders*, 40(3-4), 199-209. doi: 10.1159/000433421

- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.
- Summers, M. J., & Saunders, N. J. (2012). Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. *Neuropsychology*, *26*(4), 498-508. doi: 10.1037/a0028576
- Swinson, R., Hodges, J. R., Galton, C. J., Semple, J., Michael, A., Dunn, B. D., ... Sahakian, B. J. (2001). Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks. *Dementia and Geriatric Cognitive Disorders*, *12*(4), 265-280. doi: 10.1159/000051269
- Taffe, M., Weed, M., Gutierrez, T., Davis, S., & Gold, L. (2002). Differential muscarinic and NMDA contributions to visuo-spatial paired-associate learning in rhesus monkeys. *Psychopharmacology*, *160*(3), 253-262. doi: 10.1007/s00213-001-0954-5
- Talpos, J. C., Winters, B. D., Dias, R., Saksida, L. M., & Bussey, T. J. (2009). A novel touchscreen-automated paired-associate learning (PAL) task sensitive to pharmacological manipulation of the hippocampus: A translational rodent model of cognitive impairments in neurodegenerative disease. *Psychopharmacology*, *205*(1), 157-168. doi: 10.1007/s00213-009-1526-3
- Turner, D. C., Clark, L., Dowson, J., Robbins, T. W., & Sahakian, B. J. (2004). Modafinil improves cognition and response inhibition in adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *55*(10), 1031-1040. doi: 10.1016/j.biopsych.2004.02.008
- Turner, D. C., Robbins, T. W., Clark, L., Aron, A. R., Dowson, J., & Sahakian, B. J. (2003). Cognitive enhancing effects of modafinil in healthy volunteers. *Psychopharmacology*, *165*, 260-269. doi: 10.1007/s00213-002-1250-8
- United Nations. (2015). *World population prospects: The 2015 revision, world population 2015 wallchart, ST/ESA/SER.A/378*. New York, NY: United Nations, Department of Economic and Social Affairs, Population Division. Retrieved from http://esa.un.org/unpd/wpp/publications/Files/World_Population_2015_Wallchart.pdf.
- Wechsler, D. (2008a). *Escala de Inteligência de Wechsler para Adultos – Terceira edição (WAIS-III): Manual técnico*. Lisboa, Portugal: Cegoc-Tea.
- Wechsler, D. (2008b). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual de administração e cotação*. Lisboa, Portugal: Cegoc-Tea.
- Wechsler, D. (2008c). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual técnico*. Lisboa, Portugal: Cegoc-Tea.
- Wild, K., Howieson, D., Webbe, F., Seelye, A., & Kaye, J. (2008). Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's & Dementia*, *4*(6), 428-437. doi: 10.1016/j.jalz.2008.07.003
- Wimo, A., Jönsson, L., Bond, J., Prince, M., & Winblad, B. (2013). The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia*, *9*(1), 1-11. doi: 10.1016/j.jalz.2012.11.006
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*(1), 37-49. doi: 10.1016/0022-3956(82)90033-4

Zygouris, S., & Tsolaki, M. (2014). Computerized cognitive testing for older adults: A review. *American Journal of Alzheimer's Disease and Other Dementias*, 30(1), 13-28. doi: 10.1177/1533317514522852

APÊNDICES

Apêndice A. Instruções de administração do teste Rastreio Motor (*Motor Screening, MOT*) para a versão ou modo de alta visibilidade.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
Com a imagem de início do MOT a ser exibida no ecrã, diga isto:	« <i>Aqui está o computador que iremos usar. Vou começar por lhe mostrar como é que ele deve ser usado. Está pronto(a)?</i> »
Prima a tecla “SPACE” para iniciar o teste e comece a apontar as cruzes que forem aparecendo no ecrã.	
Demonstre.	« <i>A ideia é tocar nas cruzes rosa e verde que vão aparecendo sucessivamente, como eu vou fazer agora. Se tocar nelas correctamente, ouvirá este som. Se tocar nelas incorrectamente, não ouvirá nada.</i> » « <i>Para tocar correctamente no ecrã, deve usar a ponta do dedo indicador da mão com que escreve. Se tocar correctamente na cruz, ela desaparecerá. Então, deverá afastar a mão do ecrã e esperar pela próxima cruz.</i> »
Após três cruzes o computador exibirá a mensagem “ <i>Aguarde</i> ”.	
Agora siga as seguintes instruções:	« <i>Tente agora, por favor.</i> »
Prima a tecla “SPACE”.	
Depois diga:	« <i>Lembre-se de tocar nas cruzes com a ponta do seu dedo indicador, quando elas aparecerem.</i> »
Se a pessoa examinada tocar correctamente nas cruzes mas elas não desaparecerem, diga:	« <i>Toque com mais firmeza.</i> »
Se a pessoa examinada ficar confusa com o procedimento ou não tocar nas cruzes adequadamente, pode dar as seguintes ajudas:	« <i>Toque com a ponta do seu dedo indicador. Agora procure a próxima cruz (que aparecer).</i> » « <i>Toque apenas, não há necessidade de premir com força.</i> » « <i>Toque apenas com um dedo (o indicador).</i> »
Se o ecrã não responder, peça à pessoa examinada para afastar o dedo do ecrã e depois tocar novamente. Tocar com força não fará com que o ecrã funcione.	« <i>Afaste o dedo do ecrã. Agora toque novamente.</i> » « <i>Não é necessário tocar com força.</i> »
O teste termina após a pessoa examinada tocar na décima cruz.	

Nota. Se o(a) examinador(a) não utilizar teclado, ao invés de premir “SPACE”, deve tocar uma vez em qualquer um dos cantos inferiores do ecrã. Se o(a) examinador(a) necessitar de interromper ou abortar o teste, deve premir “ESCAPE” ou dar três toques rápidos em qualquer um dos cantos superiores do ecrã.

Apêndice B. Instruções de administração do teste Processamento Rápido de Informação Visual (*Rapid Visual Information Processing*, RVP) para a versão ou modo 3-5-7.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
FASE DE TREINO	
Quando a imagem de início do RVP estiver a ser exibida no ecrã, prima a tecla “SPACE”. O computador irá apresentar a palavra “ <i>Aguarde</i> ”, com a sequência 3-5-7 à direita.	
Agora dê a seguinte instrução:	<i>«Dentro de uma caixa branca, que vai estar no centro do ecrã, vai ver alguns números a aparecerem, um a um. O que tem de fazer é procurar uma sequência alvo de três dígitos e premir o botão sempre que detecte essa sequência.»</i>
	<i>«A sequência alvo vai ser um “3”, imediatamente seguido de um “5”, imediatamente seguido de um “7”. Só deve premir o botão quando tiver visto o último número dessa sequência.»</i>
	<i>«Para começar, vai saber que a sequência começou porque a sequência vai aparecer a vermelho e sublinhada a amarelo. Se pressionar o botão correctamente, também vai ouvir um som tipo apito. À medida que a sessão de treino for avançando, verá que essas pistas vão desaparecendo progressivamente.»</i>
Prima a tecla “SPACE” para começar a fase de treino.	
Se a pessoa examinada responder muito cedo, dizer:	<i>«Espere até ver o último número da sequência, antes de pressionar o botão.»</i>
O canto superior esquerdo do ecrã apresenta quatro caixinhas coloridas de inspecção. As caixas vermelha e verde desaparecem ao mesmo tempo quando as sequências deixam de ser sublinhadas a vermelho e quando a ajuda “ <i>Prima agora</i> ” já não for exibida.	
Quando isto acontecer, diga:	<i>«Agora as sequências alvo vão aparecer apenas sublinhadas a amarelo.»</i>
As caixinhas amarela e azul desaparecem ao mesmo tempo quando as sequências deixam de ser sublinhadas, os apitos param de ser emitidos e as mensagens de <i>feedback</i> “ <i>Bom</i> ”, “ <i>Errado</i> ” ou “ <i>Demasiado tarde</i> ” já não forem exibidas.	

(O Apêndice B continua)

Apêndice B (continuação). Instruções de administração do teste Processamento Rápido de Informação Visual (*Rapid Visual Information Processing*, RVP) para a versão ou modo 3-5-7.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
Quando as caixinhas de inspeção amarela e azul também desaparecerem, diga: Ao fim de dois minutos, a fase de treino terminará e o computador exibirá novamente a mensagem “ <i>Aguarde</i> ”.	« <i>Agora terá de detectar as sequências sozinho(a). Não vai haver mais sublinhados nem apitos.</i> »

FASE DE TESTE

Quando no ecrã se ler “*Aguarde*”, dê as seguintes instruções:

«*Desta vez vamos fazer a mesma coisa. Sempre que veja a sequência 3-5-7, deve premir o botão quando vir o último número da sequência a aparecer.*»

«*O teste vai durar quatro minutos, por favor, tente concentrar-se até ao fim. A sequência alvo vai continuar no ecrã, para o(a) ajudar a recordar-se dela. Contudo, procure concentrar-se na caixa onde os números estão sempre a mudar.*»

«*Por favor, responda o mais depressa que conseguir, ao mesmo tempo tentando evitar cometer erros. Lembre-se que não vai haver nem cor, nem sublinhado, nem apitos.*»

Quando a pessoa examinada estiver pronta, certifique-se que o bloco de interruptores ainda está em posição e depois prima a tecla “SPACE” para começar a correr esta fase de teste.

Após quatro minutos, o teste terminará.

Nota. Este teste requer a utilização do bloco de interruptores. Certifique-se que o bloco está em frente ao ecrã, no sentido perpendicular, e que a pessoa examinada consegue alcançá-lo confortavelmente, com o dedo indicador da mão dominante apoiado com suavidade no botão de resposta. Se o(a) examinador(a) não utilizar teclado, ao invés de premir “SPACE”, deve tocar uma vez em qualquer um dos cantos inferiores do ecrã. Se o(a) examinador(a) necessitar de interromper ou abortar o teste, deve premir “ESCAPE” ou dar três toques rápidos em qualquer um dos cantos superiores do ecrã.

Apêndice C. Instruções de administração do teste Aprendizagem de Pares Associados (*Paired Associates Learning*, PAL) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
UM PADRÃO (ESTÁDIOS 1 E 2)	
Entre neste teste antes de explicar as instruções.	
Com a imagem de início do PAL a ser exibida no ecrã, prima a tecla “SPACE” para começar o primeiro estágio de um padrão.	
Agora diga rapidamente:	<i>«Neste teste irá ver seis caixas brancas que se irão abrir, numa ordem ao acaso.»</i>
	<i>«Dentro de uma das caixas, vai encontrar um padrão. Terá de recordar em que caixa é que ele estava.»</i>
Quando o padrão aparecer diga:	<i>«Aqui está o padrão. Tente recordar-se em que caixa estava.»</i>
Certifique-se que a pessoa examinada olha para cada caixa. Quando todas as caixas se tiverem aberto, o padrão aparecerá no centro do ecrã.	
Diga:	<i>«Toque na caixa onde viu o padrão aparecer.»</i>
<ul style="list-style-type: none">• No caso da pessoa examinada acertar:	
Se a pessoa examinada tocar na caixa correcta, o ecrã apresentará a mensagem “ <i>Tudo correcto</i> ”.	
Diga:	<i>«Agora vamos fazer a mesma coisa, mas desta vez com um padrão diferente.»</i>
Após um momento, passe para o estágio de um padrão seguinte, premindo “SPACE”.	
<ul style="list-style-type: none">• No caso da pessoa examinada errar:	
Se a pessoa examinada tiver cometido um erro durante um estágio, o ecrã <i>não</i> exibirá “ <i>Tudo correcto</i> ” logo de imediato; em vez disso, o computador irá fazer uma pausa e depois repetir a apresentação do(s) padrão(ões).	
Durante a pausa, diga:	<i>«Não era bem essa. O computador vai abrir as caixas outra vez para o(a) lembrar onde é que os padrões estavam realmente.»</i>

(O Apêndice C continua)

Apêndice C (continuação 1). Instruções de administração do teste Aprendizagem de Pares Associados (*Paired Associates Learning*, PAL) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada																						
<p>Para ajudar o(a) examinador(a) a saber quantos ensaios a pessoa examinada já completou naquele estágio, aparece no canto inferior direito do ecrã um número de dois dígitos, a cinzento, no momento em que o padrão é exibido ao centro. A diferença entre os dois dígitos corresponde ao número do ensaio em curso:</p> <table><thead><tr><th>Número exibido</th><th>Número de ensaio</th></tr></thead><tbody><tr><td>Nenhum</td><td>1</td></tr><tr><td>53</td><td>2</td></tr><tr><td>96</td><td>3</td></tr><tr><td>84</td><td>4</td></tr><tr><td>72</td><td>5</td></tr><tr><td>93</td><td>6</td></tr><tr><td>70</td><td>7</td></tr><tr><td>91</td><td>8</td></tr><tr><td>90</td><td>9</td></tr><tr><td>00</td><td>10</td></tr></tbody></table>	Número exibido	Número de ensaio	Nenhum	1	53	2	96	3	84	4	72	5	93	6	70	7	91	8	90	9	00	10	
Número exibido	Número de ensaio																						
Nenhum	1																						
53	2																						
96	3																						
84	4																						
72	5																						
93	6																						
70	7																						
91	8																						
90	9																						
00	10																						
<p>(Se a pessoa examinada falhar as dez tentativas de qualquer um dos estádios, o teste termina).</p> <p>Terminado o segundo estágio de um padrão com sucesso, o computador exibirá “Tudo correcto. 2 novos padrões”.</p> <p>DOIS PADRÕES (ESTÁDIOS 3 E 4)</p> <p>Informe a pessoa examinada que neste estágio há dois padrões, dizendo:</p> <p>Prima “SPACE” para continuar. À medida que os padrões forem aparecendo, aponte-os e diga:</p> <p>Quando todas as caixas se tiverem aberto, os padrões irão aparecer no centro do ecrã, um a um.</p> <p>Assim que cada um dos dois padrões apareça no centro do ecrã, dê a seguinte instrução:</p> <p>Pode incentivar a pessoa examinada com a seguinte frase:</p> <ul style="list-style-type: none">• No caso da pessoa examinada errar: Tal como antes, se a pessoa examinada escolher uma caixa errada, diga:	<p>«<i>Desta vez vai haver dois padrões a recordar. Observe cada caixa a abrir-se. Terá de recordar que padrão estava em que caixa.</i>»</p> <p>«<i>Aqui está um padrão... e aqui está o outro padrão.</i>»</p> <p>«<i>Então, em que caixa estava este padrão? E em que caixa estava este padrão?</i>»</p> <p>«<i>Toque simplesmente na caixa em que pensa que o padrão estava.</i>»</p> <p>«<i>Não era bem essa. O computador vai abrir as caixas outra vez para o(a) relembrar onde é que os padrões estavam realmente.</i>»</p>																						

(O Apêndice C continua)

Apêndice C (continuação 2). Instruções de administração do teste Aprendizagem de Pares Associados (*Paired Associates Learning*, PAL) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
<ul style="list-style-type: none">• No caso da pessoa examinada acertar: Se a pessoa examinada tiver escolhido ambas as caixas correctamente, o ecrã exibirá “Tudo correcto”.	<p>«Agora vamos fazer a mesma coisa, mas desta vez com dois padrões diferentes.»</p>
TRÊS PADRÕES (ESTÁDIOS 5 E 6)	
Se a pessoa examinada tiver terminado o segundo estágio de dois padrões com sucesso, o ecrã exibirá “Tudo correcto. 3 novos padrões” .	
Prima “SPACE” para continuar e diga:	«Desta vez vai haver três padrões.»
<ul style="list-style-type: none">• No caso da pessoa examinada errar: Tal como antes, se a pessoa examinada escolher uma caixa errada, diga:	«Não era bem essa. O computador vai abrir as caixas outra vez para o(a) lembrar onde é que os padrões estavam realmente.»
<ul style="list-style-type: none">• No caso da pessoa examinada acertar: Se o primeiro estágio de três padrões for terminado com sucesso, prima “SPACE” para passar ao segundo estágio de três padrões. Se o segundo estágio de três padrões for terminado com sucesso, o ecrã exibirá “Tudo correcto. 6 novos padrões”.	«Mais uma vez, vai haver três padrões diferentes para ver onde estão.»
SEIS PADRÕES (ESTÁDIO 7)	
Neste estágio de seis padrões (há apenas um estágio para este nível), é possível que até mesmo os indivíduos muito bons comecem a cometer erros. A pessoa examinada deve ser avisada que tem de recordar a localização de todos os padrões num mesmo ensaio (ou tentativa), embora disponha de mais do que um ensaio para o fazer.	
Agora diga:	«Desta vez, todas as seis caixas irão conter um padrão. Terá que recordar que padrão estava em que caixa. É possível que não acerte tudo logo à primeira vez, mas terá mais oportunidades, à medida que for tentando. Procure fazer o melhor que for capaz.»
Agora prima “SPACE”.	

(O Apêndice C continua)

Apêndice C (continuação 3). Instruções de administração do teste Aprendizagem de Pares Associados (*Paired Associates Learning*, PAL) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
<ul style="list-style-type: none">• No caso da pessoa examinada errar: Tal como antes, se a pessoa examinada escolher uma caixa errada, diga:	<i>«Não era bem essa. O computador vai abrir as caixas outra vez para o(a) lembrar onde é que os padrões estavam realmente.»</i>
<ul style="list-style-type: none">• No caso da pessoa examinada acertar: Se o estágio for terminado com sucesso, o ecrã exibirá “Tudo correcto. Último grupo de padrões”.	
OITO PADRÕES (ESTÁDIO 8)	
Diga isto:	<i>«Agora vai haver oito caixas. Cada caixa irá conter um padrão. Terá que recordar que padrão estava em que caixa. É possível que não acerte tudo logo à primeira vez, mas terá mais oportunidades, à medida que for tentando. Procure fazer o melhor que for capaz.»</i>
Agora prima “SPACE”.	
<ul style="list-style-type: none">• No caso da pessoa examinada errar: Tal como antes, se a pessoa examinada escolher uma caixa errada, diga:	<i>«Não era bem essa. O computador vai abrir as caixas outra vez para o(a) lembrar onde é que os padrões estavam realmente.»</i>
O teste termina no final deste estágio ou mais cedo, se a pessoa examinada tiver cometido dez erros em qualquer estágio anterior.	

Nota. Se o(a) examinador(a) não utilizar teclado, ao invés de premir “SPACE”, deve tocar uma vez em qualquer um dos cantos inferiores do ecrã. Se o(a) examinador(a) necessitar de interromper ou abortar o teste, deve premir “ESCAPE” ou dar três toques rápidos em qualquer um dos cantos superiores do ecrã.

Apêndice D. Instruções de administração do teste Tempo de Reacção (*Reaction Time*, RTI) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
ESTÁDIO 1: APONTAR O CÍRCULO (FASE DE TREINO)	
Com a imagem de início do RTI a ser exibida no ecrã, diga:	<i>«Dentro de um círculo branco vai aparecer uma bola amarela. Toque no círculo, o mais depressa que conseguir, logo após a bola amarela aparecer e depois volte a colocar a sua mão na mesa, em frente ao ecrã.»</i>
Prima a tecla “SPACE” para começar o primeiro estágio.	
<ul style="list-style-type: none">• Se a pessoa examinada tocar <i>demasiado cedo</i>, ajude-o(a) dizendo:• Se a pessoa examinada tocar <i>demasiado tarde</i>, ajude-o(a) dizendo:	<i>«Procure esperar até que a bola amarela apareça.»</i> <i>«Procure tocar um pouco mais depressa.»</i>
Todas as vezes que a pessoa examinada tocar correctamente no círculo, o computador irá emitir um apito e exibir a palavra “ Bom ”. Se a pessoa examinada tocar fora do círculo, o ecrã exibirá “ Impreciso ”.	
No final deste estágio o computador fará uma pausa e exibirá “ Aguarde ”.	
ESTÁDIO 2: APONTAR COM CINCO ESCOLHAS (FASE DE TREINO)	
Diga isto:	<i>«Agora vamos continuar a tocar no ecrã mas com cinco escolhas. A bola amarela pode agora aparecer em qualquer um de cinco círculos. Toque no círculo onde vir a bola amarela aparecer.»</i>
Prima “SPACE” para começar o estágio 2.	
Dê a seguinte ajuda se a pessoa examinada parecer não ter compreendido:	<i>«A bola amarela pode aparecer em qualquer um dos cinco círculos. Toque no círculo em que a bola amarela aparecer, o mais depressa que conseguir.»</i>
<ul style="list-style-type: none">• Se a pessoa examinada apontar o círculo errado, o ecrã exibirá “Errado”.• Se a pessoa examinada tocar fora do círculo, o ecrã exibirá “Impreciso”.	
No final deste estágio o computador fará uma pausa e exibirá “ Aguarde ”.	

(O Apêndice D continua)

Apêndice D (continuação 1). Instruções de administração do teste Tempo de Reacção (Reaction Time, RTI) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
ESTÁDIO 3: ESCOLHA SIMPLES COM LARGAR O BOTÃO (FASE DE TREINO)	
Dê o bloco de interruptores à pessoa examinada e mostre-lhe como funciona. Deixe-o(a) manipulá-lo até que se sinta confortável com ele. Certifique-se que a pessoa examinada está a premir o botão com a mesma mão com que tocou no ecrã.	
Depois diga:	<i>«A bola amarela irá aparecer dentro do círculo, instantes após ter premido o botão. Logo que veja a bola amarela aparecer, largue o botão, o mais depressa que conseguir. Não largue o botão antes de ver a bola amarela aparecer. Pode começar agora.»</i>
Prima “SPACE” para começar o estádio 3.	
Pode ajudar a pessoa examinada dizendo:	<i>«Pressione o botão, para baixo.»</i> <i>«Use a mesma mão para premir o botão e tocar no ecrã.»</i>
Se a pessoa examinada largar o botão antes de aparecer a bola amarela, o ecrã exibirá “Demasiado cedo”. Diga:	<i>«Tente não largar o botão antes de ver a bola amarela.»</i>
No final deste estádio o computador fará uma pausa e exibirá “Aguarde”.	
ESTÁDIO 4: ESCOLHA SIMPLES COM LARGAR O BOTÃO E APONTAR (FASE DE TESTE)	
Diga isto:	<i>«A partir de agora, quando largar o botão, toque no círculo o mais depressa que conseguir. Lembre-se de não largar o botão antes de ver a bola amarela e que desta vez tem de tocar no círculo.»</i>
Prima “SPACE” para começar o estádio 4.	
Certifique-se que a pessoa examinada está a usar a mesma mão para premir o botão e tocar no ecrã.	
No final deste estádio o computador fará uma pausa e exibirá “Aguarde”.	

(O Apêndice D continua)

Apêndice D (continuação 2). Instruções de administração do teste Tempo de Reacção (Reaction Time, RTI) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
ESTÁDIO 5: LARGAR O BOTÃO COM CINCO ESCOLHAS E APONTAR (FASE DE TESTE)	
Diga isto:	<i>«Bom. Agora vai haver novamente cinco círculos. Lembre-se, toque no círculo onde vir a bola amarela aparecer, o mais depressa que conseguir, mas não largue o botão antes de ver a bola amarela.»</i>
Prima “SPACE” para começar o estágio 5.	
O teste termina no final deste estágio.	
<i>Nota.</i> Este teste requer a utilização do bloco de interruptores. Certifique-se que o bloco está em frente ao ecrã, no sentido perpendicular, e que a pessoa examinada consegue alcançá-lo confortavelmente, com o dedo indicador da mão dominante apoiado com suavidade no botão de resposta. Neste teste, a mão que toca no botão de resposta do bloco de interruptores É A MESMA que deve tocar no ecrã. Se o(a) examinador(a) não utilizar teclado, ao invés de premir “SPACE”, deve tocar uma vez em qualquer um dos cantos inferiores do ecrã. Se o(a) examinador(a) necessitar de interromper ou abortar o teste, deve premir “ESCAPE” ou dar três toques rápidos em qualquer um dos cantos superiores do ecrã.	

Apêndice E. Instruções de administração do teste Memória de Trabalho Espacial (*Spatial Working Memory, SWM*) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
Quando o teste se inicia, são apresentadas no ecrã três caixas vermelhas e uma coluna negra intitulada “casa”, situada no lado direito do ecrã. O(A) examinador(a) deve referir-se aos quadrados coloridos como sendo “caixas” para dar a ideia que as fichas azuis podem estar escondidas dentro delas. Demonstre a tarefa com o primeiro ensaio e depois, no segundo ensaio, deixe a pessoa examinada praticar.	
CONJUNTO 1: TRÊS CAIXAS COLORIDAS	
Com a imagem de início do SWM a ser exibida no ecrã, prima a tecla “SPACE” (no teclado) para apresentação da tarefa do SWM e diga:	<i>«Neste teste irá ver algumas caixas coloridas no ecrã. O que tem de fazer é procurar, em cada uma delas, uma ficha azul que o computador escondeu numa das caixas. Em cada busca que faça, haverá apenas uma ficha escondida. Terá de recolher um número suficiente de fichas azuis para preencher esta “casa”, no lado direito do ecrã. Para procurar no interior de uma caixa tem simplesmente de lhe tocar assim.»</i>
Toque na caixa situada no centro do ecrã. (Use esta caixa para evitar alertar a pessoa examinada para qualquer tipo de estratégia de organização sequencial).	
	<i>«Esta caixa não tem uma ficha azul lá dentro, pelo que vou tentar uma outra caixa.»</i>
Toque noutra caixa.	
	<i>«Esta caixa tem uma ficha azul lá dentro pelo que agora vou movê-la para aqui, para encher a casa.»</i>
Toque na coluna negra, no lado direito do ecrã.	
A instrução seguinte é de importância extrema e deve ser fortemente enfatizada.	
	<i>«Se voltar a procurar uma ficha azul nesta caixa, como agora vou fazer...»</i>
Toque na caixa onde encontrou a ficha azul.	

(O Apêndice E continua)

Apêndice E (continuação 1). Instruções de administração do teste Memória de Trabalho Espacial (*Spatial Working Memory, SWM*) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
	<i>«...nunca mais aqui irei encontrar uma nova ficha azul, pelo que não devo cá voltar. Há ainda mais duas fichas azuis para procurar, mas o computador nunca usa, mais de uma vez, uma mesma caixa para esconder uma ficha azul. Logo, devo sempre tocar numa outra caixa que não a que já teve uma ficha azul.»</i>
Toque numa outra caixa.	<i>«Não há qualquer ficha dentro desta caixa pelo que vou tentar outra.»</i>
Toque na caixa onde ainda não encontrou uma ficha, para mostrar a ficha azul, e depois toque no lado direito do ecrã.	<i>«Agora que já encontrei duas fichas, a última tem de estar aqui.»</i>
Toque na única caixa onde ainda não encontrou uma ficha, para mostrar a ficha azul, e depois toque no lado direito do ecrã.	<i>«Agora encontrei todas as três fichas e acabei porque enchi completamente a “casa” com as fichas azuis.»</i>
O computador irá exibir “ Concluído ” e emitir uma breve melodia. Após uma pausa, as palavras “ Novo conjunto ” irão surgir no ecrã.	
Agora diga:	<i>«Tente agora, por favor.»</i>
CONJUNTO 2	
Prima a tecla “SPACE” e três caixas amarelas irão aparecer no ecrã.	<i>«Toque numa das caixas para procurar a ficha azul.»</i>
A pessoa examinada deverá tocar numa das caixas que se irá abrir para revelar a ficha azul.	<i>«Agora tem uma ficha azul para encher a “casa”.»</i>
A pessoa examinada deverá transferir a ficha tocando na “casa”.	<i>«Procure outra ficha mas lembre-se que, desta vez, ela irá estar numa caixa diferente.»</i>

(O Apêndice E continua)

Apêndice E (continuação 2). Instruções de administração do teste Memória de Trabalho Espacial (*Spatial Working Memory*, SWM) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
A pessoa examinada deverá tocar numa das outras caixas e, desta vez, não irá encontrar a ficha azul.	
	« <i>Experimente noutra.</i> »
A pessoa examinada deverá tocar na caixa que falta, a qual ainda não foi pesquisada, e desta vez irá encontrar a ficha azul.	
	« <i>Bom. Agora encontrou a segunda ficha azul.</i> »
Se necessário, diga:	« <i>Agora coloque a ficha na “casa”.</i> » « <i>Agora já só tem mais uma ficha para procurar.</i> »
A pessoa examinada deverá agora tocar na caixa na qual ainda não foi encontrada uma ficha azul e usar essa ficha para encher a coluna negra.	
Se a pessoa examinada tocar na caixa correcta, dê-lhe encorajamento.	
Se a pessoa examinada cometer um erro, diga:	« <i>Lembre-se de não voltar a uma caixa onde já encontrou uma ficha azul, pois o computador nunca mais volta a escondê-la nessa caixa.</i> »
CONJUNTO 3	
Prima “SPACE” para passar ao ensaio seguinte.	
Para o conjunto 3, repita o essencial das instruções usadas nos ensaios de treino, conforme se segue (contudo, se a pessoa examinada tiver entendido claramente as instruções, deixe-o[a] prosseguir sem ajudas).	
Ajude a pessoa examinada, se esta regressar a uma caixa onde já tinha encontrado uma ficha azul, dizendo:	« <i>Lembre-se de não voltar a uma caixa onde já encontrou uma ficha azul, pois o computador nunca mais volta a escondê-la nessa caixa.</i> »
Agora diga:	« <i>Procure uma ficha azul... Experimente noutra... Aí está – coloque-a na “casa” e procure outra ficha azul... Ela vai estar numa caixa diferente daquela onde encontrou a outra ficha... Coloque-a à direita, para encher a “casa”.</i> »
Dê a seguinte ajuda, se a pessoa examinada abrir uma caixa vazia:	« <i>Pouca sorte, tente outra vez.</i> »
Dê a seguinte ajuda se a pessoa examinada abrir uma caixa com uma ficha azul:	« <i>Bom. Agora ponha-a na “casa” e procure as outras duas fichas.</i> »

O Apêndice E continua)

Apêndice E (continuação 3). Instruções de administração do teste Memória de Trabalho Espacial (*Spatial Working Memory*, SWM) para a versão ou modo clínico.

Instruções para o(a) examinador(a)	Instruções a ler à pessoa examinada
Haverá ainda mais um conjunto de três caixas, para treino, e depois o computador passará para os ensaios de teste que incluirão quatro ensaios com 4 caixas, quatro ensaios com 6 caixas e quatro ensaios com 8 caixas.	
	«Agora há <i>N</i> caixas e <i>N</i> fichas azuis a procurar.»
A seguir à primeira ficha azul ser descoberta, dê a instrução:	«Ainda tem <i>N-1</i> fichas para procurar.»
A partir daqui, evite relembrar a pessoa examinada das fichas que ainda tem para procurar.	
Use as seguintes ajudas:	«Coloque-a à direita/na “casa”.» «Procure outra.»
Após o quarto ensaio com 8 caixas, o teste termina.	

Nota. Se o(a) examinador(a) não utilizar teclado, ao invés de premir “SPACE”, deve tocar uma vez em qualquer um dos cantos inferiores do ecrã. Se o(a) examinador(a) necessitar de interromper ou abortar o teste, deve premir “ESCAPE” ou dar três toques rápidos em qualquer um dos cantos superiores do ecrã.

Apêndice F. Dados demográficos da amostra de pessoas mais velhas a viver em residências geriátricas constituída para o cálculo dos valores normativos dos testes Fluência Verbal Semântica (animais) e Nomeação de Objectos e Dedos.

<i>N</i> = 151	<i>n</i>	%	<i>M</i>	<i>DP</i>	Intervalo
Idade (anos)			81.98	5.65	70-96
Escolaridade (anos de estudo)			8.65	4.91	4-22
Género					
Feminino	93	62			
Masculino	58	38			
Consumo de psicotrópicos					
Não consumidores	79	52			
Consumidores	72	48			
Tempo de institucionalização (meses)			45.99	44.85	3-232
ACE-R					
Pontuação MMSE			28.48	1.38	25-30
Pontuação total			86.07	8.54	59-100
GDS-30			8.79	5.26	0-20
	<i>N</i>	<i>M</i>	<i>EPM</i>	<i>DP</i>	Intervalo
Fluência Verbal ^a (total de nomes de animais)					
Grupo “idade 70-79, escolaridade 4-9”	30	13.63	0.69	3.79	9-25
Grupo “idade 70-79, escolaridade 10+”	24	19.13	0.83	4.05	10-26
Grupo “idade 80+, escolaridade 4-9”	62	12.69	0.45	3.52	6-22
Grupo “idade 80+, escolaridade 10+”	35	14.71	0.78	4.60	6-26
Amostra total	151	14.37	0.36	4.48	6-26
Nomeação de Objectos e Dedos ^b (erros)					
Grupo “idade 70-79, escolaridade 4-9”	28	1.14	0.18	0.97	0-3
Grupo “idade 70-79, escolaridade 10+”	24	0.46	0.12	0.59	0-2
Grupo “idade 80+, escolaridade 4-9”	59	1.19	0.14	1.06	0-4
Grupo “idade 80+, escolaridade 10+”	35	0.51	0.11	0.66	0-2
Amostra total	146	0.90	0.08	0.94	0-4

Nota. ^aItem do Exame Cognitivo de Addenbrooke Revisto (Addenbrooke’s Cognitive Examination Revised, ACE-R). ^bTarefa da Escala de Avaliação da Doença de Alzheimer (Alzheimer’s Disease Assessment Scale, ADAS). Legenda: EPM = erro padrão da média; GDS-30 = Geriatric Depression Scale (Escala de Depressão Geriátrica, 30 itens); MMSE = Mini-Mental State Examination (Exame Breve do Estado Mental).

Capítulo 1.

TEST-RETEST RELIABILITY ANALYSIS OF THE CAMBRIDGE NEUROPSYCHOLOGICAL AUTOMATED TESTS FOR THE ASSESSMENT OF DEMENTIA IN OLDER PEOPLE LIVING IN RETIREMENT HOMES

Marta Matos Gonçalves, Maria Salomé Pinho, and Mário R. Simões¹

ABSTRACT

The validity of the Cambridge Neuropsychological Automated Tests has been widely studied, but their reliability has not. This study aimed to estimate the test-retest reliability of these tests in a sample of 34 older adults, aged 69 to 90 years old, without neuropsychiatric diagnoses and living in retirement homes in the district of Lisbon, Portugal. The battery was administered twice, with a 4-week interval between sessions. The Paired Associates Learning (PAL), Spatial Working Memory (SWM), Rapid Visual Information Processing, and Reaction Time tests revealed measures with high-to-adequate test-retest correlations (.71-.89), although several PAL and SWM measures showed susceptibility to practice effects. Two estimated standardized regression-based methods were found to be more efficient at correcting for practice effects than a method of fixed correction. We also found weak test-retest correlations (.56-.68) for several measures. These results suggest that some, but not all, measures are suitable for cognitive assessment and monitoring in this population.

Keywords: Alzheimer disease; CANTAB; practice effects; reliable change index; standardized regression-based models.

¹ Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016). Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, 23(4), 251-263. doi: 10.1080/23279095.2015.1053889. Artigo transcrito segundo as normas da revista.

INTRODUCTION

The majority of neuropsychological tests for the assessment of older people have been validated for individuals living in the community. However, dementia is a chronic illness that is strongly associated with entrance into institutional care (Nihtilä et al., 2007). Because institutions for older people can provide high-to-low levels of assistance and care (Ribbe et al., 1997), they are also a late-life option for healthy individuals who, for example, wish to continue living with their spouse who is in need of care, want to avoid living alone, or simply want to ensure future assistance. Therefore, validated neuropsychological tests for older people living in institutions such as retirement homes are required to ensure the proper detection and monitoring of cognitive deficits in this population.

Computerized neuropsychological tests provide advantages over the more traditional pencil-and-paper tests, such as automatic administration, recording, and scoring, which increase intraexaminer and interexaminer reliability and data accuracy. Computers can also generate multiple alternate test forms and help to reduce practice effects. Thus, they are more suitable for monitoring than pencil-and-paper tests with only one or two test forms. Although lower familiarity with computers can influence test performance (Iverson, Brooks, Ashton, Johnson, & Gualtieri, 2009), empirical evidence has shown that older people report higher acceptance and lower stress for computerized tests than for pencil-and-paper tests (Collerton et al., 2007). In Portugal, the National Institute of Statistics (Instituto Nacional de Estatística, 2008, 2013) reported that computer use increased from 6% in 2008 to 20% in 2013 among those aged 65 to 74 years old. This rapid technological adoption suggests that pencil-and-paper tests will soon be replaced by computerized tests, making the study of computerized tests especially relevant.

The Cambridge Neuropsychological Test Automated Battery (CANTAB[®], <http://www.cambridgecognition.com>) is one of the most commonly used computerized batteries worldwide. Created in 1987 (Morris, Evenden, Sahakian, & Robbins, 1987), it has continuously incorporated new tests (Cambridge Cognition, 2006a, 2006b). Four of these tests have been specifically recommended for the assessment of mild cognitive impairment (MCI) and dementia, including Alzheimer disease (AD): the Paired Associates Learning (PAL), Spatial Working Memory (SWM), Rapid Visual

Information Processing (RVP), and Reaction Time (RTI) tests. These nonverbal, visual tests based on tasks and paradigms of experimental cognitive psychology (Sahakian & Owen, 1992) have strong theoretical foundations, unlike tests resulting from factor analysis. Information about the cognitive functions that they assess and their neurological correlates is also available (e.g., Coull, Frith, Frackowiak, & Grasby, 1996; Owen, Sahakian, Semple, Polkey, & Robbins, 1995).

The cited CANTAB tests have measures showing high sensitivity to cognitive decline that distinguish between the performances of individuals with mild AD, those with MCI, and those who are healthy (e.g., PAL total errors adjusted, SWM total errors, RVP A', and RTI simple reaction time; Fowler, Saling, Conway, Semple, & Louis, 1995; Saunders & Summers, 2010). Some tests also include measures exhibiting high predictive value for identifying individuals who will later develop AD (e.g., PAL total errors adjusted and RVP A'; Summers & Saunders, 2012); these tests are among the most advocated tests for prodromal AD research (Cummings, Gould, & Zhong, 2012).

Such test features and empirical evidence suggest that these tests are a good option for clinical use with older people living in retirement homes. Studies of CANTAB's reliability and validity in this population are also of marked interest for researchers.

It is known that when using a test in a population for which it has never been studied, its reliability needs to be estimated (Calamia, Markon, & Tranel, 2013; Green, Chen, Helms, & Henze, 2011). Although the CANTAB is in use worldwide, scarce test-retest reliability studies with older adults have been conducted for the PAL and SWM tests (and those that have been done solely included participants from the community; Fowler et al., 1995; Lowe & Rabbitt, 1998), but no such published work exists for the RVP and RTI tests. Existing studies report test-retest correlations of .64 to .86 for a 4-week interval (e.g., .64 for PAL total errors adjusted; .86 for PAL mean trials to success; .68 for PAL first trial memory score and SWM total errors; Fowler et al., 1995; Lowe & Rabbitt, 1998). According to the classification by Strauss, Sherman, and Spreen (2006), such test-retest correlations are far from desirable. However, they are similar to those reported for other computerized tests: .31 to .87 (8-week interval) for the Central Nervous System Vital Signs (Gualtieri & Johnson, 2006); .40 to .84 (4-week interval) for the NeuroTrax (Schweiger, Doniger, Dwolatzky, Jaffe, & Simon, 2003);

and .38 to .83 (4-week interval) for the Computer-Administered Neuropsychological Screen for MCI (CANS-MCI; Tornatore, Hill, Laboff, & McGann, 2005).

In addition to the relativity of test-retest coefficient classifications (Calamia et al., 2013), it has been argued that cognitive tests are more prone to exhibit low test-retest correlations because they are more affected by health status, fatigue, mood, alertness, age, years of education, and general ability level (Gualtieri & Johnson, 2006). Test-retest correlations also tend to vary according to the cognitive domain assessed (Duff, 2012) and are generally weaker for tests in which a performance improvement with task familiarity or discovery of a strategy (e.g., executive function tests; Lowe & Rabbitt, 1998) is likely. Alternate forms also tend to show lower test-retest correlations than single forms (Calamia et al., 2013).

The CANTABeclipse™ Version 3.0 software (Cambridge Cognition, 2006a) automatically and randomly displays the PAL, RVP, and RTI tests in multiple alternate forms. The SWM test is a self-ordered task; thus, it is similar to a test administered in alternate forms. Therefore, the test-retest reliability of these tests should be lower than that of tests with single forms (e.g., CANS-MCI tests; Tornatore et al., 2005) or two alternate forms presented in a fixed manner (e.g., NeuroTrax tests; Schweiger et al., 2003) because multiple alternate forms introduce more interindividual variability. Curiously, the test-retest coefficients cited for the PAL and SWM tests (Fowler et al., 1995; Lowe & Rabbitt, 1998) were similar to those of other test types, suggesting that if they were administered in a single form or in a fixed manner, they would possibly show higher test-retest reliability than their counterparts.

The equivalence of multiple alternate forms is difficult to demonstrate, especially for forms that are automatically generated by a program such as CANTABeclipse™ 3.0, which does not allow for the selection of a form for fixed administration. To ensure their equivalence in difficulty to the extent possible, this software shifts the stimuli presentation conditions, but not the task itself or the test structure, between forms. For example, in the PAL test running in clinical mode, the software only shifts the location where stimuli (visual patterns) appear and their appearance order while maintaining stimuli features and number. In the RVP test's 3-5-7 mode, the software changes the digits and the length of the distractor stimuli (digit sequences) displayed between the

target sequences. In the RTI test's clinical mode, the software changes the time before the target stimulus's appearance and its location in the five-choice condition.

The major advantage of multiple alternate forms is that they help reduce practice effects across two or more assessments. Indeed, practice effects have been reported for the measures PAL total errors adjusted (Fowler et al., 1995), PAL mean trials to success, and SWM total errors (Lowe & Rabbitt, 1998), although no effect sizes to express the relative magnitude of these changes were reported. Furthermore, no studies of these tests with older adults used methods of individual change analysis that account for practice effects. Such methods are quite useful for clinical practice and research because they consider the variability associated with measurement error and practice effects (if appropriate) before judging whether a real change has occurred. Yielding a Reliable Change Index (RCI), which represents a z score, these methods assume that a change occurs if the RCI falls outside a chosen confidence interval (CI).

Some of these methods advocate the use of a fixed adjustment for practice effects (e.g., Chelune, Naugle, Lüders, Sedlak, & Awad, 1993; Iverson, 2001), the Iverson's (2001) method being stricter because it includes measurement error for both assessment times. Other methods, known as standardized regression-based (SRB) methods, consider the regression to the mean effect by building a regression equation to predict the retest scores from the initial scores, along with other possible predictors (e.g., time interval, demographic variables). Such equations can be classically derived (McSweeney, Naugle, Chelune, & Lüders, 1993) from multiple linear regression (MLR) or estimated by simpler procedures using sample descriptive statistics (e.g., Crawford & Garthwaite, 2007; Duff, 2012). The SRB-estimated methods are more suitable for individuals who obtain extreme initial scores and for modest sample sizes because MLR is more prone to Type I error inflation in such cases (Crawford & Garthwaite, 2007). Nonetheless, the SRB-estimated methods must still fulfill the assumptions of MLR. Duff (2012) proposed simpler procedures than Crawford and Garthwaite (2007), although these latter procedures included sample size in their formulas, suggesting that this method is more suited for modest samples. Empirical evidence also has shown that the classical SRB method (which uses MLR) identifies more true and false positives, that the SRB-estimated method is more conservative, and that the fixed correction method falls

between the classical SRB method and the SRB-estimated method (Maassen, Bossema, & Brand, 2009).

Thus, to the best of our knowledge, the current study is the first to analyze the test-retest reliability of the PAL, SWM, RVP, and RTI tests in older people who do not have neuropsychiatric diagnoses and who live in retirement homes. We examined changes in test performance during a 4-week interval. Furthermore, using three methods of individual change analysis (Crawford & Garthwaite, 2007; Duff, 2012; Iverson, 2001), we sought to provide easier and faster ways for clinicians or researchers to analyze whether a meaningful individual change has occurred. We expected to find marginal-to-high test-retest correlations (i.e., .60-.89; Strauss et al., 2006), especially for the PAL and SWM tests, as shown in previous studies (Fowler et al., 1995; Lowe & Rabbitt, 1998). Practice effects were expected mainly for the SWM test because a performance improvement with strategy discovery is likely for this test (Lowe & Rabbitt, 1998).

METHOD

Participants

Thirty-four adults (14 men), aged 69 to 90 years ($M = 79.18$, $SD = 5.29$), who were right-handed and spoke Portuguese as their first language (none of them were bilingual), voluntarily participated in this study. All were living in retirement homes in the district of Lisbon for 3 or more months ($M = 38.18$, $SD = 27.01$). Participants had 4 to 20 years of education ($M = 10.41$, $SD = 4.92$) and premorbid IQ scores, as estimated by the Irregular Word Reading Test (Teste de Leitura de Palavras Irregulares [TeLPI]; Alves, Simões, & Martins, 2012), within the normative range for their education level. According to scores on the Addenbrooke's Cognitive Examination Revised (ACE-R; Firmino, Simões, Pinho, Cerejeira, & Martins, 2008; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006), none of the participants were cognitively impaired at baseline or at retest.

As assessed by the 30-item Geriatric Depression Scale (GDS-30; Barreto, Leuschner, Santos, & Sobral, 2008; Yesavage et al., 1982), participants were not severely depressed. The participants also had no neuropsychiatric diagnoses, according to self-reports that were provided in the clinical interview with a neuropsychologist

(M.M.G.) and evaluation by a physician. However, 16 participants (47%) reported psychotropic drug intake, possibly due to the prescription of potentially inappropriate medications. Indeed, psychoactive drugs are the most inappropriately prescribed drug category (Manias et al., 2015; Maust, Oslin, & Marcus, 2014; Miquel et al., 2010; Ubeda et al., 2012). Periquito et al. (2014) recently found that 75% of older people living in nursing homes mainly located in Lisbon were prescribed with potentially inappropriate medications, with the majority consuming duplicated antidepressants and anxiolytics or psychotropic drugs such as benzodiazepines. Due to this high rate, which was also found in the retirement homes included in this study, we were not able to exclude these participants.

The participants were uniformly distributed according to gender, $\chi^2(1, N = 34) = 1.06, p = .30$, psychotropic drug intake, $\chi^2(1, N = 34) = 0.12, p = .73$, and computer experience, $\chi^2(1, N = 34) = 1.06, p = .30$. For the latter variable, the participants were classified as users ($n = 14$) if they were using computers autonomously or attending computer classes for at least 1 month, and they were classified as computer-naïve ($n = 20$) if they had ever used a computer or were dependent on others to perform even the most basic operations (e.g., open or close a program). Men and women also did not significantly differ in age, $t(31) = 1.05, p = .30$, TeLPI-IQ, $t(32) = 1.82, p = .08$, computer experience, $\chi^2(1, N = 34) = 2.51, p = .11$, psychotropic drug intake, $\chi^2(1, N = 34) = 1.23, p = .27$, or GDS-30 scores, $t(32) = -0.54, p = .59$. Exclusion criteria consisted of a history of neurological, psychiatric, or other medical conditions with a serious impact on cognition; non-corrected-to-normal sensory deficits or motor problems in the dominant upper limb; and cognitive deficits (according to scores on pencil-and-paper tests that were selected for a brief neuropsychological assessment). The sample's scores on the pencil-and-paper tests are listed in Table 1.

Participants were recruited via oral and written advertisement, and they all provided written informed consent prior to participation. As compensation, they received brief reports of their performance on the pencil-and-paper tests free of charge. This study was approved by the Portuguese Foundation for Science and Technology, the Scientific Council of the Faculty of Psychology and Educational Sciences of the University of Coimbra, and the Principals and Ethics Committees of the retirement homes where the participants were living.

Table 1. Descriptive statistics of the raw scores on the pencil-and-paper tests ($N = 34$).

Test measures	<i>M</i>	<i>SD</i>	Range
TeLPI-IQ score	109.74	12.14	83-129
ACE-R			
Total score (first session)	89.62	6.43	76-100
Total score (retest session)	91.44	6.04	72-100
Verbal Fluency (animals), total of correct words	16.24	5.18	9-26
Digit Span (WAIS-III)			
Forward	8.97	1.98	5-13
Backward	5.41	1.92	3-10
Word Lists (WMS-III)			
Immediate recall	24.53	5.88	14-38
Delayed recall	5.15	2.61	1-12
Delayed recognition	21.68	2.25	16-24
RCFT			
Copy	33.32	2.73	25-36
Immediate recall	16.85	5.81	9-30
Naming Objects and Fingers (ADAS), total errors	0.68	0.81	0-3
Similarities (WAIS-III)	18.26	5.51	8-27
GDS-30	8.00	5.10	0-19

Note. TeLPI-IQ = Irregular Word Reading Test (Teste de Leitura de Palavras Irregulares); ACE-R = Addenbrook's Cognitive Examination-Revised; WAIS-III = Wechsler Adult Intelligence Scale-Third Edition; WMS-III = Wechsler Memory Scale-Third Edition; RCFT = Rey Complex Figure Test; ADAS = Alzheimer's Disease Assessment Scale; GDS-30 = 30-item Geriatric Depression Scale.

Design

Three sessions were held with each participant. The first session was conducted for screening purposes. To exclude those with possible abnormal cognitive functioning, we assessed the following variables: (a) cognitive deterioration (ACE-R; Firmino et al., 2008; Mioshi et al., 2006); (b) premorbid intelligence (TeLPI; Alves et al., 2012); (c) depressive symptoms (GDS-30; Barreto et al., 2008; Yesavage et al., 1982); (d) attention and working memory (Wechsler Adult Intelligence Scale-Third Edition [WAIS-III] Digit Span subtest; Wechsler, 2008a); (e) episodic memory (Rey Complex Figure Test [RCFT] Immediate Memory task and Wechsler Memory Scale-Third Edition [WMS-III] Word Lists I and II subtests; Rey, 2002; Wechsler, 2008b, 2008c); (f) naming (Alzheimer's Disease Assessment Scale [ADAS] Naming Objects and Fingers task; Guerreiro, Fonseca, Barreto, & Garcia, 2008; Mohs, Rosen, & Davis, 1983); (g) executive functioning (ACE-R semantic verbal fluency item; Firmino et al., 2008; Mioshi et al., 2006); (h) abstraction (WAIS-III Similarities subtest; Wechsler, 2008a); (i) constructive praxis (RCFT Copy task; Rey, 2002); and (j) cognitive functional capacity for instrumental activities of daily living (Adults and Older Adults

Functional Assessment Inventory [Inventário de Avaliação Funcional de Adultos e Idosos, IAFAI]; Sousa, Prieto, Vilar, Firmino, & Simões, 2015).

Scores greater than 1.5 standard deviations below the normative mean were the cutoff for all cognitive tests except for the RCFT (for which scores below the 5th percentile were the cutoff because norms are in percentiles) and the TeLPI. For this latter test, scores greater than 2 standard deviations away from the mean were the cutoff (as suggested by the authors; Alves, Simões, Martins, & Freitas, 2015) because we had a sample with a wide range of years of education and the TeLPI is highly correlated with this variable. For the GDS-30, we used the cutoff of severe depression (i.e., scores greater than 20; Barreto et al., 2008) because of the adverse impact of severe depression on cognitive functioning, particularly on memory and mental flexibility (e.g., Elderkin-Thompson et al., 2003; Murphy & O'Leary, 2009).

The second session (first CANTAB administration) was held no longer than 1 week later. The third session (CANTAB readministration) was held 1 month later (day range: 25-33, $M = 28.26$, $SD = 1.38$); this time interval was selected on the basis of previous studies (Fowler et al., 1995; Lowe & Rabbitt, 1998). To ensure that the participants suffered no change in their global cognitive state, all of them completed the ACE-R at the beginning of the third session. The dependent variables were the CANTAB output measures. For the PAL and RTI tests (run in clinical mode), we ensured that no participant completed the same test form for both the test and retest sessions by analyzing the data sheets that detailed the stimuli presentation conditions.

Apparatus and materials

The CANTAB was administered with the CANTABeclipse™ version 3.0 software (Cambridge Cognition, 2006a) run in a Windows® XP environment. A Sahara Slate PC® i412T Tablet PC (Intel® Celeron® M ULV 423 Processor, 1.0 GB DDR2 RAM, 120 GB SATA Hard Disk Drive) with a 12.1-inch TFT XGA LCD passive resistive touch screen (monitor resolution, 1024 × 768; monitor size, 300 mm × 220 mm × 26 mm) was used. Regarding peripheral devices, a press pad (Cambridge Cognition 2-Button Press Pad version 2.0) was used for the RVP and RTI tests. An integrated touchpad keyboard (KeySonic™ ACK-540U+Mini-Keyboard) was also employed by the examiner. According to Cambridge Cognition, the timing accuracy of the Sahara

Slate PC i412T was tested with software that is equivalent to CANTABeclipse™ 3.0 using external chronometry. The timing accuracy was measured (in milliseconds [ms]) for video latency ($M = 7.40$, $SD = 0.64$), audio latency, after calibration of the software ($M = 2.50$, $SD = 0.86$), and touch latency ($M = 63.65$, $SD = 5.00$).

Motor screening test

This Motor Screening (MOT) test is intended to familiarize participants with the touch screen and to detect motor, visual, or comprehension problems that may compromise test performance. This test was run in high visibility mode. Participants were asked to touch crosses that were displayed one at a time at different locations on the screen, and they took approximately 3 min to complete the task.

Rapid visual information processing test

This visual sustained attention test is sensitive to frontal- and parietal-lobe dysfunctions (Coull et al., 1996). It was administered in 3-5-7 mode (three participants performed the test in clinical mode and were therefore excluded from the analysis) and was presented in multiple test forms. Digits 2 to 9 were displayed on the screen one at a time at a rate of 100 stimuli per minute. Participants were asked to detect the 3-5-7 sequence by pressing a button after seeing the last digit of the target sequence. A training stage was given before the test. Participants took 10 min to perform the RVP task. The computed measures were A' (sensitivity to target sequences, ranging from 0.00 [bad] to 1.00 [good]) and mean latency (mean time [ms] to correctly detect a target sequence, within a response window of 1,800 ms).

Paired associates learning test

This visuospatial learning and memory test is associated with entorhinal cortex activation (Owen et al., 1995). This test was run in clinical mode with alternate forms. The test includes eight stages of increasing difficulty (one to eight patterns). White boxes were displayed on the screen and opened up one by one, at random, to show their content. Participants were asked to memorize the pattern that appeared for each box. Next, they were invited to touch the location where each pattern had appeared. The participants had a maximum of 10 trials to solve each stage; if they failed all trials, the test ended. The task was performed in approximately 17 min. The computed measures were total errors adjusted (total errors for all stages attempted plus an adjustment for

each stage not reached), mean errors to success (mean errors made before successfully completing a stage), and first-trial memory score (sum of patterns correctly located in the first trial for all stages completed).

Reaction time test

This test of attention and reaction and movement time is similar to the five-choice serial reaction time task, which has been shown to be sensitive to lesions on the medial prefrontal cortex and nucleus basalis in rats (McGaughy, Dalley, Morrison, Everitt, & Robbins, 2002). It was run in clinical mode with alternate forms. Both in the simple (one circle) and five-choice (five circles) conditions, participants were asked to press a button until they saw a yellow spot and to then release the button and touch the circle where the yellow spot had just appeared. Training stages were held before the test stage. There was an achievement criterion of five out of six correct responses; if this criterion was not met, the test ended. One participant did not reach the criterion for the five-choice condition. Participants took approximately 7 min to complete this test. The measures that were analyzed were simple reaction time and five-choice reaction time (time [ms] to release the button after yellow spot onset) and simple movement time and five-choice movement time (time [ms] to touch the circle after button release).

Spatial working memory test

This test of ability to retain and manipulate spatial information in working memory has been shown to be sensitive to frontal-lobe dysfunctions (Owen et al., 1995), especially in the dorsolateral prefrontal cortex (Manes et al., 2002). It was run in clinical mode and encompassed four levels of increasing difficulty (three to eight boxes), with the first level serving as training. Participants were asked to search for blue tokens hidden within boxes and to fill an empty house with these tokens. Each token was automatically hidden after the previous one had been found. Participants were also instructed to not open a box in which a blue token had already been encountered. The test was completed in approximately 13 min. The computed measures were total errors (number of times a box is touched when it is sure to not have a token) and strategy (number of times a participant starts a new search by touching a different box for problems with six to eight boxes, with higher scores reflecting less strategy use).

Procedure

The sessions were conducted by a trained neuropsychologist (M.M.G.). The first session lasted 90 min to 120 min and began with a semistructured interview to collect demographic and clinical data. Then, participants were asked to perform pencil-and-paper tests in the following order: ACE-R, Digit Span (WAIS-III), Word Lists I (WMS-III), RCFT, TeLPI, Naming Objects and Fingers (ADAS), Similarities (WAIS-III), Word Lists II (WMS-III), GDS-30, and IAFAI. A semistructured interview was also carried out with the participants' physicians (with the participants' consent). A close formal caregiver (e.g., physician, nurse, social worker) was asked to complete the IAFAI. The CANTAB tests were administered in the second session in the following order: MOT, RVP, PAL, RTI, and SWM. The participants took 43 min to 70 min to complete the battery ($M = 53.71$, $SD = 7.30$). The third session started with the ACE-R. Then, the CANTAB tests were administered in the same order as previously. For this session, the participants took 34 min to 59 min to complete the CANTAB ($M = 44.95$, $SD = 6.04$). When needed, breaks were provided between tests to minimize fatigue effects. All procedures recommended for software use (Cambridge Cognition, 2006a, 2006b) were followed.

Statistical analysis

Statistical analyses were conducted with Statistical Package for the Social Sciences Statistics version 20.0 software (IBM[®], New York, USA). Chi-square tests were used to compare frequencies. Group means were compared using independent-samples Student's t tests, or if distributions were not normal (Shapiro-Wilk tests, $\alpha = .05$), they were compared using Mann-Whitney U tests. Because psychotropic drug consumers were included in the sample, we compared the CANTAB scores between consumers and nonconsumers to determine the effects of the psychotropic drugs on test performance. These two groups, which did not significantly differ in age, $t(25) = 0.45$, $p = .66$, TeLPI-IQ, $t(32) = -1.03$, $p = .31$, gender, $\chi^2(1, N = 34) = 1.23$, $p = .27$, or computer experience, $\chi^2(1, N = 34) = 3.27$, $p = .07$, only exhibited a significant difference in performance on the PAL first-trial memory score, $U = 88.50$, $p = .05$, $r = .33$, and solely at test session (consumers, $Mdn = 12$; non-consumers, $Mdn = 15$). The effect size statistic r was estimated according to Field's (2005) formula.

Test-retest coefficients were calculated using Pearson's parametric correlations, or if normality of the distributions was not met (Shapiro-Wilk tests, $\alpha = .05$), they were calculated using Spearman's nonparametric correlations. CIs at 95% were set for each test-retest correlation using Fisher's z transformations. Test-retest scores were compared with paired-samples Student's t tests. Effect sizes were estimated with the d statistic using Cohen's (1988) formula for paired samples. Additionally, we calculated the test-retest correlations and compared the test-retest scores of men and women. When distributions were not normal, we used paired-samples Wilcoxon's T tests because of the small size of groups.

For individual change analyses, we followed Iverson's (2001) procedures for calculating standard errors of the difference (SED s): $SED = \sqrt{([SD_1\sqrt{(1 - r_{12})}]^2 + [SD_2\sqrt{(1 - r_{12})}]^2)}$. We determined RCIs for measures that showed significant changes ($p < .05$) with the practice effect correction formula $RCI_{PE-Iverson} = ([T_2 - T_1] - [M_2 - M_1])/SED$ and for other measures with $RCI_{Iverson} = (T_2 - T_1)/SED$, with T_1 and T_2 as the individual observed scores at Times 1 and 2.

For the SRB methods, we first verified the assumptions of linearity (scatterplot graphs), normality (Shapiro-Wilk tests), homoscedasticity (White tests), and independence of residuals (Durbin-Watson statistic). The measures RVP A' and RTI five-choice movement time did not meet the homoscedasticity criterion and were therefore excluded from the SRB analyses. We then performed MLRs using the stepwise method (.05 for entry and .15 for removal, as in McSweeney et al., 1993) and included the Time 1 observed score, age, years of education, gender (0 = women; 1 = men), interval days, psychotropic drug intake (0 = nonconsumers; 1 = consumers), and computer experience (0 = naive; 1 = users) as predictors. Absence of multicollinearity was met (variance inflation factor). For the SWM total errors measure, the optimal model removed all predictors except the Time 1 observed score, $\beta = .72$, $t(31) = 6.82$, $p < .001$, and gender, $\beta = -.24$, $t(31) = -2.29$, $p = .03$, yielding a slightly better model, $R_a^2 = .75$, $F(2, 31) = 50.98$, $p < .001$, than that with only the Time 1 observed score as a predictor, $R_a^2 = .72$, $F(1, 32) = 85.43$, $p < .001$. For the remaining measures, the optimal model included the Time 1 observed scores as the only significant predictor ($\alpha = .05$).

After these preliminary analyses and because of our small sample size, we estimated the SRB equations with only the Time 1 observed score as a predictor. We

followed Duff's (2012) procedures of estimating the regression coefficients with the formulas $B_{\text{est}} = SD_2/SD_1$ and $C_{\text{est}} = M_2 - B_{\text{est}}M_1$. Predicted time 2 scores (T_2') were calculated with the regression equation $T_2' = B_{\text{est}}T_1 + C_{\text{est}}$, and RCIs were calculated as $RCI_{\text{Duff}} = (T_2 - T_2')/SED$. Following the Crawford and Garthwaite (2007) method, we estimated the regression equations with the formulas $B_{\text{est}} = r_{12} (SD_2/SD_1)$, $C_{\text{est}} = M_2 - B_{\text{est}}M_1$, and $T_2' = B_{\text{est}}T_1 + C_{\text{est}}$. The standard error of the estimate ($SEE_{\text{C\&G}}$) and RCIs were computed with the formulas $SEE_{\text{C\&G}} = SD_2\sqrt{([1 - r_{12}^2] [(N - 1)/(N - 2)])}$ and $RCI_{\text{C\&G}} = (T_2 - T_2')/(SEE_{\text{C\&G}} \sqrt{[1 + (1/N) + ((T_1 - M_1)^2/[SD_1^2(N - 1)])])}$, respectively.

The percentage of participants whose performance improved or declined was determined by RCIs using a 90% CI (± 1.645) only for measures with a normal distribution (Shapiro-Wilk tests, $\alpha = .05$) of differences between observed scores ($T_2 - T_1$) or between observed and predicted Time 2 scores ($T_2 - T_2'$). This lower CI was chosen for an easier interpretation of the curve extremes.

RESULTS

Test-retest reliability

Test-retest correlations varied from .56 to .89 in the total sample (see Table 2). According to the classification system of Strauss et al. (2006), high test-retest correlations were observed for SWM total errors, RVP mean latency, RTI simple movement time, and RTI five-choice movement time. Adequate test-retest coefficients were found for PAL mean errors to success and RVP A'. Test-retest correlations were marginal or low for the other measures. However, at a 95% CI, test-retest coefficients varied from adequate to very high only for SWM total errors, RVP mean latency, and RTI five-choice movement time. For the remaining measures, test-retest correlation ranges included various categories, including marginal and low.

Comparing the test-retest correlations between men and women (see Table 2), remarkable differences were found for the SWM strategy and RVP A' measures (with test-retest coefficients being nonsignificant in the male sample but marginal and high in the female sample) and for the RTI reaction time measures (with such correlations being adequate in the male sample but low in the female sample).

Table 2. Test-retest correlations and standard errors of the CANTAB tests.

Test measures	Test-retest correlations						Standard errors		
	Men		Women		Total sample		SED	SEE _{C&G}	
	<i>n</i>	<i>r</i> (<i>r_s</i>)	<i>n</i>	<i>r</i> (<i>r_s</i>)	<i>N</i>	<i>r</i> (<i>r_s</i>)			95% CI
PAL total errors adjusted	14	(.67) ^b	20	(.68) ^b	34	(.68) ^a	.44-.83	21.28	18.76
PAL mean errors to success	14	(.65) ^c	20	(.74) ^a	34	(.73) ^a	.52-.86	2.14	1.81
PAL first-trial memory score	14	.62 ^c	20	(.57) ^b	34	(.56) ^b	.27-.76	3.22	3.04
SWM total errors	14	.81 ^a	20	.74 ^a	34	.85 ^a	.72-.92	8.84	9.77
SWM strategy	14	.35	20	.69 ^b	34	.67 ^a	.43-.82	3.24	3.46
RVP A'	13	.46	18	(.81) ^a	31	(.71) ^a	.48-.85	.04	N/A
RVP mean latency	13	.93 ^a	18	.84 ^a	31	(.89) ^a	.78-.95	47.53	39.88
RTI simple reaction time	14	(.75) ^b	20	(.47) ^c	34	(.59) ^a	.32-.77	101.32	55.00
RTI five-choice reaction time	14	(.71) ^b	19	(.54) ^c	33	(.66) ^a	.41-.82	57.10	53.69
RTI simple movement time	14	.79 ^b	20	.77 ^a	34	(.81) ^a	.65-.90	109.85	104.75
RTI five-choice movement time	14	(.90) ^a	19	.88 ^a	33	(.86) ^a	.73-.93	82.56	N/A

Note. CANTAB = Cambridge Neuropsychological Test Automated Battery; PAL = Paired Associates Learning; SWM = Spatial Working Memory; RVP = Rapid Visual Information Processing; RTI = Reaction Time. Spearman's correlations (*r_s*): within brackets. Pearson's correlations (*r*): without brackets. N/A = not applicable due to heteroscedasticity of the residuals; SED = standard error of the difference; SEE_{C&G} = estimated standard error of the estimate (Crawford & Garthwaite, 2007). Time unit: milliseconds.

^a *p* < .001.

^b *p* < .01.

^c *p* < .05.

Changes in test performance and practice effects

Statistical significance of tests and effect sizes

According to paired-samples Student's t tests, significant differences ($\alpha = .05$) between test-retest means were found in the total sample for PAL total errors adjusted, PAL mean errors to success, and SWM total errors, with these changes being small in magnitude (as measured with d statistics) and trending toward improvement (see Table 3). For RTI simple movement time, with a unilateral test ($p/2 < .05$), the retest mean was significantly worse than the test mean, $t(33) = -1.95$, $p = .03$, showing a change of a small magnitude (see Table 3). Ceiling effects (i.e., achieving the highest possible score) were observed for RVP A', for 6% of the participants in the test session and 10% in the retest session.

Comparing the test performance of men and women, significant differences of large magnitude were found for SWM total errors in the test session, $t(32) = -3.87$, $p < .001$, $d = -1.35$ (men, $M = 44.00$, $SD = 12.37$; women, $M = 59.65$, $SD = 11.03$), and in the retest session, $t(32) = -4.78$, $p < .001$, $d = -1.67$ (men, $M = 34.79$, $SD = 15.32$; women, $M = 58.60$, $SD = 13.54$), and for SWM strategy in the test session, $t(32) = -3.66$, $p = .001$, $d = -1.28$ (men, $M = 35.64$, $SD = 2.47$; women, $M = 39.20$, $SD = 2.98$), and in the retest session, $t(32) = -3.15$, $p = .004$, $d = -1.10$ (men, $M = 35.00$, $SD = 4.37$; women, $M = 39.45$, $SD = 3.83$), with men outperforming women on all of these SWM measures. Examining test-retest scores by gender, men's performance significantly improved for PAL mean errors to success, $T = -2.29$, $p = .02$, $r = -.61$ (test session, $Mdn = 6.53$; retest session, $Mdn = 3.50$), and for SWM total errors, $t(13) = 3.84$, $p = .002$, $d = 0.66$. We found no significant changes in women's performance between test-retest sessions.

Table 3. Descriptive statistics for the CANTAB tests and Iverson's reliable change confidence intervals after correction for practice effects.

Test measures	<i>N</i>	<i>M</i> ₁ (<i>SD</i> ₁)	<i>M</i> ₂ (<i>SD</i> ₂)	<i>M</i> ₂ - <i>M</i> ₁ (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i>	<i>d</i>
PAL total errors adjusted	34	43.62 (27.89)	33.62 (25.18)	-10.00 (25.62)	-2.28 (33)	.029	-0.38
PAL mean errors to success	34	5.31 (3.16)	3.89 (2.59)	-1.42 (2.83)	-2.92 (33)	.006	-0.49
PAL first-trial memory score	34	14.38 (3.27)	15.03 (3.63)	0.65 (3.46)	1.09 (33)	.284	0.19
SWM total errors	34	53.21 (13.84)	48.79 (18.43)	-4.41 (9.80)	-2.62 (33)	.013	-0.27
SWM strategy	34	37.74 (3.27)	37.62 (4.57)	-0.12 (3.41)	-0.20 (33)	.842	-0.03
RVP A'	31	.95 (.06)	.96 (.04)	.01 (.04)	1.59 (30)	.123	0.23
RVP mean latency	31	424.15 (112.29)	416.32 (84.38)	-7.83 (64.82)	-0.67 (30)	.506	-0.08
RTI simple reaction time	34	386.48 (142.74)	369.03 (66.89)	-17.45 (135.48)	-0.75 (33)	.458	-0.16
RTI five-choice reaction time	33	416.27 (68.59)	411.32 (70.66)	-4.95 (67.22)	-0.42 (32)	.675	-0.07
RTI simple movement time	34	505.98 (178.56)	545.86 (174.37)	39.89 (119.25)	1.95 (33)	.060	0.23
RTI five-choice movement time	33	476.21 (176.31)	484.88 (137.03)	8.66 (96.21)	0.52 (32)	.608	0.05

Note. CANTAB = Cambridge Neuropsychological Test Automated Battery; PAL = Paired Associates Learning; SWM = Spatial Working Memory; RVP = Rapid Visual Information Processing; RTI = Reaction Time. Lower scores indicate better performance for all test measures except PAL first-trial memory score and RVP A'. Time unit: milliseconds. *M*₁ = observed Time 1 mean score; *M*₂ = observed Time 2 mean score.

Models for individual change analysis and case examples

Table 4 lists the estimated regression equations resulting from the selected SRB models (Crawford & Garthwaite, 2007; Duff, 2012). Table 5 shows the equations used to compute RCIs according to the three models of analysis of change. The RCI formulas of Crawford and Garthwaite (2007) include numbers up to three decimal places to decrease rounding errors.

Table 4. Estimated regression equations for predicting Time 2 CANTAB test scores.

Test measures	(Duff, 2012)	(Crawford & Garthwaite, 2007)
	$T_2' = B_{\text{est}} T_1 + C_{\text{est}}$	$T_2' = B_{\text{est}} T_1 + C_{\text{est}}$
PAL total errors adjusted	$T_2' = 0.90T_1 - 5.76$	$T_2' = 0.61T_1 + 6.87$
PAL mean errors to success	$T_2' = 0.82T_1 - 0.45$	$T_2' = 0.59T_1 + 0.74$
PAL first-trial memory score	$T_2' = 1.11T_1 - 0.96$	$T_2' = 0.63T_1 + 6.01$
SWM total errors	$T_2' = 1.33T_1 - 22.07$	$T_2' = 1.14T_1 - 11.65$
SWM strategy	$T_2' = 1.40T_1 - 15.17$	$T_2' = 0.93T_1 + 2.35$
RVP mean latency	$T_2' = 0.75T_1 + 97.58$	$T_2' = 0.67T_1 + 134.08$
RTI simple reaction time	$T_2' = 0.47T_1 + 187.93$	$T_2' = 0.27T_1 + 262.75$
RTI five-choice reaction time	$T_2' = 1.03T_1 - 17.50$	$T_2' = 0.68T_1 + 126.68$
RTI simple movement time	$T_2' = 0.98T_1 + 51.76$	$T_2' = 0.79T_1 + 147.49$

Note. CANTAB = Cambridge Neuropsychological Test Automated Battery; PAL = Paired Associates Learning; SWM = Spatial Working Memory; RVP = Rapid Visual Information Processing; RTI = Reaction Time; B_{est} = estimated unstandardized beta coefficient (slope of the regression model); C_{est} = estimated constant (intercept of the regression model); T_1 = observed Time 1 test score; T_2' = predicted Time 2 test score. Time unit: milliseconds.

As an example, consider a participant (man, age 79 years) who performed the SWM test for a first time (Time 1) and 4 weeks later (Time 2). Suppose that his SWM total errors scores were $T_1 = 33$ and $T_2 = 44$. Following Iverson's (2001) method and searching for the respective RCI equation in Table 5, $\text{RCI}_{\text{PE-Iverson}} = [(44 - 33) + 4.41]/8.84 = 1.74$. For a 90% CI (the clinician or researcher may choose another CI), whose limits are ± 1.645 , the RCI exceeds the upper limit ($1.74 > 1.645$). Therefore, we can conclude that for a CI of 90%, the participant's performance declined (because the participant committed more errors in the second time than in the first time) and that this change was real (i.e., it cannot simply be explained by measurement errors as we would suppose it could if the RCI value was greater than -1.645 but lower than 1.645). Analyzing the change by Duff's (2012) method, the predicted Time 2 score (T_2') must first be calculated. According to its formula in Table 4, $T_2' = 1.33 \times 33 - 22.07 = 21.82$. According to the RCI equation in Table 5, $\text{RCI}_{\text{SRB-Duff}} = (44 - 21.82)/8.84 = 2.51$, which

also surpasses the upper limit of a 90% CI ($2.51 > 1.645$). Finally, with the method of Crawford and Garthwaite (2007), $T_2' = 1.14 \times 33 - 11.65 = 25.97$ and $RCI_{SRB-C\&G} = (44 - 25.97)/(9.768 \sqrt{[1.029 + (33 - 53.206)^2/6317.559]}) = 1.77$ (see formulas in Tables 4 and 5). Again, the RCI exceeds the upper limit of a 90% CI. Thus, for this example, the conclusions of the three methods of measuring change were similar. If the RCI value had been less than -1.645 , then the participant would have committed fewer SWM errors in the second time than in the first time and we could conclude that the participant's performance had reliably improved.

Upon analyzing the percentages of individuals in the sample whose performance reliably improved or declined, at a 90% CI, the methods that better fit to a normal curve (i.e., 5% declining and 5% improving) were the SRB models, especially Crawford and Garthwaite's (2007) method (see Table 6). The highest percentages of reliable change were found for Iverson's (2001) method – namely, for PAL mean errors to success (a high percentage of participants improving, even after practice effect correction) and SWM strategy (a high percentage of participants showing performance decline).

DISCUSSION

We found high-to-low test-retest correlations (classifications by Strauss et al., 2006), a range that included weaker values than expected, considering previous findings of the PAL and SWM tests (Fowler et al., 1995; Lowe & Rabbitt, 1998). The measures that exhibited high-to-adequate test-retest reliability in the studied sample across a 4-week interval were as follows: PAL mean errors to success, SWM total errors, RVP A', RVP mean latency, RTI simple movement time, and RTI five-choice movement time. These results suggest that the four CANTAB tests had one measure with at least adequate temporal stability. However, at a 95% CI, only the test-retest correlations of SWM total errors, RVP mean latency, and RTI five-choice movement time were at least acceptable. For the other measures, we found broader ranges of test-retest correlations at this CI, which included lower coefficients. These intervals could become narrower with a larger sample size. If so, more precise conclusions could be made about the measures' temporal stability and adequacy for cognitive assessment of this population.

Table 5. Equations to estimate Reliable Change Indexes (RCIs) and obtain z scores, according to the three methods.

Test measures	$RCI_{PE-Iverson}$	$RCI_{SRB-Duff}$	$RCI_{SRB-Crawford \& Garthwaite}$
PAL total errors adjusted	$([T_2 - T_1] + 10)/21.28$	$(T_2 - T_2')/21.28$	$(T_2 - T_2')/(18.765 \sqrt{[1.029 + (T_1 - 43.618)^2/25668.029]})$
PAL mean errors to success	$([T_2 - T_1] + 1.42)/2.14$	$(T_2 - T_2')/2.14$	$(T_2 - T_2')/(1.809 \sqrt{[1.029 + (T_1 - 5.310)^2/329.715]})$
PAL first-trial memory score	$(T_2 - T_1)/3.22$	$(T_2 - T_2')/3.22$	$(T_2 - T_2')/(3.044 \sqrt{[1.029 + (T_1 - 14.382)^2/352.029]})$
SWM total errors	$([T_2 - T_1] + 4.41)/8.84$	$(T_2 - T_2')/8.84$	$(T_2 - T_2')/(9.768 \sqrt{[1.029 + (T_1 - 53.206)^2/6317.559]})$
SWM strategy	$(T_2 - T_1)/3.24$	$(T_2 - T_2')/3.24$	$(T_2 - T_2')/(3.455 \sqrt{[1.029 + (T_1 - 37.735)^2/352.618]})$
RVP A'	$(T_2 - T_1)/0.04$	N/A	N/A
RVP mean latency	$(T_2 - T_1)/47.53$	$(T_2 - T_2')/47.53$	$(T_2 - T_2')/(39.880 \sqrt{[1.032 + (T_1 - 424.150)^2/378252.879]})$
RTI simple reaction time	$(T_2 - T_1)/101.32$	$(T_2 - T_2')/101.32$	$(T_2 - T_2')/(54.996 \sqrt{[1.029 + (T_1 - 386.483)^2/672373.796]})$
RTI five-choice reaction time	$(T_2 - T_1)/57.10$	$(T_2 - T_2')/57.10$	$(T_2 - T_2')/(53.693 \sqrt{[1.030 + (T_1 - 416.266)^2/150539.472]})$
RTI simple movement time	$(T_2 - T_1)/109.85$	$(T_2 - T_2')/109.85$	$(T_2 - T_2')/(104.751 \sqrt{[1.029 + (T_1 - 505.977)^2/1052207.183]})$
RTI five-choice movement time	$(T_2 - T_1)/82.56$	N/A	N/A

Note. PAL = Paired Associates Learning; SWM = Spatial Working Memory; RVP = Rapid Visual Information Processing; RTI = Reaction Time. N/A = not applicable due to heteroscedasticity of the residuals; PE = practice effect correction; SRB = standardized regression-based model.

Table 6. Percentage of individuals classified as having reliably changed in performance with a 90% confidence interval.

Test Measures	N	$RCI_{PE-Iverson}$		$RCI_{SRB-Duff}$		$RCI_{SRB-Crawford \& Garthwaite}$	
		Declined	Improved	Declined	Improved	Declined	Improved
PAL total errors adjusted	34	6	9	6	6	N/A	N/A
PAL mean errors to success	34	3	15	3	9	N/A	N/A
PAL first-trial memory score	34	3	6	3	3	6	3
SWM total errors	34	6	6	9	6	6	0
SWM strategy	34	12	6	9	6	6	3
RVP mean latency	31	N/A	N/A	0	3	6	3
RTI simple movement time	34	3	6	3	6	3	6

Note. PAL = Paired Associates Learning; SWM = Spatial Working Memory; RVP = Rapid Visual Information Processing; RTI = Reaction Time. N/A = not applicable because normality of the change distribution ($T_2 - T_1$ for RCI_{PE} and $T_2 - T_2'$ for RCI_{SRB}) cannot be assumed (Shapiro-Wilk tests, $p < .05$); PE = practice effect correction; RCI = reliable change index; SRB = standardized regression-based model.

The PAL total errors adjusted and PAL first-trial memory score measures had previously shown weak (i.e., lower than .70) test-retest correlations (Fowler et al., 1995; Lowe & Rabbitt, 1998). PAL mean errors to success, which showed an acceptable test-retest coefficient in the current study, was not previously studied. However, this measure reflects the average difficulty of achieving success, as does PAL mean trials to success, a measure that previously showed a high test-retest correlation (Lowe & Rabbitt, 1998). This measure was not included in our analyses because it exhibited lower data variability than PAL mean errors to success. However, in this sense, a higher test-retest correlation was expected for PAL mean errors to success. Similarly, SWM total errors, which had previously shown unacceptable test-retest reliability (Lowe & Rabbitt, 1998), now exhibited a high test-retest correlation. Such discrepancies may be explained by differences in sample size, data variability, demographic characteristics (Calamia et al., 2013) such as gender (in the current study, we indeed identified gender-based differences in test-retest reliability), and the administration of different test versions or use of different running modes. These discrepancies also reveal the degree to which the results are dependent on the samples and highlight the importance of undertaking test-retest reliability studies when a test is to be used in a new population.

Of the measures that showed high-to-acceptable test-retest reliability, those measures less affected by practice effects (for $\alpha = .05$) were RVP A', RVP mean latency, RTI simple movement time, and RTI five-choice movement time. However, ceiling effects were observed for RVP A', not excluding the possible masked practice effects. The measures with high-to-acceptable test-retest reliability that revealed more susceptibility to practice effects (for $\alpha = .05$) were PAL mean errors to success and SWM total errors, although the changes were of a small magnitude (according to *d* effect sizes). Men's performance most contributed to the practice effects observed in these two measures.

Both PAL total errors adjusted and SWM total errors had previously shown susceptibility to practice effects (Fowler et al., 1995; Lowe & Rabbitt, 1998), as did PAL mean trials to success, which is similar to PAL mean errors to success. Thus, the current findings are consistent with previous studies and with the hypothesis that tests that are more prone to a performance enhancement with task familiarity or the discovery of a strategy are more susceptible to practice effects. The SWM test, which assesses

executive function, is an example of such a test. Interestingly, the observed change for SWM strategy was not significant in the total sample or within genders. It is possible that the participants employed other effective strategies that are not assessed by the SWM strategy measure. The PAL test, despite being a visuospatial learning and memory test, has been shown to be sensitive both to temporal-lobe and frontal-lobe dysfunctions (Owen et al., 1995). As a result, performance on this test can be markedly improved by using strategies that require mental flexibility to switch between variables (e.g., shape, color, location) and to elaborate visual, semantic, or other types of associations between them. If the practice effects found for PAL and SWM tests are related to strategy use, then it is possible that men have greater or more effective use of strategies than women, thereby explaining their greater contribution to the observed practice effects. In this study, men exhibited better performance than women in terms of SWM strategy.

Because the PAL test is administered in alternate forms, practice effects should have been attenuated. However, it is possible that the alternate forms were less effective in reducing practice effects because PAL problems (as well as SWM problems) can be solved by one same rule. Nonetheless, it should be noted that although practice effects were found for such PAL measures, the effect sizes were smaller than those recently found for two other memory tests administered in a single form (Brief Visuospatial Memory Test, 0.73-1.04; Hopkins Verbal Learning Test Revised, 0.76-0.80; Duff, 2014).

The main drawbacks of practice effects are that they hamper the detection of cognitive decline (which can be underestimated or confounded with a “stable” performance), and blur the distinction between a cognitive improvement due to therapeutic intervention or to test familiarity. In the current study, the methods of individual change analysis with correction for practice effects could not feasibly be applied to all measures due to nonfulfillment of the required criteria. For the measures on which they were applied, the SRB-estimated methods were more conservative and efficient in correcting for practice effects than the RCI_{PE} , as previously observed (Maassen et al., 2009). The method of Crawford and Garthwaite (2007) was slightly more conservative than Duff’s (2012) method, possibly due to the small sample size and the fact that the former takes the number of participants into account. However, the two

methods may behave similarly with a larger sample; therefore, further studies that include both healthy and clinical groups are needed.

Despite the nonverbal nature of the studied CANTAB tests, the generalizability of our findings to other cultures needs to be verified. An additional limitation of this study is that it was designed for only two test administrations, which did not permit the analysis of smoothing of practice effects and strengthening of test-retest correlations from the second session onward. Such a design could allow for a better understanding of the suitability of CANTAB measures for cognitive assessment and monitoring and provide regression equations for subsequent assessments. Additionally, a 4-week interval between the test-retest sessions was used (as in many studies, Fowler et al., 1995; Lowe & Rabbitt, 1998; Wechsler, 2008a), whereas in clinical practice, longer time intervals are generally used. Studies that include more than one time interval may be useful to understand the impact of time between sessions on test-retest correlations and changes in performance. Nonetheless, as the intervals between sessions lengthen, the risk of cognitive deficit onset in older people increases.

Other limitations of this study are the small sample size, which reduced the data's variability and widened the CIs for the test-retest correlations, and the inclusion of psychotropic drug consumers, which may have led to performance fluctuations, although only a slight significant effect of psychotropic drugs was found for PAL first-trial memory score. It is likely that low data variability, performance fluctuations, and the use of multiple alternate test forms may have weakened the test-retest correlations. Although CANTAB alternate test forms were carefully programmed to maintain the same task and test structure between forms, the shifts in stimuli presentation conditions can still endanger their equivalence in difficulty, especially for tests favored by the use of strategies based on stimuli.

This study was the first to determine the test-retest reliability of CANTAB tests conducted with a Portuguese-speaking sample. Our findings suggest that the SWM, RVP, and RTI tests – and possibly the PAL test – have measures with high-to-adequate test-retest reliability. However, some of these measures revealed susceptibility to practice effects across a 4-week interval. We highly recommend the use of SRB-estimated methods for the correction of practice effects, both in clinical practice and in research, for the judgment of individual change. Given that scarce test-retest reliability

studies with the CANTAB tests have been conducted, this work highlights their importance because a cognitive test that is temporally unreliable in a sample without cognitive deficits generates serious problems in terms of the interpretation of a patient's results and thus is unhelpful for the detection of cognitive improvements or declines.

Finally, it should be noted that the identified range of test-retest correlations is similar or even slightly better than that of other computerized and pencil-and-paper cognitive tests (e.g., Calamia et al., 2013; Gualtieri & Johnson, 2006; Schweiger et al., 2003; Tornatore et al., 2005; Wechsler, 2008c). However, it is important to note that the test-retest coefficients resulting from Pearson's correlations are consistently lower than those derived from intraclass correlation coefficients (ICCs) for the same tests and data (Schatz & Ferris, 2013). Thus, when comparing the CANTAB test-retest reliability (as estimated by Pearson's and Spearman's correlations) with that of tests using ICCs as the estimation method, this condition has to be accounted for beyond coefficient values. One last note is that the majority of the changes in test performance observed for the analyzed CANTAB tests displayed in alternate forms were small and smaller than the changes found for some single forms (Duff, 2014). As such, the studied CANTAB tests remain a good option for the neurocognitive assessment of older adults living in retirement homes, offering some temporally reliable measures for the detection and monitoring of cognitive deficits.

ACKNOWLEDGEMENTS

The authors would like to thank all of the principals and professionals at the institutions (retirement homes; José de Mello Residences and Services and Oeiras Social Support Center) who gave permission to conduct our study and helped in the recruitment of older volunteers and who kindly collaborated in this research by completing interviews for formal caregivers. They would also like to thank all of the older volunteers who kindly participated in this study and without whom this research would be impossible.

DISCLOSURE STATEMENT

All of the authors declare that they have no financial interest or benefit arising from the direct application of this research. There is no potential conflict of interest in this study.

FUNDING

The first author was supported by the Fundação para a Ciência e a Tecnologia [Portuguese Foundation for Science and Technology] and financed by the European Social Fund and national funds of the Portuguese Ministry of Science, Technology and Higher Education under Grant No. SFRH/BD/75281/2010.

REFERENCES

- Alves, L., Simões, M. R., & Martins, C. (2012). The estimation of premorbid intelligence levels among Portuguese speakers: The Irregular Word Reading Test (TeLPI). *Archives of Clinical Neuropsychology*, 27, 58-68. doi: 10.1093/arclin/acr103
- Alves, L., Simões, M. R., Martins, C., & Freitas, S. (2015). *The Irregular Word Reading Test (TeLPI): Normative study for the Portuguese population*. Manuscript submitted for publication.
- Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de Depressão Geriátrica (GDS) [Geriatric Depression Scale (GDS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (2nd ed., pp. 69-72). Lisbon, Portugal: Novartis.
- Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological measures: Meta-analyses of test-retest correlations. *The Clinical Neuropsychologist*, 27, 1077-1105. doi: 10.1080/13854046.2013.809795
- Cambridge Cognition. (2006a). *CANTABeclipse™ Version 3: Software user guide*. Cambridge, UK: Author.
- Cambridge Cognition. (2006b). *CANTABeclipse™ Version 3: Test administration guide*. Cambridge, UK: Author.
- Chelune, G. J., Naugle, R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, 7, 41-52. doi: 10.1037/0894-4105.7.1.41
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, 55, 1630-1635. doi: 10.1111/j.1532-5415.2007.01379.x
- Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia*, 34, 1085-1095. doi: 10.1016/0028-3932(96)00029-2
- Crawford, J. R., & Garthwaite, P. H. (2007). Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology*, 21, 611-620. doi: 10.1037/0894-4105.21.5.611

- Cummings, J., Gould, H., & Zhong, K. (2012). Advances in designs for Alzheimer's disease clinical trials. *American Journal of Neurodegenerative Disease*, *1*, 205-216.
- Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology*, *27*, 248-261. doi: 10.1093/arclin/acr120
- Duff, K. (2014). One-week practice effects in older adults: Tools for assessing cognitive change. *The Clinical Neuropsychologist*, *28*, 714-725. doi: 10.1080/13854046.2014.920923
- Elderkin-Thompson, V., Kumar, A., Bilker, W., Dunkin, J., Mintz, J., Moberg, P., ... Gur, R. (2003). Neuropsychological deficits among patients with late-onset minor and major depression. *Archives of Clinical Neuropsychology*, *18*, 529-549. doi: 10.1016/s0887-6177(03)00022-2
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London, UK: Sage.
- Firmino, H., Simões, M. R., Pinho, M. S., Cerejeira, L., & Martins, C. (2008). *O Exame Cognitivo de Addenbrooke - Revisto* [The Addenbrooke's Cognitive Examination - Revised]. Coimbra, Portugal: Serviço de Avaliação Psicológica da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. *Applied Neuropsychology*, *2*, 72-78. doi: 10.1207/s15324826an0202_4
- Green, C. E., Chen, C. E., Helms, J. E., & Henze, K. T. (2011). Recent reliability reporting practices in Psychological Assessment: Recognizing the people behind the data. *Psychological Assessment*, *23*, 656-669. doi: 10.1037/a0023089
- Gualtieri, C., & Johnson, L. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology*, *21*, 623-643. doi: 10.1016/j.acn.2006.05.007
- Guerreiro, M., Fonseca, S., Barreto, J., & Garcia, C. (2008). Escala de Avaliação da Doença de Alzheimer (ADAS) [Alzheimer's Disease Assessment Scale (ADAS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (2nd ed., pp. 41-68). Lisbon, Portugal: Novartis.
- Instituto Nacional de Estatística. (2008). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2008* [Survey on the use of information and communication technologies by households 2008]. Lisbon, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Retrieved from http://www.anacom.pt/streaming/inq_familias_ineumic2008.pdf?contentId=726858&field=ATTACHED_FILE
- Instituto Nacional de Estatística. (2013). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2013* [Survey on the use of information and communication technologies by households 2013]. Lisbon, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Retrieved from http://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=206557735&att_display=n&att_download=y.

- Iverson, G. L. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology*, *16*, 183-191. doi: 10.1093/arclin/16.2.183
- Iverson, G. L., Brooks, B. L., Ashton, V. L., Johnson, L. G., & Gualtieri, C. T. (2009). Does familiarity with computers affect computerized neuropsychological test performance? *Journal of Clinical and Experimental Neuropsychology*, *31*, 594-604. doi: 10.1080/13803390802372125
- Lowe, C., & Rabbitt, P. (1998). Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. *Neuropsychologia*, *36*, 915-923. doi: 10.1016/S0028-3932(98)00036-0
- Maassen, G. H., Bossema, E., & Brand, N. (2009). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology*, *31*, 339-352. doi: 10.1080/13803390802169059
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, *125*, 624-639. doi: 10.1093/brain/awf049
- Manias, E., Kusljic, S., Berry, C., Brown, E., Bryce, E., Cliffe, J., & Smykowsky, A. (2015). Use of the Screening Tool of Older Person's Prescriptions (STOPP) in older people admitted to an Australian hospital. *Australasian Journal on Ageing*, *34*, 15-20. doi: 10.1111/ajag.12054
- Maust, D. T., Oslin, D. W., & Marcus, S. C. (2014). Effect of age on the profile of psychotropic users: Results from the 2010 National Ambulatory Medical Care Survey. *Journal of the American Geriatrics Society*, *62*, 358-364. doi: 10.1111/jgs.12640
- McGaughy, J., Dalley, J. W., Morrison, C. H., Everitt, B. J., & Robbins, T. W. (2002). Selective behavioral and neurochemical effects of cholinergic lesions produced by intrabasal infusions of 192 IgG-saporin on attentional performance in a five-choice serial reaction time task. *The Journal of Neuroscience*, *22*, 1905-1913.
- McSweeney, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). "T scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *Clinical Neuropsychologist*, *7*, 300-312. doi: 10.1080/13854049308401901
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, *21*, 1078-1085. doi: 10.1002/gps.1610
- Miquel, M. D. C., Cuervo, M. S., Silveira, E. D., Machuca, I. S., González-Blázquez, S., Errasquin, B. M., & Cruz-Jentoft, A. J. (2010). Potentially inappropriate drug prescription in older subjects across health care settings. *European Geriatric Medicine*, *1*, 9-14. doi: 10.1016/j.eurger.2009.12.002
- Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin*, *19*, 448-450.
- Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1987). Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer-type and Parkinson's disease. In S. M. Stahl, S. D. Iversen & E. C. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21-36). Oxford, UK: Oxford University Press.

- Murphy, M., & O'Leary, E. (2009). Depression, cognitive reserve and memory performance in older adults. *International Journal of Geriatric Psychiatry*, 25, 665-671. doi: 10.1002/gps.2404
- Nihtilä, E. K., Martikainen, P. T., Koskinen, S. V. P., Reunanen, A. R., Noro, A. M., & Häkkinen, U. T. (2007). Chronic conditions and the risk of long-term institutionalization among older people. *The European Journal of Public Health*, 18, 77-84. doi: 10.1093/eurpub/ckm025
- Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, 33, 1-24. doi: 10.1016/0028-3932(94)00098-a
- Periquito, C. M. N., Silva, P. C., Oliveira, P., Carneiro, C., Fernandes, A. I., & Costa, F. A. (2014). Revisão da medicação em idosos institucionalizados: Aplicação dos critérios STOPP e START. [Review of medication in institutionalized elderly: Application of STOPP and START criteria]. *Revista Portuguesa de Farmacoterapia*, 6, 211-220.
- Rey, A. (2002). *Teste de Cópia de Figuras Complexas* [Complex Figure Copying Test] (2nd ed.). Lisbon, Portugal: Cegoc-Tea.
- Ribbe, M. W., Ljunggren, G., Steel, K., Topinková, E., Hawes, C., Ikegami, N., ... Jónnson, P. V. (1997). Nursing homes in 10 nations: A comparison between countries and settings. *Age and Ageing*, 26(Suppl. 2), 3-12. doi: 10.1093/ageing/26.suppl_2.3
- Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, 85, 399-402. doi: 10.1177/014107689208500711
- Saunders, N. L. J., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 32, 350-357. doi: 10.1080/13803390903042379
- Schatz, P., & Ferris, C. S. (2013). One-month test-retest reliability of the ImPACT test battery. *Archives of Clinical Neuropsychology*, 28, 499-504. doi: 10.1093/arclin/act034
- Schweiger, A., Doniger, G. M., Dwolatzky, T., Jaffe, D., & Simon, E. S. (2003). Reliability of a novel computerized neuropsychological battery for mild cognitive impairment. *Acta Neuropsychologica*, 1, 407-413. doi: 10.5604/17307503/9168
- Sousa, L. B., Prieto, G., Vilar, M., Firmino, H., & Simões, M. R. (2015). The Adults and Older Adults Functional Assessment Inventory: A Rasch model analysis. *Research on Aging*, 37, 787-814. doi: 10.1177/0164027514564469
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.
- Summers, M. J., & Saunders, N. J. (2012). Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. *Neuropsychology*, 26, 498-508. doi: 10.1037/a0028576
- Tornatore, J. B., Hill, E., Laboff, J. A., & McGann, M. E. (2005). Self-administered screening for mild cognitive impairment: Initial validation of a computerized test battery. *Journal of Neuropsychiatry*, 17, 98-105. doi: 10.1176/appi.neuropsych.17.1.98

- Ubeda, A., Ferrándiz, M. L., Maicas, N., Gomez, C., Bonte, M., & Peris, J. E. (2012). Potentially inappropriate prescribing in institutionalised older patients in Spain: The STOPP-START criteria compared with the Beers criteria. *Pharmacy Practice, 10*, 83-91. doi: 10.4321/s1886-36552012000200004
- Wechsler, D. (2008a). *Escala de Inteligência de Wechsler para Adultos – Terceira edição (WAIS-III): Manual técnico* [Wechsler Adult Intelligence Scale - Third edition (WAIS-III): Technical manual]. Lisbon, Portugal: Cegoc-Tea.
- Wechsler, D. (2008b). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual de administração e cotação* [Wechsler Memory Scale - Third edition (WMS-III): Administration and scoring manual]. Lisbon, Portugal: Cegoc-Tea.
- Wechsler, D. (2008c). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual técnico* [Wechsler Memory Scale - Third edition (WMS-III): Technical manual]. Lisbon, Portugal: Cegoc-Tea.
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17*, 37-49. doi: 10.1016/0022-3956(82)90033-4

Capítulo 2.

EFFECTS OF SOCIO-DEMOGRAPHIC VARIABLES ON PERFORMANCE ON THE CAMBRIDGE NEUROPSYCHOLOGICAL AUTOMATED TESTS FOR THE ASSESSMENT OF DEMENTIA AND PORTUGUESE NORMS FOR OLDER ADULTS LIVING IN RETIREMENT HOMES

Marta Matos Gonçalves, Maria Salomé Pinho, and Mário R. Simões²

ABSTRACT

Objective: This study aimed to analyze the effects of age, education, gender, computer experience, institutionalization time, and psychotropic drug use on performance on four tests of the Cambridge Neuropsychological Test Automated Battery (CANTAB) which are recommended for the assessment of dementia (Rapid Visual Information Processing [RVP], Paired Associates Learning [PAL], Spatial Working Memory [SWM], and Reaction Time [RTI]), and to provide norms for Portuguese older persons without neuropsychiatric diagnoses who are living in retirement homes. **Method:** The normative sample included 128 adults aged 69-96 years who had no neuropsychiatric diagnosis and who had lived in retirement homes for 3-232 months. The CANTAB was administered, at the latest, one week after a screening session that comprised an interview and the administration of pencil-and-paper tests. **Results:** The simultaneous multiple linear regression models were significant ($p < .05$) for all tests except the RTI five-choice movement time measure. The total variance explained by the socio-demographic variables was smaller for the CANTAB measures (4-14%) than for the pencil-and-paper tests (10-33%). Significant effects involving age or gender were observed for RVP, PAL, and SWM. A marginally significant computer experience effect was found for the RTI simple movement time measure. We additionally observed significant effects of education, age, gender, and computer experience on several pencil-and-paper tests. **Conclusions:** Our findings suggest that different socio-demographic

² Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016). Effects of socio-demographic variables on performance on the Cambridge neuropsychological automated tests for the assessment of dementia and Portuguese norms for older adults living in retirement homes. *The Clinical Neuropsychologist*, 30(2), 284-317. doi: 10.1080/13854046.2016.1156745. Artigo transcrito segundo as normas da revista.

variables influence distinct tests and measures of the same test, and that the associations between computer experience and several pencil-and-paper tests may be mediated by possible cognitive skills developed through computer use.

Keywords: CANTAB; age; computer experience; normative data; institutionalized older persons.

INTRODUCTION

Institutions for the elderly are sometimes regarded as a final option for older persons who have become dependent. However, this understanding does not reflect the entire reality. Many of these establishments provide high to low levels of care; therefore, they are also an option for healthy, autonomous individuals who seek a lifestyle free of concerns from household chores so that they can enjoy the company of their autonomous or dependent mates while ensuring future assistance. Typically, individuals stay in retirement homes until the end of their lives. Thus, the emergence of cognitive deficits during the stay is common. Neuropsychological assessment is helpful for the early detection and monitoring of cognitive deficits. Computerized tests offer many advantages over pencil-and-paper tests, such as the ability to store a large amount of data, which is particularly useful for retirement homes due to the numerous neuropsychological assessments they must perform over time. Furthermore, contrary to what might be expected, older persons report greater acceptance of and less stress from computerized tests than from pencil-and-paper tests (Collerton et al., 2007).

Given that few neuropsychological tests with norms for institutionalized older persons exist, the performance of individuals living in retirement homes is often compared to that of community-dwelling individuals, posing questions about the appropriateness of this comparison. For example, most normative samples of community dwellers include individuals up to the age of 80 years (or an inexpressive number of individuals upward of this age limit), whereas the mean age of the institutionalized older population is generally around or above 80 years (Gonçalves, Pinho, & Simões, 2016; Periquito et al., 2014). Thus, many institutionalized older persons are compared with younger community dwellers. Furthermore, differences in cognitive performance between community dwellers and institutionalized older persons

have been reported, with several studies describing better performance or less variability across cognitive domains on the part of community dwellers (e.g. Plati, Covre, Lukasova, & Macedo, 2006; Rapp, Schnaider-Beeri, Sano, Silverman, & Haroutunian, 2005), although these studies have compared mismatched groups with regard to depressive symptoms and global cognitive states. Irrespective of these methodological limitations, there are considerable differences between the two populations' lifestyles (e.g. accessibility of leisure and support services, types of social groups), suggesting that these populations may differ from one another regarding the cognitive skills that are most stimulated or trained. Therefore, the comparison of institutionalized individuals' performance with that of community dwellers may increase the risk of incorrect assessments, with serious consequences for institutionalized older persons (e.g. being transferred from an independent living facility to a dependent living facility or being prevented from making decisions). Considering that there are cognitively healthy older persons living in institutions whose performance can serve as a benchmark for normal cognitive functioning for institutionalized older individuals, it is relevant to study test norms for people living in retirement homes, particularly norms for computerized neuropsychological tests.

The Cambridge Neuropsychological Test Automated Battery (CANTAB[®], www.cambridgecognition.com) is one of the most widely used computerized batteries in the world. First developed in 1987 (Morris, Evenden, Sahakian, & Robbins, 1987), the CANTAB has since been supplemented with new tests (Cambridge Cognition, 2006a, 2006b). The majority of these tests are visual tests that display abstract stimuli in an attempt to reduce the effects of language and culture. The tests require verbal skills to understand brief oral instructions only; hence, they are easy to adapt to other countries. The CANTAB tests are based on tasks and paradigms of experimental cognitive psychology (Sahakian & Owen, 1992) and have strong theoretical support. Four CANTAB tests have been recommended for the assessment of mild cognitive impairment (MCI) and dementia (e.g. Sahgal et al., 1992; Swainson et al., 2001): Rapid Visual Information Processing (RVP), Paired Associates Learning (PAL), Spatial Working Memory (SWM), and Reaction Time (RTI). These tests assess cognitive functions that are commonly affected in mild to moderate stages of dementia of Alzheimer's disease (AD) type, and they exhibit high sensitivity to cognitive decline.

Thus, they can differentiate between individuals with mild AD, those with MCI, and those who are healthy based on their performance (e.g. on the measures RVP A' [an index of sensitivity to detect target sequences among non-target stimuli], PAL total errors adjusted, SWM total errors, and RTI simple reaction time; Junkkila, Oja, Laine, & Karrasch, 2012; Saunders & Summers, 2010). Some tests also offer high predictive value for AD (e.g. RVP A' [sensitivity index] and PAL total errors adjusted; Summers & Saunders, 2012), making them suitable for prodromal AD research (Cummings, Gould, & Zhong, 2012). In addition, these four tests showed good test-retest reliability (.71-.89; Gonçalves et al., 2016) in a sample of 34 individuals who were included in the current research, but who took the CANTAB tests a second time after a 4-week interval, thus making them a good choice for the cognitive assessment of Portuguese older persons living in retirement homes. However, CANTAB norms for this specific population are currently lacking.

The original CANTAB norms can be accessed with the software CANTABeclipse™ version 3.0 (Cambridge Cognition, 2006a). This software allows a comparison of an adult's performance with that of the entire normative sample of community-dwelling adults or with a group selected based on age, gender, premorbid intelligence quotient (IQ, estimated by the National Adult Reading Test; Nelson & Willison, 1991), or any combination of these variables. The CANTAB guides (Cambridge Cognition, 2006a, 2006b) provide no recommendation about the demographic variables that must be considered when accessing norms for each test measure. Nevertheless, the identification of which socio-demographic variables have a significant influence on test performance allows them to be properly considered in norms.

The influence of socio-demographic variables on CANTAB test performance

Age

In general, it is well established that tests of episodic memory, working memory, and processing speed are prone to age effects (Alexander et al., 2012). Significant positive correlations between age and the measures RVP mean latency, PAL total errors adjusted, and SWM total errors have been reported (Saunders & Summers, 2010; Steves, Jackson, & Spector, 2013), and many studies show that younger groups

outperform older ones on RVP A' (sensitivity index), RTI simple movement time, and RTI five-choice movement time (e.g. Braw, Aviram, Bloch, & Levkovitz, 2011; Maylor et al., 2006). However, nonsignificant age effects have also been described for RVP A' (Saunders & Summers, 2010) and SWM total errors (Rabbitt & Lowe, 2000).

Education

Both education and premorbid IQ have been considered cognitive reserve indices (Stern, 2006), with the most educated or intelligent individuals presenting better outcomes in coping with brain damage and manifesting functional deficits more later. Education has often been noted as significantly affecting performance on tests of executive function, episodic memory, attention, and working memory (e.g. Ganguli et al., 2010; Yang et al., 2012). For example, Reeves et al. (2005) reported a significant contribution of education to the variability of a spatial working memory factor with high factor loadings of two SWM measures (i.e. between search errors, a measure indicating the number of times a box was incorrectly opened between searches for tokens, and strategy, a measure indicating the number of times the individual started a new search by touching a different box) and of the Spatial Span test's span length measure (CANTAB). However, other studies have reported nonsignificant effects of education on SWM between errors, RVP A' (sensitivity index), and PAL mean trials to success (Braw et al., 2011; Hogervorst, De Jager, Budge, & Smith, 2004).

Gender

Gender may be considered less often than age and education in norms of cognitive tests because it tends to exhibit smaller effects on test performance (Meinz & Salthouse, 1998). No significant effect of gender or difference between genders has been described for RVP A' (sensitivity index), RVP mean latency, PAL total errors adjusted, or SWM between errors (Braw et al., 2011; Jones, Sahakian, Levy, Warburton, & Gray, 1992; Simpson et al., 2013; Skolimowska, Wesierska, Lewandowska, Szymaszek, & Szlag, 2011). However, De Luca et al. (2003) found that men commit fewer errors than women do on SWM between errors. An inconsistent influence of gender on RTI simple reaction time and RTI five-choice reaction time has also been observed (Simpson et al., 2005), with men performing faster than women in an Italian sample, but both genders performing similarly in French and Irish samples.

Other socio-demographic variables

Considering that the CANTAB involves computerized tests, another variable that may influence test performance is computer experience. Indeed, older persons with more familiarity with computers have exhibited better performance on computerized tests than those with less familiarity, especially in the cognitive flexibility and reaction time domains (Iverson, Brooks, Ashton, Johnson, & Gualtieri, 2009). Among institutionalized older persons, it is possible that institutionalization time has an influence on test performance, but no studies examining the effect of this variable have been found. A problem that is commonly observed in Portuguese older adults (and that has been identified in other countries as well; Manias et al., 2015; Maust, Oslin, & Marcus, 2014; Ubeda et al., 2012) is the prescription of potentially inappropriate medications (PIMs). Periquito et al. (2014) reported that 75% of older persons living in retirement homes mainly located in Lisbon were identified as consuming PIMs which were mostly benzodiazepines and duplicated antidepressants or anxiolytics (i.e. two different antidepressants prescribed to the same person). Psychotropic drug use is a variable that may influence test performance. Several psychotropic drugs have shown a significant effect on CANTAB test performance (e.g. diazepam, a benzodiazepine that has been found to worsen performance on PAL total errors adjusted, and escitalopram and duloxetine, two antidepressants that have been found to improve performance on RVP total hits, PAL total errors adjusted, and SWM total errors; Herrera-Guzmán et al., 2009; Herrera-Guzmán et al., 2010; Robbins et al., 1997).

This is the first normative study of the CANTAB tests in Portugal and the first to include persons living in retirement homes. This study aims to analyze the influence of age, education, gender, computer experience, institutionalization time, and psychotropic drug use (compatible with PIM prescription) on performance on RVP, PAL, SWM, and RTI in older persons without neuropsychiatric diagnoses living in retirement homes. It also aims to provide norms for older Portuguese adults living in institutions such as retirement homes in order to increase the number of cognitive tests available for the assessment of this specific population and to enable a more suitable comparison of the performance of an institutionalized person with the performance of healthy individuals with similar living arrangements. We expect to find age effects, primarily on PAL and RTI (e.g. Maylor et al., 2006; Skolimowska et al., 2011), possible education effects on

SWM (Reeves et al., 2005), eventual gender effects on SWM and RTI (De Luca et al., 2003; Simpson et al., 2005), computer experience effects mainly on RTI (Iverson et al., 2009), and possible psychotropic drug effects on RVP, PAL, and SWM (Herrera-Guzmán et al., 2009; Herrera-Guzmán et al., 2010; Robbins et al., 1997).

METHOD

Participants and procedure

This work is part of a larger research project examining the validity and psychometric properties of the aforementioned CANTAB tests for their future clinical use with Portuguese older persons living in institutions. The current study was conducted in five retirement homes in the district of Lisbon. These institutions were simultaneously an independent living facility for autonomous individuals and an assisted living facility for dependent persons. In two of the institutions, these facilities were integrated into a single edifice, whereas in the other institutions, these facilities (and an additional continuing health care unit) operated in different buildings. The participants stayed in single rooms, couple rooms, or shared rooms with an average of two residents. All of the retirement homes had a dining room, living rooms, and recreational areas (e.g. gym, game room, craft room, informatics room, library, or gardens) as shared spaces.

A total of 171 older adults who spoke Portuguese as their first language and who had no established neuropsychiatric diagnosis, voluntarily participated in this study. They were recruited via oral and written advertisement in the retirement homes where they lived, and all of them gave written informed consent before participating in the study. A trained neuropsychologist (M.M.G.) conducted an initial session with each participant to ensure that he or she met the inclusion criteria. This 90-120 min session included a semi-structured interview and the administration of pencil-and-paper tests (see the order of test administration, the variables assessed by the tests and the cut-off scores in Table 1). A semi-structured interview was also conducted with the physician of each participant, with the consent of the latter. A close formal caregiver (e.g. physician, nurse, social worker) was asked to complete the Adults and Older Adults Functional Assessment Inventory (Inventário de Avaliação Funcional de Adultos e Idosos [IAFAI]; Sousa, Prieto, Vilar, Firmino, & Simões, 2015).

We used a less stringent cut-off for the Irregular Word Reading Test (Teste de Leitura de Palavras Irregulares [TeLPI]; Alves, Simões, & Martins, 2012) because we intended to include a wide range of years of education in our sample and because the TeLPI is highly correlated with this variable (education is taken into account in the TeLPI-IQ formula). For the 30-item Geriatric Depression Scale (GDS-30; Barreto, Leuschner, Santos, & Sobral, 2008; Yesavage et al., 1982), we used the severe depression cut-off score due to the negative impact of this condition on cognitive performance (e.g. Elderkin-Thompson et al., 2003). After checking the participants who met the inclusion criteria (see the criteria in Appendix A), 43 participants (25%) were excluded for the following reasons: 17 (40% of those excluded) due to neuropsychiatric problems (e.g. MCI, stroke, dementia, bipolar disease, major depression, Parkinson's disease, cerebral aneurysm, or epilepsy), 2 (5%) due to sensory or motor problems, 3 (7%) due to illiteracy, 12 (28%) due to abandonment of the study, and 9 (21%) due to the participant's failure to meet the cognitive test conditions (see the numbered conditions [1] to [3], Appendix A).

All of the 128 participants included in the normative sample were right handed (1 was ambidextrous). According to scores on Addenbrooke's Cognitive Examination Revised (ACE-R; Firmino, Simões, Pinho, Cerejeira, & Martins, 2008; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006), they were not cognitively impaired (Table 2). The percentage of men was smaller than that of women (see chi-square test of homogeneity also in Table 2). We classified as computer users the participants who used computers autonomously or who had attended computer classes for one or more months. We classified as computer-naïve the participants who had never used a computer or who were dependent on others to perform basic operations (e.g. opening or closing a program). Psychotropic drug users were participants with daily intake of at least one psychotropic drug, and psychotropic drug non-users were those who had not consumed such drugs within the range of at least one month before the participant's first session to the end of his or her participation in the study. According to GDS-30 scores (Table 2), the participants were not severely depressed. They also had no history of neuropsychiatric diagnosis, according to them and their physicians. Therefore, the psychotropic drug users' medication consumption was consistent with the prescription of PIMs (see medication types and drugs, Appendix B).

Table 1. Sequence of pencil-and-paper tests administered for screening.

Pencil-and-paper test	Description	Assessed variables (measures)	Portuguese norms	Cut-off
ACE-R ^a	A cognitive screening test that includes the MMSE ⁱ . Assesses attention/orientation, memory, verbal fluency, language, and visuospatial function	Global cognitive state (ACE-R and MMSE scores); Executive function (total animal names correctly produced on the semantic verbal fluency item)	Set by age and education, with community (ACE-R) and institution dwellers (semantic verbal fluency item) ^{j,k}	SS<77.5
Digit Span (WAIS-III) ^b	A task of repeating aloud digit sequences forward and backward	Attention and working memory (total score)	Set by age, with community dwellers	
Word Lists I (WMS-III) ^c	A 12-word list to learn and immediately recall aloud	Verbal episodic memory (immediate recall score)	Set by age, with community dwellers	
RCFT ^d	Involves copying an abstract figure and reproducing it from memory after a 3-min delay	Constructive praxis (copy); Visual episodic memory (immediate recall score)	Set by education, with community dwellers	
Irregular Word Reading Test (TeLPI) ^e	A 46-word list of irregular words to read aloud. Errors and education (years) are entered in the TeLPI-IQ conversion formula	Premorbid IQ (TeLPI-IQ score)	Set by education, with community dwellers	SS<70
Naming Objects and Fingers (ADAS) ^f	Consists of naming 12 familiar objects and the hand's fingers	Naming (total of errors)	Set by age and education, with institution dwellers	SS<77.5
Similarities (WAIS-III) ^b	Requires finding abstract concepts that associate two words	Abstraction and conceptualization (total score)	Set by age, with community dwellers	
Word Lists II (WMS-III) ^c	A 12-word list learned 25-35 min before the words are recalled and recognized	Verbal episodic memory (delayed recall and delayed recognition scores)	Set by age, with community dwellers	
GDS-30 ^g	A checklist (yes/no) of 30 items	Depressive symptoms (total score)	Sample of community dwellers	RS>20
Adults and Older Adults Functional Assessment Inventory (IAFAI) ^h	Assesses physical, cognitive, and emotional functional capacity for basic and instrumental ADLs	Cognitive functional capacity (global cognitive incapacity score)	Set by gender and education, with community dwellers	RS>P ₅ ¹

Note. We used the Portuguese versions and norms of the tests listed above. ^aAddenbrooke's Cognitive Examination Revised (Firmino et al., 2008; Mioshi et al., 2006). ^bWechsler Adult Intelligence Scale-III (Wechsler, 2008a). ^cWechsler Memory Scale-III (Wechsler, 2008b, 2008c). ^dRey Complex Figure Test (Espírito-Santo et al., 2015; Rey, 2002). ^eTeste de Leitura de Palavras Irregulares (Alves et al., 2012). ^fAlzheimer's Disease Assessment Scale (Gonçalves, Pinho, & Simões, 2012; Guerreiro, Fonseca, Barreto, & Garcia, 2008; Mohs, Rosen, & Davis, 1983). ^gGeriatric Depression Scale (Barreto et al., 2008; Yesavage et al., 1982). ^hInventário de Avaliação Funcional de Adultos e Idosos (Sousa et al., 2015).

ⁱMini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). ^jGonçalves et al. (2012). ^kThe MMSE scores were interpreted using the norms set by education specifically with Portuguese community dwellers of Lisbon (Morgado, Rocha, Maruta, Guerreiro, & Martins, 2009). ^lThe IAFAI norms are expressed solely in percentiles, and the raw scores corresponding to the 5th percentile (P₅) of each normative group were as follows: education 1-4 years (men: P₅ = 10; women: P₅ = 15); education 5-9 years (men: P₅ = 6; women: P₅ = 10); education 10+ years (men: P₅ = 5; women: P₅ = 9). Legend: ADLs = activities of daily living; RS = raw score; SS = standardized score.

Table 2. Normative sample's demographic characteristics and pencil-and-paper test scores.

<i>N</i> = 128	<i>n</i>	%	<i>M</i>	<i>SD</i>	Range	χ^2	<i>df</i>	<i>p</i>
Age (years)			81.38	6.05	69-96			
Education (years)			9.08	4.93	4-22			
Gender								
Women	77	60				5.28	1	.022
Men	51	40						
Computer experience								
Computer naïve	91	71						
Computer users	37	29						
Psychotropic drug use								
Psychotropic drug non-users	67	52						
Psychotropic drug users	61	48						
Institutionalization time (months)			44.47	45.31	3-232			
			Raw scores			Standardized scores		
<i>Pencil-and-paper test</i>			<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range
ACE-R								
Verbal fluency – animals			14.90	4.59	6-26	100.91	14.93	66-145
MMSE score			28.68	1.30	25-30	101.50	12.04	76-118
Total score			87.73	7.58	63-100	110.88	9.82	86-130
Digit Span (WAIS-III)								
Total score			13.26	3.50	7-26	115.94	14.39	85-145
Word Lists (WMS-III)								
Immediate recall			23.14	5.86	10-38	111.41	13.00	85-145
Delayed recall			4.57	2.28	1-12	112.77	12.69	90-145
Delayed recognition			21.05	2.21	14-24	108.67	12.27	80-130
RCFT								
Copy			32.76	3.23	19-36	107.74	8.27	76-121
Immediate recall			15.13	5.57	6-30	103.56	13.42	59-139
Irregular Word Reading Test (TeLPI-IQ)			106.48	12.98	77-132	100.31	14.08	70-133
Naming Objects and Fingers (ADAS) – errors			0.83	0.90	0-4	100.98	13.86	60-118
Similarities (WAIS-III)			16.57	5.53	5-28	121.60	12.64	95-145
GDS-30			8.25	5.21	0-20	N/A	N/A	N/A
IAFAI – cognitive incapacity								
Participants' test form			0.09	0.49	0-3	N/A	N/A	N/A
Caregivers' test form			0.02	0.27	0-3	N/A	N/A	N/A

Note. We used the Portuguese test versions and norms of the tests listed above. Proportions by categories of a single variable were compared with chi-square tests of homogeneity. Legend: ACE-R = Addenbrooke's cognitive examination revised; ADAS = Alzheimer's disease assessment scale; GDS-30 = Geriatric depression scale; IAFAI = Adults and older adults functional assessment inventory; MMSE = Mini-mental state examination; N/A = Not applicable; RCFT = Rey complex figure test; TeLPI-IQ = Premorbid intelligence quotient estimated by the Irregular Word Reading Test; WAIS-III = Wechsler adult intelligence scale-III; WMS-III = Wechsler memory scale-III.

No more than one week after the first session, the same neuropsychologist held a second session with each participant to administer the CANTAB tests in the following order: Motor Screening test (MOT), RVP, PAL, RTI, and SWM. The battery was performed in 39-94 min ($M = 54.01$, $SD = 8.85$). Because we found specifications regarding the order of administration of some CANTAB tests (e.g. MOT test; Cambridge Cognition, 2006b), a counterbalancing task procedure was not used. This procedure was also not used with the pencil-and-paper tests because there were delayed tasks in which a standard time interval needed to be observed. To minimize possible fatigue effects, the participants were allowed to take breaks between tests when needed. Regarding technology, all of the procedures recommended in the software guide (Cambridge Cognition, 2006a) were followed. As compensation, the participants received brief reports of their performance on the pencil-and-paper tests, free of charge. This study was approved by the Portuguese Foundation for Science and Technology, the Scientific Council of the Faculty of Psychology and Educational Sciences of the University of Coimbra, and the Principals and Ethics Committees of the retirement homes where the participants were living.

Apparatus and materials

The CANTAB was administered through CANTABeclipse™ version 3.0 software (Cambridge Cognition, 2006a) run in a Windows® XP environment. We used a Sahara Slate PC® i412T Tablet PC (Intel® Celeron® M ULV 423 Processor, 1.0 GB DDR2 RAM, 120 GB SATA Hard Disk Drive) with a 12.1-inch (307 mm) TFT XGA LCD passive resistive touch screen (monitor resolution: 1024 × 768; monitor size: 300 mm × 220 mm × 26 mm). A button press pad (Cambridge Cognition 2-Button Press Pad version 2.0) was used for RVP and RTI. An integrated touchpad keyboard (KeySonic™ ACK-540U+ Mini-Keyboard) was also employed (only by the examiner). According to Cambridge Cognition, the timing accuracy of the Sahara Slate i412T PC was tested with software equivalent to CANTABeclipse™ 3.0 using external chronometry. It was measured (in milliseconds [ms]) for video latency ($M = 7.40$, $SD = 0.64$), audio latency after calibration in the software ($M = 2.50$, $SD = 0.86$), and touch latency ($M = 63.65$, $SD = 5.00$).

Motor screening test

This test is intended to allow participants to become acquainted with the touch screen and to simultaneously track motor, visual or comprehension problems that may hinder test performance. It was run in high-visibility mode (which is recommended for older persons because the stimuli are presented in bold font). The participants were asked to touch a cross that appeared at different points on the screen. The MOT test was completed in approximately 4 min.

Rapid visual information processing test

This test is a visual sustained attention measure that has been shown to be sensitive to frontal and parietal lobe dysfunctions (Coull, Frith, Frackowiak, & Grasby, 1996). We administered the version recommended for individuals with AD (i.e. the RVP 3-5-7 version) in multiple alternate forms. In this test, digits appeared on the screen, one by one, at a rate of 100 stimuli per minute. After completing a training stage, the participants were asked to detect the 3-5-7 sequence by pressing a button after the final digit of the target sequence appeared. The test was completed in approximately 11 min. Because seven participants performed the most demanding version of RVP (which requires them to detect the target sequences 2-4-6, 4-6-8, and 3-5-7 instead of 3-5-7 only; Cambridge Cognition, 2006b), they were excluded from the analysis of the results for this test. The measures analyzed were A' (sensitivity to detect target sequences among non-target stimuli, ranging from 0 [bad] to 1 [good]; see RVP A' formula in Roiser et al., 2005) and mean latency (mean time to correctly detect a target sequence within the response window of 1800 ms).

Paired associates learning test

This is a visuospatial learning and memory test associated with entorhinal cortex activation (Owen, Sahakian, Semple, Polkey, & Robbins, 1995). It was administered in multiple alternate forms comprising 8 problems of increasing difficulty (1-8 patterns). The participants were asked to observe white boxes opening one by one, at random, and to memorize their contents (empty or with a specific pattern). Then, they were invited to touch the box where each pattern had appeared. There were a maximum of 10 trials per stage. The test ended if the participants failed all these trials. The test was completed in

approximately 18 min. The measure analyzed was mean errors to success (the mean number of errors made before the participant successfully completed a stage).

Reaction time test

This attention and reaction time measure is analogous to the five-choice serial reaction time task, which is sensitive to lesions on the medial prefrontal cortex and nucleus basalis in rats (McGaughy, Dalley, Morrison, Everitt, & Robbins, 2002). It was administered in multiple alternate forms. In both the simple and five-choice conditions (i.e. 1 circle and 5 circles, respectively), the participants were invited to press a button until a yellow spot appeared inside a circle and to then immediately release the button and touch that circle. They completed training stages before the test (i.e. two stages for touching the circle and one stage for simply releasing the button). If a criterion of 5 of 6 correct responses was not achieved, the test ended. One participant did not satisfy the criterion with the five-choice condition. The test was completed in approximately 7 min. The measures analyzed were simple movement time and five-choice movement time (time [ms] to touch the circle after button release).

Spatial working memory test

This test assesses the ability to retain and manipulate spatial information in working memory. It has been shown to be sensitive to dysfunctions of the frontal lobe (Owen et al., 1995) and dorsolateral prefrontal cortex (Manes et al., 2002). The test covers four levels of increasing difficulty, with the first being for training. The participants were asked to search for a blue token hidden within one of several possible boxes and to then place the token in an empty house. After that token was found, another token was automatically hidden in another box. The participants were invited to search for as many tokens as needed to fill the house and instructed not to reopen boxes that already had a blue token. The test was completed in approximately 14 min. The measure analyzed was total errors (number of times a box was touched when it clearly did not contain a token).

Statistical analysis

Statistical analyses were conducted with SPSS[®] Statistics, version 22.0 (IBM[®], New York, USA), and Microsoft[®] Excel[®] 2010. The significance level adopted was $\alpha = .05$.

Multiple linear regressions (MLRs) were used to examine the effects of the socio-demographic variables on each CANTAB measure. The socio-demographic variables were simultaneously entered into the MLRs as continuous variables (i.e. age and education in years and institutionalization time in months) and dichotomous variables (i.e. gender: 0 = woman, 1 = man; computer experience: 0 = computer naïve, 1 = computer user; psychotropic drug use: 0 = psychotropic drug non-user, 1 = psychotropic drug user). Education was analyzed instead of TeLPI-IQ because the two variables are strongly correlated, $r_s(126) = .94, p < .001$, and because the TeLPI is a very recent test that has yet to be widely used in Portugal.

The analyzed CANTAB measures were those that previously showed test-retest correlations equal to or above .70 in older Portuguese persons living in retirement homes (Gonçalves et al., 2016). The RVP A', PAL mean errors to success, and RTI five-choice movement time measures were transformed (due to non-normal distributions of the residuals) using the following formulas that were found to be most appropriate for the data (Marôco, 2014): $tRVP A' = (1 - RVP A')^{1/2} - 1/3 (1 - RVP A')^{3/2}$; $tPAL \text{ mean errors to success} = (PAL \text{ mean errors to success})^{1/2} + (PAL \text{ mean errors to success} + 1)^{1/2}$; and $tRTI \text{ five-choice movement time} = RTI \text{ five-choice movement time}^{(1-3/2)}$. The formulas for tRVP A' and tRTI five-choice movement time reflected the data; therefore, lower scores on tRVP A' and higher scores on tRTI five-choice movement time indicate good performance. One residual outlier was found for SWM total errors and was extracted to meet the criterion of normality for the distribution of residuals. The extraction of this outlier did not change the conclusions about the effects of the socio-demographic variables on SWM total errors.

To obtain a broader perspective on the results, we also performed MLRs (simultaneous entry method) to examine the effects of the socio-demographic variables on six pencil-and-paper test measures assessing cognitive functions similar to those assessed by the CANTAB measures. Because institutionalization time did not (marginally or significantly) influence the CANTAB and pencil-and-paper test measures (data not shown), we performed all of the MLRs a second time but without entering the institutionalization time variable, to increase the statistical power of the MLRs. The extraction of this variable did not change the conclusions about the effects of the other socio-demographic variables on test performance. The final MLR models

had centered leverage maximum values under .13 and met the assumptions of linearity (scatterplot graphs), normality (Kolmogorov-Smirnov or Shapiro-Wilk tests), homoscedasticity (scatterplot graphs), and independence of residuals (Durbin-Watson statistic value range: 1.70-2.16). The criterion of an absence of multicollinearity was also met according to the tolerance and variance inflation factor (VIF) values (tolerance value range: .79-.96; VIF value range: 1.04-1.27). Power and f^2 effect sizes were calculated using Cohen's (1988) formulas.

The Portuguese CANTAB norms were set for groups defined by no more than two socio-demographic variables to ensure that these groups did not become too small. We selected the variables presenting larger effects on test performance according to semi-partial correlations (sr) from the MLRs. Raw scores (X) were converted into standardized scores (SS) using the mean (M) and standard deviation (SD) of each normative group, with the formula $SS = 15 [(X - M)/SD] + 100$. For measures whose lower raw scores were indicative of better performance, this formula was reversed (i.e. $SS = 15 [(M - X)/SD] + 100$) so that for all measures, higher SS corresponded to better scores.

RESULTS

Analyses of the effects of the socio-demographic variables on the CANTAB measures

The MLR models were significant ($p < .05$) for all of the CANTAB measures except for tRTI five-choice movement time (see F statistics, Table 3). The total variance of this CANTAB measure was not significantly explained by any of the socio-demographic variables (see semi-partial correlations also in Table 3). Analyzing the CANTAB measures whose MLR models were significant, we found that the socio-demographic variables collectively explained 7-14% of the total variance of each CANTAB measure (see adjusted R -squared values), corresponding to f^2 effect sizes of small to medium magnitude (i.e. small: .02-.14, medium: .15-.34; Cohen, 1988). The highest percentages of variance explained, corresponding to medium effect sizes, were identified for tRVP A' and SWM total errors. The strongest semi-partial correlation found between a socio-demographic variable and a CANTAB measure was .25; therefore, the significant associations identified were weak, at most (i.e. weak: .10-.29; Cohen, 1988).

According to the semi-partial correlations (Table 3), age significantly influenced tRVP A' and SWM total errors and had a marginally significant effect on RVP mean latency, with younger participants showing better performance than older participants. No marginally or significant effect of education on the CANTAB measures was found. Gender significantly influenced RVP mean latency, tPAL mean errors to success, and SWM total errors, with women performing better than men on tPAL mean errors to success and men performing better on the other two measures. No significant effect of computer experience on the CANTAB measures was found, but a marginally significant effect on RTI simple movement time was identified, with the computer users performing faster than the computer-naïve participants. Significant effects of psychotropic drug use on tRVP A' and tPAL mean errors to success as well as a marginally significant effect on RTI simple movement time were found, with the psychotropic drug non-users performing better than psychotropic drug users.

The socio-demographic variables that contributed most substantially to the total variance of each CANTAB measure were age ($sr^2 = 4.84\%$) followed by psychotropic drug use ($sr^2 = 3.24\%$) on tRVP A', gender ($sr^2 = 6.25\%$) followed by age ($sr^2 = 2.56\%$) on RVP mean latency, psychotropic drug use ($sr^2 = 4.41\%$) followed by gender ($sr^2 = 2.89\%$) on tPAL mean errors to success, gender ($sr^2 = 6.25\%$) followed by age ($sr^2 = 3.24\%$) on SWM total errors, and psychotropic drug use ($sr^2 = 2.56\%$) followed by computer experience ($sr^2 = 2.25\%$) on RTI simple movement time.

Table 3. Effects of the socio-demographic variables on the automated test measures.

Measure	Predictor	<i>B</i>	<i>SE_B</i>	<i>sr</i>	<i>p</i>	<i>N</i>	Model statistics
tRVP A' (sensitivity index)	Constant	-0.119	0.129			121	$F(5, 115) = 4.95; p < .001$
	Age	0.004	0.002	.22	.009		Power = .99
	Education	-0.003	0.002	-.12	.143		$R^2_a = .14; f^2 = .22$
	Gender	-0.021	0.020	-.09	.286		$SEE = 0.10$
	Computer exp.	-0.022	0.022	-.08	.319		
	Psych. drug use	0.042	0.020	.18	.033		
RVP mean latency	Constant	233.03	147.40			121	$F(5, 115) = 3.30; p = .008$
	Age	3.17	1.78	.16	.077		Power = .92
	Education	-0.22	2.37	-.01	.925		$R^2_a = .09; f^2 = .14$
	Gender	-65.30	22.91	-.25	.005		$SEE = 114.24$
	Computer exp.	-41.49	25.56	-.14	.107		
	Psych. drug use	-22.62	22.47	-.09	.316		
tPAL mean errors to success	Constant	2.64	1.39			128	$F(5, 122) = 2.81; p = .019$
	Age	0.02	0.02	.13	.142		Power = .83
	Education	-0.04	0.02	-.14	.104		$R^2_a = .07; f^2 = .12$
	Gender	0.43	0.21	.17	.045		$SEE = 1.11$
	Computer exp.	-0.06	0.24	-.02	.814		
	Psych. drug use	0.51	0.21	.21	.015		
SWM total errors	Constant	25.85	18.87			127	$F(5, 121) = 4.24; p = .001$
	Age	0.48	0.23	.18	.036		Power = .97
	Education	-0.36	0.30	-.10	.233		$R^2_a = .11; f^2 = .18$
	Gender	-8.78	2.94	-.25	.003		$SEE = 15.14$
	Computer exp.	-4.11	3.33	-.10	.219		
	Psych. drug use	-3.37	2.87	-.10	.242		
RTI simple movement time	Constant	919.78	246.59			128	$F(5, 122) = 2.84; p = .018$
	Age	-4.00	2.96	-.12	.179		Power = .83
	Education	-5.08	3.92	-.11	.197		$R^2_a = .07; f^2 = .12$
	Gender	0.57	38.21	.00	.988		$SEE = 197.82$
	Computer exp.	-75.50	43.42	-.15	.085		
	Psych. drug use	67.99	37.23	.16	.070		
tRTI five-choice movement time	Constant	0.0357	0.0097			127	$F(5, 121) = 1.96; p = .090$
	Age	0.0001	0.0001	.10	.278		Power = .64
	Education	0.0001	0.0002	.06	.508		$R^2_a = .04; f^2 = .08$
	Gender	0.0013	0.0015	.08	.376		$SEE = 0.01$
	Computer exp.	0.0020	0.0017	.10	.242		
	Psych. drug use	-0.0023	0.0015	-.14	.123		

Note. The predictors were simultaneously entered in the MLRs as dichotomous variables (gender: 0 = woman, 1 = man; psychotropic drug use: 0 = psychotropic drug non-user, 1 = psychotropic drug user; computer experience: 0 = computer naïve, 1 = computer user) and continuous variables (age: years; education: years). Due to non-normal distributions of the residuals, the three CANTAB measures were transformed with the following formulas: $tRVP A' = (1 - RVP A')^{1/2} - 1/3 (1 - RVP A')^{3/2}$; $tPAL \text{ mean errors to success} = (PAL \text{ mean errors to success})^{1/2} + (PAL \text{ mean errors to success} + 1)^{1/2}$; and $tRTI \text{ five-choice movement time} = RTI \text{ five-choice movement time}^{(1-3/2)}$. For all of the CANTAB measures listed except for tRTI five-choice movement time, lower values indicate better performance. The *p* values lower than .05 are in bold font. Legend: CANTAB = Cambridge neuropsychological test automated battery; exp. = experience; PAL = Paired associates learning; Psych. = psychotropic; RTI = Reaction time; RVP = Rapid visual information processing; SWM = Spatial working memory.

Comparative analysis of the effects of the socio-demographic variables on the CANTAB and pencil-and-paper test measures

The MLR models for the six pencil-and-paper test measures were all significant (see F statistics, Table 4). The socio-demographic variables collectively explained 10-33% of the total variance of each pencil-and-paper test measure (R_a^2 values, Table 4), equivalent to f^2 effect sizes of medium to large magnitude (medium: .15-.34, large: .35 or higher; Cohen, 1988). The highest percentages of variance explained, corresponding to large effect sizes, were identified for the Word Lists immediate recall (Wechsler Memory Scale-III [WMS-III]; Wechsler, 2008b, 2008c) and semantic verbal fluency (ACE-R) scores. The strongest semi-partial correlation found between a socio-demographic variable and a pencil-and-paper test measure was .33, indicating that the associations were at most moderate (small: .10-.29, moderate: .30-.49; Cohen, 1988). Thus, the influence of the socio-demographic variables on the pencil-and-paper test measures was, in general, greater than on the CANTAB measures.

Analyzing the semi-partial correlations for the pencil-and-paper test measures (Table 4), we found significant effects of age on the semantic verbal fluency (ACE-R) and Word Lists immediate recall (WMS-III) scores, with younger participants performing better than older participants. The unique contribution (sr^2) of age to the total variance of each of these two pencil-and-paper test measures (i.e. 4.84-5.29%, Table 4) was similar to that observed for the two CANTAB measures that were significantly influenced by age (i.e. 3.24-4.84%, Table 3).

Although we found no significant effects of education on the CANTAB measures (Table 3), this socio-demographic variable significantly influenced five of the six pencil-and-paper test measures – namely, the semantic verbal fluency score (ACE-R), Digit Span forward and backward scores (Wechsler Adult Intelligence Scale-III [WAIS-III]; Wechsler, 2008a), RCFT immediate recall, and Word Lists immediate recall (WAIS-III), with more educated participants exhibiting better performance than less educated participants (see semi-partial correlations, Table 4). The unique contribution (sr^2) of education to the total variance of each of these pencil-and-paper test measures was 7.29-9.00%.

Table 4. Effects of the socio-demographic variables on the pencil-and-paper test measures.

Measure	Predictor	<i>B</i>	<i>SE_B</i>	<i>sr</i>	<i>p</i>	Model statistics (<i>N</i> = 128)
Verbal Fluency (animals)	Constant	26.59	4.86			$F(5, 122) = 10.81; p < .001$
	Age	-0.17	0.06	-.22	.004	Power = 1.00
	Education	0.31	0.08	.30	<.001	$R^2_a = .28; f^2 = .44$
	Gender	-0.22	0.75	-.02	.770	$SEE = 3.90$
	Computer exp.	1.79	0.86	.16	.038	
Digit Span (WAIS-III) forward	Psych. drug use	-1.72	0.73	-.18	.021	
	Constant	7.57	2.52			$F(5, 122) = 5.31; p < .001$
	Age	-0.01	0.03	-.02	.819	Power = .99
	Education	0.14	0.04	.29	.001	$R^2_a = .15; f^2 = .22$
	Gender	-0.21	0.39	-.05	.583	$SEE = 2.02$
Digit Span (WAIS-III) backward	Computer exp.	0.74	0.44	.14	.096	
	Psych. drug use	-0.35	0.38	-.08	.359	
	Constant	6.30	2.14			$F(5, 122) = 4.01; p = .002$
	Age	-0.03	0.03	-.10	.251	Power = .95
	Education	0.11	0.03	.28	.001	$R^2_a = .11; f^2 = .16$
RCFT immediate recall	Gender	-0.07	0.33	-.02	.839	$SEE = 1.72$
	Computer exp.	0.36	0.38	.08	.342	
	Psych. drug use	-0.01	0.32	.00	.976	
	Constant	19.71	6.60			$F(5, 122) = 3.72; p = .004$
	Age	-0.09	0.08	-.10	.260	Power = .93
Word Lists (WMS-III) immediate recall	Education	0.34	0.10	.27	.002	$R^2_a = .10; f^2 = .15$
	Gender	-0.05	1.02	.00	.964	$SEE = 5.30$
	Computer exp.	0.62	1.16	.05	.592	
	Psych. drug use	-1.09	1.00	-.09	.277	
	Constant	39.92	5.99			$F(5, 122) = 13.32; p < .001$
Word Lists (WMS-III) delayed recall	Age	-0.23	0.07	-.23	.002	Power = 1.00
	Education	0.37	0.10	.28	<.001	$R^2_a = .33; f^2 = .55$
	Gender	-4.21	0.93	-.33	<.001	$SEE = 4.81$
	Computer exp.	2.58	1.06	.18	.016	
	Psych. drug use	-1.43	0.91	-.11	.117	
Word Lists (WMS-III) delayed recall	Constant	7.56	2.70			$F(5, 122) = 3.69; p = .004$
	Age	-0.04	0.03	-.11	.215	Power = .93
	Education	0.07	0.04	.13	.113	$R^2_a = .10; f^2 = .15$
	Gender	-1.06	0.42	-.21	.012	$SEE = 2.17$
	Computer exp.	0.90	0.48	.16	.060	
	Psych. drug use	-0.34	0.41	-.07	.404	

Note. The predictors were simultaneously entered in the MLRs as dichotomous variables (gender: 0 = woman, 1 = man; psychotropic drug use: 0 = psychotropic drug non-user, 1 = psychotropic drug user; computer experience: 0 = computer naïve, 1 = computer user) and continuous variables (age: years; education: years). For all of the pencil-and-paper test measures listed, higher values indicate better performance. The *p* values lower than .05 are in bold font. Legend: Psych. = psychotropic; RCFT = Rey complex figure test; WAIS-III = Wechsler adult intelligence scale-III; WMS-III = Wechsler memory scale-III.

We found a significant effect of gender on Word Lists immediate and delayed recall scores (WMS-III), with women performing better than men (see semi-partial correlations, Table 4), as was observed for tPAL mean errors to success (CANTAB, see

Table 3). Gender uniquely contributed (sr^2) to 4.41-10.89% of the total variance of each of these pencil-and-paper test measures, a slightly larger contribution than that observed for each of the three CANTAB measures that were also significantly influenced by gender (i.e. 2.89-6.25%, Table 3).

Although we found only a marginally significant effect of computer experience on one of the CANTAB measures (i.e. RTI simple movement time, Table 3), this variable had a significant effect on the semantic verbal fluency (ACE-R) and Word Lists immediate recall (WMS-III) scores and a marginally significant effect on the Digit Span forward (WAIS-III) and Word Lists delayed recall (WMS-III) scores, with computer users performing better than computer-naïve participants (see semi-partial correlations, Table 4). Nevertheless, the unique contribution (sr^2) of computer experience to the total variance of each of these four pencil-and-paper test measures (i.e. 1.96-3.24%, Table 4) was almost the same as that found for the RTI simple movement time (CANTAB) measure (i.e. 2.25%, Table 3).

Psychotropic drug use significantly influenced performance on the semantic verbal fluency task (ACE-R), with the psychotropic drug non-users producing more correct animal names than the psychotropic drug users (see semi-partial correlations, Table 4). The unique contribution (sr^2) of psychotropic drug use to the total variance of this pencil-and-paper test measure (i.e. 3.24%) was similar to that observed for each of the three CANTAB measures that were marginally to significantly influenced by this socio-demographic variable (i.e. 2.56-4.41%, Table 3).

Portuguese normative values

Based on the MLR results, we stratified the normative sample into groups defined by age and psychotropic drug use (for RVP A'), gender and age (for RVP mean latency and SWM total errors), psychotropic drug use and gender (for PAL mean errors to success), and psychotropic drug use and computer experience (for RTI simple movement time). Therefore, the marginally significant effects of the socio-demographic variables on RVP mean latency and RTI simple movement time were also accounted for in the sample stratification, to enable more stringent performance comparisons. Table 5 shows the raw score means and standard deviations calculated for the Portuguese normative groups. We found a ceiling effect (i.e. where the maximum possible score is

reached) for the RVP A' measure and no other ceiling or floor effects (i.e. where the minimum possible score is obtained) for the other CANTAB measures.

Table 5. Portuguese sample's raw scores on the automated tests.

CANTAB measure	<i>N</i>	<i>M</i>	<i>SEM</i>	<i>SD</i>	Range
RVP A' (sensitivity index)					
Age 69-79 years, psych. drug non-users	26	.974	.006	.028	.852-1.000
Age 69-79 years, psych. drug users	21	.945	.010	.048	.820-1.000
Age 80+ years, psych. drug non-users	37	.956	.008	.050	.731-1.000
Age 80+ years, psych. drug users	37	.921	.012	.076	.704-1.000
RVP mean latency					
Age 69-79 years, men	20	394.32	22.74	101.71	279.95-683.05
Age 69-79 years, women	27	430.42	21.78	113.16	269.25-795.81
Age 80+ years, men	30	411.32	18.30	100.23	228.00-631.60
Age 80+ years, women	44	484.68	19.76	131.05	249.56-789.55
PAL mean errors to success					
Men, psych. drug non-users	37	5.69	0.49	2.98	0.75-11.57
Men, psych. drug users	14	5.95	0.59	2.21	1.25-8.86
Women, psych. drug non-users	30	4.16	0.40	2.20	1.00-9.40
Women, psych. drug users	47	5.79	0.39	2.67	1.75-11.29
SWM total errors					
Age 69-79 years, men	21	43.95	4.34	19.91	4-81
Age 69-79 years, women	29	54.66	2.13	11.49	24-72
Age 80+ years, men	30	54.90	3.01	16.47	15-97
Age 80+ years, women	48	60.00	2.21	15.32	14-85
RTI simple movement time					
Computer users, psych. drug non-users	23	472.79	38.24	183.40	198.13-880.44
Computer users, psych. drug users	14	514.89	45.60	170.63	290.29-857.78
Computer-naïve, psych. drug non-users	44	544.80	30.07	199.49	240.25-977.50
Computer-naïve, psych. drug users	47	627.47	30.96	212.27	270.38-1153.56
RTI five-choice movement time					
Total sample	127	493.96	15.76	177.65	200.71-1249.63

Note. Time is in milliseconds. For all of the CANTAB measures listed except for RVP A', lower values indicate better performance. Legend: CANTAB = Cambridge neuropsychological test automated battery; PAL = Paired associates learning; psych. = psychotropic; RTI = Reaction time; RVP = Rapid visual information processing; SWM = Spatial working memory.

Tables that convert the raw scores into SS are provided in Appendices (C)-(H) (for RVP A', RVP mean latency, PAL mean errors to success, SWM total errors, RTI simple movement time, and RTI five-choice movement time, respectively). To convert a raw score into a standardized score, consider the example of a 74-year-old woman with a score of 63 on the SWM total errors measure. This raw score corresponds to the

standardized score of 89 (see Appendix F); therefore, it is less than 1 *SD* below the normative mean (i.e. $M = 100$, $SD = 15$), indicating that this women exhibited normal performance for a healthy retirement home dweller on the SWM test. In contrast, a standardized score of 77 or less could be suggestive of a spatial working memory function deficit if corroborated by other cognitive tests and clinical information.

DISCUSSION

After examining the effects of age, education, gender, computer experience, institutionalization time, and psychotropic drug use on four CANTAB tests performed by older Portuguese persons with no neuropsychiatric diagnosis and living in retirement homes, we found that age significantly influenced performance on tRVP A' (sensitivity index) and SWM total errors and had a marginally significant effect on RVP mean latency. Education did not influence any of the CANTAB measures. Gender significantly influenced performance on RVP mean latency, tPAL mean errors to success, and SWM total errors. Computer experience had only a marginally significant effect on RTI simple movement time. Psychotropic drug use significantly influenced performance on tRVP A' and tPAL mean errors to success and had a marginally significant effect on RTI simple movement time. No significant effect of institutionalization time was observed on the CANTAB measures. Therefore, none of the socio-demographic variables significantly influenced performance on tRTI five-choice movement time.

Furthermore, the analysis of the effects of the socio-demographic variables on six pencil-and-paper test measures revealed a significant effect of age on the semantic verbal fluency (ACE-R) and Word Lists immediate recall (WMS-III) scores, with age uniquely contributing to the total variance of each of these measures ($sr^2 = 4.84\%$ and $sr^2 = 5.29\%$, respectively) in a similar percentage to that observed for tRVP A' and SWM total errors (CANTAB; $sr^2 = 4.84\%$ and $sr^2 = 3.24\%$, respectively). For all of these measures, younger participants outperformed older participants. Education, which had no significant effect on the CANTAB measures, significantly influenced performance on all the pencil-and-paper test measures except Word Lists delayed recall (WMS-III), with more educated participants outperforming less educated participants (sr^2 value range: 7.29-9.00%). Gender significantly influenced Word Lists immediate

and delayed recall (WMS-III) scores, with women obtaining better memory scores than men, as observed for tPAL mean errors to success (CANTAB). However, the unique contribution of gender to the total variance of each of these pencil-and-paper test measures ($sr^2 = 10.89\%$ and $sr^2 = 4.41\%$, respectively) was greater than that for tPAL mean errors to success (CANTAB; $sr^2 = 2.89\%$) as well as for RVP mean latency and SWM total errors ($sr^2 = 6.25\%$ for these two CANTAB measures, on which men outperformed women). Notably, computer experience significantly influenced the semantic verbal fluency (ACE-R) and Word Lists immediate recall (WMS-III) scores but uniquely contributed to a similar percentage of the total variance of each of these measures ($sr^2 = 2.56\%$ and $sr^2 = 3.24\%$, respectively) as observed for RTI simple movement time (CANTAB; $sr^2 = 2.25\%$), for which a marginally significant effect was found. For all of these measures, computer users performed better than computer-naïve individuals. Psychotropic drug use significantly influenced the semantic verbal fluency (ACE-R) score, where the unique contribution of this variable to the total variance of this measure ($sr^2 = 3.24\%$) was similar to that observed for tRVP A' and RTI simple movement time (CANTAB; $sr^2 = 3.24\%$ and $sr^2 = 2.56\%$, respectively). For all of these measures, psychotropic drug non-users outperformed psychotropic drug users. Finally, institutionalization time did not influence the pencil-and-paper test measures, as observed for the CANTAB measures.

Effect of age

Concerning the influence of age on the CANTAB measures, our findings were consistent with previous studies reporting significant age effects on RVP A', RVP mean latency, and SWM total errors, with younger individuals showing better performance (Braw et al., 2011; Maylor et al., 2006; Saunders & Summers, 2010). The age effect that we found on RVP mean latency was nevertheless marginally significant. However, some studies report no effect of age on RVP A' or SWM total errors (Rabbitt & Lowe, 2000; Saunders & Summers, 2010). Moreover, in contrast with previous studies (e.g. Maylor et al., 2006; Skolimowska et al., 2011) we found no significant effect of age on the tPAL mean errors to success and RTI measures. These contradictory results may be explained by differences in sample characteristics (such as the age ranges included), sample sizes, or statistical methods (e.g. analyzing correlations instead of semi-partial

correlations, treating age as a categorical instead of continuous variable). Regarding the age ranges, Robbins et al. (1998) found that only groups with a large age gap (e.g. 75-79 years olds compared with 60-year-old or younger individuals) differ significantly on SWM between errors (a measure included in SWM total errors), indicating that performance on this test slightly decreases with age. Thus, studies that include wider age ranges may more easily detect age effects on SWM than studies that include smaller age ranges. Nevertheless, the specific age range included may also contribute to explaining the discrepancies in the reported result. For example, the decline in normal age-related performance on PAL and RTI may be more pronounced among individuals younger than those included in our sample and smoother at more advanced ages. This normal aging process has already been described for cognitive domains such as memory, reasoning, and processing speed (Salthouse, 2004). Such a pattern of performance decline might therefore explain the nonsignificant age effects that we found on PAL and RTI and on two pencil-and-paper test measures assessing memory (i.e. RCFT immediate recall and Word Lists delayed recall [WMS-III]).

Effect of education

Regarding the influence of education on the CANTAB measures, the nonsignificant effects that we found on tRVP A', tPAL mean errors to success, and SWM total errors are consistent with studies reporting no significant effect of education on these or other related CANTAB measures (i.e. RVP A', PAL mean trials to success, and SWM between errors; Braw et al., 2011; Hogervorst et al., 2004) but are inconsistent with a study that found a significant influence of education on the variability of a factor with high loadings of SWM measures (i.e. SWM between errors and SWM strategy; Reeves et al., 2005). As mentioned, the presence of different sample characteristics or the use of distinct statistical procedures may explain the discrepancies between studies. However, we also found significant effects of education on performance on almost all of the analyzed pencil-and-paper test measures in our sample. Therefore, the nonsignificant effects of education on the CANTAB measures may also be explained by the nature of the CANTAB tests, which involve tasks and stimuli that are perhaps less culturally biased than those of the analyzed pencil-and-paper tests.

Effect of gender

With respect to the influence of gender on the CANTAB measures, our results are consistent with a study that found a significant effect of gender on SWM between errors, with men performing better than women (De Luca et al., 2003). However, our results are inconsistent with studies that have reported no significant effect of gender on RVP mean latency, PAL or SWM (Braw et al., 2011; Jones et al., 1992; Simpson et al., 2013; Skolimowska et al., 2011). Moreover, we did not find a significant effect of gender on RTI, in line with the results described by Simpson et al. (2005) for a French and an Irish sample but in contrast to the results these authors reported for an Italian sample. Gender-related cultural differences may explain these contradictory results without excluding other hypothesis regarding differences in the statistical methods used and in sample sizes, as already noted.

Effect of computer experience

With regard to the influence of computer experience on the CANTAB measures, we expected to find significant effects on RTI. However, only a marginally significant effect on RTI simple movement time was observed, whereas no significant effect on tRTI five-choice movement time was found. Because the five-choice condition follows the simple condition, a motor training effect might have occurred during RTI such that no significant differences in performance between computer users and computer-naïve individuals were detected in this last condition. Interestingly, we found significant effects of computer experience on two pencil-and-paper test measures (i.e. semantic verbal fluency [ACE-R] and Word Lists immediate recall [WMS-III]) and marginally significant effects on two other measures (i.e. Digit Span forward [WAIS-III] and Word Lists delayed recall [WMS-III]), with computer users performing better. This result suggests that computer experience may involve more than the ability to use computer technology. For example, it may be related to specific cognitive skills acquired through experience using computer software (e.g. attention, mental flexibility, or associative thinking). Thus, these cognitive skills may have helped the computer users, for example, organize or create associations between words in Word Lists (WMS-III). However, such cognitive skills would be expected to also benefit the performance of computer users on the CANTAB tests. Considering that all of the pencil-and-paper tests

that were significantly influenced by computer experience were verbal tests, we may conjecture that the cognitive skills that were developed by computer software use and favor performance on these tests might be verbal. Hence, these verbal skills would be less effective for successful completion of visual tests such as the CANTAB tests. Even so, these are intriguing results in view of studies suggesting the eventual involvement of verbal skills during performance on RVP (e.g. sub-vocal rehearsal of digits; Coull et al., 1996) and PAL (e.g. creating semantic associations between visual patterns and locations; Gonçalves et al., 2016). Ultimately, it seems that more important than simply knowing whether an individual is a computer user is knowing which programs he or she uses or which tasks he or she performs to better understand the nature of the cognitive skills that could be developed through these tasks and that would favor performance on verbal pencil-and-paper tests but not on visual automated tests.

Effect of psychotropic drug use

Concerning the influence of psychotropic drug use on test performance, our results revealed that psychotropic drug users (84% ingesting benzodiazepines, 47% antidepressants, and 7% anxiolytics) exhibited poorer performance on three CANTAB measures and one pencil-and-paper test measure. The most sensitive measure of psychotropic drugs was tPAL mean errors to success, followed by tRVP A' (CANTAB), semantic verbal fluency (ACE-R), and RTI simple movement time (CANTAB). Considering that benzodiazepines were the most commonly used type of medication, our results were consistent with those of Robbins et al. (1997), which revealed that diazepam impairs PAL test performance. Both psychotropic drug users and non-users were individuals without neuropsychiatric diagnoses or severe depressive symptoms. Therefore, the performance differences observed between these groups may be explained by psychotropic drug effects. We found that more CANTAB measures than pencil-and-paper test measures were sensitive to psychotropic drug use. Given that three of the four measures that were sensitive to psychotropic drug use involved a speed component (i.e. tRVP A' [CANTAB], semantic verbal fluency [ACE-R], and RTI simple movement time [CANTAB]), the greater sensitivity of the CANTAB tests to psychotropic drugs could be explained by CANTAB tests comprising more psychomotor speed measures than the pencil-and-paper tests, and the psychomotor

speed function may be impaired by psychotropic drugs. However, the RVP mean latency and RTI five-choice movement time (CANTAB) measures also involve psychomotor speed but were not sensitive to psychotropic drug use. Thus, additional cognitive functions may be impaired by psychotropic drugs (e.g. selective attention, a component involved in tRVP A' [CANTAB] and semantic verbal fluency [ACE-R]). The most sensitive measure of psychotropic drugs was tPAL mean errors to success (CANTAB). The PAL test has shown an association with attention measures but primarily with episodic memory and learning measures (Fowler, Saling, Conway, Semple, & Louis, 1995). Nevertheless, episodic memory and learning measures such as RCFT immediate recall and Word Lists immediate and delayed recall (WMS-III) were not sensitive to psychotropic drug use. These pencil-and-paper test measures, unlike tPAL mean errors to success (CANTAB), have the ability to allow an individual to commit mistakes that are not counted in the total scores. For example, the Word Lists (WMS-III) immediate and delayed recall scores are simply the sum of the correctly recalled words. Similarly, in the RCFT, an individual can commit a mistake and fix it (crossing out or drawing over it). However, in the PAL test, once a wrong answer is given, the error is recorded and affects the error scores. Hence, the psychotropic drugs may have impaired performance on tPAL mean errors to success by scattering attention.

Effect of institutionalization time

With respect to institutionalization time, no significant effect on any of the CANTAB or pencil-and-paper test measures was found. Because our sample included a large institutionalization time range, we conjecture that the nonsignificant effects were not due to a short time range that would impede detection of changes in performance. It can be assumed that institutionalization time may influence cognition through the cumulative impact of institutionalization, but that its impact may differ according to institutional conditions and individual characteristics (e.g. motivation for institutionalization, health status). In the same way, considering that institutionalization represents a change in life conditions, it can be assumed that the impact of institutionalization is greater in the first months but that this impact progressively decreases from a normal process of adaptation. The effects in any of these hypotheses would contribute to weakening the association between institutionalization time and

cognitive test performance, thus hampering the detection of significant effects through MLRs. Notably, however, the impact of institutionalization on cognition remains unclear, not only because of the methodological problems of many studies (e.g. Plati et al., 2006; Rapp et al., 2005), but also because of the complexity of this variable. For example, it is difficult to distinguish the impact of institutionalization from the impact of the transition to a new place with new neighbors, an aspect involved in the institutionalization process but not specific to it (i.e. this aspect can also occur among those living in the community). The positive or negative impact of institutionalization may also depend on the quality of the accommodations, care services, and environmental conditions provided by institutions, which are difficult to operationalize. Therefore, although institutionalization time is not equivalent to the institutionalization variable, the former also involves complexities that render studying its influence on cognition difficult.

Overall, our findings suggest that the analyzed socio-demographic variables have a different influence on different test measures. We have noted that the majority of CANTAB normative studies (conducted with community dwellers) provide norms for groups defined by the same socio-demographic variables for all the tests considered (e.g. De Luca et al., 2003; Luciana & Nelson, 2002; Roque et al., 2011). In these and other types of CANTAB studies (e.g. Robbins et al., 1998; Simpson et al., 2013; Skolimowska et al., 2011), groups are usually defined by age or by age and gender. Therefore, future studies should take into consideration that different test measures may be influenced by different socio-demographic variables. Our results also suggest that the impact of unfamiliarity with computers on computerized test performance is negligible for tests requiring simple motor responses, such as touching a screen or pressing a single button; however, computer experience may involve the development of cognitive skills that affect verbal pencil-and-paper tests.

Study limitations, strengths, and future directions

A limitation of this study is that it includes in a normative sample participants who reported psychotropic drug use consistent with the prescription of PIM. Typically, normative studies exclude such drug users due to the well-known effects of psychotropic drugs in bolstering or hampering test performance (e.g. Herrera-Guzmán

et al., 2010; Robbins et al., 1997). Given the high rate of PIM prescription identified in Portuguese retirement homes (Periquito et al., 2014), the exclusion of these individuals would result in a much smaller sample size that would be less representative of this population, thereby jeopardizing the attainment of the study's aims. Therefore, we analyzed psychotropic drug use as another socio-demographic variable, simultaneously attempting to control for its effect on the remaining variables using MLRs. However, the effects of the psychotropic drugs were not analyzed separately due to the high number of active substances found, the consumption of multiple substances, and the small number of participants for each active substance, medication type (i.e. benzodiazepine, antidepressant, or anxiolytic) or consumption type (i.e. single or multiple). Consequently, the psychotropic drug users in our sample may have consumed substances with opposite effects on test performance, thus hampering the detection of a general psychotropic drug effect. For the CANTAB measures whose norms were stratified according to this variable (i.e. RVP A', PAL mean errors to success, and RTI simple movement time), we note that a comparison of the score of an individual who consumed a specific psychotropic drug (e.g. lorazepam) with the normative group of psychotropic drug users (who ingested this and/or other medications) is not as rigorous as a comparison with a normative group of persons who consumed only that specific drug.

Because our sample was recruited in retirement homes located only in the Lisbon district, studies that consider additional geographic areas are needed. Studies that include Portuguese persons living in the community and younger age ranges are also needed to examine the lifelong age effects on cognitive test performance. Studies examining the effects of groups of medications and consumption types (e.g. single, double, triple) on test performance are still needed.

This study makes an original contribution to the literature by providing Portuguese norms for RVP, PAL, SWM, and RTI for older adults without neuropsychiatric diagnoses and residing in retirement homes. We hope to thereby improve clinical and neuropsychological practice in institutions for older persons, enabling the use of automated tests that are properly validated for this population. The future of neuropsychological assessment will likely employ computerized tests due to their advantages and benefits. Institutions for older persons have much to gain from

these tests given the high number of regular neuropsychological assessments they must perform. Examples of these gains are material savings and better conservation, ease of data storage and battery transport for the assessment of bedridden persons, and suitability for several aphasia. Through this work, we also hope to promote greater use of computerized tests and further studies with older persons residing in retirement homes, contributing to increased knowledge on the influence of socio-demographic variables on CANTAB test performance.

ACKNOWLEDGEMENTS

The authors would like to thank all of the Principals of the following institutions where this study was conducted for their permission to perform the study: José de Mello Residências e Serviços [José de Mello Residences and Services], Centro de Apoio Social de Oeiras do Instituto de Acção Social das Forças Armadas [Oeiras Social Support Center of the Armed Forces' Social Welfare Institute], Casa dos Professores de Carcavelos da Associação de Solidariedade Social dos Professores [Carcavelos House of Teachers of Teachers' Social Solidarity Association], Residência Assistida da Delegação da Costa do Estoril da Cruz Vermelha Portuguesa [Costa do Estoril Residential Home of the Portuguese Red Cross], and Inválidos do Comércio Residências e Lares [Invalids of Commerce Residences and Nursing Homes]. They would also like to thank all of the professionals at these institutions for their invaluable assistance in recruiting older volunteers and responding to the interviews for formal caregivers. Finally, the authors would like to express particular gratitude to all of the older volunteers who kindly participated and without whom this research could not have been conducted.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

The first author was supported by the Fundação para a Ciência e a Tecnologia [Portuguese Foundation for Science and Technology; grant number SFRH/BD/75281/2010].

ORCID

Marta Matos Gonçalves, ID: <http://orcid.org/0000-0001-7976-4465>

Maria Salomé Pinho, ID: <http://orcid.org/0000-0001-9786-2819>

Mário Rodrigues Simões, ID: <http://orcid.org/0000-0002-1311-1338>

REFERENCES

- Alexander, G. E., Ryan, L., Bowers, D., Foster, T. C., Bizon, J. L., Geldmacher, D. S., & Glisky, E. L. (2012). Characterizing cognitive aging in humans with links to animal models. *Frontiers in Aging Neuroscience*, *4*, 1-18. doi: 10.3389/fnagi.2012.00021
- Alves, L., Simões, M. R., & Martins, C. (2012). The estimation of premorbid intelligence levels among Portuguese speakers: The Irregular Word Reading Test (TeLPI). *Archives of Clinical Neuropsychology*, *27*, 58-68. doi: 10.1093/arclin/acr103
- Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de Depressão Geriátrica (GDS) [Geriatric Depression Scale (GDS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (2nd ed., pp. 69-72). Lisbon, Portugal: Novartis.
- Braw, Y., Aviram, S., Bloch, Y., & Levkovitz, Y. (2011). The effect of age on frontal lobe related cognitive functions of unmedicated depressed patients. *Journal of Affective Disorders*, *129*, 342-347. doi: 10.1016/j.jad.2010.07.032
- Cambridge Cognition. (2006a). *CANTABeclipse™ version 3: Software user guide*. Cambridge, UK: Author.
- Cambridge Cognition. (2006b). *CANTABeclipse™ version 3: Test administration guide*. Cambridge, UK: Author.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, *55*, 1630-1635. doi: 10.1111/j.1532-5415.2007.01379.x
- Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained

- attention and working memory. *Neuropsychologia*, 34, 1085-1095. doi: 10.1016/0028-3932(96)00029-2
- Cummings, J., Gould, H., & Zhong, K. (2012). Advances in designs for Alzheimer's disease clinical trials. *American Journal of Neurodegenerative Disease*, 1, 205-216.
- De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the CANTAB. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 25, 242-254. doi: 10.1076/jcen.25.2.242.13639
- Elderkin-Thompson, V., Kumar, A., Bilker, W., Dunkin, J., Mintz, J., Moberg, P., ... Gur, R. (2003). Neuropsychological deficits among patients with late-onset minor and major depression. *Archives of Clinical Neuropsychology*, 18, 529-549. doi: 10.1016/s0887-6177(03)00022-2
- Espírito-Santo, H., Lemos, L., Ventura, L., Moitinho, S., Pinto, A. L., Rodrigues, F., ... Daniel, F. (2015). Teste da Figura Complexa de Rey-Osterrieth-A [Rey-Osterrieth A-Complex Figure Test]. In M. R. Simões & I. Santana (Eds.), *Escalas e testes na demência [Scales and tests in dementia]* (3rd ed., pp. 100-107). Porto Salvo, Portugal: Novartis.
- Firmino, H., Simões, M. R., Pinho, M. S., Cerejeira, L., & Martins, C. (2008). *O Exame Cognitivo de Addenbrooke - Revisto [The Addenbrooke's Cognitive Examination - Revised]*. Coimbra, Portugal: Serviço de Avaliação Psicológica da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State. *Journal of Psychiatric Research*, 12, 189-198. doi: 10.1016/0022-3956(75)90026-6
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. *Applied Neuropsychology*, 2, 72-78. doi: 10.1207/s15324826an0202_4
- Ganguli, M., Snitz, B. E., Lee, C.-W., Vanderbilt, J., Saxton, J. A., & Chang, C.-C. H. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: The Monongahela–Youghiogheny Healthy Aging Team (MYHAT). *Aging & Mental Health*, 14, 100-107. doi: 10.1080/13607860903071014
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2012). *Portuguese norms for the Naming Objects and Fingers Task of the Alzheimer's Disease Assessment Scale performed by older persons living in retirement homes*. Unpublished manuscript.
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016). Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, 23, 251-263. doi: 10.1080/23279095.2015.1053889
- Guerreiro, M., Fonseca, S., Barreto, J., & Garcia, C. (2008). Escala de Avaliação da Doença de Alzheimer (ADAS) [Alzheimer's Disease Assessment Scale (ADAS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência [Scales and tests in dementia]* (2nd ed., pp. 41-68). Lisbon, Portugal: Novartis.
- Herrera-Guzmán, I., Gudayol-Ferré, E., Herrera-Guzmán, D., Guàrdia-Olmos, J., Hinojosa-Calvo, E., & Herrera-Abarca, J. E. (2009). Effects of selective serotonin reuptake and dual serotonergic–noradrenergic reuptake treatments on memory and

- mental processing speed in patients with major depressive disorder. *Journal of Psychiatric Research*, 43, 855-863. doi: 10.1016/j.jpsychires.2008.10.015
- Herrera-Guzmán, I., Herrera-Abarca, J. E., Gudayol-Ferré, E., Herrera-Guzmán, D., Gómez-Carbajal, L., Peña-Olvira, M., ... Joan, G.-O. (2010). Effects of selective serotonin reuptake and dual serotonergic–noradrenergic reuptake treatments on attention and executive functions in patients with major depressive disorder. *Psychiatry Research*, 177, 323-329. doi: 10.1016/j.psychres.2010.03.006
- Hogervorst, E., De Jager, C., Budge, M., & Smith, A. D. (2004). Serum levels of estradiol and testosterone and performance in different cognitive domains in healthy elderly men and women. *Psychoneuroendocrinology*, 29, 405-421.
- Iverson, G. L., Brooks, B. L., Ashton, V. L., Johnson, L. G., & Gualtieri, C. T. (2009). Does familiarity with computers affect computerized neuropsychological test performance? *Journal of Clinical and Experimental Neuropsychology*, 31, 594-604. doi: 10.1080/13803390802372125
- Jones, G. M. M., Sahakian, B. J., Levy, R., Warburton, D. M., & Gray, J. A. (1992). Effects of acute subcutaneous nicotine on attention, information processing and short-term memory in Alzheimer's disease. *Psychopharmacology*, 108, 485-494. doi: 10.1007/bf02247426
- Junkkila, J., Oja, S., Laine, M., & Karrasch, M. (2012). Applicability of the CANTAB-PAL computerized memory test in identifying amnesic mild cognitive impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 34, 83-89. doi: 10.1159/000342116
- Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology*, 22, 595-624. doi: 10.1207/s15326942dn2203_3
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, 125, 624-639. doi: 10.1093/brain/awf049
- Manias, E., Kusljic, S., Berry, C., Brown, E., Bryce, E., Cliffe, J., & Smykowsky, A. (2015). Use of the Screening Tool of Older Person's Prescriptions (STOPP) in older people admitted to an Australian hospital. *Australasian Journal on Ageing*, 34, 15-20. doi: 10.1111/ajag.12054
- Marôco, J. (2014). *Análise estatística com o SPSS statistics* [Statistical analysis with the SPSS statistics] (6th ed.). Pêro Pinheiro, Portugal: Report Number - Análise e Gestão de Informação, Lda.
- Maust, D. T., Oslin, D. W., & Marcus, S. C. (2014). Effect of age on the profile of psychotropic users: Results from the 2010 National Ambulatory Medical Care Survey. *Journal of the American Geriatrics Society*, 62, 358-364. doi: 10.1111/jgs.12640
- Maylor, E. A., Simpson, E. E., Secker, D. L., Meunier, N., Andriollo-Sanchez, M., Polito, A., ... Coudray, C. (2006). Effects of zinc supplementation on cognitive function in healthy middle-aged and older adults: The ZENITH study. *British Journal of Nutrition*, 96, 752-760. doi: 10.1079/BJN20061911
- McGaughy, J., Dalley, J. W., Morrison, C. H., Everitt, B. J., & Robbins, T. W. (2002). Selective behavioral and neurochemical effects of cholinergic lesions produced by intrabasal infusions of 192 IgG-saporin on attentional performance in a five-choice serial reaction time task. *The Journal of Neuroscience*, 22, 1905-1913.

- Meinz, E. J., & Salthouse, T. A. (1998). Is age kinder to females than to males? *Psychonomic Bulletin & Review*, 5, 56-70. doi: 10.3758/bf03209457
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21, 1078-1085. doi: 10.1002/gps.1610
- Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin*, 19, 448-450.
- Morgado, J., Rocha, C. S., Maruta, C., Guerreiro, M., & Martins, I. P. (2009). Novos valores normativos do Mini-Mental State Examination [New normative values of Mini-Mental State Examination]. *Sinapse*, 2, 10-16.
- Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1987). Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer-type and Parkinson's disease. In S. M. Stahl, S. D. Iversen & E. C. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21-36). Oxford, UK: Oxford University Press.
- Nelson, H. E., & Willison, J. (1991). *The National Adult Reading Test (NART): Test manual* (2nd ed.). Windsor, UK: NFER-Nelson Publishing Company.
- Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, 33, 1-24. doi: 10.1016/0028-3932(94)00098-a
- Periquito, C. M. N., Silva, P. C., Oliveira, P., Carneiro, C., Fernandes, A. I., & Costa, F. A. (2014). Revisão da medicação em idosos institucionalizados: Aplicação dos critérios STOPP e START. [Review of medication in institutionalized elderly: Application of STOPP and START criteria]. *Revista Portuguesa de Farmacoterapia*, 6, 211-220.
- Plati, M. C. F., Covre, P., Lukasova, K., & Macedo, E. C. (2006). Depressive symptoms and cognitive performance of the elderly: Relationship between institutionalization and activity programs. *Revista Brasileira de Psiquiatria*, 28, 118-121. doi: 10.1590/s1516-44462006000200008
- Rabbitt, P., & Lowe, C. (2000). Patterns of cognitive ageing. *Psychological Research*, 63, 308-316. doi: 10.1007/s004269900009
- Rapp, M. A., Schnaider-Beri, M., Sano, M., Silverman, J. M., & Haroutunian, V. (2005). Cross-domain variability of cognitive performance in very old nursing home residents and community dwellers: Relationship to functional status. *Gerontology*, 51, 206-212. doi: 10.1159/000083995
- Reeves, S. J., Grasby, P. M., Howard, R. J., Bantick, R. A., Asselin, M.-C., & Mehta, M. A. (2005). A positron emission tomography (PET) investigation of the role of striatal dopamine (D2) receptor availability in spatial cognition. *Neuroimage*, 28, 216-226. doi: 10.1016/j.neuroimage.2005.05.034
- Rey, A. (2002). *Teste de Cópia de Figuras Complexas* [Complex Figure Copying Test] (2nd ed.). Lisbon, Portugal: Cegoc-Tea.
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging.

- Journal of the International Neuropsychological Society*, 4, 474-490. doi: 10.1017/s1355617798455073
- Robbins, T. W., Semple, J., Kumar, R., Truman, M. I., Shorter, J., Ferraro, A., ... Matthews, K. (1997). Effects of scopolamine on delayed-matching-to-sample and paired associates tests of visual memory and learning in human subjects: comparison with diazepam and implications for dementia. *Psychopharmacology*, 134, 95-106. doi: 10.1007/s002130050430
- Roiser, J. P., McLean, A., Ogilvie, A. D., Blackwell, A. D., Bamber, D. J., Goodyer, I., ... Sahakian, B. J. (2005). The subjective and cognitive effects of acute phenylalanine and tyrosine depletion in patients recovered from depression. *Neuropsychopharmacology*, 30, 775-785. doi: 10.1038/sj.npp.1300659
- Roque, D. T., Teixeira, R. A., Zachi, E. C., & Ventura, D. F. (2011). The use of the Cambridge Neuropsychological Test Automated Battery (CANTAB) in neuropsychological assessment: Application in Brazilian research with control children and adults with neurological disorders. *Psychology and Neuroscience*, 4, 255-265. doi: 10.3922/j.psns.2011.2.011
- Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, 85, 399-402. doi: 10.1177/014107689208500711
- Sahgal, A., Lloyd, S., Wray, C. J., Galloway, P. H., Robbins, T. W., Sahakian, B. J., ... Edwardson, J. A. (1992). Does visuospatial memory in senile dementia of the Alzheimer type depend on the severity of the disorder? *International Journal of Geriatric Psychiatry*, 7, 427-436. doi: 10.1002/gps.930070608
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, 13, 140-144. doi: 10.1111/j.0963-7214.2004.00293.x
- Saunders, N. L. J., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 32, 350-357. doi: 10.1080/13803390903042379
- Simpson, E. E. A., Hodkinson, C. F., Maylor, E. A., McCormack, J. M., Rae, G., Strain, S., ... Wallace, J. M. W. (2013). Intracellular cytokine production and cognition in healthy older adults. *Psychoneuroendocrinology*, 38, 2196-2208. doi: 10.1016/j.psyneuen.2013.04.007
- Simpson, E. E. A., Maylor, E. A., Rae, G., Meunier, N., Andriollo-Sanchez, M., Catasta, G., ... Coudray, C. (2005). Cognitive function in healthy older European adults: The ZENITH study. *European Journal of Clinical Nutrition*, 59(Suppl. 2), S26-S30. doi: 10.1038/sj.ejcn.1602294
- Skolimowska, J., Wesierska, M., Lewandowska, M., Szymaszek, A., & Szlag, E. (2011). Divergent effects of age on performance in spatial associative learning and real idiothetic memory in humans. *Behavioural Brain Research*, 218, 87-93. doi: 10.1016/j.bbr.2010.11.035
- Sousa, L. B., Prieto, G., Vilar, M., Firmino, H., & Simões, M. R. (2015). The Adults and Older Adults Functional Assessment Inventory: A Rasch model analysis. *Research on Aging*, 37, 787-814. doi: 10.1177/0164027514564469
- Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 20, 112-117. doi: 10.1097/01.wad.0000213815.20177.19
- Steves, C. J., Jackson, S. H. D., & Spector, T. D. (2013). Cognitive change in older women using a computerised battery: A longitudinal quantitative genetic twin study. *Behavior Genetics*, 43, 468-479. doi: 10.1007/s10519-013-9612-z

- Summers, M. J., & Saunders, N. J. (2012). Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. *Neuropsychology*, 26, 498-508. doi: 10.1037/a0028576
- Swainson, R., Hodges, J. R., Galton, C. J., Semple, J., Michael, A., Dunn, B. D., ... Sahakian, B. J. (2001). Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks. *Dementia and Geriatric Cognitive Disorders*, 12, 265-280. doi: 10.1159/000051269
- Ubeda, A., Ferrándiz, M. L., Maicas, N., Gomez, C., Bonte, M., & Peris, J. E. (2012). Potentially inappropriate prescribing in institutionalised older patients in Spain: The STOPP-START criteria compared with the Beers criteria. *Pharmacy Practice (Internet)*, 10, 83-91.
- Wechsler, D. (2008a). *Escala de Inteligência de Wechsler para Adultos – Terceira edição (WAIS-III): Manual técnico* [Wechsler Adult Intelligence Scale - Third edition (WAIS-III): Technical manual]. Lisbon, Portugal: Cegoc-Tea.
- Wechsler, D. (2008b). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual de administração e cotação* [Wechsler Memory Scale - Third edition (WMS-III): Administration and scoring manual]. Lisbon, Portugal: Cegoc-Tea.
- Wechsler, D. (2008c). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual técnico* [Wechsler Memory Scale - Third edition (WMS-III): Technical manual]. Lisbon, Portugal: Cegoc-Tea.
- Yang, L., Unverzagt, F. W., Jin, Y., Hendrie, H. C., Liang, C., Hall, K. S., ... Gao, S. (2012). Normative data for neuropsychological tests in a rural elderly Chinese cohort. *The Clinical Neuropsychologist*, 26, 641-653. doi: 10.1080/13854046.2012.666266
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37-49. doi: 10.1016/0022-3956(82)90033-4

APPENDICES

Appendix A. Inclusion criteria for the normative sample.

Normal premorbid intelligence quotient

A neuropsychological profile suggestive of absence of cognitive deficits. Participants must not exceed the cut-off score on more than two cognitive tests and must strictly adhere to the following conditions:

- (1) The exceeded cut-off score must not be the ACE-R total score;
- (2) The two exceeded cut-off scores must be scores for tests measuring different cognitive functions;
- (3) If more than one test assesses a cognitive function, the exceeded cut-off score must be the score for the less comprehensive measure of the cognitive function assessed (this condition only applies if the previous condition is met), namely,
 - The score for the ACE-R semantic verbal fluency item instead of the score for the ACE-R verbal fluency subtotal (which also includes a phonemic verbal fluency item);
 - The score for the Naming Objects and Fingers task (Alzheimer's Disease Assessment Scale) instead of the score for ACE-R language subtotal (which covers a similar number of naming items and other language domains);
 - The score for the ACE-R visuospatial subtotal instead of the score for the copy task of the Rey Complex Figure Test (which is a more demanding measure of constructive praxis).

Cognitive functional capacity for activities of daily living

No history of neurological, psychiatric or other medical condition with serious impact on cognition^a

Corrected-to-normal or no sensory deficits

No motor problems in the dominant upper limb

Absence of severe depressive symptoms

Note. ^aParticipants with clinical conditions such as hypertension, medically controlled endocrine, respiratory, or heart diseases, or osteoarticular problems not affecting the dominant upper limb were not excluded from the normative sample. Legend: ACE-R = Addenbrooke's Cognitive Examination Revised.

Appendix B. Psychotropic drugs and consumption types among the psychotropic drug users ($N = 61$).

Medication type	Psychotropic drugs	Psychotropic drug users	
		<i>n</i>	%
<i>Benzodiazepines</i>		51	84
	Alprazolam	11	18
	Bromazepam	14	23
	Brotizolam	2	3
	Clorazepate dipotassium	1	2
	Diazepam	2	3
	Estazolam	4	7
	Ethyl loflazepate	1	2
	Flurazepam	1	2
	Halazepam	1	2
	Loprazolam	1	2
	Lorazepam	11	18
	Mexazolam	6	10
	Midazolam	1	2
	Oxazepam	1	2
	Zolpidem	2	3
<i>Antidepressants</i>		28	47
	Clomipramine	1	2
	Dosulepin	1	2
	Escitalopram	7	11
	Fluoxetine	1	2
	Mirtazapine	4	7
	Nortriptyline	1	2
	Paroxetine	4	7
	Sertraline	10	16
	Trazodone	3	5
	Venlafaxine	1	2
<i>Anxiolytics</i>		4	7
	Pregabalin	1	2
	Valerian	3	5
Consumption type and number of drugs			
	One benzodiazepine	27	44
	Two benzodiazepines	5	8
	One benzodiazepine and one antidepressant	12	20
	Two benzodiazepines and one antidepressant	3	5
	One benzodiazepine and two antidepressants	3	5
	One benzodiazepine and one anxiolytic	1	2
	One antidepressant	5	8
	Two antidepressants	2	3
	One antidepressant and one anxiolytic	3	5

Appendix C. Standardized scores for the RVP A' (sensitivity index) measure.

SS	Raw score		SS	Raw score		SS	Raw score	
	Age 69-79, PD-NU (n = 26)	Age 69-79, PD-U (n = 21)		Age 69-79, PD-NU (n = 26)	Age 69-79, PD-U (n = 21)		Age 69-79, PD-NU (n = 26)	Age 69-79, PD-U (n = 21)
117		.999-1.000	92	.959-.960	.919-.921	67	.911-.912	.838-.841
116		.996-.998	91	.957-.958	.915-.918	66	.909-.910	.835-.837
115		.992-.995	90	.955-.956	.912-.914	65	.907-.908	.832-.834
114	1.000	.989-.991	89	.953-.954	.909-.911	64	.906	.829-.831
113	.998-.999	.986-.988	88	.951-.952	.906-.908	63	.904-.905	.826-.828
112	.997	.983-.985	87	.949-.950	.903-.905	62	.902-.903	.822-.825
111	.995-.996	.980-.982	86	.947-.948	.899-.902	61	.900-.901	.819-.821
110	.993-.994	.976-.979	85	.945-.946	.896-.898	60	.898-.899	.816-.818
109	.991-.992	.973-.975	84	.943-.944	.893-.895	59	.896-.897	.813-.815
108	.989-.990	.970-.972	83	.942	.890-.892	58	.894-.895	.810-.812
107	.987-.988	.967-.969	82	.940-.941	.887-.889	57	.892-.893	.806-.809
106	.985-.986	.964-.966	81	.938-.939	.883-.886	56	.890-.891	.803-.805
105	.983-.974	.960-.963	80	.936-.937	.880-.882	55	.889	.800-.802
104	.981-.982	.957-.959	79	.934-.935	.877-.879	<55	<.889	<.800
103	.979-.980	.954-.956	78	.932-.933	.874-.876			
102	.978	.951-.953	77	.930-.931	.871-.873			
101	.976-.977	.948-.950	76	.928-.929	.867-.870			
100	.974-.975	.944-.947	75	.926-.927	.864-.866			
99	.972-.973	.941-.943	74	.925	.861-.863			
98	.970-.971	.938-.940	73	.923-.924	.858-.860			
97	.968-.969	.935-.937	72	.921-.922	.855-.857			
96	.966-.967	.931-.934	71	.919-.920	.851-.854			
95	.964-.965	.928-.930	70	.917-.918	.848-.850			
94	.962-.963	.925-.927	69	.915-.916	.845-.847			
93	.961	.922-.924	68	.913-.914	.842-.844			

(Appendix C continues)

Appendix C (continued). Standardized scores for the RVP A' (sensitivity index) measure.

SS	Raw score		SS	Raw score		SS	Raw score	
	Age 80+, PD-NU (n = 37)	Age 80+, PD-U (n = 37)		Age 80+, PD-NU (n = 37)	Age 80+, PD-U (n = 37)		Age 80+, PD-NU (n = 37)	Age 80+, PD-U (n = 37)
117			92	.928-.930	.878-.882	67	.845-.847	.752-.757
116		.999-1.000	91	.925-.927	.873-.877	66	.842-.844	.747-.751
115		.994-.998	90	.921-.924	.868-.872	65	.838-.841	.742-.746
114		.989-.993	89	.918-.920	.863-.867	64	.835-.837	.737-.741
113	.998-1.000	.984-.988	88	.915-.917	.858-.862	63	.832-.834	.732-.736
112	.994-.997	.979-.983	87	.911-.914	.853-.857	62	.828-.831	.727-.731
111	.991-.993	.974-.978	86	.908-.910	.848-.852	61	.825-.827	.722-.726
110	.988-.990	.969-.973	85	.905-.907	.843-.847	60	.822-.824	.717-.721
109	.984-.987	.964-.968	84	.901-.904	.838-.842	59	.818-.821	.712-.716
108	.981-.983	.959-.963	83	.898-.900	.833-.837	58	.815-.817	.707-.711
107	.978-.980	.954-.958	82	.895-.897	.828-.832	57	.812-.814	.702-.706
106	.974-.977	.949-.953	81	.891-.894	.823-.827	56	.808-.811	.697-.701
105	.971-.973	.944-.948	80	.888-.890	.818-.822	55	.805-.807	.692-.696
104	.968-.970	.939-.943	79	.885-.887	.813-.817	<55	<.805	<.692
103	.964-.967	.934-.938	78	.881-.884	.808-.812			
102	.961-.963	.929-.933	77	.878-.880	.803-.807			
101	.958-.960	.924-.928	76	.875-.877	.798-.802			
100	.954-.957	.919-.923	75	.871-.874	.793-.797			
99	.951-.953	.914-.918	74	.868-.870	.788-.792			
98	.948-.950	.909-.913	73	.865-.867	.783-.787			
97	.945-.947	.904-.908	72	.861-.864	.778-.782			
96	.941-.944	.899-.903	71	.858-.860	.773-.777			
95	.938-.940	.894-.898	70	.855-.857	.768-.772			
94	.935-.937	.888-.893	69	.852-.854	.763-.767			
93	.931-.934	.883-.887	68	.848-.851	.758-.762			

Note. Age is in years. Legend: PD-NU = Psychotropic drug non-users; PD-U = Psychotropic drug users; RVP = Rapid visual information processing; SS = Standardized score.

Appendix D. Standardized scores for the RVP mean latency measure.

SS	Raw score		SS	Raw score		SS	Raw score		SS	Raw score	
	Age 69-79, Men (n = 20)	Age 69-79, Women (n = 27)		Age 69-79, Men (n = 20)	Age 69-79, Women (n = 27)		Age 69-79, Men (n = 20)	Age 69-79, Women (n = 27)		Age 69-79, Men (n = 20)	Age 69-79, Women (n = 27)
>145	<86	<88	122	242-248	261-268	98	405-411	442-449	74	568-574	623-630
145	86-92	88-94	121	249-255	269-275	97	412-418	450-456	73	575-580	631-637
144	93-99	95-102	120	256-262	276-283	96	419-424	457-464	72	581-587	638-645
143	100-106	103-109	119	263-268	284-290	95	425-431	465-471	71	588-594	646-652
142	107-112	110-117	118	269-275	291-298	94	432-438	472-479	70	595-601	653-660
141	113-119	118-124	117	276-282	299-305	93	439-445	480-486	69	602-607	661-668
140	120-126	125-132	116	283-289	306-313	92	446-451	487-494	68	608-614	669-675
139	127-133	133-139	115	290-295	314-321	91	452-458	495-502	67	615-621	676-683
138	134-140	140-147	114	296-302	322-328	90	459-465	503-509	66	622-628	684-690
137	141-146	148-155	113	303-309	329-336	89	466-472	510-517	65	629-635	691-698
136	147-153	156-162	112	310-316	337-343	88	473-479	518-524	64	636-641	699-705
135	154-160	163-170	111	317-323	344-351	87	480-485	525-532	63	642-648	706-713
134	161-167	171-177	110	324-329	352-358	86	486-492	533-539	62	649-655	714-720
133	168-173	178-185	109	330-336	359-366	85	493-499	540-547	61	656-662	721-728
132	174-180	186-192	108	337-343	367-373	84	500-506	548-554	60	663-668	729-735
131	181-187	193-200	107	344-350	374-381	83	507-512	555-562	59	669-675	736-743
130	188-194	201-207	106	351-357	382-388	82	513-519	563-569	58	676-682	744-751
129	195-201	208-215	105	358-363	389-396	81	520-526	570-577	57	683-689	752-758
128	202-207	216-222	104	364-370	397-404	80	527-533	578-585	56	690-696	759-766
127	208-214	223-230	103	371-377	405-411	79	534-540	586-592	55	697-702	767-773
126	215-221	231-238	102	378-384	412-419	78	541-546	593-600	<55	>702	>773
125	222-228	239-245	101	385-390	420-426	77	547-553	601-607			
124	229-234	246-253	100	391-397	427-434	76	554-560	608-615			
123	235-241	254-260	99	398-404	435-441	75	561-567	616-622			

(Appendix D continues)

Appendix D (continued). Standardized scores for the RVP mean latency measure.

SS	Raw score		SS	Raw score		SS	Raw score		SS	Raw score	
	Age 80+, Men (n = 30)	Age 80+, Women (n = 44)		Age 80+, Men (n = 30)	Age 80+, Women (n = 44)		Age 80+, Men (n = 30)	Age 80+, Women (n = 44)		Age 80+, Men (n = 30)	Age 80+, Women (n = 44)
>145	<108	<88	122	261-267	289-296	98	422-428	498-506	74	582-588	708-716
145	108-113	88-95	121	268-274	297-305	97	429-434	507-515	73	589-595	717-724
144	114-120	96-104	120	275-281	306-314	96	435-441	516-523	72	596-601	725-733
143	121-127	105-113	119	282-287	315-323	95	442-448	524-532	71	602-608	734-742
142	128-134	114-122	118	288-294	324-331	94	449-454	533-541	70	609-615	743-751
141	135-140	123-130	117	295-301	332-340	93	455-461	542-550	69	616-621	752-759
140	141-147	131-139	116	302-307	341-349	92	462-468	551-558	68	622-628	760-768
139	148-154	140-148	115	308-314	350-357	91	469-474	559-567	67	629-635	769-777
138	155-160	149-157	114	315-321	358-366	90	475-481	568-576	66	636-641	778-786
137	161-167	158-165	113	322-327	367-375	89	482-488	577-585	65	642-648	787-794
136	168-174	166-174	112	328-334	376-384	88	489-494	586-593	64	649-655	795-803
135	175-180	175-183	111	335-341	385-392	87	495-501	594-602	63	656-661	804-812
134	181-187	184-191	110	342-347	393-401	86	502-508	603-611	62	662-668	813-821
133	188-194	192-200	109	348-354	402-410	85	509-514	612-620	61	669-675	822-829
132	195-200	201-209	108	355-361	411-419	84	515-521	621-628	60	676-681	830-838
131	201-207	210-218	107	362-367	420-427	83	522-528	629-637	59	682-688	839-847
130	208-214	219-226	106	368-374	428-436	82	529-534	638-646	58	689-695	848-855
129	215-220	227-235	105	375-381	437-445	81	535-541	647-655	57	696-702	856-864
128	221-227	236-244	104	382-387	446-454	80	542-548	656-663	56	703-708	865-873
127	228-234	245-253	103	388-394	455-462	79	549-554	664-672	55	709-715	874-882
126	235-240	254-261	102	395-401	463-471	78	555-561	673-681	<55	>716	>882
125	241-247	262-270	101	402-407	472-480	77	562-568	682-689			
124	248-254	271-279	100	408-414	481-489	76	569-575	690-698			
123	255-260	280-288	99	415-421	490-497	75	576-581	699-707			

Note. Age is in years. Time is in milliseconds. Legend: RVP = Rapid visual information processing; SS = Standardized score.

Appendix E. Standardized scores for the PAL mean errors to success measure.

SS	Raw score		SS	Raw score		SS	Raw score	
	Men, PD-NU (n = 37)	Men, PD-U (n = 14)		Men, PD-NU (n = 37)	Men, PD-U (n = 14)		Men, PD-NU (n = 37)	Men, PD-U (n = 14)
140		0.00-0.11	111	3.41-3.60	4.26-4.39	82	9.18-9.37	8.54-8.68
139		0.12-0.26	110	3.61-3.80	4.40-4.54	81	9.38-9.57	8.69-8.82
138		0.27-0.41	109	3.81-4.00	4.55-4.69	80	9.58-9.76	8.83-8.97
137		0.42-0.55	108	4.01-4.20	4.70-4.84	79	9.77-9.96	8.98-9.12
136		0.56-0.70	107	4.21-4.39	4.85-4.98	78	9.97-10.16	9.13-9.27
135		0.71-0.85	106	4.40-4.59	4.99-5.13	77	10.17-10.36	9.28-9.41
134		0.86-1.00	105	4.60-4.79	5.14-5.28	76	10.37-10.56	9.42-9.56
133		1.01-1.14	104	4.80-4.99	5.29-5.43	75	10.57-10.76	9.57-9.71
132		1.15-1.29	103	5.00-5.19	5.44-5.57	74	10.77-10.96	9.72-9.86
131		1.30-1.44	102	5.20-5.39	5.58-5.72	73	10.97-11.16	9.87-10.00
130		1.45-1.59	101	5.40-5.59	5.73-5.87	72	11.17-11.36	10.01-10.15
129	0.00-0.02	1.60-1.74	100	5.60-5.79	5.88-6.02	71	11.37-11.55	10.16-10.30
128	0.03-0.22	1.75-1.88	99	5.80-5.99	6.03-6.17	70	11.56-11.75	10.31-10.45
127	0.23-0.42	1.89-2.03	98	6.00-6.18	6.18-6.31	69	11.76-11.95	10.46-10.60
126	0.43-0.62	2.04-2.18	97	6.19-6.38	6.32-6.46	68	11.96-12.15	10.61-10.74
125	0.63-0.82	2.19-2.33	96	6.39-6.58	6.47-6.61	67	12.16-12.35	10.75-10.89
124	0.83-1.01	2.34-2.47	95	6.59-6.78	6.62-6.76	66	12.36-12.55	10.90-11.04
123	1.02-1.21	2.48-2.62	94	6.79-6.98	6.77-6.90	65	12.56-12.75	11.05-11.19
122	1.22-1.41	2.63-2.77	93	6.99-7.18	6.91-7.05	64	12.76-12.95	11.20-11.33
121	1.42-1.61	2.78-2.92	92	7.19-7.38	7.06-7.20	63	12.96-13.14	11.34-11.48
120	1.62-1.81	2.93-3.06	91	7.39-7.58	7.21-7.35	62	13.15-13.34	11.49-11.63
119	1.82-2.01	3.07-3.21	90	7.59-7.78	7.36-7.49	61	13.35-13.54	11.64-11.78
118	2.02-2.21	3.22-3.36	89	7.79-7.97	7.50-7.64	60	13.55-13.74	11.79-11.92
117	2.22-2.41	3.37-3.51	88	7.98-8.17	7.65-7.79	59	13.75-13.94	11.93-12.07
116	2.42-2.60	3.52-3.66	87	8.18-8.37	7.80-7.94	58	13.95-14.14	12.08-12.22
115	2.61-2.80	3.67-3.80	86	8.38-8.57	7.95-8.09	57	14.15-14.34	12.23-12.37
114	2.81-3.00	3.81-3.95	85	8.58-8.77	8.10-8.23	56	14.35-14.54	12.38-12.52
113	3.01-3.20	3.96-4.10	84	8.78-8.97	8.24-8.38	55	14.55-14.75	12.53-12.66
112	3.21-3.40	4.11-4.25	83	8.98-9.17	8.39-8.53	<55	>14.75	>12.66

(Appendix E continues)

Appendix E (continued). Standardized scores for the PAL mean errors to success measure.

SS	Raw score		SS	Raw score		SS	Raw score	
	Women, PD-NU (n = 30)	Women, PD-U (n = 47)		Women, PD-NU (n = 30)	Women, PD-U (n = 47)		Women, PD-NU (n = 30)	Women, PD-U (n = 47)
140			111	2.48-2.61	3.74-3.91	82	6.73-6.87	8.91-9.08
139			110	2.62-2.76	3.92-4.09	81	6.88-7.01	9.09-9.26
138			109	2.77-2.91	4.10-4.27	80	7.02-7.16	9.27-9.43
137			108	2.92-3.05	4.28-4.44	79	7.17-7.31	9.44-9.61
136			107	3.06-3.20	4.45-4.62	78	7.32-7.45	9.62-9.79
135			106	3.21-3.35	4.63-4.80	77	7.46-7.60	9.80-9.97
134			105	3.36-3.49	4.81-4.98	76	7.61-7.75	9.98-10.15
133			104	3.50-3.64	4.99-5.16	75	7.76-7.89	10.16-10.33
132		0.00-0.17	103	3.65-3.79	5.17-5.33	74	7.90-8.04	10.34-10.50
131		0.18-0.34	102	3.80-3.93	5.34-5.51	73	8.05-8.19	10.51-10.68
130		0.35-0.52	101	3.94-4.08	5.52-5.69	72	8.20-8.33	10.69-10.86
129		0.53-0.70	100	4.09-4.23	5.70-5.87	71	8.34-8.48	10.87-11.04
128	0.00-0.12	0.71-0.88	99	4.24-4.37	5.88-6.05	70	8.49-8.63	11.05-11.22
127	0.13-0.26	0.89-1.06	98	4.38-4.52	6.06-6.23	69	8.64-8.77	11.23-11.40
126	0.27-0.41	1.07-1.24	97	4.53-4.67	6.24-6.40	68	8.78-8.92	11.41-11.57
125	0.42-0.56	1.25-1.41	96	4.68-4.81	6.41-6.58	67	8.93-9.07	11.58-11.75
124	0.57-0.70	1.42-1.59	95	4.82-4.96	6.59-6.76	66	9.08-9.21	11.76-11.93
123	0.71-0.85	1.60-1.77	94	4.97-5.11	6.77-6.94	65	9.22-9.36	11.94-12.11
122	0.86-1.00	1.78-1.95	93	5.12-5.25	6.95-7.12	64	9.37-9.51	12.12-12.29
121	1.01-1.14	1.96-2.13	92	5.26-5.40	7.13-7.30	63	9.52-9.66	12.30-12.46
120	1.15-1.29	2.14-2.30	91	5.41-5.55	7.31-7.47	62	9.67-9.80	12.47-12.64
119	1.30-1.44	2.31-2.48	90	5.56-5.69	7.48-7.65	61	9.81-9.95	12.65-12.82
118	1.45-1.58	2.49-2.66	89	5.70-5.84	7.66-7.83	60	9.96-10.10	12.83-13.00
117	1.59-1.73	2.67-2.84	88	5.85-5.99	7.84-8.01	59	10.11-10.24	13.01-13.18
116	1.74-1.88	2.85-3.02	87	6.00-6.13	8.02-8.19	58	10.25-10.39	13.19-13.36
115	1.89-2.02	3.03-3.20	86	6.14-6.28	8.20-8.37	57	10.40-10.54	13.37-13.53
114	2.03-2.17	3.21-3.37	85	6.29-6.43	8.38-8.54	56	10.55-10.68	13.54-13.71
113	2.18-2.32	3.38-3.55	84	6.44-6.57	8.55-8.72	55	10.69-10.83	13.72-13.89
112	2.33-2.47	3.56-3.73	83	6.58-6.72	8.73-8.90	<55	>10.83	>13.89

Note. Legend: PAL = Paired associates learning; PD-NU = Psychotropic drug non-users; PD-U = Psychotropic drug users; SS = Standardized score.

Appendix F. Standardized scores for the SWM total errors measure.

SS	Raw score		SS	Raw score		SS	Raw score		SS	Raw score	
	Age 69-79, Men (n = 21)	Age 69-79, Women (n = 29)		Age 69-79, Men (n = 21)	Age 69-79, Women (n = 29)		Age 69-79, Men (n = 21)	Age 69-79, Women (n = 29)		Age 69-79, Men (n = 21)	Age 69-79, Women (n = 29)
>145		<20	122	15	38	98	46-47	56	74	78-79	
145		20	121	16		97	48	57	73	80	75
144		21	120	17-18	39	96	49	58	72	81	76
143		22	119	19	40	95	50-51		71	82-83	77
142			118	20	41	94	52	59	70	84	78
141		23	117	21-22	42	93	53	60	69	85	
140		24	116	23		92	54-55	61	68	86-87	79
139		25	115	24	43	91	56		67	88	80
138			114	25-26	44	90	57	62	66	89	81
137		26	113	27	45	89	58-59	63	65	90-91	
136		27	112	28		88	60	64	64	92	82
135		28	111	29-30	46	87	61		63	93	83
134			100	31	47	86	62-63	65	62	94-95	84
133	0	29	109	32	48	85	64	66	61	96	
132	1-2	30	108	33		84	65	67	60	97	85
131	3	31	107	34-35	49	83	66-67	68	59	98-99	86
130	4	32	106	36	50	82	68		58	100	87
129	5-6		105	37	51	81	69	69	57	101	
128	7	33	104	38-39		80	70-71	70	56	102-103	88
127	8	34	103	40	52	79	72	71	55	104	89
126	9-10	35	102	41	53	78	73		<55	>104	>89
125	11		101	42-43	54	77	74-75	72			
124	12	36	100	44	55	76	76	73			
123	13-14	37	99	45		75	77	74			

(Appendix F continues)

Appendix F (continued). Standardized scores for the SWM total errors measure.

SS	Raw score		SS	Raw score		SS	Raw score		SS	Raw score	
	Age 80+, Men (n = 30)	Age 80+, Women (n = 48)		Age 80+, Men (n = 30)	Age 80+, Women (n = 48)		Age 80+, Men (n = 30)	Age 80+, Women (n = 48)		Age 80+, Men (n = 30)	Age 80+, Women (n = 48)
>145	<5	<14	122	31	38	98	57	62	74	83	87
145	5-6	14	121	32	39	97	58	63	73	84-85	88
144	7	15	120	33	40	96	59	64	72	86	89
143	8	16	119	34	41	95	60	65	71	87	90
142	9	17	118	35	42	94	61-62	66	70	88	91
141	10	18	117	36	43	93	63	67	69	89	92
140	11	19	116	37	44	92	64	68	68	90	93
139	12	20	115	38	45	91	65	69	67	91	94
138	13	21	114	39-40	46	90	66	70	66	92	95
137	14	22	113	41	47	89	67	71	65	93	96
136	15	23	112	42	48	88	68	72	64	94	97
135	16-17	24	111	43	49	87	69	73	63	95-96	98
134	18	25	100	44	50	86	70	74	62	97	99
133	19	26	109	45	51	85	71	75	61	98	100
132	20	27	108	46	52	84	72-73	76	60	99	101
131	21	28	107	47	53	83	74	77	59	100	102
130	22	29	106	48	54	82	75	78	58	101	103
129	23	30	105	49	55	81	76	79	57	102	104
128	24	31	104	50-51	56	80	77	80	56	103	105
127	25	32	103	52	57	79	78	81	55	104	106
126	26	33	102	53	58	78	79	82	<55	>104	>106
125	27-28	34	101	54	59	77	80	83-84			
124	29	35	100	55	60	76	81	85			
123	30	36-37	99	56	61	75	82	86			

Note. Age is in years. Legend: SS = Standardized score; SWM = Spatial working memory.

Appendix G. Standardized scores for the RTI simple movement time measure.

SS	Raw score		SS	Raw score		SS	Raw score		SS	Raw score	
	C-Users, PD-NU (n = 23)	C-Users, PD-U (n = 14)		C-Users, PD-NU (n = 23)	C-Users, PD-U (n = 14)		C-Users, PD-NU (n = 23)	C-Users, PD-U (n = 14)		C-Users, PD-NU (n = 23)	C-Users, PD-U (n = 14)
145		0-8	122	198-209	259-270	99	479-491	521-531	76	761-772	783-793
144		9-20	121	210-222	271-281	98	492-503	532-543	75	773-784	794-804
143		21-31	120	223-234	282-293	97	504-515	544-554	74	785-796	805-816
142		32-42	119	235-246	294-304	96	516-527	555-566	73	797-809	817-827
141		43-54	118	247-258	305-315	95	528-540	567-577	72	810-821	828-839
140		55-65	117	259-271	316-327	94	541-552	578-588	71	822-833	840-850
139	0-2	66-76	116	272-283	328-338	93	553-564	589-600	70	834-845	851-861
138	3-14	77-88	115	284-295	339-349	92	565-576	601-611	69	846-857	862-873
137	15-26	89-99	114	296-307	350-361	91	577-588	612-622	68	858-870	874-884
136	27-38	100-111	113	308-319	362-372	90	589-601	623-634	67	871-882	885-895
135	39-50	112-122	112	320-332	273-384	89	602-613	635-645	66	883-894	896-907
134	51-63	123-133	111	333-344	385-395	88	614-625	646-657	65	895-906	908-918
133	64-75	134-145	100	345-356	396-406	87	626-637	658-668	64	907-919	919-930
132	76-87	146-156	109	357-368	407-418	86	638-650	669-679	63	920-931	931-941
131	88-99	157-167	108	369-381	419-429	85	651-662	680-691	62	932-943	942-952
130	100-112	168-179	107	382-393	430-440	84	663-674	692-702	61	944-955	953-964
129	113-124	180-190	106	394-405	441-452	83	675-686	703-713	60	956-967	965-975
128	125-136	191-202	105	406-417	453-463	82	687-698	714-725	59	968-980	976-986
127	137-148	203-213	104	418-429	464-475	81	699-711	726-736	58	981-992	987-998
126	149-161	214-224	103	430-442	476-486	80	712-723	737-748	57	993-1004	999-1009
125	162-173	225-236	102	443-454	487-497	79	724-735	749-759	56	1005-1016	1010-1021
124	174-185	237-247	101	455-466	498-509	78	736-747	760-770	55	1017-1029	1022-1032
123	186-197	248-258	100	467-478	510-520	77	748-760	771-782	<55	>1029	>1032

(Appendix G continues)

Appendix G (continued). Standardized scores for the RTI simple movement time measure.

SS	Raw score		SS	Raw score		SS	Raw score		SS	Raw score	
	C-Naïve, PD-NU (n = 44)	C-Naïve, PD-U (n = 47)		C-Naïve, PD-NU (n = 44)	C-Naïve, PD-U (n = 47)		C-Naïve, PD-NU (n = 44)	C-Naïve, PD-U (n = 47)		C-Naïve, PD-NU (n = 44)	C-Naïve, PD-U (n = 47)
145			122	246-258	310-323	99	552-564	635-648	76	858-870	961-974
144		0-11	121	259-272	324-337	98	565-578	649-662	75	871-883	975-988
143		12-26	120	273-285	338-351	97	579-591	663-677	74	884-897	989-1002
142		27-40	119	286-298	352-365	96	592-604	678-691	73	898-910	1003-1016
141	0-6	41-54	118	299-312	366-379	95	605-617	692-705	72	911-923	1017-1030
140	7-19	55-68	117	313-325	380-393	94	618-631	706-719	71	924-937	1031-1044
139	20-32	69-82	116	326-338	394-408	93	632-644	720-733	70	938-950	1045-1059
138	33-46	83-96	115	339-351	409-422	92	645-657	734-747	69	951-963	1060-1073
137	47-59	97-110	114	352-365	423-436	91	658-671	748-761	68	964-977	1074-1087
136	60-72	111-125	113	366-378	437-450	90	672-684	762-776	67	978-990	1088-1101
135	73-85	126-139	112	379-391	451-464	89	685-697	777-790	66	991-1003	1102-1115
134	86-99	140-153	111	392-405	465-478	88	698-711	791-804	65	1004-1016	1116-1129
133	100-112	154-167	100	406-418	479-493	87	712-724	805-818	64	1017-1030	1130-1144
132	113-125	168-181	109	419-431	494-507	86	725-737	819-832	63	1031-1043	1145-1158
131	126-139	182-195	108	432-445	508-521	85	738-750	833-846	62	1044-1056	1159-1172
130	140-152	196-210	107	446-458	522-535	84	751-764	847-860	61	1057-1070	1173-1186
129	153-165	211-224	106	459-471	536-549	83	765-777	861-875	60	1071-1083	1187-1200
128	166-179	225-238	105	472-484	550-563	82	778-790	876-889	59	1084-1096	1201-1214
127	180-192	239-252	104	485-498	564-577	81	791-804	890-903	58	1097-1110	1215-1228
126	193-205	253-266	103	499-511	578-592	80	805-817	904-917	57	1111-1123	1229-1243
125	206-218	267-280	102	512-524	593-606	79	818-830	918-931	56	1124-1136	1244-1257
124	219-232	281-294	101	525-538	607-620	78	831-844	932-945	55	1137-1149	1258-1271
123	233-245	295-309	100	539-551	621-634	77	845-857	946-960	<55	>1149	>1271

Note. Time is in milliseconds. Legend: C-Naïve = Computer naïve; C-Users = Computer users; PD-NU = Psychotropic drug non-users; PD-U = Psychotropic drug users; RTI = Reaction time; SS = Standardized score.

Appendix H. Standardized Scores for the RTI five-choice movement time measure.

SS	Raw score	SS	Raw score	SS	Raw score
142	0-2	112	346-357	82	702-713
141	3-14	111	358-369	81	714-724
140	15-26	110	370-381	80	725-736
139	27-37	109	382-393	79	737-748
138	38-49	108	394-405	78	749-760
137	50-61	107	406-416	77	761-772
136	62-73	106	417-428	76	773-784
135	74-85	105	429-440	75	785-795
134	86-97	104	441-452	74	796-807
133	98-109	103	453-464	73	808-819
132	110-120	102	465-476	72	820-831
131	121-132	101	477-488	71	832-843
130	133-144	100	489-499	70	844-855
129	145-156	99	500-511	69	856-867
128	157-168	98	512-523	68	868-878
127	169-180	97	524-535	67	879-890
126	181-191	96	536-547	66	891-902
125	192-203	95	548-559	65	903-914
124	204-215	94	560-570	64	915-926
123	216-227	93	571-582	63	927-938
122	228-239	92	583-594	62	939-949
121	240-251	91	595-606	61	950-961
120	252-263	90	607-618	60	962-973
119	264-274	89	619-630	59	974-985
118	275-286	88	631-642	58	986-997
117	287-298	87	643-653	57	998-1009
116	299-310	86	654-665	56	1010-1020
115	311-322	85	666-677	55	1021-1032
114	323-334	84	678-689	<55	>1032
113	335-345	83	690-701		

Note. $N = 127$. Time is in milliseconds. Legend: RTI = Reaction time; SS = Standardized score.

Capítulo 3.

CONSTRUCT AND CONCURRENT VALIDITY OF THE CAMBRIDGE NEUROPSYCHOLOGICAL AUTOMATED TESTS FOR THE ASSESSMENT OF DEMENTIA IN PORTUGUESE OLDER ADULTS

Marta Matos Gonçalves, Maria Salomé Pinho, and Mário R. Simões³

ABSTRACT

Introduction: This study aimed to examine the factorial structure of the Cambridge Neuropsychological Test Automated Battery (CANTAB) for the assessment of dementia, which includes the Rapid Visual Information Processing (RVP), Paired Associates Learning (PAL), Reaction Time (RTI), and Spatial Working Memory (SWM) tests, in Portuguese older adults without neuropsychiatric diagnoses (WND) and with mild-to-moderate Alzheimer's disease (AD) dementia. It also compares these tests with pencil-and-paper tests assessing similar cognitive functions. **Method:** Inclusion criteria were checked in a first session. The CANTAB and additional pencil-and-paper tests were administered within one-week. The total sample [TS] consisted of 174 adults (WND group: $n = 137$; AD group: $n = 37$) aged 69-96 years. Comparisons were made between the CANTAB tests and between these tests and the Rey Complex Figure Test (RCFT), Verbal Fluency (VF) test, and some Wechsler Memory Scale-III and Wechsler Adult Intelligence Scale-III subtests through correlational and exploratory factor analyses. **Results:** The CANTAB tests loaded highly onto distinct factors of selective attention and processing speed (RVP), visuospatial episodic memory and learning (PAL), psychomotor speed (RTI), and visuospatial working memory and planning (SWM). Most intra-test correlations were stronger than the inter-test correlations. The RVP correlated more with VF animals (WND = $-.49$; TS = $-.67$); the PAL with RCFT immediate recall (WND = $-.52$; TS = $-.70$); the RTI with RVP mean latency (WND = $.42$; TS = $.56$); and the SWM with VF animals (WND = $-.47$; TS = $-.59$) and Spatial

³ Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016). *Construct and concurrent validity of the Cambridge neuropsychological automated tests for assessment of dementia in Portuguese older adults*. Manuscript submitted for publication.

Span backward (WND = $-.37$; TS = $-.53$). **Conclusions:** Our findings indicate that the CANTAB tests assess distinct cognitive functions but share other cognitive functions. The strongest correlations with the pencil-and-paper tests indicated good-to-adequate concurrent validity. This study provides further knowledge about the cognitive functions assessed by these tests and, thereby, promotes a substantiated use and interpretation of them.

Keywords: CANTAB; Alzheimer's disease; older persons; computerized tests; pencil-and-paper tests.

INTRODUCTION

Computerized neuropsychological assessments offer many advantages over pencil-and-paper tests, such as automatic administration, registration and scoring and the ease of transporting various materials and storing a large amount of data. Furthermore, we are witnessing very rapid technological adherence worldwide, even among older persons. In Portugal, the computer use of adults aged 65 to 74 years increased from 6% in 2008 to 29% in 2015, according to the National Institute of Statistics (Instituto Nacional de Estatística, 2008, 2015). Furthermore, Collerton et al. (2007) found that older persons report lower levels of stress on automated tests than on pencil-and-paper tests. Therefore, we expect that pencil-and-paper tests will be replaced by computerized tests in the near future. However, for such replacement to be achieved, researchers and clinicians must know which cognitive functions are assessed by the automated tests, or the extent to which these tests assess such cognitive functions (construct validity), and the degree of concordance between computerized tests and pencil-and-paper tests assessing similar cognitive functions (concurrent validity) (American Psychological Association, 2015).

The Cambridge Neuropsychological Test Automated Battery (CANTAB[®], www.cambridgecognition.com) was developed in 1987 (Morris, Evenden, Sahakian, & Robbins, 1987). Since then, new tests have been added to this battery (Cambridge Cognition, 2006a, 2006b). The CANTAB is used worldwide, partly because nearly all of the tests are non-verbal and contain abstract stimuli; thus the battery can be easily adapted to other countries. Several discriminating power and predictive validity studies

have suggested that four CANTAB tests are particularly suitable for assessing mild cognitive impairment (MCI) and Alzheimer's disease (AD) dementia: the Rapid Visual Information Processing (RVP), Paired Associates Learning (PAL), Reaction Time (RTI), and Spatial Working Memory (SWM) tests (e.g., Fowler, Saling, Conway, Semple, & Louis, 1995; Saunders & Summers, 2010; Summers & Saunders, 2012). Recently, these tests have been applied in studies of Portuguese older adults living in retirement homes. Several measures showed good to adequate test-retest reliability (i.e., .71-.89) in a study of 34 persons without neuropsychiatric problems, and norms were established based on a larger dataset including 128 individuals (Gonçalves, Pinho, & Simões, 2016a, 2016b). Nevertheless, validity studies with Portuguese older adults are still needed, particularly studies examining the four CANTAB tests' construct validity and concurrent validity, which can be determined through comparing them with pencil-and-paper tests commonly used in clinical practice.

Correlations between the CANTAB tests and pencil-and-paper tests

Previous studies comparing the CANTAB tests with pencil-and-paper tests generally found correlations below .69 (e.g., Fowler et al., 1995; Smith, Need, Cirulli, Chiba-Falek, & Attix, 2013), which denote weak to moderate associations (weak: .10-.39; moderate: .40-.69; classification of Dancey & Reidy, 2004). According to Elwood (2001), these weak associations may not necessarily indicate that the tests have poor psychometric properties. Rather, they may reflect important differences between the two test types, such as their reliance on different sensory modalities for stimuli presentation and responding. For example, the majority of the automated tests display visual stimuli and require motor responses, while most of the pencil-and-paper tests present auditory stimuli and request oral responses. Furthermore, the CANTAB tests involve tasks based on paradigms developed to test animal models of dementia and other disorders (Robbins et al., 1994); hence, their characteristics are distinct from those of most pencil-and-paper tests (e.g., less familiar tasks and stimuli; Cambridge Cognition, 2006b), which may also explain the weak associations found between the two types of tests.

Inter-test correlations of the Rapid Visual Information Processing test

The RVP test is a measure of visual sustained attention and continuous performance and requires participants to detect a target sequence of three digits (Cambridge Cognition,

2006b). This test also involves working memory (keeping the 3-digit sequence mentally online, possibly through sub-vocal rehearsal) and selective attention (distinguishing the target sequence from non-target stimuli) and is associated with bilateral activations in the frontal and parietal cortices and fusiform gyrus (Coull, Frith, Frackowiak, & Grasby, 1996).

Smith et al. (2013) reported weak correlations (.12-.26) between the RVP total errors measure and the Stroop test (Stroop, 1935) color-word score, Trail Making Test (TMT; Reitan, 1958) A and B scores, Wechsler Adult Intelligence Scale-III (WAIS-III) Digit Span (forward and backward scores) and Digit Symbol (total score) subtests (Wechsler, 2008a), and Semantic Verbal Fluency (Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006) animal names score. These measures assess selective attention and response inhibition (e.g., Stroop test), processing speed (e.g., TMT-A and Digit Symbol), working memory (e.g., Digit Span backward), and cognitive flexibility (e.g., TMT-B and Verbal Fluency animals).

Inter-test correlations of the Paired Associates Learning test

The PAL is a visuospatial associative memory and learning test that requires participants to encode, retain, and retrieve visuospatial associations (i.e., between a visual pattern and a spatial location) that were learned immediately before the assessment (Cambridge Cognition, 2006b). The task involves bilateral activations in the hippocampus and parahippocampal gyrus (middle temporal lobe) and additional activations in occipital, parietal and frontal areas (e.g., lingual gyrus, cuneus, precuneus, anterior cingulate gyrus, and frontal gyrus; de Rover et al., 2011). Interestingly, dysfunctions in the parahippocampal gyrus and precuneus have recently been identified in preclinical AD (Khan et al., 2014), which may explain why the PAL test highly predicts AD (Summers & Saunders, 2012).

Several studies (Fowler et al., 1995; Kim et al., 2009; Torgersen, Flaatten, Engelsen, & Gramstad, 2012) have observed moderate correlations (.47-.69) between the PAL test (total errors adjusted, total trials adjusted, and first trial memory score measures) and the Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987) Visual Memory, Verbal Memory, Delayed Recall, and Attention/Concentration indexes; the WMS-R Visual Memory Span subtest (backward score); and the Rey Complex Figure Test (RCFT; Rey, 2002) immediate and delayed recall scores. Most of these pencil-and-

paper tests assess episodic memory, but the Visual Memory Span subtest involves visuospatial attention and working memory.

Inter-test correlations of the Reaction Time test

The RTI test is a measure of visual sustained attention (to the appearance of a target stimulus on the screen, first at a predictable location and then at one of five possible locations), reaction time (i.e., time taken to release a button), and movement time (i.e., time to move the arm and touch the screen after releasing the button; Cambridge Cognition, 2006b). This test is sensitive to dysfunctions in the basal ganglia, genu of the corpus callosum, and medial prefrontal cortex (McGaughy, Dalley, Morrison, Everitt, & Robbins, 2002; Sexton et al., 2012).

Sexton et al. (2012) reported a high internal consistency coefficient (.80) for a group of processing speed tests, including the RTI, Digit Symbol (WAIS-III), and TMT-A.

Inter-test correlations of the Spatial Working Memory test

The SWM test is a self-ordered task that requires participants to monitor and manipulate spatial information in working memory (i.e., information regarding the boxes in which a token had previously been found and the boxes that were previously opened in the search for a token) but also involves a strategy component (Cambridge Cognition, 2006b). The successful completion of this test requires the integrity of the prefrontal cortex (especially the right inferior frontal gyrus and dorsolateral prefrontal cortex), lateral premotor cortex, visual areas, posterior parietal lobe, medial temporal lobe, and cingulate region (Chase, Clark, Sahakian, Bullmore, & Robbins, 2008; Manes et al., 2002; Oosterman et al., 2008; Owen, Evans, & Petrides, 1996).

Smith et al. (2013) reported weak correlations (.16-.28) between the SWM between search errors measure and the Digit Symbol (WAIS-III) total score, Stroop test color-word score, Verbal Fluency animals score, and Digit Span (WAIS-III) backward score and between the SWM strategy measure and the Verbal Fluency animals score. These pencil-and-paper tests assess executive functions and working memory.

Correlations between CANTAB tests

Studies analyzing the correlations between the CANTAB tests (Robbins et al., 1998; Singer, MacGregor, Cherkas, & Spector, 2006) have generally described stronger correlations between measures of the same test (e.g., .53-.72 for SWM; .67 for RTI) than between measures of different tests. Concerning inter-test correlations, these studies have found significant correlations between PAL and SWM (.11-.44) and between RTI and PAL (.16-.18) and SWM (.13-.22). To the best of our knowledge, no study has examined the RVP intra-test correlations or the correlations between this test and the PAL, RTI, and SWM tests.

Factor analytic studies

The first CANTAB studies using factor analysis reported solutions in which the PAL and SWM tests had high loadings onto separate factors, namely, a visual memory and learning factor (PAL) and a planning and executive function factor (SWM) (Robbins et al., 1998; Robbins et al., 1994). Subsequent studies have corroborated this factorial division (e.g., Rabbitt & Lowe, 2000; Stilley, Bender, Dunbar-Jacob, Sereika, & Ryan, 2010; Stoffers et al., 2007), with the PAL test loading highly onto factors that other memory and learning tests also loaded onto (e.g., CANTAB Delayed Matching to Sample, Pattern Recognition Memory, and Spatial Recognition Memory tests, and RCFT immediate and delayed recall) and the SWM test loading highly onto factors that other executive function, working memory, and planning tests loaded onto (e.g., CANTAB Stockings of Cambridge and Spatial Span tests, Semantic Verbal Fluency, and TMT).

A factor analytic study (Stilley et al., 2010) employing the RVP, PAL, SWM, and other cognitive tests reported that the RVP test loaded highly onto a factor of attention and psychomotor speed that included the Digit Symbol (WAIS-III), Digit Vigilance (Lewis & Rennick, 1979), and Grooved Pegboard (Kløve, 1963) tests. However, using principal component analysis, Smith et al. (2013) described a component that incorporated measures of RVP mean latency, PAL total errors (8-boxes problem), and SWM between search errors, suggesting that these tests may share some cognitive functions. Moreover, Lenehan, Summers, Saunders, Summers, and Vickers (2016) examined the RTI, RVP, PAL, SWM, and other CANTAB tests using a procedure that

was similar to exploratory factor analysis but forced a solution with a pre-specified number of factors. In this study, the RTI test showed a high factor loading onto a processing speed factor, the PAL onto an episodic memory factor, and the SWM onto a working memory factor, but the RVP did not load onto any of these factors. After adding pencil-and-paper tests to this set of CANTAB tests, the authors reported a solution in which the RTI and RVP tests most highly loaded (loadings below .40) onto a processing speed factor including the Digit Symbol (WAIS-III), TMT, and Stroop tests.

Any solution resulting from factor analysis depends largely on the set of variables introduced and the analyzed samples. Most validity studies of the CANTAB tests have been conducted with healthy samples (e.g., Reeves et al., 2005; Robbins et al., 1998; Smith et al., 2013); however, some studies have included mixed samples of healthy individuals and patients with AD, Parkinson's disease or other clinical conditions (e.g., diabetes and cancer) (Fowler et al., 1995; Stille et al., 2010; Stoffers et al., 2007). Studies involving clinical groups are very important because they allow researchers to test hypotheses regarding the consistency of the factorial structure of the battery and the cognitive functions assessed by the tests. Thus, such studies increase our understanding of which cognitive functions may be impaired in people with specific clinical conditions.

Therefore, our study aims to analyze the construct and concurrent validity of the RVP, PAL, RTI, and SWM tests by examining the correlations between these tests, the correlations between these tests and pencil-and-paper tests commonly used in clinical practice and the latent factorial structure of these CANTAB tests in a sample of Portuguese older persons without neuropsychiatric diagnoses (WND) and with mild to moderate AD dementia. We expect to find weak to moderate correlations between the CANTAB and pencil-and-paper tests. In particular, we expect the RVP test to correlate with selective attention and executive functioning measures, the PAL test to correlate primarily with episodic memory measures, the RTI test to correlate with attention and psychomotor speed measures, and the SWM test to correlate with working memory and executive function measures (e.g., Sexton et al., 2012; Smith et al., 2013; Torgersen et al., 2012). We also anticipate that we will find higher correlations in the total sample than in the WND and AD groups, due to the total sample's greater heterogeneity

(Anastasi, 1988), and that at least the PAL and SWM measures will load highly onto two distinct factors (Robbins et al., 1998; Stillel et al., 2010).

METHOD

Participants and procedure

A total of 271 older adults (including those who were enrolled in the Portuguese studies of the CANTAB test norms and test-retest reliability; Gonçalves et al., 2016a, 2016b) voluntarily participated in this research. The participants were recruited from retirement homes (both independent and dependent living facilities), adult day care centers, and mental health clinics in the district of Lisbon via oral and written advertisements. All of the participants gave written informed consent prior to participating in the study. For participants whose judgment capacity could be compromised, a formal or informal caregiver provided written informed consent.

A trained neuropsychologist (M.M.G.) conducted a first session with each participant to determine whether he or she met the inclusion criteria for the AD group (composed of participants with mild to moderate amnesic probable AD dementia; McKhann et al., 2011) or the WND group (see inclusion criteria, Table 1). This session included a semi-structured interview and the administration of pencil-and-paper tests (see the sequence of test administration, Table 2). In addition, the same neuropsychologist conducted a semi-structured interview with each participant's physician, and a close caregiver (e.g., physician, nurse, social worker, psychologist, or relative) completed the Adults and Older Adults Functional Assessment Inventory (Inventário de Avaliação Funcional de Adultos e Idosos [IAFAI]; Sousa, Prieto, Vilar, Firmino, & Simões, 2015). The neuropsychologist completed the Global Deterioration Scale (GDS; Leitão, Nina, & Monteiro, 2008; Reisberg, Ferris, de Leon, & Crook, 1982) based on all of the collected information.

Standardized scores more than 1.5 *SD* below the normative group mean ($M = 100$, $SD = 15$) were the cut-off points used for all of the cognitive tests administered for screening purposes except for the Irregular Word Reading Test (Teste de Leitura de Palavras Irregulares [TeLPI]; Alves, Simões, & Martins, 2012). We applied a less stringent cut-off score for the TeLPI because we intended to obtain a sample encompassing a large range of years of education and the TeLPI Intelligence Quotient (TeLPI-IQ) score is

strongly correlated with this variable. For the 30-item Geriatric Depression Scale (GDS-30; Barreto, Leuschner, Santos, & Sobral, 2008; Yesavage et al., 1982), we used the Portuguese severe depression cut-off score (more than 20 points) because of this condition's negative impact on cognitive test performance (Elderkin-Thompson et al., 2003).

Table 1. Inclusion criteria for the groups without neuropsychiatric diagnoses (WND) and with Alzheimer's disease (AD) dementia.

WND group	AD group
(1) No history of neurological, psychiatric or other medical condition with serious impact on cognition.	(1) Criteria for amnesic probable AD (McKhann et al., 2011). Insidious progressive cognitive impairment not explained by a medical condition other than AD. No history of neurological, psychiatric or medical condition other than AD.
(2) Normal global cognitive state.	(2) Global cognitive impairment. Mild to moderate stages of AD dementia.
(3) Normal premorbid intelligence quotient.	(3) Normal premorbid intelligence quotient.
(4) A neuropsychological profile suggestive of absence of cognitive deficits. Participants must not exceed the cut-off score on more than two cognitive tests and must strictly adhere to the following conditions: <ul style="list-style-type: none"> • The exceeded cut-off score must not be the ACE-R total score; • The two exceeded cut-off scores must be scores for tests assessing different cognitive functions. 	(4) Episodic memory deficit (verbal and/or visual) and deficits in one or more of the cognitive domains below: <ul style="list-style-type: none"> • Attention/working memory; • Constructive praxis; • Abstract thinking; • Mental flexibility; • Naming.
(5) Cognitive functional capacity for activities of daily living.	(5) Cognitive functional incapacity for activities of daily living.
(6) No severe depressive symptoms.	(6) No severe depressive symptoms.
(7) Corrected-to-normal or no sensory deficits.	(7) Corrected-to-normal or no sensory deficits.
(8) No motor problems in the dominant upper limb.	(8) No motor problems in the dominant upper limb.
(9) Three or more years of education and 65 or more years of age.	(9) Three or more years of education and 65 or more years of age.

Note. Clinical conditions such as medically controlled endocrine, respiratory, or heart diseases; hypertension; or osteoarticular problems not affecting the dominant upper limb were not exclusion criteria for any of the groups. Legend: ACE-R = Addenbrooke's Cognitive Examination Revised.

Table 2. Sequence of pencil-and-paper tests administered in the first session.

Pencil-and-paper test	Description	Cognitive domains (measures)	Cut-off score
Addenbrooke's Cognitive Examination Revised ^a	A cognitive screening test that includes the MMSE ^b and further items and cognitive domains (e.g., the semantic verbal fluency item, which is an executive function measure that requires participants to produce the maximum possible number of animal names aloud in 60 seconds).	Global cognitive state (MMSE ^b and total scores). Attention and orientation, memory, verbal fluency, language, visuospatial functions (subtotals). Mental flexibility (total animal names).	standardized score < 77.5
Digit Span (WAIS-III) ^c	This test involves repeating aloud digit sequences forward and backward.	Attention and working memory (total score).	
Word Lists I (WMS-III) ^d	A test that requires learning and recalling aloud as many words as possible from a 12-word list.	Verbal episodic memory (immediate recall).	
Rey Complex Figure Test ^e	This test requires copying an abstract figure and reproducing it (without seeing the model) after a 3-minute delay.	Constructive praxis (copy). Visual episodic memory (immediate recall).	
Irregular Word Reading Test (TeLPI) ^f	This test consists in reading aloud a 46-irregular word list. Total errors and years of education are taken into account in the TeLPI-IQ conversion formula.	Premorbid IQ (TeLPI-IQ score).	standardized score < 70
Naming Objects and Fingers (ADAS) ^g	This task involves naming 12 familiar objects and the hand's fingers.	Naming (total errors).	standardized score < 77.5
Similarities (WAIS-III) ^c	This test requires finding abstract concepts that relate two words.	Abstract thinking (total score).	
Word Lists II (WMS-III) ^d	A test that involves recalling and recognizing a 12-word list learned 25-35 minutes before.	Verbal episodic memory (delayed recall and recognition).	
Geriatric Depression Scale ^h	A 30-item checklist (yes/no).	Depressive symptoms (total score).	raw score > 20
Adults and Older Adults Functional Assessment Inventory (IAFAI) ⁱ	This inventory assesses physical, cognitive, and emotional functional capacity for basic and instrumental activities of daily living.	Cognitive functional capacity (global cognitive incapacity score).	raw score > P ₅ ^j
Global Deterioration Scale ^k	This scale describes seven stages of cognitive function (from absence of cognitive decline to severe AD dementia).	Dementia severity ^l (stage score).	raw scores > 2 (WND group); < 4 or > 5 (AD group)

Note. The Portuguese versions and norms of the tests listed above were used. ^a(Firmino, Simões, Pinho, Cerejeira, & Martins, 2008; Gonçalves, Pinho, & Simões, 2012; Mioshi et al., 2006); norms set by age and education. ^bMini-Mental State Examination (Folstein, Folstein, & McHugh, 1975; Morgado, Rocha, Maruta, Guerreiro, & Martins, 2009); norms set by education. ^cWechsler Adult Intelligence Scale-III

(Wechsler, 2008a); norms set by age. ^dWechsler Memory Scale-III (Wechsler, 2008b, 2008c); norms set by age. ^e(Espírito-Santo et al., 2015; Rey, 2002); norms set by education. ^fTeste de Leitura de Palavras Irregulares (Alves et al., 2012); norms set by education. ^gAlzheimer's Disease Assessment Scale (Gonçalves et al., 2012; Guerreiro, Fonseca, Barreto, & Garcia, 2008; Mohs, Rosen, & Davis, 1983); norms set by age and education. ^h(Barreto et al., 2008; Yesavage et al., 1982). ⁱInventário de Avaliação Funcional de Adultos e Idosos (Sousa et al., 2015); norms set by gender and education. ^jThe IAFAI norms are only expressed in percentiles; the raw scores corresponding to the fifth percentile (P_5) of each normative group are as follows: 1-4 years of education (men, $P_5 = 10$; women, $P_5 = 15$); 5-9 years of education (men, $P_5 = 6$; women, $P_5 = 10$); 10+ years of education (men, $P_5 = 5$; women, $P_5 = 9$). ^k(Leitão et al., 2008; Reisberg et al., 1982). ^lThe severe AD dementia cut-off score proposed by the National Institute for Health and Care Excellence (2011) (i.e., MMSE score < 10) was also used as an exclusion criterion. Legend: AD = Alzheimer's disease; IQ = intelligence quotient; WND = without neuropsychiatric diagnoses.

Psychotropic or anti-dementia drug intake was not an exclusion criterion for the AD group for ethical reasons. Specifically, we did not want to compromise the medical treatment. In regard to the participants who fulfilled the inclusion criteria for the WND group, we did not exclude those who reported psychotropic drug use consistent with the prescription of potentially inappropriate medication. The high incidence of such drug use among older people is problematic and psychotropic drugs are the most frequently improperly prescribed medications (Manias et al., 2015; Periquito et al., 2014; Ubeda et al., 2012). However, to avoid decreasing the representativeness of the sample and the ecological validity of data, we did not exclude these participants. After examining the participants who met the inclusion criteria for the WND or AD groups, we excluded 97 individuals (36% of the initial sample) for the following reasons: 46% due to neuropsychiatric problems other than AD (e.g., MCI, stroke, frontotemporal or vascular dementia, Parkinson's disease, traumatic brain injury, cerebral aneurysm, epilepsy, bipolar disease, or major depression), 6% due to sensory or motor problems, 3% due to illiteracy, 31% due to study abandonment, and 13% due to failure to meet the cognitive test conditions described in Table 1 (issues 3 and 4).

Therefore, our final sample consisted of 174 participants (WND group: $n = 137$; AD group: $n = 37$) who were right-handed (1 was ambidextrous) and spoke Portuguese as their first language. Participants were 65 to 96 years of age (see demographic characteristics, Table 3). The proportion of men was smaller than that of women, $\chi^2(1, N = 174) = 16.76, p < .001$.

Table 3. Demographic characteristics of the sample.

	WND group (<i>n</i> = 137)		AD group (<i>n</i> = 37)			Total sample (<i>N</i> = 174)			
	<i>f</i>	%	<i>f</i>	%		<i>f</i>	%		
Gender									
Men	52	38	8	22		60	34		
Women	85	62	29	78		114	66		
Living arrangements									
Community dwellers	9	7	4	11		13	7		
Retirement home dwellers	128	93	33	89		161	93		
Computer experience									
Computer naïve	100	73	36	97		136	78		
Computer users	37	27	1	3		38	22		
Psychotropic drug use									
Psychotropic drug non-users	71	52	4	11		75	43		
Psychotropic drug users	66	48	33	89		99	57		
	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range
Age (years)	80.78	6.57	65-96	83.73	5.82	72-94	81.41	6.52	65-96
Education (years)	8.82	4.89	4-22	6.41	3.27	3-16	8.30	4.69	3-22
Premorbid intelligence quotient (TeLPI)	105.97	12.96	77-132	94.98	12.83	70-120	103.64	13.66	70-132
Global cognitive state (ACE-R)									
MMSE score	28.69	1.29	25-30	19.11	4.56	10-27	26.66	4.60	10-30
Total score	87.64	7.66	63-100	51.27	13.29	23-73	79.90	17.48	23-100
Depressive symptoms (GDS-30)	8.32	5.21	0-20	11.51	5.58	1-20	9.00	5.44	0-20
Cognitive incapacity (IAFAI)									
Participants' test form	0.09	0.50	0-3	9.73	10.30	0-45	2.14	6.16	0-45
Caregivers' test form	0.02	0.27	0-3	35.95	16.45	8-64	8.33	17.09	0-64

Note. The Portuguese versions and norms of the tests listed above were used. Computer users were participants who used computers independently or who had attended computer classes for one or more months. Computer naïve were participants who had never used a computer or who were dependent on others to perform basic operations. Psychotropic drug users were those who reported daily intake of psychotropic drugs. Psychotropic drug non-users were those who were not consuming such drugs for at least one month before their first session in the study. Legend: ACE-R = Addenbrooke's Cognitive Examination Revised; AD = Alzheimer's disease; GDS-30 = Geriatric Depression Scale; IAFAI = Adults and Older Adults Functional Assessment Inventory; MMSE = Mini-Mental State Examination; TeLPI = Irregular Word Reading Test; WND = without neuropsychiatric diagnoses.

The same neuropsychologist held a final session with each participant at most one week after the first session to administer the CANTAB tests (Cambridge Cognition, 2006a, 2006b) and Wechsler Memory Scale-III (WMS-III; Wechsler, 2008b) subtests (see order of test administration, Table 4).

When needed, participants were given breaks between tests to minimize fatigue effects. For participants with a more severe clinical status, the sessions were partitioned into two halves but also completed within one week. The participants received brief reports of their performance on the pencil-and-paper tests (free of charge) as compensation for their collaboration. The participants with AD and their informal caregivers provided informed consent for us to give their reports to their physicians. This study was approved by the Portuguese Foundation for Science and Technology, by the Scientific Council of the Faculty of Psychology and Educational Sciences of the University of Coimbra, and by the Principals and Ethics Committees of the institutions where the participants were recruited.

Apparatus and materials

The CANTAB was administered with the CANTABeclipse™ version 3.0 software (Cambridge Cognition, 2006a), which was run in a Windows® XP environment. A Sahara Slate PC® i412T Tablet PC (Intel® Celeron® M ULV 423 Processor, 1.0 GB DDR2 RAM, 120 GB SATA Hard Disk Drive) with a 12.1-inch (307 mm) TFT XGA LCD passive resistive touch screen was employed (monitor resolution: 1024 × 768). For the RVP and RTI tests, we also used a press pad (Cambridge Cognition 2-Button Press Pad version 2.0). An integrated touchpad keyboard (KeySonic™ ACK-540U+ Mini-Keyboard) was handled only by the examiner. Cambridge Cognition tested the timing accuracy of Sahara Slate PC i412T with software equivalent to CANTABeclipse™ 3.0 using external chronometry. The timing accuracy was measured (in milliseconds) for video latency ($M = 7.40$, $SD = 0.64$), audio latency after calibration in the software ($M = 2.50$, $SD = 0.86$), and touch latency ($M = 63.65$, $SD = 5.00$).

Table 4. Sequence of tests administered in the second session.

Test	Description	Test measures
<i>CANTAB tests</i>		
Motor Screening (MOT)	This task requires participants to touch a cross displayed at different points on screen and is intended to familiarize examinees with the touch screen.	Mean latency (mean time to touch the cross after it has appeared on the screen).
Rapid Visual Information Processing (RVP)	Digits 2 to 9 are displayed one at a time on the screen (100 stimuli per minute). The task is to detect the 3-5-7 sequence and press a button after the target sequence's last digit appears.	Probability of hit (probability of correctly detecting the target sequence ^a). Probability of false alarm (probability of responding outside the response window of the target sequence ^b). A' (sensitivity to target sequences ^c , ranging from 0 [bad] to 1 [good]). Mean latency (mean time to detect the target sequence within a response window of 1800 ms).
Paired Associates Learning (PAL)	This task requires participants to observe boxes opening up one by one at random and to memorize which pattern appeared in which box and then touch the box where each pattern appeared. There are eight stages of increasing difficulty (1 to 8 patterns) and a maximum of ten trials to solve each stage.	Mean errors to success (mean errors made before successfully completing a stage). Mean trials to success (mean trials needed to successfully complete a stage). First trial memory score (sum of patterns correctly located on the first trial, for all of the stages completed).
Reaction Time (RTI)	For simple (one circle) and five-choice (five circles) conditions, the task is to press a button until a yellow spot appears in a circle and then to release the button and touch that circle. If the criterion of five of six correct responses is not achieved, the test ends.	Simple reaction time and five-choice reaction time (time to release the button). Simple movement time and five-choice movement time (time to touch the circle after button release).
Spatial Working Memory (SWM)	This task consists of searching for blue tokens hidden one at a time inside boxes and with the understanding that a box in which a blue token has already been found should not be touched again. There are four levels of increasing difficulty (three to eight boxes). Each level (or problem) ends after a house is filled with the tokens.	Total errors (number of times a box is touched when it surely does not have a token). Strategy (number of times a new search begins by touching a different box, for problems of six to eight boxes; higher scores indicate less strategy use). Mean time to last response (mean time between displaying the problem and opening the box with the final token for the problem).

(Table 4 continues)

Table 4 (continued). Sequence of tests administered in the second session.

Test	Description	Test measures
<i>WMS-III tests</i>		
Logical Memory I	This episodic memory test requires learning and recalling aloud as many details as possible from two stories.	Immediate recall score.
Family Pictures I	This episodic memory test involves looking at four drawings and recalling as much information as possible. Then, the examinee is asked to name the characters that were in each picture, where they were (location) and what were they doing (activity).	Immediate recall score.
Verbal Paired Associates I	An episodic associative memory test that requires learning and recalling as many pairs of semantically-unrelated words as possible from an eight word-pair list.	Immediate recall score.
Spatial Span	This visuospatial attention and working memory test consists in reproducing sequences of taps on cubes, forward and backward.	Forward and backward scores.
Logical Memory II	This test involves recalling and recognizing as many details as possible from two stories learned 25-35 minutes before.	Delayed recall and recognition scores.
Family Pictures II	This test requires recalling as much information as possible from four family pictures observed 25-35 minutes before.	Delayed recall score.
Verbal Paired Associates II	This test requires recalling and recognizing as many word pairs as possible from a list of eight word-pairs learned 25-35 minutes before.	Delayed recall and recognition scores.

Note. ^aThe probability of hit is calculated by the formula $p(\text{hit}) = \text{total hits}/(\text{total hits} + \text{total misses})$. ^bThe probability of false alarm is calculated with the formula $p(\text{false alarm}) = \text{total false alarms}/(\text{total false alarms} + \text{total correct rejections of non-target stimuli})$. ^cThe A' measure is calculated by the formula $A' = 0.5 + \{p(\text{hit}) - p(\text{false alarm}) + [p(\text{hit}) - p(\text{false alarm})]^2\} / \{4 [p(\text{hit}) (1 - p(\text{false alarm}))]\}$, which is based on the principles of the Signal Detection Theory (Roiser et al., 2005). Legend: CANTAB = Cambridge Neuropsychological Test Automated Battery (Cambridge Cognition, 2006a, 2006b); ms = milliseconds; WMS-III = Wechsler Memory Scale-III (Wechsler, 2008b, 2008c).

To analyze the concurrent validity of the RVP, PAL, RTI, and SWM (CANTAB) tests, we selected pencil-and-paper tests of episodic memory (i.e., RCFT immediate recall and WMS-III Family Pictures, Logical Memory, and Verbal Paired Associates [VPA] subtests), attention and working memory (i.e., Spatial Span [WMS-III] and Digit Span [WAIS-III]), and executive functions such as mental flexibility and selective attention (i.e., Verbal Fluency animals score from the Addenbrooke's Cognitive

Examination Revised [ACE-R]; Firmino, Simões, Pinho, Cerejeira, & Martins, 2008; Mioshi et al., 2006). See descriptions of these tests and measures in Tables 2 and 4.

Statistical analysis

Statistical analyses were conducted with SPSS[®] Statistics version 22.0 software (IBM[®], New York, USA), and a significance level of $p < .05$ was adopted. We used Kolmogorov-Smirnov tests (if $N > 50$) or Shapiro-Wilk tests (if $N \leq 50$) to check the normality of the distributions (Marôco, 2014). The associations between tests were calculated with Pearson parametric correlations or Spearman nonparametric correlations (when the assumption of normality of the distribution was not met). Because some participants had missing values (the percentage of such participants per variable was less than 5% of the total sample, less than 6% of the WND group, and less than 17% of the AD group), we used the pairwise comparison method for correlations and exploratory factor analyses (EFAs). We also performed these analyses using the listwise method (which excludes cases with missing values); however, these results are not shown because they were similar to those of the statistical analyses employing the pairwise method.

The EFAs were calculated using the principal component extraction method followed by varimax rotation with Kaiser Normalization to analyze the latent factorial structure of the set of CANTAB measures and, then, of the set of CANTAB and pencil-and-paper tests' measures in the WND group and in the total sample. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy varied between .72 and .90, which is above the minimum recommended value (i.e., .60; Tabachnick & Fidell, 2001) and indicates that the sample sizes were adequate for the number of analyzed variables. The criteria used to determine the number of factors to be retained were scree plot observation, initial eigenvalues greater than 1, and factors collectively explain more than 50% of the total variance in the original variables. The percentage of non-redundant residuals greater than .05 varied from 28% to 35% (i.e., less than 50%), suggesting that all of the factor models presented a good fit between the observed and reproduced correlations (Marôco, 2014). The retained factors explained approximately 50% or more of the variance in each test measure (communalities range: .48-.93).

RESULTS

With respect to the raw scores on the CANTAB and pencil-and-paper tests selected for validity analysis (Table 5), both the WND and AD groups reached the maximum possible score for RVP probability of hit (CANTAB) and the WND group reached the maximum score for RVP A' (i.e., sensitivity index to detect the target 3-5-7 sequence from non-target stimuli) (CANTAB), VPA delayed recall (WMS-III), and Digit Span forward (WAIS-III). The minimum possible score was obtained in both the WND and AD groups for VPA immediate and delayed recall (WMS-III); in the WND group for RVP probability of false alarm (CANTAB); and in the AD group for RCFT immediate recall, Family Pictures immediate and delayed recall, Logical Memory immediate and delayed recall, and Spatial Span forward and backward (WMS-III). Student's independent samples *t*-tests indicated that the WND group scored significantly better than the AD group on all of the test measures listed in Table 5 ($p < .01$; $|d| \geq .40$).

Correlational analysis

Correlations between the CANTAB tests and the pencil-and-paper tests

The correlations between the CANTAB and pencil-and-paper tests ranged from .00 to .52 in the WND group (Table 6), from .00 to .64 in the AD group (Table 7), and from .02 to .70 in the total sample (Table 8); therefore, they were strong at most (weak: .10-.39; moderate: .40-.69; strong: .70-.89; classification of Dancey & Reidy, 2004). The correlations between tests were generally higher in the total sample than in the two mentioned groups.

Table 5. Raw scores on the tests selected for concurrent validity analysis.

Test measure	WND group				AD group				Total sample		
	<i>n</i>	<i>M</i>	<i>SD</i>	Range	<i>n</i>	<i>M</i>	<i>SD</i>	Range	<i>N</i>	<i>M</i>	<i>SD</i>
RVP ^a probability of hit	130	.822	.168	.250-1.000	37	.652	.212	.250-1.000	167	.785	.191
RVP ^a probability of false alarm	130	.017	.030	.000-.166	37	.073	.056	.004-.210	167	.029	.044
RVP ^a A' (sensitivity index)	130	.947	.057	.704-1.000	37	.870	.089	.604-.982	167	.930	.073
RVP ^a mean latency	130	438.53	117.30	228.00-795.81	37	647.52	202.40	294.58-1199.83	167	484.83	164.79
PAL ^a mean errors to success	137	5.35	2.65	0.75-11.57	37	8.97	2.41	4.67-14.29	174	6.12	2.99
PAL ^a mean trials to success	137	2.72	0.87	1.25-5.67	37	5.03	1.27	2.83-7.50	174	3.21	1.35
PAL ^a first trial memory score	137	13.31	3.81	2-21	37	5.30	3.58	1-17	174	11.61	4.99
RTI ^a simple reaction time	137	383.48	106.74	245.57-1027.91	34	535.06	206.23	293.63-1189.86	171	413.62	145.10
RTI ^a five-choice reaction time	136	421.41	84.01	277.13-734.00	31	605.90	266.46	326.43-1640.13	167	455.66	154.10
RTI ^a simple movement time	137	563.73	208.16	198.13-1169.00	34	961.82	517.26	348.89-2273.50	171	642.88	334.66
RTI ^a five-choice movement time	136	490.84	174.99	200.71-1249.63	31	768.48	446.63	245.86-2268.29	167	542.38	269.58
SWM ^a total errors	137	54.49	16.58	4-97	37	73.32	14.76	40-126	174	58.49	17.92
SWM ^a strategy	137	38.40	3.76	26-46	37	39.81	2.38	36-47	174	38.70	3.56
SWM ^a mean time to last response	137	43541.93	9608.76	25891.75-79812.58	37	60628.25	18363.00	36233.67-133940.00	174	47175.23	13853.23
RCFT immediate recall	137	15.13	5.58	6-30	37	2.04	3.71	0-15	174	12.34	7.50
Family Pictures ^b imm. recall	136	29.75	8.57	10-51	37	9.22	6.37	0-27	173	25.36	11.72
Family Pictures ^b delayed recall	136	29.17	8.70	13-49	37	6.54	7.21	0-26	173	24.33	12.53
Logical Memory ^b imm. recall	136	36.49	8.61	13-57	37	10.73	10.38	0-35	173	30.98	13.89
Logical Memory ^b delayed recall	136	23.40	6.32	3-37	37	3.24	5.06	0-17	173	19.09	10.27
VPA ^b immediate recall	136	9.67	6.51	0-27	36	0.50	1.03	0-4	172	7.75	6.90
VPA ^b delayed recall	136	3.15	2.35	0-8	36	0.08	0.28	0-1	172	2.51	2.44
Spatial Span ^b forward	136	6.34	1.65	3-11	37	3.78	1.47	0-6	173	5.79	1.92
Spatial Span ^b backward	136	5.10	1.66	2-9	37	2.43	1.57	0-6	173	4.53	1.97
Digit Span ^c forward	137	8.17	2.19	4-16	37	6.84	1.69	4-10	174	7.89	2.16
Digit Span ^c backward	137	4.98	1.81	2-10	37	3.30	1.39	1-6	174	4.62	1.86
Verbal Fluency ^d animals	137	14.91	4.48	6-26	37	5.70	2.34	1-13	174	12.95	5.58

Note. Time is in milliseconds. Lower scores reflect better performance for all of the CANTAB measures except RVP probability of hit, RVP A', and PAL first trial memory score. ^aCambridge Neuropsychological Test Automated Battery. ^bWechsler Memory Scale-III. ^cWechsler Adult Intelligence Scale-III. ^dAddenbrooke's

Cognitive Examination Revised. Legend: imm. = immediate; PAL = Paired Associates Learning; RCFT = Rey Complex Figure Test; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory; VPA = Verbal Paired Associates; WND = without neuropsychiatric diagnoses.

Table 6. Correlations between the Cambridge neuropsychological automated tests and pencil-and-paper tests in the WND group.

CANTAB tests	RCFT immediate recall	Family Pictures ^a immediate recall	Family Pictures ^a delayed recall	Logical Memory ^a immediate recall	Logical Memory ^a delayed recall	VPA ^a immediate recall	VPA ^a delayed recall	Spatial Span ^a forward	Spatial Span ^a backward	Digit Span ^b forward	Digit Span ^b backward	Verbal Fluency ^c animals
RVP probability of hit	.18*	.14	.12	.19*	.25**	.20*	.15	.24**	.33***	.26**	.14	.40***
RVP probability of false alarm	-.23**	-.20*	-.19*	-.24**	-.26**	-.21*	-.09	-.26**	-.32***	-.37***	-.25**	-.49***
RVP A' (sensitivity index)	.20*	.16	.14	.20*	.26**	.21*	.15	.26**	.34***	.28**	.16	.44***
RVP mean latency	-.07	-.25**	-.21*	-.11	-.14	-.16	-.10	-.12	-.25**	-.13	-.28**	-.38***
PAL mean errors to success	-.52***	-.42***	-.41***	-.35***	-.33***	-.34***	-.33***	-.24**	-.21*	-.18*	-.10	-.25**
PAL mean trials to success	-.51***	-.40***	-.42***	-.37***	-.34***	-.28***	-.27**	-.25**	-.25**	-.17*	-.12	-.31***
PAL first trial memory score	.42***	.38***	.33***	.24**	.21*	.31***	.36***	.26**	.34***	.19*	.11	.38***
RTI simple reaction time	-.18*	-.26**	-.27**	.00	.00	-.04	-.06	-.17	-.17	-.08	-.04	-.35***
RTI five-choice reaction time	-.10	-.25**	-.25**	-.01	-.04	.08	.09	-.08	-.07	-.04	-.03	-.37***
RTI simple movement time	-.17*	-.24**	-.25**	-.02	.01	-.10	-.12	-.28***	-.36***	-.25**	-.19*	-.33***
RTI five-choice movement time	-.19*	-.14	-.14	-.00	-.01	-.08	-.06	-.16	-.28***	-.20*	-.14	-.34***
SWM total errors	-.23**	-.19*	-.17	-.16	-.10	-.15	-.14	-.32***	-.37***	-.14	-.30***	-.27**
SWM strategy	-.15	-.12	-.10	.07	.02	-.05	-.05	-.26**	-.34***	-.29***	-.18*	-.02
SWM mean time to last response	-.18*	-.19*	-.17*	-.17*	-.14	-.19*	-.14	-.23**	-.34***	-.14	-.20*	-.47***

Note. Pearson correlations are italicized. Other coefficients are Spearman correlations. Coefficients equal to or higher than .40 are presented in bold. The pairwise method was used. *Ns* range: 129-137. ^aWechsler Memory Scale-III. ^bWechsler Adult Intelligence Scale-III. ^cAddenbrooke's Cognitive Examination Revised. Legend: CANTAB = Cambridge Neuropsychological Test Automated Battery; PAL = Paired Associates Learning; RCFT = Rey Complex Figure Test; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory; VPA = Verbal Paired Associates; WND = without neuropsychiatric diagnoses. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7. Correlations between the Cambridge neuropsychological automated tests and pencil-and-paper tests in the AD group.

CANTAB tests	RCFT immediate recall	Family Pictures ^a immediate recall	Family Pictures ^a delayed recall	Logical Memory ^a immediate recall	Logical Memory ^a delayed recall	VPA ^a immediate recall	VPA ^a delayed recall	Spatial Span ^a forward	Spatial Span ^a backward	Digit Span ^b forward	Digit Span ^b backward	Verbal Fluency ^c animals
RVP probability of hit	.06	<i>.11</i>	.12	-.11	-.24	-.01	-.03	-.11	<i>.31</i>	-.19	-.07	-.19
RVP probability of false alarm	-.27	-.13	-.22	-.18	-.02	-.17	-.26	-.43**	-.12	-.16	-.13	-.33*
RVP A' (sensitivity index)	.12	.09	.14	-.07	-.20	.05	.03	.02	.27	-.13	-.09	-.07
RVP mean latency	-.36*	-.10	-.41*	-.42**	-.17	-.06	-.10	-.34*	-.50**	-.08	-.22	-.08
PAL mean errors to success	.16	.27	.14	-.03	.01	-.04	.39*	-.02	-.05	.03	-.29	.06
PAL mean trials to success	-.38*	-.20	-.33*	-.64***	-.34*	-.29	-.12	-.51**	-.46**	-.24	-.39*	-.25
PAL first trial memory score	.45*	.19	.44**	.61***	.42**	.26	.26	.46**	.32	.21	.23	.28
RTI simple reaction time	-.05	-.29	-.22	-.24	-.20	-.16	-.27	-.13	-.29	.01	-.26	-.28
RTI five-choice reaction time	-.12	-.09	-.25	-.52**	-.27	-.32	-.12	-.16	-.34	-.21	-.33	-.20
RTI simple movement time	.02	-.12	-.15	-.40*	-.21	-.38*	-.22	-.35*	-.44**	-.32	-.33	-.24
RTI five-choice movement time	-.04	-.10	-.19	-.37*	-.23	-.42*	-.15	-.41*	-.37*	-.48**	-.58***	-.25
SWM total errors	-.02	-.02	-.13	-.24	-.07	-.07	.20	-.47**	-.59***	-.07	-.17	-.00
SWM strategy	-.05	<i>.13</i>	.02	.12	.10	-.03	-.03	-.15	-.13	-.16	-.45**	-.02
SWM mean time to last response	-.35*	-.27	-.22	-.43**	-.14	.18	-.01	-.32	-.45**	-.18	-.27	.00

Note. Pearson correlations are italicized. Other coefficients are Spearman correlations. Coefficients equal to or higher than .40 are presented in bold. The pairwise method was used. *Ns* range: 30-37. ^aWechsler Memory Scale-III. ^bWechsler Adult Intelligence Scale-III. ^cAddenbrooke's Cognitive Examination Revised. Legend: AD = Alzheimer's disease; CANTAB = Cambridge Neuropsychological Test Automated Battery; PAL = Paired Associates Learning; RCFT = Rey Complex Figure Test; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory; VPA = Verbal Paired Associates. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 8. Correlations between the Cambridge neuropsychological automated tests and pencil-and-paper tests in the total sample.

CANTAB tests	RCFT immediate recall	Family Pictures ^a immediate recall	Family Pictures ^a delayed recall	Logical Memory ^a immediate recall	Logical Memory ^a delayed recall	VPA ^a immediate recall	VPA ^a delayed recall	Spatial Span ^a forward	Spatial Span ^a backward	Digit Span ^b forward	Digit Span ^b backward	Verbal Fluency ^c animals
RVP probability of hit	<i>.34</i> ^{***}	<i>.32</i> ^{***}	<i>.31</i> ^{***}	<i>.32</i> ^{***}	<i>.38</i> ^{***}	<i>.36</i> ^{***}	<i>.31</i> ^{***}	<i>.32</i> ^{***}	.42 ^{***}	<i>.24</i> ^{**}	<i>.19</i> [*]	.47 ^{***}
RVP probability of false alarm	-.50 ^{***}	-.49 ^{***}	-.48 ^{***}	-.50 ^{***}	-.52 ^{***}	-.48 ^{***}	-.41 ^{***}	-.48 ^{***}	-.50 ^{***}	-.41 ^{***}	-.38 ^{***}	-.67 ^{***}
RVP A' (sensitivity index)	.42 ^{***}	.40 ^{***}	<i>.39</i> ^{***}	.41 ^{***}	.46 ^{***}	.43 ^{***}	<i>.38</i> ^{***}	.41 ^{***}	.48 ^{***}	<i>.29</i> ^{***}	<i>.25</i> ^{***}	.57 ^{***}
RVP mean latency	-.37 ^{***}	-.48 ^{***}	-.46 ^{***}	-.39 ^{***}	-.41 ^{***}	-.42 ^{***}	-.36 ^{***}	-.35 ^{***}	-.45 ^{***}	-.26 ^{***}	-.42 ^{***}	-.54 ^{***}
PAL mean errors to success	-.63 ^{***}	-.55 ^{***}	-.57 ^{***}	-.53 ^{***}	-.52 ^{***}	-.53 ^{***}	-.50 ^{***}	-.43 ^{***}	-.41 ^{***}	-.27 ^{***}	-.29 ^{***}	-.48 ^{***}
PAL mean trials to success	-.70 ^{***}	-.65 ^{***}	-.66 ^{***}	-.64 ^{***}	-.62 ^{***}	-.58 ^{***}	-.56 ^{***}	-.52 ^{***}	-.52 ^{***}	-.31 ^{***}	-.36 ^{***}	-.60 ^{***}
PAL first trial memory score	.67 ^{***}	.62 ^{***}	.62 ^{***}	.55 ^{***}	.55 ^{***}	.58 ^{***}	.59 ^{***}	.51 ^{***}	.55 ^{***}	<i>.30</i> ^{***}	<i>.33</i> ^{***}	.63 ^{***}
RTI simple reaction time	-.36 ^{***}	-.43 ^{***}	-.43 ^{***}	-.26 ^{***}	-.27 ^{***}	-.31 ^{***}	-.30 ^{***}	-.34 ^{***}	-.37 ^{***}	<i>-.18</i> [*]	<i>-.23</i> ^{**}	-.48 ^{***}
RTI five-choice reaction time	-.31 ^{***}	-.41 ^{***}	-.41 ^{***}	-.26 ^{***}	-.28 ^{***}	-.21 ^{**}	-.18 [*]	-.25 ^{***}	-.27 ^{***}	<i>-.17</i> [*]	<i>-.21</i> ^{**}	-.49 ^{***}
RTI simple movement time	-.33 ^{***}	-.39 ^{***}	-.39 ^{***}	-.26 ^{***}	-.24 ^{**}	-.30 ^{***}	-.30 ^{***}	-.42 ^{***}	-.49 ^{***}	-.33 ^{***}	-.32 ^{***}	-.45 ^{***}
RTI five-choice movement time	-.31 ^{***}	-.27 ^{***}	-.28 ^{***}	-.19 [*]	-.20 ^{**}	-.27 ^{***}	-.23 ^{**}	-.31 ^{***}	-.41 ^{***}	-.30 ^{***}	-.31 ^{***}	-.40 ^{***}
SWM total errors	-.44 ^{***}	-.43 ^{***}	-.42 ^{***}	-.41 ^{***}	-.38 ^{***}	-.41 ^{***}	-.39 ^{***}	-.49 ^{***}	-.53 ^{***}	<i>-.24</i> ^{**}	-.41 ^{***}	-.48 ^{***}
SWM strategy	<i>-.18</i> [*]	<i>-.16</i> [*]	<i>-.15</i> [*]	<i>-.02</i>	<i>-.07</i>	<i>-.11</i>	<i>-.13</i>	<i>-.27</i> ^{***}	<i>-.33</i> ^{***}	<i>-.27</i> ^{***}	<i>-.25</i> ^{**}	<i>-.09</i>
SWM mean time to last response	-.43 ^{***}	-.44 ^{***}	-.43 ^{***}	-.43 ^{***}	-.41 ^{***}	-.41 ^{***}	-.38 ^{***}	-.42 ^{***}	-.52 ^{***}	<i>-.24</i> ^{**}	-.36 ^{***}	-.59 ^{***}

Note. Pearson correlations are italicized. Other coefficients are Spearman correlations. Coefficients equal to or higher than .40 are presented in bold. The pairwise method was used. *N*s range: 165-174. ^aWechsler Memory Scale-III. ^bWechsler Adult Intelligence Scale-III. ^cAddenbrooke's Cognitive Examination Revised. Legend: CANTAB = Cambridge Neuropsychological Test Automated Battery; PAL = Paired Associates Learning; RCFT = Rey Complex Figure Test; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory; VPA = Verbal Paired Associates. * $p < .05$. ** $p < .01$. *** $p < .001$.

The strongest RVP inter-test correlations observed in the WND group (Table 6) were those between the four RVP measures and the executive function measures of Verbal Fluency animals (ACE-R) and Spatial Span backward (WMS-III) (moderate to weak associations). In the AD group (Table 7), only RVP probability of false alarm and RVP mean latency revealed moderate to weak correlations with these executive function tests (i.e., RVP probability of false alarm with Verbal Fluency animals [ACE-R] and Spatial Span forward [WMS-III] and RVP mean latency with Spatial Span forward and backward [WMS-III]). In addition, we found moderate to weak correlations between RVP mean latency and the episodic memory measures of Logical Memory immediate recall, Family Pictures delayed recall (WMS-III), and RCFT immediate recall in the AD group. As in the WND group, in the total sample (Table 8), the strongest RVP inter-test correlations were with Verbal Fluency animals (ACE-R) and Spatial Span backward (WMS-III) (moderate associations).

The highest PAL inter-test correlations identified in the WND group (Table 6) were between the three PAL measures and the episodic memory measures of RCFT immediate recall and Family Pictures (WMS-III) (moderate to weak associations). In the AD group (Table 7), only PAL mean trials to success and PAL first trial memory score revealed moderate to weak correlations with RCFT immediate recall and Family Pictures delayed recall (WMS-III), but these two PAL measures' strongest inter-test correlations were with Logical Memory immediate recall and Spatial Span forward (WMS-III) (moderate to weak associations). In the total sample (Table 8), the PAL showed the highest inter-test correlations with RCFT immediate recall, Family Pictures, Logical Memory (WMS-III) and Verbal Fluency animals (ACE-R) (strong to moderate associations).

Regarding the strongest RTI inter-test correlations, in the WND group (Table 6), the correlations between the four RTI measures and Verbal Fluency animals (ACE-R) and between the RTI movement time measures and Spatial Span backward (WMS-III) were significant but weak. By contrast, in the AD group (Table 7), the correlations between the RTI measures and Verbal Fluency animals (ACE-R) were nonsignificant. However, we identified moderate to weak associations between the RTI movement time measures and Spatial Span (WMS-III) and between RTI five-choice movement time and Digit Span (WAIS-III). Moreover, we observed that the RTI five-choice reaction time

and RTI movement time measures correlated moderately to weakly with Logical Memory immediate recall (WMS-III). In the total sample (Table 8), the highest RTI inter-test correlations were with Verbal Fluency animals (ACE-R), Spatial Span, and Family Pictures (WMS-III) (moderate to weak associations).

In the WND group (Table 6), the strongest inter-test correlations registered for the SWM test were moderate to weak, particularly the correlations between the three SWM measures and Spatial Span (WMS-III) (mostly with Spatial Span backward) and Digit Span backward (WAIS-III), and between two SWM measures (i.e., SWM total errors and SWM mean time to last response) and Verbal Fluency animals (ACE-R). In the AD group (Table 7), only SWM total errors and SWM mean time to last response correlated with Spatial Span backward, and SWM strategy correlated with Digit Span backward (WAIS-III) (moderate associations). All of the SWM inter-test correlations with Verbal Fluency animals (ACE-R) were nonsignificant. In the total sample (Table 8), the highest SWM inter-test correlations were between the three SWM measures and Spatial Span (WMS-III) and between two SWM measures (i.e., SWM total errors and SWM mean time to last response) and Verbal Fluency animals (ACE-R) (moderate to weak associations).

CANTAB intra-test correlations

Considering that both RVP probability of hit and RVP probability of false alarm are taken into account in the RVP A' calculation formula, not surprisingly, the correlations between the first two measures and RVP A' were high. Specifically, these intra-test correlations were very strong to strong (strong: .70-.89; very strong: .90-.99; classification of Dancey & Reidy, 2004) in the WND group (Table 9), very strong to moderate in the AD group (Table 10), and very strong to strong in the total sample (Table 11). Examining the intra-test correlations among the RVP measures that were calculated independently from each other (i.e., RVP probability of hit, RVP probability of false alarm, and RVP mean latency), we identified a strong correlation between RVP probability of hit and RVP probability of false alarm in the WND group (Table 9), but these measures revealed weak associations (although significant) with RVP mean latency. In the AD group (Table 10), we observed a nonsignificant correlation between RVP probability of hit and RVP probability of false alarm, with both measures showing

moderate correlations with RVP mean latency. In the total sample (Table 11), we found a similar pattern of RVP intra-test correlations to that identified in the WND group.

The PAL intra-test correlations were very strong to moderate in the WND group (Table 9), strong to nonsignificant in the AD group (Table 10) and strong in the total sample (Table 11). The nonsignificant intra-test correlations identified in the AD group were those involving PAL mean errors to success.

The RTI intra-test correlations were strong to moderate in the WND group (Table 9), AD group (Table 10), and total sample (Table 11).

The SWM intra-test correlations were moderate to weak (but significant) in the WND group (Table 9), moderate to nonsignificant in the AD group (Table 10), and moderate to weak (but significant) in the total sample (Table 11). The SWM total errors measure revealed a moderate correlation with SWM mean time to last response in the WND group, AD group, and the total sample. However, the correlation between SWM total errors and SWM strategy was moderate in the WND group, nonsignificant in the AD group, and weak (but significant) in the total sample. In addition, the correlation between SWM strategy and SWM mean time to last response was weak in the WND group and total sample and nonsignificant in the AD group.

Table 9. Correlations between the Cambridge neuropsychological automated tests in the WND group.

	RVP probability of hit	RVP probability of false alarm	RVP A' (sensitivity index)	RVP mean latency	PAL mean errors to success	PAL mean trials to success	PAL first trial memory score	RTI simple reaction time	RTI five-choice reaction time	RTI simple movement time	RTI five-choice movement time	SWM total errors	SWM strategy	SWM mean time to last response
RVP probability of hit														
RVP probability of false alarm	<i>-.73</i> ^{***}													
RVP A'	<i>.99</i> ^{***}	<i>-.81</i> ^{***}												
RVP mean latency	<i>-.26</i> ^{**}	<i>.29</i> ^{***}	<i>-.29</i> ^{***}											
PAL mean errors to success	<i>-.25</i> [*]	<i>.27</i> ^{**}	<i>-.26</i> ^{**}	<i>.08</i>										
PAL mean trials to success	<i>-.21</i> [*]	<i>.27</i> ^{**}	<i>-.23</i> ^{**}	<i>.17</i>	<i>.93</i> ^{***}									
PAL first trial memory score	<i>.26</i> ^{**}	<i>-.30</i> ^{***}	<i>.28</i> ^{**}	<i>-.14</i>	<i>-.64</i> ^{***}	<i>-.68</i> ^{***}								
RTI simple reaction time	<i>-.33</i> ^{***}	<i>.25</i> ^{**}	<i>-.33</i> ^{***}	<i>.26</i> ^{**}	<i>.15</i>	<i>.17</i>	<i>-.05</i>							
RTI five-choice reaction time	<i>-.32</i> ^{***}	<i>.30</i> ^{***}	<i>-.33</i> ^{***}	<i>.42</i> ^{***}	<i>.19</i> [*]	<i>.24</i> ^{**}	<i>-.15</i>	<i>.69</i> ^{***}						
RTI simple movement time	<i>-.24</i> [*]	<i>.19</i> [*]	<i>-.24</i> ^{**}	<i>.22</i> [*]	<i>.17</i> [*]	<i>.18</i> [*]	<i>-.19</i> [*]	<i>.54</i> ^{***}	<i>.48</i> ^{***}					
RTI five-choice movement time	<i>-.22</i> [*]	<i>.21</i> [*]	<i>-.23</i> [*]	<i>.22</i> [*]	<i>.15</i>	<i>.16</i>	<i>-.20</i> [*]	<i>.45</i> ^{***}	<i>.50</i> ^{***}	<i>.76</i> ^{***}				
SWM total errors	<i>-.14</i>	<i>.39</i> ^{***}	<i>-.19</i> [*]	<i>.32</i> ^{***}	<i>.19</i> [*]	<i>.24</i> ^{**}	<i>-.23</i> ^{**}	<i>.16</i>	<i>.18</i> [*]	<i>.22</i> [*]	<i>.28</i> ^{***}			
SWM strategy	<i>-.11</i>	<i>.33</i> ^{***}	<i>-.16</i>	<i>.12</i>	<i>.06</i>	<i>.05</i>	<i>-.11</i>	<i>.13</i>	<i>.11</i>	<i>.11</i>	<i>.12</i>	<i>.45</i> ^{***}		
SWM mean time to last response	<i>-.27</i> ^{**}	<i>.37</i> ^{***}	<i>-.30</i> ^{***}	<i>.37</i> ^{***}	<i>.16</i>	<i>.20</i> [*]	<i>-.30</i> ^{**}	<i>.23</i> ^{**}	<i>.30</i> ^{***}	<i>.19</i> [*]	<i>.20</i> [*]	<i>.47</i> ^{***}	<i>.26</i> ^{**}	

Note. Pearson correlations are italicized. Other coefficients are Spearman correlations. Coefficients equal to or higher than .40 are presented in bold. The pairwise method was used. *Ns* range: 129-137. Legend: PAL = Paired Associates Learning; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory; WND = without neuropsychiatric diagnoses. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 10. Correlations between the Cambridge neuropsychological automated tests in the AD group.

	RVP probability of hit	RVP probability of false alarm	RVP A' (sensitivity index)	RVP mean latency	PAL mean errors to success	PAL mean trials to success	PAL first trial memory score	RTI simple reaction time	RTI five-choice reaction time	RTI simple movement time	RTI five-choice movement time	SWM total errors	SWM strategy	SWM mean time to last response
RVP probability of hit														
RVP probability of false alarm	<i>-.20</i>													
RVP A'	.94^{***}	-.47^{**}												
RVP mean latency	-.49^{**}	.44^{**}	-.49^{**}											
PAL mean errors to success	<i>.06</i>	<i>-.23</i>	<i>.17</i>	<i>.11</i>										
PAL mean trials to success	<i>-.03</i>	<i>.28</i>	<i>-.06</i>	.61^{***}	<i>.32</i>									
PAL first trial memory score	<i>.04</i>	<i>-.33[*]</i>	<i>.12</i>	-.61^{***}	<i>.03</i>	-.83^{***}								
RTI simple reaction time	<i>-.16</i>	.42[*]	<i>-.20</i>	.42[*]	<i>-.07</i>	<i>.31</i>	<i>-.30</i>							
RTI five-choice reaction time	<i>-.09</i>	<i>.11</i>	<i>-.06</i>	.64^{***}	<i>.20</i>	.57^{***}	-.59^{***}	.59^{***}						
RTI simple movement time	<i>-.03</i>	<i>.39[*]</i>	<i>-.10</i>	.43[*]	<i>.06</i>	.63^{***}	-.48^{**}	.60^{***}	.56^{**}					
RTI five-choice movement time	<i>.12</i>	<i>.34</i>	<i>.08</i>	.44[*]	<i>.14</i>	.64^{***}	-.51^{**}	.65^{***}	.68^{***}	.81^{***}				
SWM total errors	<i>-.31</i>	<i>.02</i>	<i>-.25</i>	<i>.38[*]</i>	<i>.10</i>	<i>.33[*]</i>	<i>-.29</i>	<i>.20</i>	<i>.14</i>	<i>.20</i>	<i>.16</i>			
SWM strategy	<i>.23</i>	<i>-.02</i>	<i>.28</i>	<i>.03</i>	<i>.23</i>	<i>.06</i>	<i>-.06</i>	<i>-.03</i>	<i>-.14</i>	<i>.00</i>	<i>.21</i>	<i>-.17</i>		
SWM mean time to last response	<i>-.09</i>	<i>.14</i>	<i>-.08</i>	<i>.37[*]</i>	<i>-.07</i>	.49^{**}	-.47^{**}	<i>.13</i>	<i>.09</i>	<i>.30</i>	<i>.07</i>	.56^{***}	<i>-.08</i>	

Note. Pearson correlations are italicized. Other coefficients are Spearman correlations. Coefficients equal to or higher than .40 are presented in bold. The pairwise method was used. *Ns* range: 31-37. Legend: AD = Alzheimer's disease; PAL = Paired Associates Learning; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 11. Correlations between the Cambridge neuropsychological automated tests in the total sample.

	RVP probability of hit	RVP probability of false alarm	RVP A' (sensitivity index)	RVP mean latency	PAL mean errors to success	PAL mean trials to success	PAL first trial memory score	RTI simple reaction time	RTI five-choice reaction time	RTI simple movement time	RTI five-choice movement time	SWM total errors	SWM strategy	SWM mean time to last response
RVP probability of hit														
RVP probability of false alarm	-.68 ^{***}													
RVP A'	.97 ^{***}	-.82 ^{***}												
RVP mean latency	-.38 ^{***}	.50 ^{***}	-.45 ^{***}											
PAL mean errors to success	-.34 ^{***}	.42 ^{***}	-.38 ^{***}	.29 ^{***}										
PAL mean trials to success	-.35 ^{***}	.52 ^{***}	-.43 ^{***}	.45 ^{***}	.89 ^{***}									
PAL first trial memory score	.38 ^{***}	-.54 ^{***}	.45 ^{***}	-.43 ^{***}	-.70 ^{***}	-.81 ^{***}								
RTI simple reaction time	-.35 ^{***}	.40 ^{***}	-.40 ^{***}	.43 ^{***}	.30 ^{***}	.38 ^{***}	-.30 ^{***}							
RTI five-choice reaction time	-.35 ^{***}	.42 ^{***}	-.40 ^{***}	.56 ^{***}	.35 ^{***}	.42 ^{***}	-.35 ^{***}	.71 ^{***}						
RTI simple movement time	-.27 ^{***}	.36 ^{***}	-.33 ^{***}	.39 ^{***}	.30 ^{***}	.37 ^{***}	-.35 ^{***}	.61 ^{***}	.56 ^{***}					
RTI five-choice movement time	-.20 [*]	.33 ^{***}	-.26 ^{***}	.39 ^{***}	.27 ^{***}	.33 ^{***}	-.33 ^{***}	.55 ^{***}	.60 ^{***}	.78 ^{***}				
SWM total errors	-.28 ^{***}	.49 ^{***}	-.35 ^{***}	.50 ^{***}	.36 ^{***}	.48 ^{***}	-.46 ^{***}	.34 ^{***}	.31 ^{***}	.33 ^{***}	.38 ^{***}			
SWM strategy	-.07	.31 ^{***}	-.13	.11	.14	.12	-.15 [*]	.14	.11	.13	.15 [*]	.36 ^{***}		
SWM mean time to last response	-.35 ^{***}	.49 ^{***}	-.42 ^{***}	.52 ^{***}	.33 ^{***}	.46 ^{***}	-.52 ^{***}	.36 ^{***}	.40 ^{***}	.33 ^{***}	.30 ^{***}	.59 ^{***}	.23 ^{**}	

Note. The coefficients are Spearman correlations and were calculated using the pairwise method (*N*s range: 160-174). Coefficients equal to or higher than .40 are presented in bold. Legend: PAL = Paired Associates Learning; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory. * $p < .05$. ** $p < .01$. *** $p < .001$.

Correlations between CANTAB tests

The correlations between CANTAB tests varied from .05 to .42 in the WND group (Table 9), from .00 to .64 in the AD group (Table 10), and from .07 to .56 in the total sample (Table 11), indicating that the associations were moderate at most. Once again, the inter-test correlations were generally higher in the total sample than in the WND and AD groups.

In the WND group (Table 9), only a moderate inter-test correlation between RVP mean latency and RTI five-choice reaction time was found; all of the other inter-test correlations were weak to nonsignificant. In the AD group (Table 10), we identified moderate inter-test correlations between RVP probability of false alarm and RTI simple reaction time; between RVP mean latency and PAL mean trials to success, PAL first trial memory score, and the RTI measures; and between PAL mean trials to success and PAL first trial memory score and RTI five-choice reaction time, the RTI movement time measures, and SWM mean time to last response. For the total sample (Table 11), we observed moderate inter-test correlations among the four CANTAB tests (i.e., between RVP probability of false alarm and PAL mean errors to success; between RVP probability of false alarm, RVP A', and RVP mean latency and PAL mean trials to success, PAL first trial memory score, the RTI reaction time measures, and SWM mean time to last response; between RVP probability of false alarm and RVP mean latency and SWM total errors; between PAL mean trials to success and RTI five-choice reaction time; between PAL mean trials to success and PAL first trial memory score and SWM total errors and SWM mean time to last response; and between RTI five-choice reaction time and SWM mean time to last response).

Both in the WND group (Table 9) and the total sample (Table 11), the correlations between CANTAB tests were weaker than the intra-test correlations observed for RVP (all except those involving RVP mean latency), PAL, RTI, and SWM (all except the correlation between SWM strategy and SWM mean time to last response). In the AD group (Table 10), the CANTAB inter-test correlations were weaker than all of the RVP intra-test correlations (except those involving RVP mean latency), the correlation between PAL mean trials to success and PAL first trial memory score, all of the RTI intra-test correlations, and the correlation between SWM total errors and SWM mean time to last response.

Factorial analysis

Table 12 illustrates the EFAs (pairwise comparison method) performed with the CANTAB measures as variables. Four factors were retained, and these factors collectively explained 68% (WND group) and 75% (total sample) of the total variance in the original variables. Both factor models divided the CANTAB tests into four distinct domains. Thus, the RVP measures displayed high loadings onto a factor of selective attention and processing speed (WND group: factor 4; total sample: factor 3), the PAL measures onto a factor of visuospatial episodic memory and learning (WND group and total sample: factor 2), the RTI measures onto a factor of psychomotor speed (WND group and total sample: factor 1), and the SWM measures onto a factor of visuospatial working memory and planning (WND group: factor 3; total sample: factor 4). However, in the total sample, the SWM mean time to last response measure also loaded onto the visuospatial episodic memory and learning factor (i.e., factor 2), with both factor loadings having been lower than .50.

The EFAs (pairwise comparison method) calculated with the CANTAB and pencil-and-paper test measures for the WND group and the total sample (Table 13) resulted in factor models that, once again, consistently divided the CANTAB tests into four distinct domains.

In the WND group, seven factors were retained, and this factor set explained 69% of the total variance in the original variables. The PAL (CANTAB) and RCFT immediate recall exhibited high loadings onto factor 1 (visuospatial episodic memory and learning); Logical Memory and VPA (WMS-III) onto factor 2 (verbal episodic memory and learning); RTI (CANTAB) onto factor 3 (psychomotor speed); SWM (CANTAB) onto factor 4 (visuospatial working memory and planning); RVP (CANTAB) onto factor 5 (selective attention and processing speed); Digit Span (WAIS-III) onto factor 6 (verbal working memory); and Family Pictures (WMS-III) onto factor 7 (visual episodic memory). The factor loadings of Spatial Span (WMS-III) and Verbal Fluency animals (ACE-R) were lower than .50. For Verbal Fluency animals (ACE-R), the highest loading was onto the selective attention and processing speed factor, and Spatial Span backward (WMS-III) had high loadings onto the working memory factors 4 and 6.

Table 12. Factor loadings from exploratory factor analysis (after varimax rotation) for the Cambridge neuropsychological automated tests.

Test measure	WND group				Total sample			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
RVP probability of hit				-.77				-.84
RVP probability of false alarm				.85				.75
RVP mean latency				.59	.46			.61
PAL mean errors to success		.91				.87		
PAL mean trials to success		.92			.39	.83		
PAL first trial memory score		-.83				-.80	-.31	
RTI simple reaction time	.74				.79			
RTI five-choice reaction time	.76				.83			
RTI simple movement time	.82				.83			
RTI five-choice movement time	.74				.77			
SWM total errors			.84			.31		.74
SWM strategy			.81					.87
SWM mean time to last response			.62			.44		.48
Eigenvalue	4.24	1.87	1.54	1.18	6.11	1.46	1.12	1.03
% of variance explained	32.59	14.37	11.81	9.09	46.97	11.25	8.61	7.96

Note. Loadings < .30 were suppressed. Loadings \geq .50 are presented in bold. The pairwise method was used. *Ns* range: 130-137 (WND group) and 167-174 (total sample). Legend: PAL = Paired Associates Learning; RTI = Reaction Time; RVP = Rapid Visual Information Processing; SWM = Spatial Working Memory; WND = without neuropsychiatric diagnoses.

Table 13. Factor loadings from exploratory factor analysis (after varimax rotation) for the automated and pencil-and-paper tests.

Test measure	Without neuropsychiatric diagnoses (WND) group							Total sample				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
RVP ^a probability of hit					-.62					.77		
RVP ^a probability of false alarm					.78					-.72		
RVP ^a mean latency				.32	.51				.45	-.58		
PAL ^a mean errors to success	-.88							-.72				
PAL ^a mean trials to success	-.86							-.73	.45		.30	
PAL ^a first trial memory score	.77							.73	-.30			
RTI ^a simple reaction time			.74					.73				
RTI ^a five-choice reaction time			.70		.37			.83				
RTI ^a simple movement time			.82					.81				
RTI ^a five-choice movement time			.74					.78				
SWM ^a total errors				.80							.75	
SWM ^a strategy				.74							.76	
SWM ^a mean time to last response				.63				-.39	.30		.52	
Rey Complex Figure Test immediate recall	.58					.30		.72				
Family Pictures ^b immediate recall	.31						.82	.77				
Family Pictures ^b delayed recall	.33						.81	.77	-.30			
Logical Memory ^b immediate recall		.77			-.30			.76				
Logical Memory ^b delayed recall		.80			-.31			.80		.31		
Verbal Paired Associates ^b immediate recall		.80						.73				.41
Verbal Paired Associates ^b delayed recall		.76						.73				.39
Spatial Span ^b forward				-.34		.32	-.39	.38			-.47	
Spatial Span ^b backward				-.45		.45		.34		.33	-.48	.36
Digit Span ^c forward						.80						.77
Digit Span ^c backward						.72						.74
Verbal Fluency ^d animals		.43	-.33		-.48			.62		.44		
Eigenvalue	6.92	2.79	2.05	1.59	1.42	1.34	1.20	11.87	1.96	1.60	1.15	1.04
% of variance explained	27.67	11.14	8.19	6.35	5.68	5.34	4.81	47.46	7.86	6.40	4.61	4.17

Note. Loadings < .30 were suppressed. Loadings ≥ .50 are presented in bold. The pairwise method was used. Ns range: 130-137 (WND group) and 167-174 (total sample). ^aCambridge Neuropsychological Test Automated Battery; ^bWechsler Memory Scale-III. ^cWechsler Adult Intelligence Scale-III. ^dAddenbrooke's Cognitive Examination Revised. Legend: PAL = Paired associates learning; RTI = Reaction time; RVP = Rapid visual information processing; SWM = Spatial working memory.

In the total sample, five factors were retained and collectively explained 71% of the total variance in the initial variables. The PAL (CANTAB), RCFT immediate recall, Family Pictures, Logical Memory, VPA (WMS-III), and Verbal Fluency animals (ACE-R) most highly loaded onto factor 1 (visuospatial and verbal memory); RTI (CANTAB) onto factor 2 (psychomotor speed); RVP (CANTAB) onto factor 3 (selective attention and processing speed); SWM (CANTAB) onto factor 4 (visuospatial working memory and planning); and Digit Span (WAIS-III) onto factor 5 (verbal working memory). The factor loadings of Spatial Span (WMS-III) were lower than .50, but this subtest's highest factor loadings were onto the factor of visuospatial working memory (i.e., factor 4).

DISCUSSION

This study aimed to examine the relationship between the RVP, PAL, RTI, and SWM (CANTAB) tests and several pencil-and-paper tests commonly used in clinical practice (concurrent validity), the relationship between the RVP, PAL, RTI, and SWM (CANTAB) tests, and their latent factorial structure (construct validity). The concurrent and construct validity of these tests were examined in a sample of older Portuguese persons without neuropsychiatric diagnoses and with mild to moderate AD dementia and in a WND group to increase our understanding of the cognitive functions assessed by these tests. Moreover, we examined the intra- and inter-test correlations in the AD group to clarify possible differences between results from the total sample and WND group and to test the consistency of the associations in the clinical group.

Concurrent validity of the CANTAB tests in the group without neuropsychiatric diagnoses

The RVP measures showed the highest inter-test correlations with Verbal Fluency animals (ACE-R) (.38-.49), Spatial Span backward (WMS-III) (.25-.34), and the RTI (CANTAB) measures (.19-.42). Therefore, the RVP test revealed significant correlations with measures of selective attention, processing speed, attention, working memory, and psychomotor speed, suggesting that this test is multidimensional in nature, which is in line with previous studies (Coull et al., 1996; Smith et al., 2013).

For the PAL measures, the strongest inter-test correlations were with the episodic memory measures of RCFT immediate recall (.42-.52) and Family Pictures (WMS-III) (.33-.42). These findings are consistent with those of other concurrent validity studies (e.g., Fowler et al., 1995; Kim et al., 2009; Torgersen et al., 2012) conducted with samples involving healthy and clinical groups. Some of these studies (Fowler et al., 1995; Torgersen et al., 2012) also reported moderate associations between the PAL and verbal episodic memory, attention, and working memory measures, as observed in our total sample. However, in the WND group, the correlations between the PAL and Logical Memory, VPA, and Spatial Span (WMS-III) tests were significant but weak, perhaps due to less heterogeneity of the data.

Regarding the RTI measures, they evidenced the highest inter-test correlations with RVP (CANTAB) (.19-.42) and Verbal Fluency animals (ACE-R) (.33-.37). Thus, the RTI test was most strongly associated with measures involving selective attention, processing speed, and psychomotor speed. Considering that both RTI simple condition and RTI five-choice condition measures correlated with the mentioned selective attention and processing speed measures, the processing speed component may have contributed more to these associations than the selective attention component. Our results corroborated those of a previous study that described a strong association between the RTI test and visual sustained attention and psychomotor speed measures (i.e., Digit Symbol [WAIS-III] and TMT-A; Sexton et al., 2012).

The three SWM measures showed strong inter-test correlations with RVP probability of false alarm (CANTAB) (.33-.39), Spatial Span (WMS-III) (.23-.37), and Digit Span backward (WAIS-III) (.18-.30). The SWM total errors and SWM mean time to last response measures displayed the strongest correlations with Verbal Fluency animals (ACE-R) (.27-.47) and RVP mean latency (CANTAB) (.32-.37). Hence, the SWM test revealed stronger correlations with working memory, selective attention, and processing speed measures. These results are in line with those of a previous study conducted with healthy young adults (Smith et al., 2013), except that these authors reported a significant correlation between SWM strategy and Verbal Fluency animals, and this correlation was nonsignificant in our WND group.

Therefore, the correlations between the CANTAB tests and the pencil-and-paper tests were moderate at most, as expected (e.g., Sexton et al., 2012; Smith et al., 2013).

According to the guidelines of the European Federation of Psychologists' Associations (EFPA) for interpreting correlations in concurrent validity analysis (i.e., inadequate: < .20; adequate: .20-.34; good: .35-.49; excellent: \geq .50; Evers et al., 2013), our results indicated good-to-adequate concordance between the CANTAB tests and several pencil-and-paper tests selected as their counterparts.

Construct validity of the CANTAB tests in the group without neuropsychiatric diagnoses

CANTAB intra-test correlations

The associations between measures of the same test were all significant, but some intra-test correlations were not as high as expected and were not stronger than some inter-test correlations. In particular, the RVP intra-test correlations involving RVP mean latency were weak and weaker than the inter-test correlations between RVP mean latency and other latency measures (e.g., RTI five-choice reaction time and SWM mean time to last response). The correlation between SWM strategy and SWM mean time to last response was also weak and weaker than some SWM inter-test correlations. In line with previous studies analyzing the PAL, RTI, and SWM tests in healthy samples (e.g., Robbins et al., 1998; Singer et al., 2006), all of the remaining CANTAB intra-test correlations were very strong to moderate (RVP: .73; PAL: .64-.93; RTI: .45-.76; SWM: .44-.47) and higher than these tests' correlations with other tests.

Weak intra-test correlations may be problematic from the perspective of a test's convergent validity (i.e., a construct validity subtype for which responses on conceptually similar domains of a test, or on conceptually similar tests, are expected to be strongly correlated; American Psychological Association, 2015). Concerning the RVP test, we can assume that higher response latency may lead to more correctly detected target sequences but also to more responses given outside the response window; this process might explain the weak correlations between RVP mean latency and the other RVP measures. However, for the SWM test, the use of strategy (i.e., starting a new token search following the same sequence of boxes) is expected to decrease the number of search errors and, consequently, the time needed to solve the problem (i.e., the time to last response). In this sense, a stronger intra-test correlation between SWM strategy and SWM mean time to last response should be expected. In our

WND group, most correlations between SWM strategy and the other CANTAB measures were nonsignificant. This result is in contrast to that of Singer et al. (2006), who identified significant (although weak) correlations between SWM strategy and the PAL and RTI reaction time measures. Singer et al. (2006) also included younger participants (range: 18-76 years of age) in their study; thus, these individuals may have used the SWM strategy more than our participants. Indeed, research has shown that individuals aged 75 years or above show a decreased use of the SWM strategy (Robbins et al., 1998). Likewise, our participants' limited use of the SWM strategy might explain why we found moderate to weak intra-test correlations between SWM strategy and the other SWM measures, while Singer et al. (2006) reported a strong intra-test correlation between SWM strategy and SWM between search errors.

Models from the exploratory factor analysis

The EFAs calculated with the CANTAB measures only and with the CANTAB and pencil-and-paper test measures pointed to factor models in which the RVP, PAL, RTI, and SWM measures consistently loaded highly onto separate factors. Therefore, our results suggest that these CANTAB tests are primarily associated with distinct cognitive functions. The RVP measures showed high loadings onto a factor of selective attention and processing speed, which Verbal Fluency animals (ACE-R) loaded onto moderately. The PAL measures loaded highly onto a factor of visuospatial episodic memory and learning along with RCFT immediate recall. Notably, the Family Pictures (WMS-III) measures displayed high loadings onto a distinct factor even though it also assessed visuospatial episodic memory. The Family Pictures (WMS-III) stimuli are easier to name than the abstract stimuli of PAL or RCFT. In fact, the Family Pictures task requires participants to label characters and verbally describe the activities that these characters are engaged in. Thus, the Family Pictures (WMS-III) subtest relies heavily on not only visuospatial skills but also verbal skills, which may explain the results found. The RTI measures had high loadings onto a factor of psychomotor speed, but none of the pencil-and-paper test measures loaded highly onto this factor. Finally, the SWM measures loaded highly onto a visuospatial working memory factor, which the Spatial Span (WMS-III) backward loaded onto moderately, perhaps because the SWM test is more than a measure of visuospatial span and involves additional executive functions such as problem solving and planning.

Previous studies with adult samples of healthy individuals and patients with breast cancer (e.g., Lenehan et al., 2016; Rabbitt & Lowe, 2000; Robbins et al., 1998; Stille et al., 2010) have also described factor solutions in which the RVP, PAL, and SWM tests (or the RTI, PAL, and SWM tests) loaded highly onto separate factors. However, using principal component analysis on a sample of healthy young adults, Smith et al. (2013) reported a solution in which the measures RVP mean latency, PAL total errors (8-boxes problem), and SWM between search errors were combined into a single component; SWM strategy was into a second component; and RVP total false alarms (a measure included in the RVP probability of false alarm) was into a third component. Forcing a four-factor solution (by pre-selecting this number of factors), Lenehan et al. (2016) described a factor including the RVP A' and RTI five-choice reaction time measures. These discrepancies in results may be explained by different sample characteristics, statistical methods, and sets of analyzed variables in the different studies.

Comparative analysis of the concurrent and construct validity of the CANTAB tests on a sample of patients with mild to moderate Alzheimer's disease dementia

Inter-test correlations between the CANTAB tests and between these tests and the pencil-and-paper tests

In general, as expected, the inter-test correlations were higher in the total sample than in the WND and AD groups (Anastasi, 1988). We identified a mismatch between the WND and AD groups regarding the highest inter-test correlations. For example, we found moderate inter-test correlations between two PAL measures (i.e., PAL mean trials to success and PAL first trial memory score) and psychomotor speed measures (i.e., RVP mean latency and most of the RTI measures) in the AD group, while such associations were weak to nonsignificant in the WND group. One explanation for this finding is that the memory deficit or the slower processing speed (which implies longer reaction and movement times) exhibited by the AD group may have interfered with the group's performance on these tests, thereby strengthening the associations between them. Such interference may have had a lesser impact in the WND group due to the group's normal cognitive functioning. Moreover, for the AD group, we found almost no significant correlation between RVP A', PAL mean errors to success (CANTAB), and

Verbal Fluency animals (ACE-R) and other CANTAB or pencil-and-paper test measures. By contrast, in the WND group, these measures exhibited several significant inter-test correlations (especially Verbal Fluency animals [ACE-R]). Both RVP A' (CANTAB) and Verbal Fluency animals (ACE-R) have shown high sensitivity to MCI and mild AD dementia (Gomez & White, 2006; Saunders & Summers, 2010); therefore, poor scores or sparse score variability (due to impairment in one or more of the cognitive functions involved in these tasks) may have resulted in the nonsignificant inter-test correlations in the AD group.

CANTAB intra-test correlations

When compared with the WND group, the total sample revealed several weaker CANTAB intra-test correlations that were explained by the nonsignificant intra-test correlations found in the AD group, namely, the correlations between RVP probability of hit and RVP probability of false alarm, between PAL mean errors to success and PAL mean trials to success, between SWM strategy and SWM total errors, and between SWM strategy and SWM mean time to last response.

Some possible explanations for the nonsignificant correlation between RVP probability of hit and RVP probability of false alarm are as follows: the participants in the AD group may have responded arbitrarily to stimuli, may have been perseverative in pressing the button, or may have been contained in responding due to their slower processing speed.

The nonsignificant association between PAL mean errors to success and PAL mean trials to success may be explained by the small number of stages completed by most of the participants in the AD group (65% of whom completed less than six stages) in contrast to the large number of stages completed by most participants in the WND group (97% completed six or more stages). A small number of stages completed indicates that a small range of possible errors were made because the PAL test involves the presentation of an increasing number of patterns (i.e., one pattern in stages 1 and 2, two patterns in stages 3 and 4, three patterns in stages 5 and 6, six patterns in stage 7, and eight patterns in stage 8), with each participant always having a maximum of ten trials per stage. Thus, because the AD group completed few stages, a small range of errors may have weakened the association with PAL mean trials to success.

The SWM strategy measure is calculated for problems with six and eight boxes based on the number of times the participant started a new token search by touching on a different box, with lower scores suggesting better strategy use (i.e., indicating that the participant employed the strategy of beginning a new search by the same box following a predetermined sequence of boxes). All of the nonsignificant SWM intra-test correlations found in the AD group involved SWM strategy. Therefore, some of the participants in the AD group may have started new token searches by touching on the same box that they previously touched due to perseveration rather than due to SWM strategy use. Consequently, these participants would reach better SWM strategy scores than patients without perseveration and with poor strategy use, even though they all committed a high number of search errors. This process would thereby disrupt the linearity of the association between SWM strategy and the other two SWM measures, weakening the correlations.

Models from the exploratory factor analysis

Despite the discrepancies in the results of the WND and AD groups, the EFAs calculated for the total sample corroborated the division of the RVP, PAL, RTI, and SWM measures into four separated factors, as observed in the WND group. However, the EFA calculated with the CANTAB and pencil-and-paper test measures resulted in five factors in the total sample, in contrast to the seven factors identified in the WND group. Both of these models included the psychomotor speed factor (including the RTI [CANTAB] measures); the visuospatial working memory and planning factor (incorporating the SWM [CANTAB] measures); the selective attention and processing speed factor (encompassing the RVP [CANTAB] measures), and the verbal working memory factor (grouping the Digit Span [WAIS-III] measures). The other three factors identified in the WND group (i.e., visuospatial episodic memory and learning, verbal episodic memory and learning, and visual episodic memory) were combined into a single factor of visuospatial and verbal memory in the total sample. This factor also included the Verbal Fluency animals (ACE-R) measure, possibly due to the semantic component that it shares with the other verbal measures. The AD group exhibited episodic memory deficits, and these deficits may have contributed to the four memory factors (identified in the WND group) loading onto a single factor (in total sample). In the WND group and in the total sample, the highest factor loadings of Spatial Span

(WMS-III) were moderate; however, in total sample, the highest loadings were found mainly onto the visuospatial working memory factor (which included the SWM measures).

In summary, our findings revealed consistency regarding the factors that the RVP, PAL, RTI, and SWM tests showed high loadings onto in both the WND group and the total sample. Our results suggest that these tests assess distinct cognitive functions, namely, selective attention and processing speed (RVP), visuospatial episodic memory and learning (PAL), psychomotor speed (RTI), and visuospatial working memory and planning (SWM). The very strong to moderate intra-test correlations found indicate good-to-adequate convergent validity. However, we also identified weak intra-test correlations involving RVP mean latency and SWM mean time to last response, with these measures showing stronger inter-test correlations with other psychomotor speed measures (e.g., RTI measures). In addition, weak but significant inter-test correlations between PAL and some RVP, RTI, and SWM measures were also observed, even in the WND group. Thus, our results indicate that the four CANTAB tests may share additional cognitive functions even though they assess distinct cognitive functions.

Study limitations, strengths, and future directions

One limitation of this study is that it did not include tests that specifically assess continuous performance and psychomotor speed, which may have improved the concurrent validity analysis of the RVP and RTI tests. We selected the neuropsychological tests that are most commonly used in Portugal and that have Portuguese norms available for older persons. Unfortunately, most of the reaction time tests used in Portugal have norms only for younger adults. Some continuous performance tests that are similar to the RVP test have not yet been studied in Portugal. Another limitation is the small size of the AD group, which compelled us to perform the EFAs for the total sample. A larger AD group size would have allowed a simpler comparative analysis and would have avoided the problem of the results of the AD group being attenuated by those of the WND group in the total sample.

The majority of the construct and concurrent validity studies on the CANTAB tests were conducted with healthy samples or simply with heterogeneous samples including individuals with clinical conditions (e.g., Fowler et al., 1995; Robbins et al.,

1998; Smith et al., 2013). The original contribution of our study is the analysis of the construct and concurrent validity of the CANTAB tests in a heterogeneous sample including a WND group and an AD group as well as the comparative analysis between the two groups. Overall, our findings suggest that the RVP, PAL, RTI, and SWM tests are basically measures of different cognitive functions, although they share additional functions. Our results also indicated good-to-adequate concordance between automated and pencil-and-paper tests assessing similar functions. However, some correlations were weak, albeit significant. Therefore, the correlations were not strong enough to confidently replace the specific pencil-and-paper test with its automated counterpart. These weak inter-test correlations do not necessarily signify that such tests do not measure the same cognitive function. Performance on the CANTAB tests has been studied using neuroimaging techniques that provide knowledge about the cerebral areas of activation and, therefore, about the cognitive functions that may be involved during task performance (e.g., Coull et al., 1996). This is an advantage that the CANTAB tests have over many pencil-and-paper tests, as such tests are not frequently used in this type of study.

Concerning future directions, more studies including instruments of continuous performance and reaction time are needed to better analyze the concurrent validity of the RVP and RTI tests. Concurrent validity studies comparing the CANTAB tests with other automated tests relying on the same sensory modalities for stimuli presentation and responding may also be useful in testing whether their associations are stronger than those between tests based on different sensory modalities. Discriminating power studies of these CANTAB and pencil-and-paper tests in Portuguese older persons are also needed to increase our understanding of the clinical relevance of these tests and to provide further orientation regarding the best tests to employ based on the clinical or assessment aims.

We hope that this study has increased our knowledge of the cognitive functions assessed by the four CANTAB tests most recommended for the assessment of AD dementia and how they relate to each other and to pencil-and-paper tests assessing similar functions to promote a substantiated use and interpretation of these instruments.

ACKNOWLEDGEMENTS

The authors would like to thank all of the principals and professionals at the institutions who gave permission to conduct this study, aided in the recruitment of older volunteers and kindly collaborated in this research by completing interviews for formal caregivers: Casa de Repouso da Enfermagem Portuguesa [Residential Home of the Portuguese Nursing], Casa de Saúde do Telhal do Instituto S. João de Deus [Health Home of the Institute of St. João de Deus], Casa dos Professores de Carcavelos da Associação de Solidariedade Social dos Professores [Carcavelos House of Teachers of Teachers' Social Solidarity Association], Centro de Apoio Social de Oeiras do Instituto de Acção Social das Forças Armadas [Oeiras Social Support Center of the Armed Forces' Social Welfare Institute], Centro de Dia Oeiras São Julião [Oeiras St. Julião Adult Day Care Center], Centro Psicogeriátrico Nossa Senhora de Fátima das Irmãs Hospitaleiras do Sagrado Coração de Jesus [Our Lady of Fátima Psychogeriatric Center of the Hospitaller Sisters of the Sacred Heart of Jesus], Centro Social Paroquial de Oeiras [Oeiras Parochial Social Center], Corpuscare Casa de Repouso [Corpuscare Residential Home], Inválidos do Comércio Residências e Lares [Invalids of Commerce Residences and Nursing Homes], José de Mello Residências e Serviços [José de Mello Residences and Services], and Residência Assistida da Delegação da Costa do Estoril da Cruz Vermelha Portuguesa [Costa do Estoril Residential Home of the Portuguese Red Cross]. The authors would also like to thank all of the older volunteers who kindly participated in this study and without whom this research could not have been conducted.

DISCLOSURE STATEMENT

All of the authors declare that they have no financial interest or benefit arising from the direct applications of this research and, thus, that there are no conflicts of interest in this study.

FUNDING

The first author was supported by the Fundação para a Ciência e a Tecnologia [Portuguese Foundation for Science and Technology] and financed by the European

Social Fund and national funds of the Portuguese Ministry of Science, Technology and Higher Education, under Grant number SFRH/BD/75281/2010.

REFERENCES

- Alves, L., Simões, M. R., & Martins, C. (2012). The estimation of premorbid intelligence levels among Portuguese speakers: The Irregular Word Reading Test (TeLPI). *Archives of Clinical Neuropsychology*, 27, 58-68. doi: 10.1093/arclin/acr103
- American Psychological Association. (2015). *APA dictionary of psychology* (2nd ed.). Washington, DC: Author.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan Publishing Company.
- Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de Depressão Geriátrica (GDS) [Geriatric Depression Scale (GDS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (2nd ed., pp. 69-72). Lisbon, Portugal: Novartis.
- Cambridge Cognition. (2006a). *CANTABeclipse™ version 3: Software user guide*. Cambridge, UK: Author.
- Cambridge Cognition. (2006b). *CANTABeclipse™ version 3: Test administration guide*. Cambridge, UK: Author.
- Chase, H. W., Clark, L., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2008). Dissociable roles of prefrontal subregions in self-ordered working memory performance. *Neuropsychologia*, 46, 2650-2661. doi: 10.1016/j.neuropsychologia.2008.04.021
- Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, 55, 1630-1635. doi: 10.1111/j.1532-5415.2007.01379.x
- Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia*, 34, 1085-1095. doi: 10.1016/0028-3932(96)00029-2
- Dancey, C. P., & Reidy, J. (2004). *Statistics without maths for psychology: Using SPSS for Windows* (3rd ed.). Essex, England: Prentice Hall.
- de Rover, M., Pironti, V. A., McCabe, J. A., Acosta-Cabronero, J., Arana, F. S., Morein-Zamir, S., ... Sahakian, B. J. (2011). Hippocampal dysfunction in patients with mild cognitive impairment: A functional neuroimaging study of a visuospatial paired associates learning task. *Neuropsychologia*, 49, 2060-2070. doi: 10.1016/j.neuropsychologia.2011.03.037
- Elderkin-Thompson, V., Kumar, A., Bilker, W., Dunkin, J., Mintz, J., Moberg, P., ... Gur, R. (2003). Neuropsychological deficits among patients with late-onset minor and major depression. *Archives of Clinical Neuropsychology*, 18, 529-549. doi: 10.1016/s0887-6177(03)00022-2
- Elwood, R. W. (2001). MicroCog: Assessment of cognitive functioning. *Neuropsychology Review*, 11, 89-100. doi: 10.1023/a:1016671201211

- Espírito-Santo, H., Lemos, L., Ventura, L., Moitinho, S., Pinto, A. L., Rodrigues, F., ... Daniel, F. (2015). Teste da Figura Complexa de Rey-Osterrieth-A [Rey-Osterrieth A-Complex Figure Test]. In M. R. Simões & I. Santana (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (3rd ed., pp. 100-107). Porto Salvo, Portugal: Novartis.
- Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J., & Sjöberg, A. (2013). *EFPA review model for the description and evaluation of psychological and educational tests: Test review form and notes for reviewers - Version 4.2.6*. Brussels, Belgium: European Federation of Psychologists' Associations (EFPA). Retrieved from <http://www.efpa.eu/professional-development/assessment>.
- Firmino, H., Simões, M. R., Pinho, M. S., Cerejeira, L., & Martins, C. (2008). *O Exame Cognitivo de Addenbrooke – Revisto* [The Addenbrooke's Cognitive Examination – Revised]. Coimbra, Portugal: Serviço de Avaliação Psicológica da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State. *Journal of Psychiatric Research*, *12*, 189-198. doi: 10.1016/0022-3956(75)90026-6
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. *Applied Neuropsychology*, *2*, 72-78. doi: 10.1207/s15324826an0202_4
- Gomez, R., & White, D. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology*, *21*, 771-775. doi: 10.1016/j.acn.2006.06.012
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2012). *Portuguese norms of the Semantic Verbal Fluency Test and Naming Objects and Fingers Task for older persons living in retirement homes*. Unpublished manuscript.
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016a). Effects of socio-demographic variables on performance on the Cambridge neuropsychological automated tests for the assessment of dementia and Portuguese norms for older adults living in retirement homes. *The Clinical Neuropsychologist*, *30*, 284-317. doi: 10.1080/13854046.2016.1156745
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016b). Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, *23*, 251-263. doi: 10.1080/23279095.2015.1053889
- Guerreiro, M., Fonseca, S., Barreto, J., & Garcia, C. (2008). Escala de Avaliação da Doença de Alzheimer (ADAS) [Alzheimer's Disease Assessment Scale (ADAS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (2nd ed., pp. 41-68). Lisbon, Portugal: Novartis.
- Instituto Nacional de Estatística. (2008). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2008* [Survey on the use of information and communication technologies by households 2008]. Lisbon, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Retrieved from http://www.anacom.pt/streaming/inq_familias_ineumic2008.pdf?contentId=726858&field=ATTACHED_FILE.
- Instituto Nacional de Estatística. (2015). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2015* [Survey on the use of

information and communication technologies by households 2015]. Lisbon, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Retrieved from http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUE_Sdest_boui=224732374&DESTAQUESmodo=2.

- Khan, U. A., Liu, L., Provenzano, F. A., Berman, D. E., Profaci, C. P., Sloan, R., ... Small, S. A. (2014). Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nature Neuroscience*, *17*, 304-311. doi: 10.1038/nn.3606
- Kim, C., Lee, J. Y., Ha, T. H., Choi, J. H., Yu, E.-M., Lee, S. J., ... Park, M. Y. (2009). The usefulness of the Cambridge Neuropsychological Test Automated Battery (CANTAB) for assessing cognitive functions in the elderly: A pilot study. *Journal of the Korean Geriatrics Society*, *13*, 69-78. doi: 10.4235/jkgs.2009.13.2.69
- Kløve, H. (1963). Clinical neuropsychology. In F. M. Forster (Ed.), *The medical clinics of North America* (pp. 1647-1658). New York, NY: Saunders.
- Leitão, O. R., Nina, A., & Monteiro, I. (2008). Escala de Deterioração Global (GDS) [Global Deterioration Scale (GDS)]. In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* [Scales and tests in dementia] (2nd ed., pp. 12-16). Lisbon, Portugal: Novartis.
- Lenahan, M. E., Summers, M. J., Saunders, N. L., Summers, J. J., & Vickers, J. C. (2016). Does the Cambridge Automated Neuropsychological Test Battery (CANTAB) distinguish between cognitive domains in healthy older adults? *Assessment*, *23*, 163-172. doi: 10.1177/1073191115581474
- Lewis, R., & Rennick, P. M. (1979). *Manual for the Repeatable Cognitive-Perceptual-Motor Battery*. Grosse Point, MI: Axon.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, *125*, 624-639. doi: 10.1093/brain/awf049
- Manias, E., Kusljic, S., Berry, C., Brown, E., Bryce, E., Cliffe, J., & Smykowsky, A. (2015). Use of the Screening Tool of Older Person's Prescriptions (STOPP) in older people admitted to an Australian hospital. *Australasian Journal on Ageing*, *34*, 15-20. doi: 10.1111/ajag.12054
- Marôco, J. (2014). *Análise estatística com o SPSS statistics* [Statistical analysis with the SPSS statistics] (6th ed.). Pêro Pinheiro, Portugal: Report Number - Análise e Gestão de Informação, Lda.
- McGaughy, J., Dalley, J. W., Morrison, C. H., Everitt, B. J., & Robbins, T. W. (2002). Selective behavioral and neurochemical effects of cholinergic lesions produced by intrabasal infusions of 192 IgG-saporin on attentional performance in a five-choice serial reaction time task. *The Journal of Neuroscience*, *22*, 1905-1913.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., ... Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*, 263-269. doi: 10.1016/j.jalz.2011.03.005
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, *21*, 1078-1085. doi: 10.1002/gps.1610

- Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin*, *19*, 448-450.
- Morgado, J., Rocha, C. S., Maruta, C., Guerreiro, M., & Martins, I. P. (2009). Novos valores normativos do Mini-Mental State Examination [New normative values of Mini-Mental State Examination]. *Sinapse*, *2*, 10-16.
- Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1987). Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer-type and Parkinson's disease. In S. M. Stahl, S. D. Iversen & E. C. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21-36). Oxford, UK: Oxford University Press.
- National Institute for Health and Care Excellence. (2011). *Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease*. (NICE Technology Appraisal Guidance No. 217). London, UK: National Institute for Health and Care Excellence (NICE). Retrieved from <http://www.nice.org.uk/guidance/ta217/resources/guidance-donepezil-galantamine-rivastigmine-and-memantine-for-the-treatment-of-alzheimers-disease-pdf>.
- Oosterman, J. M., Vogels, R. L. C., van Harten, B., Gouw, A. A., Scheltens, P., Poggesi, A., ... Scherder, E. J. A. (2008). The role of white matter hyperintensities and medial temporal lobe atrophy in age-related executive dysfunctioning. *Brain and Cognition*, *68*, 128-133. doi: 10.1016/j.bandc.2008.03.006
- Owen, A. M., Evans, A. C., & Petrides, m. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: A positron emission tomography study. *Cerebral Cortex*, *6*, 31-38. doi: 10.1093/cercor/6.1.31
- Periquito, C. M. N., Silva, P. C., Oliveira, P., Carneiro, C., Fernandes, A. I., & Costa, F. A. (2014). Revisão da medicação em idosos institucionalizados: Aplicação dos critérios STOPP e START. [Review of medication in institutionalized elderly: Application of STOPP and START criteria]. *Revista Portuguesa de Farmacoterapia*, *6*, 211-220.
- Rabbitt, P., & Lowe, C. (2000). Patterns of cognitive ageing. *Psychological Research*, *63*, 308-316. doi: 10.1007/s004269900009
- Reeves, S. J., Grasby, P. M., Howard, R. J., Bantick, R. A., Asselin, M.-C., & Mehta, M. A. (2005). A positron emission tomography (PET) investigation of the role of striatal dopamine (D2) receptor availability in spatial cognition. *Neuroimage*, *28*, 216-226. doi: 10.1016/j.neuroimage.2005.05.034
- Reisberg, B., Ferris, S. H., de Leon, M. J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry*, *139*, 1136-1139.
- Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271-276. doi: 10.2466/pms.1958.8.3.271
- Rey, A. (2002). *Teste de Cópia de Figuras Complexas* [Complex Figure Copying Test] (2nd ed.). Lisbon, Portugal: Cegoc-Tea.
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging.

- Journal of the International Neuropsychological Society*, 4, 474-490. doi: 10.1017/s1355617798455073
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): A factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitive Disorders*, 5, 266-281. doi: 10.1159/000106735
- Roiser, J. P., McLean, A., Ogilvie, A. D., Blackwell, A. D., Bamber, D. J., Goodyer, I., ... Sahakian, B. J. (2005). The subjective and cognitive effects of acute phenylalanine and tyrosine depletion in patients recovered from depression. *Neuropsychopharmacology*, 30, 775-785. doi: 10.1038/sj.npp.1300659
- Saunders, N. L. J., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 32, 350-357. doi: 10.1080/13803390903042379
- Sexton, C. E., McDermott, L., Kalu, U. G., Herrmann, L. L., Bradley, K. M., Allan, C. L., ... Ebmeier, K. P. (2012). Exploring the pattern and neural correlates of neuropsychological impairment in late-life depression. *Psychological Medicine*, 42, 1195-1202. doi: 10.1017/s0033291711002352
- Singer, J. J., MacGregor, A. J., Cherkas, L. F., & Spector, T. D. (2006). Genetic influences on cognitive function using the Cambridge Neuropsychological Test Automated Battery. *Intelligence*, 34, 421-428. doi: 10.1016/j.intell.2005.11.005
- Smith, P. J., Need, A. C., Cirulli, E. T., Chiba-Falek, O., & Attix, D. K. (2013). A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with "traditional" neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology*, 35, 319-328. doi: 10.1080/13803395.2013.771618
- Sousa, L. B., Prieto, G., Vilar, M., Firmino, H., & Simões, M. R. (2015). The Adults and Older Adults Functional Assessment Inventory: A Rasch model analysis. *Research on Aging*, 37, 787-814. doi: 10.1177/0164027514564469
- Stilley, C. S., Bender, C. M., Dunbar-Jacob, J., Sereika, S., & Ryan, C. M. (2010). The impact of cognitive function on medication management: Three studies. *Health Psychology*, 29, 50-55. doi: 10.1037/a0016940
- Stoffers, D., Bosboom, J. L. W., Deijen, J. B., Wolters, E. C., Berendse, H. W., & Stam, C. J. (2007). Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia. *Brain*, 130, 1847-1860. doi: 10.1093/brain/awm034
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662. doi: 10.1037/h0054651
- Summers, M. J., & Saunders, N. J. (2012). Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. *Neuropsychology*, 26, 498-508. doi: 10.1037/a0028576
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Torgersen, J., Flaatten, H., Engelsen, B. A., & Gramstad, A. (2012). Clinical validation of Cambridge Neuropsychological Test Automated Battery in a Norwegian epilepsy population. *Journal of Behavioral and Brain Science*, 2, 108-116. doi: 10.4236/jbbs.2012.21013
- Ubeda, A., Ferrándiz, M. L., Maicas, N., Gomez, C., Bonte, M., & Peris, J. E. (2012). Potentially inappropriate prescribing in institutionalised older patients in Spain:

- The STOPP-START criteria compared with the Beers criteria. *Pharmacy Practice*, 10, 83-91.
- Wechsler, D. (1987). *Wechsler Memory Scale - Revised manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2008a). *Escala de Inteligência de Wechsler para Adultos – Terceira edição (WAIS-III): Manual técnico* [Wechsler Adult Intelligence Scale - Third edition (WAIS-III): Technical manual]. Lisbon, Portugal: Cegoc-Tea.
- Wechsler, D. (2008b). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual de administração e cotação* [Wechsler Memory Scale - Third edition (WMS-III): Administration and scoring manual]. Lisbon, Portugal: Cegoc-Tea.
- Wechsler, D. (2008c). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual técnico* [Wechsler Memory Scale - Third edition (WMS-III): Technical manual]. Lisbon, Portugal: Cegoc-Tea.
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37-49. doi: 10.1016/0022-3956(82)90033-4

DISCUSSÃO GLOBAL

SÍNTESE E REFLEXÃO INTEGRATIVA DOS RESULTADOS

Com este projecto de investigação procurou-se responder a questões como as seguintes: (a) qual a fiabilidade teste-reteste dos testes computadorizados RVP, PAL, RTI e SWM da CANTAB (Cambridge Cognition, 2006a, 2006b)?, (b) qual a sua susceptibilidade aos efeitos da prática, de chão e de tecto? (capítulo 1), (c) qual a influência das variáveis sócio-demográficas idade, escolaridade, género, experiência com computadores, consumo de psicotrópicos e tempo de institucionalização no desempenho nestes testes? (capítulo 2), (d) o que sugerem os resultados em termos dos constructos avaliados por estas provas? e (e) qual a relação dos resultados entre estes testes e com os obtidos em provas de papel-e-lápis que avaliam funções semelhantes? (capítulo 3).

Em termos de fiabilidade teste-reteste (Gonçalves, Pinho, & Simões, 2016a), seis das onze medidas analisadas apresentaram correlações teste-reteste altas a adequadas (i.e., .71-.89; classificação de Strauss, Sherman, & Spreen, 2006) em 34 pessoas mais velhas sem problemas neuropsiquiátricos, a residirem em instituições geriátricas e avaliadas com um intervalo de quatro semanas entre as sessões de teste e reteste. Essas medidas foram RVP A', RVP latência média, PAL média de erros até ao sucesso, RTI tempo de movimento simples, RTI tempo de movimento escolha-cinco e SWM total de erros. Contudo, também foram identificadas medidas com correlações marginais a baixas (i.e., .56-.69; PAL total de erros ajustado, PAL pontuação de memória no primeiro ensaio, RTI tempo de reacção simples, RTI tempo de reacção escolha-cinco e SWM estratégia). O intervalo de correlações teste-reteste encontrado foi semelhante ao observado para outros testes computadorizados e de papel-e-lápis (e.g., Calamia, Markon, & Tranel, 2013; Gualtieri & Johnson, 2006; Schweiger, Doniger, Dwolatzky, Jaffe, & Simon, 2003; Tornatore, Hill, Laboff, & McGann, 2005; Wechsler, 2008). Das onze medidas de teste analisadas, uma evidenciou um efeito de tecto (i.e., RVP A'), nenhuma apresentou efeitos de chão e só três revelaram susceptibilidade aos efeitos da prática (i.e., PAL total de erros ajustado, PAL média de erros até ao sucesso e SWM total de erros). Estes efeitos da prática registaram pequena magnitude ($|d| \leq .49$) e foram corrigidos de modo mais eficaz pelos métodos de estimação de equações baseadas na regressão padronizada de Crawford e Garthwaite (2007) e de Duff (2012) do que pelo

método de correcção por ajustamento fixo proposto por Iverson (2001). Os efeitos da prática observados foram também menores do que os reportados para testes com uma única forma (Duff, 2014).

Relativamente à influência das variáveis sócio-demográficas (Gonçalves, Pinho, & Simões, 2016b), analisada numa amostra de 128 pessoas mais velhas sem problemas neuropsiquiátricos e a viverem em instituições geriátricas, verificou-se que a escolaridade e o tempo de institucionalização não evidenciaram efeitos significativos no desempenho nas medidas da CANTAB consideradas (i.e., as medidas que revelaram boa a aceitável fiabilidade teste-reteste; Gonçalves et al., 2016a). A idade influenciou significativamente o desempenho em RVP A' e SWM total de erros, tendo sido observado um efeito marginalmente significativo em RVP latência média. Os participantes mais novos foram os que apresentaram melhores resultados. O género influenciou significativamente as pontuações em RVP latência média e SWM total de erros (tendo os homens obtido melhores resultados) e PAL média de erros até ao sucesso (tendo as mulheres revelado melhor desempenho). A experiência com computadores exerceu um efeito marginalmente significativo apenas em RTI tempo de movimento simples. O consumo de psicotrópicos influenciou significativamente o desempenho em RVP A' e PAL média de erros até ao sucesso e revelou um efeito marginalmente significativo em RTI tempo de movimento simples. Nenhuma das variáveis sócio-demográficas influenciou o desempenho em RTI tempo de movimento escolha-cinco. Por conseguinte, diferentes variáveis sócio-demográficas apresentaram efeitos significativos em diferentes testes e medidas de um mesmo teste.

Contrariamente ao observado para as medidas da CANTAB, a escolaridade apresentou um efeito significativo na maior parte das medidas dos testes de papel-e-lápis examinados (Gonçalves et al., 2016b). A escolaridade e o género revelaram maior influência no desempenho nos testes de papel-e-lápis do que nos testes da CANTAB. Curiosamente, a experiência com computadores apresentou um efeito significativo a marginalmente significativo em várias medidas dos testes de papel-e-lápis (i.e., Fluência Verbal animais [ACE-R], Listas de Palavras evocação imediata e diferida [WMS-III] e Memória de Dígitos sentido directo [WAIS-III]), tendo os utilizadores de computador obtido melhores resultados nestes testes de cariz verbal. Quanto ao consumo de psicotrópicos, foram mais as medidas da CANTAB sensíveis aos efeitos desta variável

(i.e., RVP A', PAL média de erros até ao sucesso e RTI tempo de movimento simples) do que os testes de papel-e-lápis (Fluência Verbal animais [ACE-R]).

No estudo de validade de constructo e validade concorrente dos testes da CANTAB (Gonçalves, Pinho, & Simões, 2016c) foram examinadas algumas medidas que ainda não tinham sido consideradas nos estudos anteriores (i.e., Gonçalves et al., 2016a, 2016b). O objectivo foi seleccionar, para cada teste, um mínimo de três medidas o mais independentes umas das outras, de modo a proceder correctamente à análise factorial. Deste modo, a medida RVP A' foi decomposta nas duas medidas que entram no seu cálculo (i.e., RVP probabilidade de sucesso e RVP probabilidade de falso alarme) e analisadas a par de RVP latência média. Como o cálculo das medidas PAL total de erros ajustado e PAL média de erros até ao sucesso envolve a medida PAL total de erros, seleccionaram-se para este estudo as medidas PAL média de erros até ao sucesso, PAL média de ensaios até ao sucesso e PAL pontuação de memória no primeiro ensaio. Para o teste RTI, consideraram-se as mesmas medidas que no estudo da fiabilidade teste-reteste (Gonçalves et al., 2016a). Quanto ao teste SWM, analisaram-se as medidas SWM total de erros, SWM estratégia e SWM média de tempo até à última resposta. Como o estudo da validade de constructo e da validade concorrente envolveu medidas da CANTAB anteriormente não consideradas e foi realizado com 137 pessoas mais velhas sem problemas neuropsiquiátricos (grupo SPN) e 37 pessoas com demência ligeira a moderada devido a DA (grupo DA), a residir em instituições geriátricas e na comunidade, decidiu-se examinar novamente os efeitos de chão e de tecto. No grupo SPN detectaram-se efeitos de tecto em RVP A' (à semelhança dos estudos anteriores; Gonçalves et al., 2016a, 2016b) e em RVP probabilidade de sucesso e um efeito de chão em RVP probabilidade de falso alarme. No grupo DA identificou-se um efeito de tecto em RVP probabilidade de sucesso.

Em termos de validade concorrente, examinada pela relação entre os resultados nos testes da CANTAB com os resultados em outros testes de funções cognitivas semelhantes, observou-se que as correlações mais fortes foram no máximo moderadas, no grupo SPN (Gonçalves et al., 2016c). As medidas do teste RVP correlacionaram-se mais fortemente (.19-.49) com medidas envolvendo atenção selectiva, memória de trabalho e velocidade psicomotora (i.e., Fluência Verbal animais [ACE-R], Localização Espacial sentido inverso [WMS-III] e medidas do RTI [CANTAB]). As medidas do

PAL apresentaram associação mais forte (.33-.52) com medidas de memória episódica (i.e., evocação imediata do Teste da Figura Complexa de Rey e evocação imediata e diferida de Cenas de Família [WMS-III]). Por sua vez, as correlações mais fortes das medidas do RTI (.19-.42) foram com medidas envolvendo velocidade no processamento da informação (i.e., Fluência Verbal animais [ACE-R] e medidas do RVP [CANTAB]). As medidas do SWM (.18-.47) revelaram associação mais forte com medidas de memória de trabalho, atenção selectiva e velocidade de processamento (i.e., RVP probabilidade de falso alarme, RVP latência média [CANTAB], Localização Espacial sentido directo e inverso [WMS-III], Memória de Dígitos sentido inverso [WAIS-III] e Fluência Verbal animais [ACE-R]). Segundo as normas da Federação Europeia de Associações de Psicólogos (Evers et al., 2013), estes resultados sugerem uma concordância boa a adequada entre os testes mencionados.

Quanto à validade de constructo dos testes da CANTAB, as correlações intra-teste observadas no grupo SPN foram, no geral, fortes a moderadas (.44-.93) e mais fortes do que as correlações entre testes (Gonçalves et al., 2016c), deste modo indicando boa a adequada validade convergente. Exceptuam-se as correlações intra-teste envolvendo a medida RVP latência média, a qual apresentou correlações mais fortes com medidas de latência de outros testes (e.g., RTI e SWM), e a correlação entre SWM estratégia e SWM média de tempo até à última resposta, que foi mais fraca do que algumas correlações destas medidas com outros testes. As soluções da análise factorial realizada com as medidas da CANTAB e com estas medidas e as dos testes de papel-e-lápis foram consistentes ao dividirem os testes RVP, PAL, RTI e SWM em factores distintos, tanto no grupo SPN como na amostra total. Assim, as medidas do teste RVP saturaram mais fortemente no factor “atenção selectiva e velocidade de processamento”, as medidas do PAL no factor “memória episódica” (visuo-espacial), as do RTI no factor “velocidade psicomotora” e as do SWM no factor “memória de trabalho visuo-espacial e planeamento”. Por conseguinte, estes resultados sugerem que os quatro testes da CANTAB avaliam diferentes funções cognitivas embora também envolvam funções cognitivas semelhantes, atendendo às correlações significativas encontradas entre algumas medidas destes testes. Um exemplo dessas funções comuns é a latência de resposta ou velocidade de processamento, avaliada por medidas específicas dos testes

RVP (latência média), RTI (tempo de reacção e de movimento) e SWM (média de tempo até à última resposta).

Em relação à fiabilidade teste-reteste, só se conheciam os valores correlacionais para algumas medidas dos testes PAL e SWM (Fowler, Saling, Conway, Semple, & Louis, 1995; Lowe & Rabbitt, 1998). Devido ao número elevado de medidas disponíveis com o *software* CANTABeclipse® 3.0 (Cambridge Cognition, 2006a, 2006b) para cada um dos testes estudados, foi necessário seleccionar as medidas a analisar por teste. No estudo da fiabilidade teste-reteste (Gonçalves et al., 2016a) seleccionaram-se as descritas na literatura como apresentando elevado poder discriminante e validade preditiva (e.g., Ahmed, Mitchell, Arnold, Nestor, & Hodges, 2008; Blackwell et al., 2004; Fowler, Saling, Conway, Semple, & Louis, 1997, 2002; Mitchell, Arnold, Dawson, Nestor, & Hodges, 2009) ou as mais relacionadas com as funções cognitivas avaliadas pelos testes (Cambridge Cognition, 2006a, 2006b). Apesar das correlações teste-reteste de algumas medidas não terem, por essa razão, sido apresentadas, é de referir que foram encontradas mais medidas com fiabilidade teste-reteste boa a aceitável e baixa a inaceitável. Atendendo a que os testes da CANTAB para avaliação da DA são apresentados em formas alternativas e que as correlações teste-reteste foram calculadas para uma pequena amostra (aspectos que poderão contribuir para diminuir a força das correlações), considera-se que os resultados em termos de consistência temporal das pontuações nestas provas são prometedores quanto à sua utilização para fins de monitorização.

Os efeitos da prática encontrados para os testes PAL e SWM (Gonçalves et al., 2016a) já tinham sido reportados anteriormente (Fowler et al., 1995; Lowe & Rabbitt, 1998). O desempenho nestes testes tende a melhorar com a utilização de estratégias podendo esta ser uma justificação para a maior susceptibilidade destes testes aos efeitos da prática (Lowe & Rabbitt, 1998). Atendendo a que não foi identificado um efeito da prática na medida SWM estratégia, sugeriu-se a possibilidade dos participantes terem utilizado outras estratégias (Gonçalves et al., 2016a). No estudo da validade concorrente e convergente (Gonçalves et al., 2016c) a medida SWM estratégia evidenciou correlações mais fracas do que as referidas em outros estudos com participantes mais jovens (Robbins et al., 1998; Singer, MacGregor, Cherkas, & Spector, 2006). Efectivamente, tem sido demonstrado que as pessoas mais velhas tendem a utilizar

menos a estratégia de seguir um mesmo percurso de caixas na procura de fichas no teste SWM (ou seja, tendem a exibir pior desempenho na medida SWM estratégia) do que os adultos mais jovens (De Luca et al., 2003; Robbins et al., 1998). Tal como esta estratégia tende a ser menos utilizada por parte dos adultos mais velhos, é possível admitir que outras estratégias também sejam menos empregues (Lemaire, 2010). Note-se que os efeitos da prática encontrados nos testes PAL e SWM foram de pequena magnitude (Gonçalves et al., 2016a). Por conseguinte, é possível que uma menor utilização de estratégias por parte dos participantes tenha contribuído para efeitos da prática de pequena magnitude.

No estudo da fiabilidade teste-reteste (Gonçalves et al., 2016a) foram também comparadas as diferenças de resultados teste-reteste por género e as diferenças entre géneros por teste. Tendo sido encontradas diferenças teste-reteste significativas apenas para os homens (em PAL média de erros até ao sucesso e SWM total de erros) e tendo os homens apresentado melhor desempenho do que as mulheres em SWM total de erros e SWM estratégia, sugeriu-se que o desempenho dos homens teria contribuído mais para os efeitos da prática por fazerem maior uso de estratégias do que as mulheres.

No estudo da influência das variáveis sócio-demográficas (Gonçalves et al., 2016b) confirmou-se a vantagem dos homens no desempenho em SWM total de erros. Porém, verificou-se também melhor desempenho por parte dos homens em RVP latência média e melhores resultados por parte das mulheres em PAL média de erros até ao sucesso. O facto do estudo da fiabilidade teste-reteste (Gonçalves et al., 2016a) ter sido conduzido com uma amostra pequena e ter recorrido a testes estatísticos que não os utilizados no estudo da influência das variáveis sócio-demográficas, poderá explicar as diferenças entre os dois estudos. Naturalmente, são mais confiáveis os resultados do estudo da influência das variáveis sócio-demográficas por a amostra ser maior e por se ter utilizado a regressão linear múltipla. Este teste estatístico permite calcular correlações semi-parciais que indicam o contributo único de uma determinada variável independente ou preditor na variação de uma medida de teste, excluindo o contributo de outros preditores para a variação dessa medida. Por conseguinte, este é um teste mais apropriado na análise dos efeitos do género, ao excluir os efeitos das outras variáveis sócio-demográficas. Assim, ainda que a vantagem das mulheres no desempenho no PAL não invalide a hipótese dos homens poderem utilizar mais estratégias visuo-

espaciais do que as mulheres, não é certo que os homens usem mais da estratégia no PAL do que as mulheres ou que a estratégia seja o único factor a beneficiar o desempenho neste teste. Em termos de validade de constructo, o PAL e o SWM saturaram em factores distintos (Gonçalves et al., 2016c): o PAL num factor de memória visuo-espacial e o SWM num factor de memória de trabalho visuo-espacial. Deste modo, é possível que pessoas de ambos os sexos usem estratégias distintas ou que apresentem aptidões diferenciadas que favoreçam mais o desempenho em testes que avaliam uma função cognitiva e não outra.

Quanto aos efeitos de tecto detectados nas medidas RVP A' e RVP probabilidade de sucesso e ao efeito de chão detectado em RVP probabilidade de falso alarme em pessoas mais velhas sem problemas neuropsiquiátricos (Gonçalves et al., 2016a, 2016b, 2016c), os mesmos poderão ser atribuíveis a um nível baixo de dificuldade do teste RVP quando executado em modo 3-5-7. De facto, o modo RVP recomendado para adultos saudáveis é o modo clínico (Sahakian, Jones, Levy, Gray, & Warburton, 1989) que requer a detecção de três sequências-alvo de três dígitos cada. Este modo chegou a ser administrado a mais de uma dezena de participantes, alguns com escolaridade elevada, mas revelou-se muito difícil. No estudo da validade de constructo e da validade concorrente (Gonçalves et al., 2016c) observou-se que, dos 130 participantes do grupo SPN que realizaram o teste RVP no modo 3-5-7, só 5.4% atingiu o resultado máximo em RVP A' e 7.7% em RVP probabilidade de sucesso, tendo 28.5% obtido o resultado mínimo em RVP probabilidade de falso alarme. No caso do efeito de tecto identificado no grupo DA para a medida RVP probabilidade de sucesso, este sugere não o alcance de um bom nível de desempenho mas uma possível consequência do comportamento perseverativo ou de uma elevada tendência de resposta por parte de alguns participantes.

Tanto a medida RVP probabilidade de sucesso como a medida SWM estratégia poderão, na presença de comportamento perseverativo, conduzir a resultados ilusórios, ou seja, resultados aparentemente sugestivos de um bom desempenho, sem que este o seja. Nestas situações, estas medidas requerem especial cuidado na sua interpretação.

Como já referido, as medidas da CANTAB analisadas no estudo da influência das variáveis sócio-demográficas (Gonçalves et al., 2016b) foram somente aquelas que revelaram boa a adequada fiabilidade teste-reteste (Gonçalves et al., 2016a). É possível que a experiência com computadores não tenha tido influência significativa ($p < .05$)

nestas medidas, uma vez serem medidas calculadas com base em respostas simples e intuitivas como premir um botão (teste RVP) ou tocar no ecrã (testes PAL e SWM e medidas RTI tempo de movimento). Menos simples é a resposta de libertar um botão que se manteve premido, a qual corresponde às medidas RTI tempo de reacção. Atendendo a que o efeito da experiência com computadores não foi analisado nestas medidas, não é de excluir a hipótese desta variável poder afectar o desempenho nessas medidas da CANTAB. A influência desta variável sócio-demográfica no desempenho em testes de papel-e-lápis de cariz verbal é intrigante (Gonçalves et al., 2016b). Se tal influência for mediada por competências verbais desenvolvidas com a experiência com o computador, é de questionar porque motivo não terão essas competências facilitado também o desempenho dos utilizadores de computador em testes da CANTAB que podem envolver uma componente verbal (e.g., RVP e PAL) (Coull, Frith, Frackowiak, & Grasby, 1996; Gonçalves et al., 2016a). Efectivamente, o estudo da validade de constructo e da validade concorrente (Gonçalves et al., 2016c) revelou associações significativas (embora moderadas a fracas) entre estes testes e provas de papel-e-lápis verbais (e.g. entre o RVP e Fluência Verbal animais [ACE-R] e entre o PAL e Memória Lógica e Pares de Palavras [WMS-III]).

Relativamente à influência significativa do consumo de psicotrópicos nas medidas RVP A', PAL média de erros até ao sucesso, RTI tempo de movimento simples (CANTAB) e Fluência Verbal animais (ACE-R), foi sugerido que esta poderia ocorrer por prejuízo de processos de atenção, em especial de atenção selectiva (Gonçalves et al., 2016b). O estudo da validade de constructo e da validade concorrente (Gonçalves et al., 2016c) revelou associações moderadas a fracas mas significativas entre estas quatro medidas, podendo a atenção ser uma função comum entre elas, ainda que os testes RVP, RTI e PAL se distingam na principal função cognitiva avaliada.

CONSIDERAÇÕES PRÁTICAS ACERCA DA UTILIZAÇÃO DOS TESTES COMPUTORIZADOS COM PESSOAS MAIS VELHAS

As vantagens referidas na introdução desta dissertação acerca dos testes computadorizados e, em particular, dos testes da CANTAB, foram observadas durante a administração dos testes MOT, RVP, PAL, RTI e SWM a pessoas mais velhas portuguesas e em operações com o *software* CANTABeclipse[®] 3.0 (Cambridge Cognition, 2006a, 2006b), como a

produção de ficheiros .csv e relatórios sumários dos resultados. Não obstante, também foram identificados alguns problemas relacionados com as plataformas de resposta (i.e., o ecrã táctil e o bloco de interruptores), com as características de algumas provas e com a base de dados normativa original integrada no *software* CANTABeclipse[®] 3.0. Alguns destes problemas foram igualmente reconhecidos na administração dos testes da CANTAB a crianças (Luciana & Nelson, 2002) ou em outras provas computadorizadas (Zygouris & Tsolaki, 2014). Ainda assim, é de referir que os testes da CANTAB foram melhor tolerados por parte dos participantes mais velhos portugueses do que os subtestes da WMS-III administrados e o Teste da Figura Complexa de Rey.

Problemas relacionados com as plataformas de resposta

Ainda que não tenham sido identificados efeitos significativos ($p < .05$) da experiência com computadores nas medidas da CANTAB analisadas (Gonçalves et al., 2016b), foram observados alguns problemas na resposta por toque no ecrã, tanto em utilizadores como não utilizadores de computadores.

Um dos problemas é nem sempre o ecrã táctil revelar sensibilidade ao toque e por isso não registar a resposta. O manual (Cambridge Cognition, 2006b) recomenda que nestas situações a pessoa examinada afaste o dedo e volte a tocar no ecrã. No entanto, este problema traduz-se em menor exactidão no registo da latência de resposta, por exemplo, nos testes MOT e RTI. Nos testes PAL e SWM, sucede algumas pessoas assumirem que o facto do toque não ter aberto uma caixa significa que a sua resposta estava errada. Deste modo, é frequente estas pessoas tocarem noutra caixa mesmo antes do examinador lhes pedir para afastarem o dedo e voltarem a tocar no ecrã. Portanto, nestes casos, a não responsividade do ecrã táctil pode gerar erros de resposta. Já a escuta de instruções verbais durante a realização dos testes pode provocar distracção ou esquecimento relativamente a informação recentemente aprendida (e.g., teste PAL) ou armazenada na memória a curto prazo (e.g., teste SWM).

Nem sempre a não responsividade do ecrã táctil se deve à tecnologia. Se uma pessoa, inadvertidamente, tocar com mais de um dedo no ecrã, este não regista a resposta (no caso dos testes MOT, PAL e SWM) ou regista-a como um erro de imprecisão (no caso do RTI, um teste em que existe um critério de interrupção para um número específico de erros). O mesmo poderá suceder se a pessoa tiver uma ponta do

dedo muito grossa ou tocar com toda a falangeta, pois tocar numa região extensa do ecrã representa tocar em vários pontos ou áreas de toque. Algumas pessoas também apresentam uma polpa do dedo muito achatada, necessitando de tocar no ecrã com mais força, para este registar a resposta. Existem ainda casos em que as pessoas apresentam unhas muito compridas pelo que, não tocando com a falangeta mas com a unha, poderão danificar o ecrã.

Muitos destes problemas são detectados logo na administração do MOT, sendo fundamental esclarecer a pessoa sobre como deve proceder. Porém, no caso de pessoas com menos experiência na utilização deste tipo de tecnologia ou de pessoas com demência devido a DA, é frequente esquecerem-se das recomendações ao longo da sessão de teste e, por exemplo, permanecerem com o dedo no ecrã por mais tempo do que o devido, deslizarem-no pela superfície ou tocarem repetidas vezes naquele local (sendo que, neste caso, o sistema pode registar tal acção como tratando-se de mais de uma resposta).

Por vezes, o sistema não detecta o toque por especificações do próprio *software* CANTABeclipse[®] 3.0. No caso do SWM, uma caixa não abre enquanto a anterior não se tiver fechado. Se a pessoa for mais rápida a tocar na caixa seguinte do que o tempo que a caixa anterior demora a fechar, essa caixa seguinte não abre. Esta é outra situação em que dar instruções à pessoa para aguardar que a caixa anterior feche antes de tocar na caixa seguinte pode gerar distração e esquecimento da informação espacial mantida na memória de trabalho.

Relativamente ao bloco de interruptores, também foram observados problemas na sua utilização. No teste RVP encontraram-se pessoas com dificuldade em entender que podiam ter o dedo poisado no botão sem carregarem nele. Ao afastarem a mão do botão e ao procurarem-no no momento em que pretendiam responder, perdiam tempo de resposta. Algumas pessoas também revelaram o problema de carregar no botão e manterem-no premido, sem estarem depois em condições de assinalarem a detecção da sequência seguinte. No teste RTI, observou-se ser difícil para algumas pessoas (especialmente as com demência devido a DA) entenderem que deviam manter o dedo no botão, carregando nele até a bola amarela aparecer e libertando-o de seguida. Esta instrução envolve alguma complexidade em termos de coordenação motora. Além disso,

é uma acção contrária à que tinham aprendido no teste RVP (i.e., tocar uma vez no botão para dar a resposta, libertando-o de seguida).

Problemas relacionados com as características das provas

A tarefa do RVP, ainda que administrada no modo 3-5-7, foi referida como difícil por alguns participantes. Os dígitos são apresentados rapidamente e algumas pessoas (sobretudo as com demência devido a DA) revelam dificuldade em compreender que só devem premir o botão ao verem o último número da sequência. Não é raro as pessoas premirem o botão sempre que visualizam o “7”, independentemente deste ter sido antecedido pelos números “3” e “5”. As pessoas com dificuldade no controlo inibitório podem apresentar elevada tendência de resposta neste teste mas, em alguns casos, pode não se tratar de uma dificuldade a este nível mas em compreender que não devem premir o botão nos dígitos “3” e “5” ou em todos os “7”. Há quem refira cansaço visual após a tarefa de desempenho continuado do RVP.

O PAL é um teste bem construído do ponto de vista da dificuldade crescente dos itens. De um modo geral, os estádios com um padrão são considerados fáceis, mesmo por pessoas com demência devido a DA. Os estádios de 6 e 8 padrões são sentidos como difíceis, mesmo por pessoas muito escolarizadas. Deste modo, o PAL é um teste em que dificilmente se encontram efeitos de chão e de tecto em pessoas mais velhas. Contudo, este teste requer demasiadas tentativas até à sua interrupção. Em alguns casos (sobretudo nos estádios com menos padrões), sucede a pessoa acertar ao acaso, tendo então de realizar mais dez tentativas, no estádio seguinte. Tantos ensaios poderão tornar-se frustrantes, sobretudo para pessoas com demência devido a DA. Na versão mais recente do PAL, conforme comunicado pela Cambridge Cognition, este problema foi atenuado reduzindo-se o número de tentativas. Porém, esta nova versão carece de estudos normativos e psicométricos.

O RTI é um teste muito bem tolerado por pessoas sem dificuldade na utilização do bloco de interruptores e do ecrã tátil. Frequentemente, a tarefa é considerada fácil, sendo muito bem recebida após o PAL, cuja tarefa é mais exigente. No entanto, muitas pessoas mais velhas revelam pouco esforço em termos de rapidez de resposta. O som de *feedback* demora a ser emitido após a resposta, o que pode promover alguma letargia.

Quanto ao SWM, a tarefa torna-se muito difícil a partir dos problemas com 6 caixas, até para pessoas muito escolarizadas. É frequente a pessoa examinada esquecer-se de que as caixas que deve memorizar são as que já tiveram uma ficha e não as que ainda não a tiveram. No caso das pessoas com demência devido a DA, o teste torna-se difícil a partir dos problemas com 4 caixas. É frequente estas pessoas não entenderem a tarefa, tocando aleatoriamente nas caixas. Algumas pessoas queixam-se ainda de dores no braço dominante após este teste e o PAL, por o manterem levantado durante muito tempo.

Problemas relacionados com a base de dados normativa original

O *software* CANTABeclipse[®] 3.0 (Cambridge Cognition, 2006a, 2006b) integra uma base de dados com os resultados normativos calculados para a população inglesa residente na comunidade. Porém, a única maneira de aceder a esta base é solicitar ao sistema informático a comparação dos resultados de um dado participante com o grupo correspondente. Assim, se for um participante adulto, este pode ser comparado com a amostra total de adultos ingleses ou com o respectivo grupo de idade, género, inteligência pré-mórbida ou um outro grupo combinando qualquer uma destas três variáveis sócio-demográficas. Nem os manuais nem o *software* CANTABeclipse[®] 3.0 descrevem as características desta amostra normativa. Os grupos etários da amostra adulta, por exemplo, estão definidos em intervalos de dez anos. Não é possível conhecer a média das idades por grupo etário. Também não existe qualquer informação quanto aos anos de estudo ou escolaridade desta amostra. A inteligência pré-mórbida é estimada pelo *National Adult Reading Test* (Nelson & Willison, 1991), disponível somente para a população que fale o inglês como primeira língua. Por conseguinte, é difícil proceder a uma comparação entre a amostra normativa original da CANTAB e amostras de outros países, de modo a testar se estas provas são efectivamente pouco influenciadas por factores culturais.

Outro problema relacionado com a amostra normativa original da CANTAB é o facto de o *software* CANTABeclipse[®] 3.0 não apresentar normas para todas as medidas de um teste. Também existem testes sem normas disponíveis. Quanto aos dados psicométricos, nem o manual nem o *software* CANTABeclipse[®] 3.0 os apresentam. Para se obter informação relativamente às propriedades psicométricas dos testes da

CANTAB é necessário pesquisar artigos científicos, podendo o utilizador não ter facilidade de acesso aos mesmos. Alguns artigos foram publicados há bastantes anos, sendo por isso difíceis de aceder, mesmo no meio académico. Assim, apesar das muitas vantagens apresentadas pelos testes da CANTAB, há muito trabalho a fazer quanto ao estudo das suas propriedades psicométricas.

Com o rápido avanço tecnológico, os testes da CANTAB são frequentemente vendidos em novos suportes informáticos. O *software* que os disponibiliza também tem sido frequentemente actualizado, por vezes alterando-se a estrutura de alguns testes (como o PAL) e os próprios estímulos. Qualquer uma destas mudanças implica que novos estudos normativos e psicométricos tenham de ser realizados. Basta uma nova tecnologia de ecrã tátil ser mais sensível para que os valores normativos disponíveis para a versão anterior possam caducar. Acompanhar a velocidade do progresso tecnológico não é fácil. Os estudos normativos e psicométricos requerem amostras de elevada dimensão, pelo que são sempre mais demorados. Não obstante, os testes computadorizados são o futuro da avaliação neuropsicológica, justificando-se o investimento e esforço na sua adaptação e validação.

LIMITAÇÕES E VANTAGENS DESTE PROJECTO DE INVESTIGAÇÃO

Muitas das limitações e vantagens deste projecto de investigação foram já abordadas para cada um dos estudos realizados (Gonçalves et al., 2016a, 2016b, 2016c). Assim, relativamente às limitações, seria desejável que o estudo da fiabilidade teste-reteste tivesse sido conduzido com uma amostra de maior dimensão, para garantir maior robustez dos dados. Como o projecto de investigação previa inicialmente a constituição de dois grupos clínicos, definiu-se como objectivo estudar a fiabilidade teste-reteste numa amostra mais pequena, de modo a garantir o cumprimento dos restantes objectivos, dentro do período de tempo estabelecido. Infelizmente, não se conseguiu constituir, com uma dimensão razoável, o grupo clínico de depressão major com défice objectivo na memória episódica, pelo que teria sido preferível que este projecto tivesse sido concebido sem este grupo clínico e com uma amostra maior para o estudo da precisão teste-reteste.

Outra limitação deste projecto de investigação é o facto de não incluir a avaliação de biomarcadores da DA por punção lombar e nem todos os participantes possuem

exames de neuroimagem. Estes dados complementares poderiam conferir maior rigor à classificação dos participantes, em especial os do grupo DA, embora estas técnicas apresentem limitações, como a falta de valores de referência consensualmente aceites para muitos biomarcadores (e.g., os avaliados por análise do líquido céfalo-raquidiano; Albert et al., 2011; Blennow, Biscetti, Eusebi, & Parnetti, 2016) e, por vezes, resultados conflitantes entre técnicas (e.g., entre a punção lombar e a PET, quanto a eventual amiloidose cerebral; McKhann et al., 2011). É ainda de referir que as técnicas de neuroimagem nem sempre revelam alterações em indivíduos com défices cognitivos, segundo a avaliação neuropsicológica (Lezak, Howieson, Bigler, & Tranel, 2012). Porém, apesar de não se terem analisado os biomarcadores, este projecto de investigação apresenta a vantagem de reunir um protocolo bastante completo para rastreio, avaliando uma gama vasta de funções cognitivas e dispondo, para algumas funções, de mais de um instrumento de avaliação. Além disso, houve o cuidado de complementar esta avaliação com o recurso a outros informadores, entre os quais o médico de cada participante. Ainda assim, é de referir a limitação de terem sido aplicadas normas para residentes na comunidade a participantes a viver em instituições geriátricas, um problema infelizmente difícil de ser contornado dada a escassez de instrumentos de avaliação com normas para pessoas mais velhas institucionalizadas, o qual se mencionou anteriormente.

Na perspectiva do diagnóstico precoce da DA, este projecto apresenta também a limitação de não incluir um grupo de DCL devido a DA, de misturar participantes com demência devido a DA em estágio ligeiro com outros em estágio moderado e de não incluir um estudo de poder discriminante analisando os pontos de corte entre estes grupos e um grupo controlo. Ainda que muito relevante para o diagnóstico precoce da DA e a prática clínica, este estudo tornaria o projecto demasiado ambicioso, pondo em causa a sua exequibilidade dentro dos limites de tempo definidos.

Devido ao número elevado de medidas de resultados passíveis de serem calculadas para cada teste da CANTAB, foi necessário proceder à sua selecção, pelo que os estudos apresentam a limitação de não apresentarem os dados para todas as medidas de teste. Esta limitação poderá, no entanto, ser resolvida com a publicação futura de artigos com os resultados obtidos para essas medidas.

O facto de se terem incluído no estudo participantes com consumo de psicotrópicos compatível com a prescrição de medicamentos potencialmente inapropriados é simultaneamente uma limitação e uma vantagem, ao conferir maior validade ecológica aos dados (cf. Periquito et al., 2014). Como os efeitos dessas substâncias não puderam ser analisados em separado, uns fármacos poderão ter acentuado ou contrariado o efeito de outros, afectando de ambos os modos os resultados observados.

Este projecto incluiu participantes com idades muito avançadas. Há quem critique os estudos realizados com pessoas com muita idade (e.g., 75 ou mais anos), justificando que a maioria destes indivíduos apresenta alterações patológicas a nível cerebral, ainda que assintomáticas, e que a sua exclusão por exames de neuroimagem conduz, por outro lado, a amostras de “super-idosos”, sem validade ecológica (Rabbitt, Lowe, & Shilling, 2001). Todavia, com o aumento da esperança de vida, este grupo etário tem vindo a ganhar maior dimensão. Como qualquer outro grupo de idade, é necessário estudá-lo de modo a que os processos de envelhecimento normal sejam conhecidos. Se excluirmos estas idades dos estudos sobre envelhecimento, a consequência na prática é a de compararmos o desempenho destas pessoas com outras mais jovens, podendo formar julgamentos inadequados, com prejuízo para as pessoas mais velhas.

Neste sentido, é considerado um ponto forte deste projecto de investigação o facto de incluir pessoas com idades avançadas e a residir em instituições geriátricas, atendendo aos poucos estudos que se encontram com esta população. Além disso, este é o primeiro projecto de investigação de adaptação e validação dos testes RVP, PAL, RTI e SWM da CANTAB para pessoas mais velhas portuguesas.

Internacionalmente, também faltavam estudos de fiabilidade teste-reteste para os testes RVP e RTI, sendo que nenhum dos estudos de precisão teste-reteste com o PAL e o SWM (Fowler et al., 1995; Lowe & Rabbitt, 1998) tinha sido realizado com pessoas mais velhas institucionalizadas nem se tinha procedido à análise de métodos de correcção dos efeitos da prática para interpretação de mudanças individuais confiáveis (i.e., uma melhoria ou declínio real no desempenho de um indivíduo). Ao responder a estas lacunas, este projecto de investigação assumiu um carácter inovador.

Simultaneamente, este projecto de investigação apresenta a vantagem de fornecer normas de testes computadorizados para pessoas mais velhas residentes em instituições

geriátricas quando são tão escassos os instrumentos de avaliação com normas para esta população. Este projecto é também o primeiro a apresentar normas para adultos mais velhos para o RVP modo 3-5-7, pois o *software* CANTABeclipse® 3.0 (Cambridge Cognition, 2006a, 2006b) só dispõe de normas para crianças para este modo de teste. O presente projecto de investigação é também o primeiro a analisar a influência da experiência com computadores no desempenho nos testes da CANTAB, contribuindo assim para o conhecimento da influência de mais variáveis sócio-demográficas do que as até então estudadas sobre o desempenho nestes testes.

Por último, este projecto tem como ponto forte o facto de estudar em conjunto os quatro testes da CANTAB recomendados para avaliação da demência devido a DA. Por exemplo, neste projecto foi analisada a validade de constructo do teste RVP em conjunto com o PAL, RTI e SWM, o que ainda não tinha sido feito antes. Além disso, esse estudo não se limitou à análise da validade de constructo numa só amostra (e.g., mista ou sem problemas neuropsiquiátricos, como em outros estudos; Fowler et al., 1995; Reeves et al., 2005; Robbins et al., 1998; Smith, Need, Cirulli, Chiba-Falek, & Attix, 2013; Stillely, Bender, Dunbar-Jacob, Sereika, & Ryan, 2010; Stoffers et al., 2007), mas em ambas as amostras, comparando-as e, desse modo, testando melhor a robustez da estrutura factorial.

DIRECÇÕES FUTURAS

Em face de todo o potencial dos testes da CANTAB para a avaliação neuropsicológica da DA e outras demências, são muitos os estudos que ainda precisam ser realizados. Este projecto de investigação representa somente um pequeno contributo para o estado do conhecimento actual, suscitando novas questões e a necessidade de estudos adicionais.

Assim, são necessários mais estudos de fiabilidade teste-reteste que apliquem métodos de correcção dos efeitos da prática e que sejam conduzidos com amostras maiores de pessoas mais velhas residentes em instituições geriátricas, residentes na comunidade e com grupos clínicos. Os estudos comparando os resultados nestas provas entre consumidores e não consumidores de psicotrópicos sem problemas neuropsiquiátricos poderá igualmente esclarecer melhor as propriedades psicométricas destes testes da CANTAB e, em termos práticos, os resultados individuais ao longo do

tempo. Neste sentido, são também necessários estudos que analisem os resultados em mais do que dois momentos de avaliação e que considerem, preferencialmente, intervalos mais alargados de tempo. Estes estudos, se conduzidos com populações de elevada dimensão, apresentam a vantagem de permitirem estabelecer normas para cada momento de avaliação (e.g., normas baseadas na regressão), sendo este um outro processo de minimização dos efeitos da prática (Attix et al., 2009).

A generalização dos resultados dos testes da CANTAB entre culturas diferentes continua a ter de ser verificada. Este trabalho implicará uma articulação maior com a Cambridge Cognition, no sentido de se ter acesso às características da amostra normativa original, para uma adequada comparação dos resultados com amostras de outras culturas. Relativamente a Portugal, são ainda necessários estudos normativos com pessoas a residir na comunidade (e incluindo grupos etários mais jovens) e com pessoas a viver em outras áreas geográficas, assim como o estabelecimento de normas para mais medidas da CANTAB do que as agora analisadas.

Relativamente às variáveis sócio-demográficas estudadas, fazem falta pesquisas que aprofundem o conhecimento acerca do impacto da utilização de computadores nos processos de treino e desenvolvimento de competências cognitivas. Compreender que tipo de competências podem ser desenvolvidas consoante a utilização que se faz do computador poderá ajudar a compreender melhor o processo de influência desta variável no desempenho em determinados testes cognitivos e o motivo da sua não influência nos resultados em outros testes cognitivos. Quanto ao impacto da institucionalização no desempenho em testes, é uma questão que continua por responder. Esta é uma variável complexa a requerer maior operacionalização. Talvez o desenvolvimento de um instrumento que permitisse caracterizar as instituições em termos dos serviços disponibilizados, estilo de vida, qualidade do ambiente institucional, entre outros aspectos envolvidos no conceito de institucionalização, pudesse ajudar a compreender melhor este processo de influência sobre o desempenho cognitivo. Já a análise do desempenho cognitivo em diferentes momentos do processo de institucionalização (e.g., admissão, três meses depois, um ano depois) poderia elucidar melhor sobre o impacto do tempo de institucionalização no desempenho em testes. De relevância são também os estudos que analisem o efeito de substâncias psicotrópicas específicas no desempenho nestes testes da CANTAB, comparando os

efeitos de diferentes tipos de psicotrópicos e de diferentes tipos de consumo (e.g., simples, duplicado ou triplicado).

Em termos de validade concorrente, são necessários mais estudos comparando os testes RVP e RTI com testes de desempenho continuado, velocidade psicomotora e tempo de reacção e comparando os quatro testes da CANTAB com outros testes que envolvam a mesma modalidade sensorial para a apresentação dos estímulos e resposta ao teste, no sentido de examinar se as relações entre eles são mais fortes do que entre testes envolvendo diferentes modalidades sensoriais, como tem sido sugerido (Elwood, 2001).

Muitíssimo importantes são os estudos de validade preditiva com estes testes da CANTAB, sendo necessário realizá-los na população mais velha portuguesa. Igualmente importantes são os estudos de poder discriminante conduzidos com grupos saudáveis, com DCL e com demência devido a DA e com outros grupos clínicos, de modo a conhecer a sua sensibilidade e especificidade e a sua utilidade para o diagnóstico diferencial. Tais estudos permitirão aumentar o conhecimento acerca da relevância clínica destes testes na população portuguesa, ajudando o examinador a fazer opções mais fundamentadas na escolha dos testes a administrar, consoante os seus objectivos de avaliação e a população alvo. Está prevista para breve a realização de um estudo de poder discriminante conduzido com o grupo DA e um grupo controlo seleccionado a partir da amostra de indivíduos mais velhos portugueses sem problemas neuropsiquiátricos, de modo a analisar a sensibilidade e especificidade destes testes da CANTAB e a se calcularem os respectivos pontos de corte.

CONCLUSÃO

Apesar das provas da CANTAB para avaliação da DA e outras demências reunirem um conjunto vasto de estudos de validade preditiva e poder discriminante, foram poucos os estudos encontrados analisando a fiabilidade teste-reteste e os efeitos da prática e nenhuns estudos examinando as quatro provas em conjunto para definição de normas e análise da validade de constructo e da validade concorrente. Por conseguinte, os estudos compreendidos no presente projecto de investigação pretenderam responder a estas lacunas. Em termos práticos, representam um contributo para a população portuguesa no sentido de tornarem disponível para utilização as provas desta bateria, com a sua

adaptação e cálculo de valores normativos para a população mais velha residente em instituições geriátricas.

Os estudos deste projecto de investigação permitiram identificar medidas da CANTAB com boa a adequada fiabilidade teste-reteste e métodos eficazes para corrigir os pequenos efeitos da prática detectados em algumas medidas, para detecção de mudanças confiáveis no desempenho individual (Gonçalves et al., 2016a). Simultaneamente, permitiram estabelecer normas em função das variáveis sócio-demográficas com maior influência e aumentar o conhecimento acerca do impacto de variáveis anteriormente ainda não estudadas em termos de influência no desempenho nestes testes (Gonçalves et al., 2016b). Relativamente a este aspecto, é de realçar o facto de diferentes variáveis sócio-demográficas terem influenciado testes distintos e medidas de um mesmo teste, implicando futuramente maior cuidado quanto às variáveis consideradas para definição de normas, visto que com frequência os estudos normativos apresentam valores definidos segundo as mesmas variáveis sócio-demográficas. Por último, os estudos permitiram acrescentar conhecimento relativamente às funções cognitivas avaliadas por estes testes e às relações entre as suas medidas e com medidas de testes de papel-e-lápis, tendo os resultados sugerido boa a adequada validade convergente e concorrente (Gonçalves et al., 2016c).

Portanto, os testes da CANTAB para avaliação da DA e outras demências apresentam-se como uma boa opção para a avaliação e monitorização do desempenho cognitivo de pessoas mais velhas portuguesas a residir em instituições geriátricas. Contudo, são ainda necessários mais estudos com esta população e com a população residente na comunidade. Com este projecto de investigação espera-se promover uma maior utilização destes testes computadorizados em Portugal, nos sectores clínico e académico e na investigação médica e farmacêutica.

REFERÊNCIAS

- Ahmed, S., Mitchell, J., Arnold, R., Nestor, P. J., & Hodges, J. R. (2008). Predicting rapid clinical progression in amnesic mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, 25(2), 170-177. doi: 10.1159/000113014
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., ... Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-

- Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270-279. doi: 10.1016/j.jalz.2011.03.008
- Attix, D. K., Story, T. J., Chelune, G. J., Ball, J. D., Stutts, M. L., Hart, R. P., & Barth, J. T. (2009). The prediction of change: Normative neuropsychological trajectories. *The Clinical Neuropsychologist*, 23(1), 21-38. doi: 10.1080/13854040801945078
- Blackwell, A. D., Sahakian, B. J., Vesey, R., Semple, J. M., Robbins, T. W., & Hodges, J. R. (2004). Detecting dementia: Novel neuropsychological markers of preclinical Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 17(1-2), 42-48. doi: 10.1159/000074081
- Blennow, K., Biscetti, L., Eusebi, P., & Parnetti, L. (2016). Cerebrospinal fluid biomarkers in Alzheimer's and Parkinson's diseases - From pathophysiology to clinical practice. *Movement Disorders*, 31(6), 836-847. doi: 10.1002/mds.26656
- Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological measures: Meta-analyses of test-retest correlations. *The Clinical Neuropsychologist*, 27(7), 1077-1105. doi: 10.1080/13854046.2013.809795
- Cambridge Cognition. (2006a). *CANTABeclipse™ version 3: Software user guide*. Cambridge, UK: Author.
- Cambridge Cognition. (2006b). *CANTABeclipse™ version 3: Test administration guide*. Cambridge, UK: Author.
- Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia*, 34(11), 1085-1095. doi: 10.1016/0028-3932(96)00029-2
- Crawford, J. R., & Garthwaite, P. H. (2007). Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology*, 21(5), 611-620. doi: 10.1037/0894-4105.21.5.611
- De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the CANTAB. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 25(2), 242-254. doi: 10.1076/jcen.25.2.242.13639
- Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology*, 27(3), 248-261. doi: 10.1093/arclin/acr120
- Duff, K. (2014). One-week practice effects in older adults: Tools for assessing cognitive change. *The Clinical Neuropsychologist*, 28(5), 714-725. doi: 10.1080/13854046.2014.920923
- Elwood, R. W. (2001). MicroCog: Assessment of cognitive functioning. *Neuropsychology Review*, 11(2), 89-100. doi: 10.1023/a:1016671201211
- Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J., & Sjöberg, A. (2013). *EFPA review model for the description and evaluation of psychological and educational tests: Test review form and notes for reviewers - Version 4.2.6*. Brussels, Belgium: European Federation of Psychologists' Associations (EFPA). Retrieved from <http://www.efpa.eu/professional-development/assessment>.
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in

- the early detection of dementia. *Applied Neuropsychology*, 2(2), 72-78. doi: 10.1207/s15324826an0202_4
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1997). Computerized neuropsychological tests in the early detection of dementia: Prospective findings. *Journal of the International Neuropsychological Society*, 3(2), 139-146.
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (2002). Paired associate performance in the early detection of DAT. *Journal of the International Neuropsychological Society*, 8(1), 58-71. doi: 10.1017/S1355617701020069
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016a). Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, 23(4), 251-263. doi: 10.1080/23279095.2015.1053889
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016b). Effects of socio-demographic variables on performance on the Cambridge neuropsychological automated tests for the assessment of dementia and Portuguese norms for older adults living in retirement homes. *The Clinical Neuropsychologist*, 30(2), 284-317. doi: 10.1080/13854046.2016.1156745
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016c). *Construct and concurrent validity of the Cambridge neuropsychological automated tests for assessment of dementia in Portuguese older adults*. Manuscript submitted for publication.
- Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology*, 21(7), 623-643. doi: 10.1016/j.acn.2006.05.007
- Iverson, G. L. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology*, 16(2), 183-191. doi: 10.1093/arclin/16.2.183
- Lemaire, P. (2010). Cognitive strategy variations during aging. *Current Directions in Psychological Science*, 19(6), 363-369. doi: 10.1177/0963721410390354
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford, NY: Oxford University Press.
- Lowe, C., & Rabbitt, P. (1998). Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. *Neuropsychologia*, 36(9), 915-923. doi: 10.1016/S0028-3932(98)00036-0
- Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology*, 22(3), 595-624. doi: 10.1207/s15326942dn2203_3
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., ... Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263-269. doi: 10.1016/j.jalz.2011.03.005
- Mitchell, J., Arnold, R., Dawson, K., Nestor, P. J., & Hodges, J. R. (2009). Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm. *Journal of Neurology*, 256(9), 1500-1509. doi: 10.1007/s00415-009-5152-0

- Nelson, H. E., & Willison, J. (1991). *The National Adult Reading Test (NART): Test manual* (2nd ed.). Windsor, UK: NFER-Nelson Publishing Company.
- Periquito, C. M. N., Silva, P. C., Oliveira, P., Carneiro, C., Fernandes, A. I., & Costa, F. A. (2014). Revisão da medicação em idosos institucionalizados: Aplicação dos critérios STOPP e START. *Revista Portuguesa de Farmacoterapia*, *6*, 211-220.
- Rabbitt, P., Lowe, C., & Shilling, V. (2001). Frontal tests and models for cognitive ageing. *The European Journal of Cognitive Psychology*, *13*(1), 5-28. doi: 10.1080/09541440042000197
- Reeves, S. J., Grasby, P. M., Howard, R. J., Bantick, R. A., Asselin, M.-C., & Mehta, M. A. (2005). A positron emission tomography (PET) investigation of the role of striatal dopamine (D2) receptor availability in spatial cognition. *NeuroImage*, *28*(1), 216-226. doi: 10.1016/j.neuroimage.2005.05.034
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society*, *4*(5), 474-490. doi: 10.1017/s1355617798455073
- Sahakian, B. J., Jones, G., Levy, R., Gray, J., & Warburton, D. (1989). The effects of nicotine on attention, information processing, and short-term memory in patients with dementia of the Alzheimer type. *The British Journal of Psychiatry*, *154*(6), 797-800. doi: 10.1192/bjp.154.6.797
- Schweiger, A., Doniger, G. M., Dwolatzky, T., Jaffe, D., & Simon, E. S. (2003). Reliability of a novel computerized neuropsychological battery for mild cognitive impairment. *Acta Neuropsychologica*, *1*(4), 407-413. doi: 10.5604/17307503/9168
- Singer, J. J., MacGregor, A. J., Cherkas, L. F., & Spector, T. D. (2006). Genetic influences on cognitive function using the Cambridge Neuropsychological Test Automated Battery. *Intelligence*, *34*(5), 421-428. doi: 10.1016/j.intell.2005.11.005
- Smith, P. J., Need, A. C., Cirulli, E. T., Chiba-Falek, O., & Attix, D. K. (2013). A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with "traditional" neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology*, *35*(3), 319-328. doi: 10.1080/13803395.2013.771618
- Stilley, C. S., Bender, C. M., Dunbar-Jacob, J., Sereika, S., & Ryan, C. M. (2010). The impact of cognitive function on medication management: Three studies. *Health Psychology*, *29*(1), 50-55. doi: 10.1037/a0016940
- Stoffers, D., Bosboom, J. L. W., Deijen, J. B., Wolters, E. C., Berendse, H. W., & Stam, C. J. (2007). Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia. *Brain*, *130*(7), 1847-1860. doi: 10.1093/brain/awm034
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.
- Tornatore, J. B., Hill, E., Laboff, J. A., & McGann, M. E. (2005). Self-administered screening for mild cognitive impairment: Initial validation of a computerized test battery. *Journal of Neuropsychiatry*, *17*(1), 98-105. doi: 10.1176/appi.neuropsych.17.1.98

- Wechsler, D. (2008). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual técnico*. Lisboa, Portugal: Cegoc-Tea.
- Zygouris, S., & Tsolaki, M. (2014). Computerized cognitive testing for older adults: A review. *American Journal of Alzheimer's Disease and Other Dementias*, 30(1), 13-28. doi: 10.1177/1533317514522852

LISTA INTEGRAL DE REFERÊNCIAS

- Ahmed, S., Mitchell, J., Arnold, R., Nestor, P. J., & Hodges, J. R. (2008). Predicting rapid clinical progression in amnesic mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, 25(2), 170-177. doi: 10.1159/000113014
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., ... Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270-279. doi: 10.1016/j.jalz.2011.03.008
- Alexander, G. E., Ryan, L., Bowers, D., Foster, T. C., Bizon, J. L., Geldmacher, D. S., & Glisky, E. L. (2012). Characterizing cognitive aging in humans with links to animal models. *Frontiers in Aging Neuroscience*, 4(21), 1-18. doi: 10.3389/fnagi.2012.00021
- Alichniewicz, K. K., Brunner, F., Klünemann, H. H., & Greenlee, M. W. (2012). Structural and functional neural correlates of visuospatial information processing in normal aging and amnesic mild cognitive impairment. *Neurobiology of Aging*, 33(12), 2782-2797. doi: 10.1016/j.neurobiolaging.2012.02.010
- Alladi, S., Arnold, R., Mitchell, J., Nestor, P. J., & Hodges, J. R. (2006). Mild cognitive impairment: Applicability of research criteria in a memory clinic and characterization of cognitive profile. *Psychological Medicine*, 36(4), 507-515. doi: 10.1017/s0033291705006744
- Alves, L., Simões, M. R., & Martins, C. (2012). The estimation of premorbid intelligence levels among Portuguese speakers: The Irregular Word Reading Test (TeLPI). *Archives of Clinical Neuropsychology*, 27(1), 58-68. doi: 10.1093/arclin/acr103
- Alves, L., Simões, M. R., Martins, C., & Freitas, S. (2015). *The Irregular Word Reading Test (TeLPI): Normative study for the Portuguese population*. Manuscript submitted for publication.
- American Psychological Association. (2015). *APA dictionary of psychology* (2nd ed.). Washington, DC: Author.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan Publishing Company.
- Attix, D. K., Story, T. J., Chelune, G. J., Ball, J. D., Stutts, M. L., Hart, R. P., & Barth, J. T. (2009). The prediction of change: Normative neuropsychological trajectories. *The Clinical Neuropsychologist*, 23(1), 21-38. doi: 10.1080/13854040801945078
- Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de Depressão Geriátrica (GDS). In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* (2ª ed., pp. 69-72). Lisboa, Portugal: Novartis.
- Bartko, S. J., Vendrell, I., Saksida, L. M., & Bussey, T. J. (2010). A computer-automated touchscreen paired-associates learning (PAL) task for mice: Impairments following administration of scopolamine or dicyclomine and improvements following donepezil. *Psychopharmacology*, 214(2), 537-548. doi: 10.1007/s00213-010-2050-1
- Beats, B. C., Sahakian, B. J., & Levy, R. (1996). Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. *Psychological Medicine*, 26(3), 591. doi: 10.1017/s0033291700035662

- Benedict, R. H. B., & Zgaljardic, D. J. (1998). Practice effects during repeated administrations of memory tests with and without alternate forms. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 20(3), 339-352. doi: 10.1076/jcen.20.3.339.822
- Blackwell, A. D., Sahakian, B. J., Vesey, R., Semple, J. M., Robbins, T. W., & Hodges, J. R. (2004). Detecting dementia: Novel neuropsychological markers of preclinical Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 17(1-2), 42-48. doi: 10.1159/000074081
- Blennow, K., Biscetti, L., Eusebi, P., & Parnetti, L. (2016). Cerebrospinal fluid biomarkers in Alzheimer's and Parkinson's diseases - From pathophysiology to clinical practice. *Movement Disorders*, 31(6), 836-847. doi: 10.1002/mds.26656
- Braw, Y., Aviram, S., Bloch, Y., & Levkovitz, Y. (2011). The effect of age on frontal lobe related cognitive functions of unmedicated depressed patients. *Journal of Affective Disorders*, 129(1-3), 342-347. doi: 10.1016/j.jad.2010.07.032
- Bruggemans, E. F., van de Vijver, F. J. R., & Huysmans, H. A. (1997). Assessment of cognitive deterioration in individual patients following cardiac surgery: Correcting for measurement error and practice effects. *Journal of Clinical and Experimental Neuropsychology*, 19(4), 543-559. doi: 10.1080/01688639708403743
- Burns, A., Lawlor, B., & Craig, S. (2004). *Assessment scales in old age psychiatry* (2nd ed.). London, UK: Martin Dunitz of the Taylor & Francis Group.
- Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological measures: Meta-analyses of test-retest correlations. *The Clinical Neuropsychologist*, 27(7), 1077-1105. doi: 10.1080/13854046.2013.809795
- Cambridge Cognition. (2006a). *CANTABeclipse™ version 3: Software user guide*. Cambridge, UK: Author.
- Cambridge Cognition. (2006b). *CANTABeclipse™ version 3: Test administration guide*. Cambridge, UK: Author.
- Cernich, A. N., Brennana, D. M., Barker, L. M., & Bleiberg, J. (2007). Sources of error in computerized neuropsychological assessment. *Archives of Clinical Neuropsychology*, 22(Suppl. 1), S39-S48. doi: 10.1016/j.acn.2006.10.004
- Chamberlain, S. R., Blackwell, A. D., Nathan, P. J., Hammond, G., Robbins, T. W., Hodges, J. R., ... Sahakian, B. J. (2011). Differential cognitive deterioration in dementia: A two year longitudinal study. *Journal of Alzheimer's Disease*, 24(1), 125-136. doi: 10.3233/jad-2010-100450
- Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2007). A neuropsychological comparison of obsessive-compulsive disorder and trichotillomania. *Neuropsychologia*, 45(4), 654-662. doi: 10.1016/j.neuropsychologia.2006.07.016
- Chase, H. W., Clark, L., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2008). Dissociable roles of prefrontal subregions in self-ordered working memory performance. *Neuropsychologia*, 46(11), 2650-2661. doi: 10.1016/j.neuropsychologia.2008.04.021
- Chelune, G. J., Naugle, R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, 7(1), 41-52. doi: 10.1037/0894-4105.7.1.41

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, *55*(10), 1630-1635. doi: 10.1111/j.1532-5415.2007.01379.x
- Collie, A., Maruff, P., & Currie, J. (2002). Behavioral characterization of mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, *24*(6), 720-733. doi: 10.1076/jcen.24.6.720.8397
- Collie, A., Maruff, P., Darby, D. G., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *Journal of the International Neuropsychological Society*, *9*(3), 419-428. doi: 10.1017/s1355617703930074
- Collie, A., Maruff, P., Shafiq-Antonacci, R., Smith, M., Hallup, M., Schofield, P. R., ... Currie, J. (2001). Memory decline in healthy older people: Implications for identifying mild cognitive impairment. *Neurology*, *56*(11), 1533-1538. doi: 10.1212/wnl.56.11.1533
- Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia*, *34*(11), 1085-1095. doi: 10.1016/0028-3932(96)00029-2
- Crawford, J. R., & Garthwaite, P. H. (2007). Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology*, *21*(5), 611-620. doi: 10.1037/0894-4105.21.5.611
- Cummings, J., Gould, H., & Zhong, K. (2012). Advances in designs for Alzheimer's disease clinical trials. *American Journal of Neurodegenerative Disease*, *1*(3), 205-216.
- Cysique, L. A., Franklin, D., Abramson, I., Ellis, R. J., Letendre, S., Collier, A., ... Heaton, R. K. (2011). Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. *Journal of Clinical and Experimental Neuropsychology*, *33*(5), 505-522. doi: 10.1080/13803395.2010.535504
- Dancey, C. P., & Reidy, J. (2004). *Statistics without maths for psychology: Using SPSS for Windows* (3rd ed.). Essex, England: Prentice Hall.
- Darby, D., Maruff, P., Collie, A., & McStephen, M. (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*, *59*(7), 1042-1046. doi: 10.1212/wnl.59.7.1042
- De Jager, C., Blackwell, A. D., Budge, M. M., & Sahakian, B. J. (2005). Predicting cognitive decline in healthy older adults. *The American Journal of Geriatric Psychiatry*, *13*(8), 735-740. doi: 10.1097/00019442-200508000-00014
- De Jager, C., & Budge, M. (2005). Stability and predictability of the classification of mild cognitive impairment as assessed by episodic memory test performance over time. *Neurocase*, *11*(1), 72-79. doi: 10.1080/13554790490896820
- De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the CANTAB. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental*

- Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 25(2), 242-254. doi: 10.1076/jcen.25.2.242.13639
- de Rover, M., Pironti, V. A., McCabe, J. A., Acosta-Cabronero, J., Arana, F. S., Morein-Zamir, S., ... Sahakian, B. J. (2011). Hippocampal dysfunction in patients with mild cognitive impairment: A functional neuroimaging study of a visuospatial paired associates learning task. *Neuropsychologia*, 49(7), 2060-2070. doi: 10.1016/j.neuropsychologia.2011.03.037
- Downes, J. J., Roberts, A. C., Sahakian, B. J., Evenden, J. L., Morris, R. G., & Robbins, T. W. (1989). Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: Evidence for a specific attentional dysfunction. *Neuropsychologia*, 27(11-12), 1329-1343. doi: 10.1016/0028-3932(89)90128-0
- Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology*, 27(3), 248-261. doi: 10.1093/arclin/acr120
- Duff, K. (2014). One-week practice effects in older adults: Tools for assessing cognitive change. *The Clinical Neuropsychologist*, 28(5), 714-725. doi: 10.1080/13854046.2014.920923
- Egger, S. A., Levy, R., & Sahakian, B. J. (1991). Tacrine in Alzheimer's disease. *The Lancet*, 337(8748), 989-992. doi: 10.1016/0140-6736(91)92656-m
- Edgin, J. O., Mason, G. M., Allman, M. J., Capone, G. T., DeLeon, I., Maslen, C., ... Nadel, L. (2010). Development and validation of the Arizona Cognitive Test Battery for Down syndrome. *Journal of Neurodevelopmental Disorders*, 2(3), 149-164. doi: 10.1007/s11689-010-9054-3
- Elderkin-Thompson, V., Kumar, A., Bilker, W., Dunkin, J., Mintz, J., Moberg, P., ... Gur, R. (2003). Neuropsychological deficits among patients with late-onset minor and major depression. *Archives of Clinical Neuropsychology*, 18(5), 529-549. doi: 10.1016/s0887-6177(03)00022-2
- Elwood, R. W. (2001). MicroCog: Assessment of cognitive functioning. *Neuropsychology Review*, 11(2), 89-100. doi: 10.1023/a:1016671201211
- Espírito-Santo, H., Lemos, L., Ventura, L., Moitinho, S., Pinto, A. L., Rodrigues, F., ... Daniel, F. (2015). Teste da Figura Complexa de Rey-Osterrieth-A. In M. R. Simões & I. Santana (Eds.), *Escalas e testes na demência* (3ª ed., pp. 100-107). Porto Salvo, Portugal: Novartis.
- Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñoz, J., & Sjöberg, A. (2013). *EFPA review model for the description and evaluation of psychological and educational tests: Test review form and notes for reviewers - Version 4.2.6*. Brussels, Belgium: European Federation of Psychologists' Associations (EFPA). Retrieved from <http://www.efpa.eu/professional-development/assessment>.
- Facal, D., Rodríguez, N., Juncos-Rabadán, O., Caamaño, J. M., & Sueiro, J. (2009). Use of the Cambridge Neuropsychological Test Automated Battery for the diagnosis of mild cognitive impairment. A pilot study in a Spanish sample. *Rev Esp Geriatr Gerontol*, 44(2), 79-84. doi: 10.1016/j.regg.2008.07.004
- Falletti, M. G., Maruff, P., Collie, A., & Darby, D. G. (2006). Practice effects associated with the repeated assessment of cognitive function using the CogState Battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, 28(7), 1095-1112. doi: 10.1080/13803390500205718

- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London, UK: Sage Publications.
- Firmino, H., Simões, M. R., Pinho, M. S., Cerejeira, L., & Martins, C. (2008). *O Exame Cognitivo de Addenbrooke - Revisto*. Coimbra, Portugal: Serviço de Avaliação Psicológica da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State. *Journal of Psychiatric Research*, *12*(3), 189-198. doi: 10.1016/0022-3956(75)90026-6
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. *Applied Neuropsychology*, *2*(2), 72-78. doi: 10.1207/s15324826an0202_4
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1997). Computerized neuropsychological tests in the early detection of dementia: Prospective findings. *Journal of the International Neuropsychological Society*, *3*(2), 139-146.
- Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (2002). Paired associate performance in the early detection of DAT. *Journal of the International Neuropsychological Society*, *8*(1), 58-71. doi: 10.1017.S1355617701020069
- Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry*, *11*, 329-336. doi: 10.1002/(sici)1099-1166(199604)11:4%3c329::aid-gps453%3e3.3.co;2-y
- Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2011). Montreal Cognitive Assessment (MoCA): Normative study for the Portuguese population. *Journal of Clinical and Experimental Neuropsychology*, *33*(9), 989-996. doi: 10.1080/13803395.2011.589374
- Ganguli, M., Snitz, B. E., Lee, C.-W., Vanderbilt, J., Saxton, J. A., & Chang, C.-C. H. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: The Monongahela–Youghiogheny Healthy Aging Team (MYHAT). *Aging & Mental Health*, *14*(1), 100-107. doi: 10.1080/13607860903071014
- Gaugler, J. E., Duval, S., Anderson, K. A., & Kane, R. L. (2007). Predicting nursing home admission in the U.S.: A meta-analysis. *BMC Geriatrics*, *7*(13), 1-14. doi: 10.1186/1471-2318-7-13
- Gomez, R., & White, D. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology*, *21*(8), 771-775. doi: 10.1016/j.acn.2006.06.012
- Gonçalves, C., Pinho, M. S., Cruz, V., Gens, H., Oliveira, F., Pais, J., ... Santos, J. M. (2016). Portuguese version of Wechsler Memory Scale–3rd edition's utility with demented elderly adults. *Applied Neuropsychology: Adult*, 1-14. doi: 10.1080/23279095.2015.1135440
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2012). *Portuguese norms of the Semantic Verbal Fluency Test and Naming Objects and Fingers Task for older persons living in retirement homes*. Unpublished manuscript.
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016a). *Construct and concurrent validity of the Cambridge neuropsychological automated tests for assessment of dementia in Portuguese older adults*. Manuscript submitted for publication.

- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016b). Effects of socio-demographic variables on performance on the Cambridge neuropsychological automated tests for the assessment of dementia and Portuguese norms for older adults living in retirement homes. *The Clinical Neuropsychologist*, 30(2), 284-317. doi: 10.1080/13854046.2016.1156745
- Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016c). Test-retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, 23(4), 251-263. doi: 10.1080/23279095.2015.1053889
- Gould, R. L., Brown, R. G., Owen, A. M., Bullmore, E. T., Williams, S. C., & Howard, R. J. (2005). Functional neuroanatomy of successful Paired Associate Learning in Alzheimer's disease. *American Journal of Psychiatry*, 162(11), 2049-2060. doi: 10.1176/appi.ajp.162.11.2049
- Green, C. E., Chen, C. E., Helms, J. E., & Henze, K. T. (2011). Recent reliability reporting practices in Psychological Assessment: Recognizing the people behind the data. *Psychological Assessment*, 23(3), 656-669. doi: 10.1037/a0023089
- Gualtieri, C. T., & Johnson, L. G. (2005). Neurocognitive testing supports a broader concept of mild cognitive impairment. *American Journal of Alzheimer's Disease and Other Dementias*, 20(6), 359-366. doi: 10.1177/153331750502000607
- Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology*, 21(7), 623-643. doi: 10.1016/j.acn.2006.05.007
- Guerreiro, M. (1998). *Contributo da neuropsicologia para o estudo das demências*. (Tese de doutoramento não publicada), Faculdade de Medicina de Lisboa, Lisboa, Portugal.
- Guerreiro, M. (2005). Avaliação neuropsicológica das demências degenerativas. In A. Castro-Caldas & A. Mendonça (Eds.), *A doença de Alzheimer e outras demências em Portugal* (pp. 83-109). Lisboa, Portugal: Lidel, edições técnicas, lda.
- Guerreiro, M., Fonseca, S., Barreto, J., & Garcia, C. (2008). Escala de Avaliação da Doença de Alzheimer (ADAS). In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* (2ª ed., pp. 41-68). Lisboa, Portugal: Novartis.
- Herrera-Guzmán, I., Gudayol-Ferré, E., Herrera-Guzmán, D., Guàrdia-Olmos, J., Hinojosa-Calvo, E., & Herrera-Abarca, J. E. (2009). Effects of selective serotonin reuptake and dual serotonergic-noradrenergic reuptake treatments on memory and mental processing speed in patients with major depressive disorder. *Journal of Psychiatric Research*, 43(9), 855-863. doi: 10.1016/j.jpsychires.2008.10.015
- Herrera-Guzmán, I., Herrera-Abarca, J. E., Gudayol-Ferré, E., Herrera-Guzmán, D., Gómez-Carbajal, L., Peña-Olvira, M., ... Joan, G.-O. (2010). Effects of selective serotonin reuptake and dual serotonergic-noradrenergic reuptake treatments on attention and executive functions in patients with major depressive disorder. *Psychiatry Research*, 177(3), 323-329. doi: 10.1016/j.psychres.2010.03.006
- Hogervorst, E., De Jager, C., Budge, M., & Smith, A. D. (2004). Serum levels of estradiol and testosterone and performance in different cognitive domains in healthy elderly men and women. *Psychoneuroendocrinology*, 29(3), 405-421.
- Huxley, T. H. (n.d.). BrainyQuote.com. Retrieved from <http://www.brainyquote.com/quotes/quotes/t/thomashux1592661.html>
- Instituto Nacional de Estatística. (2001). *Óbitos pela doença de Alzheimer 2000*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e

- Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=70753&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2005). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2005*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=73455&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2006). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2006*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=74055&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2007). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2007*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=5582190&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2008). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2008*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.anacom.pt/streaming/inq_familias_ineumic2008.pdf?contentId=726858&field=ATTACHED_FILE.
- Instituto Nacional de Estatística. (2009). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2009*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=56910276&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2010). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2010*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=83386604&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2011). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2011*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=107940220&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2012a). *Censos 2011 resultados definitivos - Portugal*. Lisboa, Portugal: Autor. Recuperado de http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=ine_censos_publicacao_det&contexto=pu&PUBLICACOESpub_boui=73212469&PUBLICACOESmodo=2&selTab=tab1&pcensos=61969554.
- Instituto Nacional de Estatística. (2012b). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2012*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaquas&DESTAQUE_Sdest_boui=133548146&DESTAQUESmodo=2.

- Instituto Nacional de Estatística. (2013). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2013*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=206557735&att_display=n&att_download=y.
- Instituto Nacional de Estatística. (2014). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2014*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUE_Sdest_boui=211422735&DESTAQUEStema=00&DESTAQUESmodo=2.
- Instituto Nacional de Estatística. (2015a). *Causas de morte 2013*. Lisboa, Portugal: Autor. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=229905695&PUBLICACOESStema=00&PUBLICACOESmodo=2.
- Instituto Nacional de Estatística. (2015b). *Estatísticas demográficas 2014*. Lisboa, Portugal: Autor. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=139659&PUBLICACOESmodo=2.
- Instituto Nacional de Estatística. (2015c). *Inquérito à utilização de tecnologias da informação e da comunicação pelas famílias 2015*. Lisboa, Portugal: Instituto Nacional de Estatística, Serviço de Comunicação e Imagem. Recuperado de http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUE_Sdest_boui=224732374&DESTAQUESmodo=2.
- Iverson, G. L. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology*, 16(2), 183-191. doi: 10.1093/arclin/16.2.183
- Iverson, G. L., Brooks, B. L., Ashton, V. L., Johnson, L. G., & Gualtieri, C. T. (2009). Does familiarity with computers affect computerized neuropsychological test performance? *Journal of Clinical and Experimental Neuropsychology*, 31(5), 594-604. doi: 10.1080/13803390802372125
- Jack, C. R., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., ... Phelps, C. H. (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 257-262. doi: 10.1016/j.jalz.2011.03.004
- Jellinger, K. A. (2006). Clinicopathological analysis of dementia disorders in the elderly - An update. *Journal of Alzheimer's Disease*, 9(Suppl. 3), 61-70.
- Jones, G. M. M., Sahakian, B. J., Levy, R., Warburton, D. M., & Gray, J. A. (1992). Effects of acute subcutaneous nicotine on attention, information processing and short-term memory in Alzheimer's disease. *Psychopharmacology*, 108(4), 485-494. doi: 10.1007/bf02247426
- Juncos-Rabadán, O., Facal, D., Pereiro, A. X., & Lojo-Seoane, C. (2014). Visual memory profiling with CANTAB in mild cognitive impairment (MCI) subtypes. *International Journal of Geriatric Psychiatry*, 29(10), 1040-1048. doi: 10.1002/gps.4095
- Juncos-Rabadán, O., Pereiro, A. X., Facal, D., Reboredo, A., & Lojo-Seoane, C. (2014). Do the Cambridge Neuropsychological Test Automated Battery episodic memory

- measures discriminate amnesic mild cognitive impairment? *International Journal of Geriatric Psychiatry*, 29(6), 602-609. doi: 10.1002/gps.4042
- Junkkila, J., Oja, S., Laine, M., & Karrasch, M. (2012). Applicability of the CANTAB-PAL computerized memory test in identifying amnesic mild cognitive impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 34(2), 83-89. doi: 10.1159/000342116
- Kahana, E. (2006). Institutionalization. In R. Schulz (Ed.), *The encyclopedia of aging* (4th ed., Vol. I, pp. 595-600). New York, NY: Springer Publishing Company, Inc.
- Khan, U. A., Liu, L., Provenzano, F. A., Berman, D. E., Profaci, C. P., Sloan, R., ... Small, S. A. (2014). Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nature Neuroscience*, 17(2), 304-311. doi: 10.1038/nn.3606
- Kiddle, S. J., Steves, C. J., Mehta, M., Simmons, A., Xu, X., Newhouse, S., ... Dobson, R. J. B. (2015). Plasma protein biomarkers of Alzheimer's disease endophenotypes in asymptomatic older twins: early cognitive decline and regional brain volumes. *Translational Psychiatry*, 5(6), 584-589. doi: 10.1038/tp.2015.78
- Kim, C., Lee, J. Y., Ha, T. H., Choi, J. H., Yu, E.-M., Lee, S. J., ... Park, M. Y. (2009). The usefulness of the Cambridge Neuropsychological Test Automated Battery (CANTAB) for assessing cognitive functions in the elderly: A pilot study. *Journal of the Korean Geriatrics Society*, 13(2), 69-78. doi: 10.4235/jkgs.2009.13.2.69
- Klekociuk, S. Z., Summers, J. J., Vickers, J. C., & Summers, M. J. (2014). Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. *European Journal of Neurology*, 21(10), 1330-1336. doi: 10.1111/ene.12488
- Klekociuk, S. Z., & Summers, M. J. (2013). The self-fulfilling prophecy of episodic memory impairment in mild cognitive impairment: Do episodic memory deficits identified at classification remain evident when later examined with different memory tests? *Neurology Research International*, 2013, 1-6. doi: 10.1155/2013/437013
- Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. London, UK: Methuen & Co. Ltd.
- Kløve, H. (1963). Clinical neuropsychology. In F. M. Forster (Ed.), *The medical clinics of North America* (pp. 1647-1658). New York, NY: Saunders.
- Larrabee, G. J., & Crook, T. H. (1996). Computers and memory. In L. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed., pp. 102-117). New York, NY: Oxford University Press.
- Lee, A. C. H., Rahman, S., Hodges, J. R., Sahakian, B. J., & Graham, K. S. (2003). Associative and recognition memory for novel objects in dementia: Implications for diagnosis. *European Journal of Neuroscience*, 18(6), 1660-1670. doi: 10.1046/j.1460-9568.2003.02883.x
- Leitão, O. R., Nina, A., & Monteiro, I. (2008). Escala de Deterioração Global (GDS). In A. Mendonça & M. Guerreiro (Eds.), *Escalas e testes na demência* (2ª ed., pp. 12-16). Lisboa, Portugal: Novartis.
- Lemaire, P. (2010). Cognitive strategy variations during aging. *Current Directions in Psychological Science*, 19(6), 363-369. doi: 10.1177/0963721410390354
- Lenahan, M. E., Summers, M. J., Saunders, N. L., Summers, J. J., & Vickers, J. C. (2016). Does the Cambridge Automated Neuropsychological Test Battery

- (CANTAB) distinguish between cognitive domains in healthy older adults? *Assessment*, 23(2), 163-172. doi: 10.1177/1073191115581474
- Lesk, V. E., Honey, T. E., & de Jager, C. A. (2009). The effect of recent consumption of caffeine-containing foodstuffs on neuropsychological tests in the elderly. *Dementia and Geriatric Cognitive Disorders*, 27(4), 322-328. doi: 10.1159/000207445
- Levy-Gigi, E., Kelemen, O., Gluck, M. A., & Kéri, S. (2011). Impaired context reversal learning, but not cue reversal learning, in patients with amnesic mild cognitive impairment. *Neuropsychologia*, 49(12), 3320-3326. doi: 10.1016/j.neuropsychologia.2011.08.005
- Lewis, R., & Rennick, P. M. (1979). *Manual for the Repeatable Cognitive-Perceptual-Motor Battery*. Grosse Point, MI: Axon.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford, NY: Oxford University Press.
- Lowe, C., & Rabbitt, P. (1998). Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. *Neuropsychologia*, 36(9), 915-923. doi: 10.1016/S0028-3932(98)00036-0
- Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology*, 22(3), 595-624. doi: 10.1207/s15326942dn2203_3
- Maassen, G. H., Bossema, E., & Brand, N. (2009). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology*, 31(3), 339-352. doi: 10.1080/13803390802169059
- Maia, L., Correia, C., & Leite, R. (2009). *Avaliação e intervenção neuropsicológica: Estudos de casos e instrumentos*. Lisboa, Portugal: Lidel - edições técnicas, lda.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, 125(3), 624-639. doi: 10.1093/brain/awf049
- Manias, E., Kusljic, S., Berry, C., Brown, E., Bryce, E., Cliffe, J., & Smykowsky, A. (2015). Use of the Screening Tool of Older Person's Prescriptions (STOPP) in older people admitted to an Australian hospital. *Australasian Journal on Ageing*, 34(1), 15-20. doi: 10.1111/ajag.12054
- Marôco, J. (2014a). *Análise de equações estruturais: Fundamentos teóricos, software e aplicações* (2ª ed.). Pêro Pinheiro, Portugal: Report Number - Análise e Gestão de Informação, Lda.
- Marôco, J. (2014b). *Análise estatística com o SPSS statistics* (6ª ed.). Pêro Pinheiro, Portugal: Report Number - Análise e Gestão de Informação, Lda.
- Maust, D. T., Oslin, D. W., & Marcus, S. C. (2014). Effect of age on the profile of psychotropic users: Results from the 2010 National Ambulatory Medical Care Survey. *Journal of the American Geriatrics Society*, 62(2), 358-364. doi: 10.1111/jgs.12640
- Maylor, E. A., Simpson, E. E., Secker, D. L., Meunier, N., Andriollo-Sanchez, M., Polito, A., ... Coudray, C. (2006). Effects of zinc supplementation on cognitive function in healthy middle-aged and older adults: The ZENITH study. *British Journal of Nutrition*, 96(4), 752-760. doi: 10.1079/BJN20061911
- McGaughy, J., Dalley, J. W., Morrison, C. H., Everitt, B. J., & Robbins, T. W. (2002). Selective behavioral and neurochemical effects of cholinergic lesions produced by

- intrabasal infusions of 192 IgG-saporin on attentional performance in a five-choice serial reaction time task. *The Journal of Neuroscience*, 22(5), 1905-1913.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., ... Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263-269. doi: 10.1016/j.jalz.2011.03.005
- McSweeney, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). "T scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *Clinical Neuropsychologist*, 7(3), 300-312. doi: 10.1080/13854049308401901
- Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114(3), 449-458. doi: 10.1037/0033-2909.114.3.449
- Meinz, E. J., & Salthouse, T. A. (1998). Is age kinder to females than to males? *Psychonomic Bulletin & Review*, 5(1), 56-70. doi: 10.3758/bf03209457
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21(11), 1078-1085. doi: 10.1002/gps.1610
- Miquel, M. D. C., Cuervo, M. S., Silveira, E. D., Machuca, I. S., González-Blázquez, S., Errasquin, B. M., & Cruz-Jentoft, A. J. (2010). Potentially inappropriate drug prescription in older subjects across health care settings. *European Geriatric Medicine*, 1(1), 9-14. doi: 10.1016/j.eurger.2009.12.002
- Mitchell, J., Arnold, R., Dawson, K., Nestor, P. J., & Hodges, J. R. (2009). Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm. *Journal of Neurology*, 256(9), 1500-1509. doi: 10.1007/s00415-009-5152-0
- Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin*, 19(3), 448-450.
- Moreira, P. I., & Oliveira, C. R. (2005). Fisiopatologia da doença de Alzheimer e de outras demências. In A. Castro-Caldas & A. Mendonça (Eds.), *A doença de Alzheimer e outras demências em Portugal* (pp. 41-60). Lisboa, Portugal: Lidel, edições técnicas, lda.
- Morgado, J., Rocha, C. S., Maruta, C., Guerreiro, M., & Martins, I. P. (2009). Novos valores normativos do Mini-Mental State Examination. *Sinapse*, 2(9), 10-16.
- Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 51(6), 757-766. doi: 10.1136/jnnp.51.6.757
- Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1987). Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer-type and Parkinson's disease. In S. M. Stahl, S. D. Iversen & E. C. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21-36). Oxford, UK: Oxford University Press.
- Murphy, F. C., Sahakian, B. J., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., & Paykel, E. S. (1999). Emotional bias and inhibitory control processes in

- mania and depression. *Psychological Medicine*, 29(6), 1307-1321. doi: 10.1017/s0033291799001233
- Murphy, M., & O'Leary, E. (2009). Depression, cognitive reserve and memory performance in older adults. *International Journal of Geriatric Psychiatry*, 25(7), 665-671. doi: 10.1002/gps.2404
- National Institute for Health and Care Excellence. (2011). *Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease*. (NICE Technology Appraisal Guidance No. 217). London, UK: Author. Retrieved from <http://www.nice.org.uk/guidance/ta217/resources/guidance-donepezil-galantamine-rivastigmine-and-memantine-for-the-treatment-of-alzheimers-disease-pdf>.
- Need, A. C., Attix, D. K., McEvoy, J. M., Cirulli, E. T., Linney, K. L., Hunt, P., ... Goldstein, D. B. (2009). A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Human Molecular Genetics*, 18(23), 4650-4661. doi: 10.1093/hmg/ddp413
- Nelson, H. E., & Willison, J. (1991). *The National Adult Reading Test (NART): Test manual* (2nd ed.). Windsor, UK: NFER-Nelson Publishing Company.
- Nihtilä, E. K., Martikainen, P. T., Koskinen, S. V. P., Reunanen, A. R., Noro, A. M., & Häkkinen, U. T. (2007). Chronic conditions and the risk of long-term institutionalization among older people. *The European Journal of Public Health*, 18(1), 77-84. doi: 10.1093/eurpub/ckm025
- Olde Dubbelink, K. T. E., Hillebrand, A., Twisk, J. W. R., Deijen, J. B., Stoffers, D., Schmand, B. A., ... Berendse, H. W. (2014). Predicting dementia in Parkinson disease by combining neurophysiologic and cognitive markers. *Neurology*, 82(3), 263-270. doi: 10.1212/wnl.0000000000000034
- Oosterman, J. M., Vogels, R. L. C., van Harten, B., Gouw, A. A., Scheltens, P., Poggesi, A., ... Scherder, E. J. A. (2008). The role of white matter hyperintensities and medial temporal lobe atrophy in age-related executive dysfunctioning. *Brain and Cognition*, 68(2), 128-133. doi: 10.1016/j.bandc.2008.03.006
- Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, 28(10), 1021-1034. doi: 10.1016/0028-3932(90)90137-d
- Owen, A. M., Evans, A. C., & Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: A positron emission tomography study. *Cerebral Cortex*, 6(1), 31-38. doi: 10.1093/cercor/6.1.31
- Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, 33(1), 1-24. doi: 10.1016/0028-3932(94)00098-a
- Periquito, C. M. N., Silva, P. C., Oliveira, P., Carneiro, C., Fernandes, A. I., & Costa, F. A. (2014). Revisão da medicação em idosos institucionalizados: Aplicação dos critérios STOPP e START. *Revista Portuguesa de Farmacoterapia*, 6, 211-220.
- Plati, M. C. F., Covre, P., Lukasova, K., & Macedo, E. C. (2006). Depressive symptoms and cognitive performance of the elderly: Relationship between institutionalization and activity programs. *Revista Brasileira de Psiquiatria*, 28(2), 118-121. doi: 10.1590/s1516-44462006000200008

- Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia*, 9(1), 63-75. doi: 10.1016/j.jalz.2012.11.007
- Rabbitt, P., & Lowe, C. (2000). Patterns of cognitive ageing. *Psychological Research*, 63(3-4), 308-316. doi: 10.1007/s004269900009
- Rabbitt, P., Lowe, C., & Shilling, V. (2001). Frontal tests and models for cognitive ageing. *The European Journal of Cognitive Psychology*, 13(1), 5-28. doi: 10.1080/09541440042000197
- Racinais, S., Gaoua, N., & Grantham, J. (2008). Hyperthermia impairs short-term memory and peripheral motor drive transmission. *The Journal of Physiology*, 586(19), 4751-4762. doi: 10.1113/jphysiol.2008.157420
- Rahman, S., Robbins, T. W., & Sahakian, B. J. (1999). Comparative cognitive neuropsychological studies of frontal lobe function: Implications for therapeutic strategies in frontal variant frontotemporal dementia. *Dementia and Geriatric Cognitive Disorders*, 10(Suppl. 1), 15-28. doi: 10.1159/000051207
- Rapp, M. A., Schnaider-Beeri, M., Sano, M., Silverman, J. M., & Haroutunian, V. (2005). Cross-domain variability of cognitive performance in very old nursing home residents and community dwellers: Relationship to functional status. *Gerontology*, 51(3), 206-212. doi: 10.1159/000083995
- Reeves, S. J., Grasby, P. M., Howard, R. J., Bantick, R. A., Asselin, M.-C., & Mehta, M. A. (2005). A positron emission tomography (PET) investigation of the role of striatal dopamine (D2) receptor availability in spatial cognition. *NeuroImage*, 28(1), 216-226. doi: 10.1016/j.neuroimage.2005.05.034
- Reisberg, B., Ferris, S. H., de Leon, M. J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry*, 139(9), 1136-1139.
- Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8(3), 271-276. doi: 10.2466/pms.1958.8.3.271
- Rey, A. (2002). *Teste de Cópia de Figuras Complexas* (2ª ed.). Lisboa, Portugal: Cegoc-Tea.
- Ribbe, M. W., Ljunggren, G., Steel, K., Topinková, E., Hawes, C., Ikegami, N., ... Jónnson, P. V. (1997). Nursing homes in 10 nations: A comparison between countries and settings. *Age and Ageing*, 26(Suppl. 2), 3-12. doi: 10.1093/ageing/26.suppl_2.3
- Riekkinen, M., Laakso, M. P., Jäkälä, P., & Riekkinen, P. (1999). Clonidine impairs sustained attention and memory in Alzheimer's disease. *Neuroscience*, 92(3), 975-982. doi: 10.1016/s0306-4522(99)00070-6
- Riekkinen, P., & Riekkinen, M. (1999). THA improves word priming and clonidine enhances fluency and working memory in Alzheimer's disease. *Neuropsychopharmacology*, 20(4), 357-364. doi: 10.1016/s0893-133x(98)00093-1
- Robbins, T. W., James, M., Lange, K. W., Owen, A. M., Quinn, N. P., & Marsden, C. D. (1992). Cognitive performance in multiple system atrophy. *Brain*, 115(1), 271-291. doi: 10.1093/brain/115.1.271
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal

- volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society*, 4(5), 474-490. doi: 10.1017/s1355617798455073
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): A factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitive Disorders*, 5(5), 266-281. doi: 10.1159/000106735
- Robbins, T. W., Semple, J., Kumar, R., Truman, M. I., Shorter, J., Ferraro, A., ... Matthews, K. (1997). Effects of scopolamine on delayed-matching-to-sample and paired associates tests of visual memory and learning in human subjects: comparison with diazepam and implications for dementia. *Psychopharmacology*, 134(1), 95-106. doi: 10.1007/s002130050430
- Roiser, J. P., McLean, A., Ogilvie, A. D., Blackwell, A. D., Bamber, D. J., Goodyer, I., ... Sahakian, B. J. (2005). The subjective and cognitive effects of acute phenylalanine and tyrosine depletion in patients recovered from depression. *Neuropsychopharmacology*, 30, 775-785. doi: 10.1038/sj.npp.1300659
- Roque, D. T., Teixeira, R. A., Zachi, E. C., & Ventura, D. F. (2011). The use of the Cambridge Neuropsychological Test Automated Battery (CANTAB) in neuropsychological assessment: Application in Brazilian research with control children and adults with neurological disorders. *Psychology and Neuroscience*, 4(2), 255-265. doi: 10.3922/j.psns.2011.2.011
- Sahakian, B. J., & Coull, J. T. (1993). Tetrahydroaminoacridine (THA) in Alzheimer's disease: An assessment of attentional and mnemonic function using CANTAB. *Acta Neurologica Scandinavica*, 88(Suppl. 149), 29-35. doi: 10.1111/j.1600-0404.1993.tb04251.x
- Sahakian, B. J., Downes, J. J., Eagger, S., Everden, J. L., Levy, R., Philpot, M. P., ... Robbins, T. W. (1990). Sparing of attentional relative to mnemonic function in a subgroup of patients with dementia of the Alzheimer type. *Neuropsychologia*, 28(11), 1197-1213. doi: 10.1016/0028-3932(90)90055-s
- Sahakian, B. J., Elliott, R., Low, N., Mehta, M., Clark, R. T., & Pozniak, A. L. (2009). Neuropsychological deficits in tests of executive function in asymptomatic and symptomatic HIV-1 seropositive men. *Psychological Medicine*, 25(6), 1233-1246. doi: 10.1017/s0033291700033201
- Sahakian, B. J., Jones, G., Levy, R., Gray, J., & Warburton, D. (1989). The effects of nicotine on attention, information processing, and short-term memory in patients with dementia of the Alzheimer type. *The British Journal of Psychiatry*, 154(6), 797-800. doi: 10.1192/bjp.154.6.797
- Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain*, 111(3), 695-718. doi: 10.1093/brain/111.3.695
- Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, 85(7), 399-402. doi: 10.1177/014107689208500711
- Sahgal, A., Lloyd, S., Wray, C. J., Galloway, P. H., Robbins, T. W., Sahakian, B. J., ... Edwardson, J. A. (1992). Does visuospatial memory in senile dementia of the Alzheimer type depend on the severity of the disorder? *International Journal of Geriatric Psychiatry*, 7(6), 427-436. doi: 10.1002/gps.930070608

- Sahgal, A., McKeith, I. G., Galloway, P. H., Tasker, N., & Steckler, T. (1995). Do differences in visuospatial ability between senile dementias of the Alzheimer and Lewy body types reflect differences solely in mnemonic function? *Journal of Clinical and Experimental Neuropsychology*, 17(1), 35-43. doi: 10.1080/13803399508406579
- Sales, F. (2005). Electroencefalograma. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 38-39). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, 13(4), 140-144. doi: 10.1111/j.0963-7214.2004.00293.x
- Santana, I. (2005a). Avaliação neuropsicológica. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 23-29). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I. (2005b). A doença de Alzheimer e outras demências: Diagnóstico diferencial. In A. Castro-Caldas & A. Mendonça (Eds.), *A doença de Alzheimer e outras demências em Portugal* (pp. 61-82). Lisboa, Portugal: Lidel, edições técnicas, lda.
- Santana, I. (2005c). Estudos de imagem. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 40-42). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I., Baldeiras, I., & Oliveira, C. (2005). Exames laboratoriais. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 35-37). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santana, I., Farinha, F., Freitas, S., Rodrigues, V., & Carvalho, A. (2015). Epidemiologia da demência e da doença de Alzheimer em Portugal: Estimativas da prevalência e dos encargos financeiros com a medicação. *Acta Médica Portuguesa*, 28(2), 182-188.
- Santana, I., & Oliveira, C. (2005). Estudos genéticos. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 43-47). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Santiago, B. (2005). Avaliação psicológica. In I. Santana & L. Cunha (Eds.), *Demência(s): Manual para médicos* (pp. 30-34). Coimbra, Portugal: Faculdade de Medicina, Universidade de Coimbra.
- Saunders, N. L. J., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 32(4), 350-357. doi: 10.1080/13803390903042379
- Saunders, N. L. J., & Summers, M. J. (2011). Longitudinal deficits to attention, executive, and working memory in subtypes of mild cognitive impairment. *Neuropsychology*, 25(2), 237-248. doi: 10.1037/a0021134
- Schatz, P., & Ferris, C. S. (2013). One-month test-retest reliability of the ImPACT test battery. *Archives of Clinical Neuropsychology*, 28(5), 499-504. doi: 10.1093/arclin/act034
- Schweiger, A., Doniger, G. M., Dwolatzky, T., Jaffe, D., & Simon, E. S. (2003). Reliability of a novel computerized neuropsychological battery for mild cognitive impairment. *Acta Neuropsychologica*, 1(4), 407-413. doi: 10.5604/17307503/9168
- Scoriels, L., Barnett, J. H., Murray, G. K., Cherukuru, S., Fielding, M., Cheng, F., ... Jones, P. B. (2011). Effects of modafinil on emotional processing in first episode

- psychosis. *Biological Psychiatry*, 69(5), 457-464. doi: 10.1016/j.biopsych.2010.09.043
- Sexton, C. E., McDermott, L., Kalu, U. G., Herrmann, L. L., Bradley, K. M., Allan, C. L., ... Ebmeier, K. P. (2012). Exploring the pattern and neural correlates of neuropsychological impairment in late-life depression. *Psychological Medicine*, 42(06), 1195-1202. doi: 10.1017/s0033291711002352
- Simões, M. R. (2012). Instrumentos de avaliação psicológica de pessoas idosas: Investigação e estudos de validação em Portugal. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 34(1), 9-33.
- Simões, M. R., & Santana, I. (2015). *Escalas e testes na demência* (3ª ed.). Porto Salvo, Portugal: Novartis.
- Simpson, E. E. A., Hodkinson, C. F., Maylor, E. A., McCormack, J. M., Rae, G., Strain, S., ... Wallace, J. M. W. (2013). Intracellular cytokine production and cognition in healthy older adults. *Psychoneuroendocrinology*, 38(10), 2196-2208. doi: 10.1016/j.psyneuen.2013.04.007
- Simpson, E. E. A., Maylor, E. A., Rae, G., Meunier, N., Andriollo-Sanchez, M., Catasta, G., ... Coudray, C. (2005). Cognitive function in healthy older European adults: The ZENITH study. *European Journal of Clinical Nutrition*, 59(Suppl. 2), S26-S30. doi: 10.1038/sj.ejcn.1602294
- Singer, J. J., MacGregor, A. J., Cherkas, L. F., & Spector, T. D. (2006). Genetic influences on cognitive function using the Cambridge Neuropsychological Test Automated Battery. *Intelligence*, 34(5), 421-428. doi: 10.1016/j.intell.2005.11.005
- Skolimowska, J., Wesierska, M., Lewandowska, M., Szymaszek, A., & Szelag, E. (2011). Divergent effects of age on performance in spatial associative learning and real idiothetic memory in humans. *Behavioural Brain Research*, 218(1), 87-93. doi: 10.1016/j.bbr.2010.11.035
- Smith, P. J., Need, A. C., Cirulli, E. T., Chiba-Falek, O., & Attix, D. K. (2013). A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with “traditional” neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology*, 35(3), 319-328. doi: 10.1080/13803395.2013.771618
- Sousa, L. B., Prieto, G., Vilar, M., Firmino, H., & Simões, M. R. (2015). The Adults and Older Adults Functional Assessment Inventory: A Rasch model analysis. *Research on Aging*, 37(8), 787-814. doi: 10.1177/0164027514564469
- Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., ... Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 280-292. doi: 10.1016/j.jalz.2011.03.003
- Stefanova, E., Ziropadja, L., Stojkovic, T., Stankovic, I., Tomic, A., Jecmenica-Lukic, M., ... Kostic, V. (2015). Mild cognitive impairment in early Parkinson's disease using the movement disorder society task force criteria: Cross-sectional study in Hoehn and Yahr stage 1. *Dementia and Geriatric Cognitive Disorders*, 40(3-4), 199-209. doi: 10.1159/000433421
- Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 20(2), 112-117. doi: 10.1097/01.wad.0000213815.20177.19

- Steves, C. J., Jackson, S. H. D., & Spector, T. D. (2013). Cognitive change in older women using a computerised battery: A longitudinal quantitative genetic twin study. *Behavior Genetics*, *43*(6), 468-479. doi: 10.1007/s10519-013-9612-z
- Stilley, C. S., Bender, C. M., Dunbar-Jacob, J., Sereika, S., & Ryan, C. M. (2010). The impact of cognitive function on medication management: Three studies. *Health Psychology*, *29*(1), 50-55. doi: 10.1037/a0016940
- Stoffers, D., Bosboom, J. L. W., Deijen, J. B., Wolters, E. C., Berendse, H. W., & Stam, C. J. (2007). Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia. *Brain*, *130*(7), 1847-1860. doi: 10.1093/brain/awm034
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662. doi: 10.1037/h0054651
- Summers, M. J., & Saunders, N. J. (2012). Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. *Neuropsychology*, *26*(4), 498-508. doi: 10.1037/a0028576
- Swanson, R., Hodges, J. R., Galton, C. J., Semple, J., Michael, A., Dunn, B. D., ... Sahakian, B. J. (2001). Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks. *Dementia and Geriatric Cognitive Disorders*, *12*(4), 265-280. doi: 10.1159/000051269
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Taffe, M., Weed, M., Gutierrez, T., Davis, S., & Gold, L. (2002). Differential muscarinic and NMDA contributions to visuo-spatial paired-associate learning in rhesus monkeys. *Psychopharmacology*, *160*(3), 253-262. doi: 10.1007/s00213-001-0954-5
- Talpos, J. C., Winters, B. D., Dias, R., Saksida, L. M., & Bussey, T. J. (2009). A novel touchscreen-automated paired-associate learning (PAL) task sensitive to pharmacological manipulation of the hippocampus: A translational rodent model of cognitive impairments in neurodegenerative disease. *Psychopharmacology*, *205*(1), 157-168. doi: 10.1007/s00213-009-1526-3
- Torgersen, J., Flaatten, H., Engelsen, B. A., & Gramstad, A. (2012). Clinical validation of Cambridge Neuropsychological Test Automated Battery in a Norwegian epilepsy population. *Journal of Behavioral and Brain Science*, *2*(1), 108-116. doi: 10.4236/jbbs.2012.21013
- Tornatore, J. B., Hill, E., Laboff, J. A., & McGann, M. E. (2005). Self-administered screening for mild cognitive impairment: Initial validation of a computerized test battery. *Journal of Neuropsychiatry*, *17*(1), 98-105. doi: 10.1176/appi.neuropsych.17.1.98
- Turner, D. C., Clark, L., Dowson, J., Robbins, T. W., & Sahakian, B. J. (2004). Modafinil improves cognition and response inhibition in adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *55*(10), 1031-1040. doi: 10.1016/j.biopsych.2004.02.008
- Turner, D. C., Robbins, T. W., Clark, L., Aron, A. R., Dowson, J., & Sahakian, B. J. (2003). Cognitive enhancing effects of modafinil in healthy volunteers. *Psychopharmacology*, *165*, 260-269. doi: 10.1007/s00213-002-1250-8

- Ubeda, A., Ferrándiz, M. L., Maicas, N., Gomez, C., Bonet, M., & Peris, J. E. (2012). Potentially inappropriate prescribing in institutionalised older patients in Spain: The STOPP-START criteria compared with the Beers criteria. *Pharmacy Practice (Internet)*, *10*(2), 83-91.
- United Nations. (2015). *World population prospects: The 2015 revision, world population 2015 wallchart, ST/ESA/SER.A/378*. New York, NY: United Nations, Department of Economic and Social Affairs, Population Division. Retrieved from http://esa.un.org/unpd/wpp/publications/Files/World_Population_2015_Wallchart.pdf.
- Wechsler, D. (1987). *Wechsler Memory Scale - Revised manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2008a). *Escala de Inteligência de Wechsler para Adultos – Terceira edição (WAIS-III): Manual técnico*. Lisboa, Portugal: Cegoc-Tea.
- Wechsler, D. (2008b). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual de administração e cotação*. Lisboa, Portugal: Cegoc-Tea.
- Wechsler, D. (2008c). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual técnico*. Lisboa, Portugal: Cegoc-Tea.
- Wild, K., Howieson, D., Webbe, F., Seelye, A., & Kaye, J. (2008). Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's & Dementia*, *4*(6), 428-437. doi: 10.1016/j.jalz.2008.07.003
- Wimo, A., Jönsson, L., Bond, J., Prince, M., & Winblad, B. (2013). The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia*, *9*(1), 1-11. doi: 10.1016/j.jalz.2012.11.006
- Yang, L., Unverzagt, F. W., Jin, Y., Hendrie, H. C., Liang, C., Hall, K. S., ... Gao, S. (2012). Normative data for neuropsychological tests in a rural elderly Chinese cohort. *The Clinical Neuropsychologist*, *26*(4), 641-653. doi: 10.1080/13854046.2012.666266
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*(1), 37-49. doi: 10.1016/0022-3956(82)90033-4
- Zygouris, S., & Tsolaki, M. (2014). Computerized cognitive testing for older adults: A review. *American Journal of Alzheimer's Disease and Other Dementias*, *30*(1), 13-28. doi: 10.1177/1533317514522852