



UC/FPCE — 2009

Universidade de Coimbra
Faculdade de Psicologia e de Ciências da Educação

Test of Memory Malingering (TOMM): Estudos de validação em adultos idosos com Declínio Cognitivo Ligeiro*

Susete Pereira Fernandes
(*e-mail*: susetefern@hotmail.com)

Dissertação de Mestrado Integrado em **Psicologia Clínica e da Saúde**, sub-especialização em **Psicogerontologia Clínica**, sob a orientação de **Professor Doutor Mário R. Simões** (Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra) e da **Dr.ª Maria Teresa Gonçalves** (Serviço de Neurologia do Hospital Geral do Centro Hospitalar de Coimbra).

* Estudo inserido no âmbito do Projecto "Validação de Provas de Memória e de Inventários de Avaliação Funcional e da Qualidade de Vida", financiado pela Fundação Calouste Gulbenkian (Proc. 74569).

“O valor das coisas não está no tempo que elas duram, mas na intensidade com que acontecem. Por isso existem momentos inesquecíveis, coisas inexplicáveis e pessoas incomparáveis.”

Fernando Pessoa

Test of Memory Malingering (TOMM): Estudos de validação em adultos idosos com Declínio Cognitivo Ligeiro

As questões da “simulação” e do “esforço insuficiente” – acto intencional de fingir ou exagerar sintomas para obter ganhos pessoais – constituem tópicos de crescente importância em Neuropsicologia em consequência dos cada vez mais frequentes casos envolvendo patologia neurológica. Em contextos médico-legais, os (neuro)psicólogos podem ser chamados a tribunal para testemunhar, como peritos, sobre a autenticidade dos sintomas comunicados, devendo, nestas circunstâncias, testemunhar a validade dos resultados obtidos nos testes (neuro)psicológicos.

É neste âmbito que surge o interesse pelos testes de validade de sintomas (TVS), instrumentos especificamente desenvolvidos para detectar o esforço insuficiente e/ou reduzido. O *Test of Memory Malingering (TOMM; Tombaugh, 1996)* é um TVS com estudos sistemáticos em diferentes países e com vários grupos clínicos.

No presente estudo é analisado o perfil das respostas ao TOMM em 2 grupos de adultos idosos: grupo de Controlo (N=20) e grupo de sujeitos com Declínio Cognitivo Ligeiro (DCL) (N=20).

Os resultados no TOMM não se associam a variáveis como idade, escolaridade, género, situação laboral, capacidade intelectual verbal, sintomatologia depressiva e capacidade funcional. Os resultados obtidos são sugestivos da utilidade do TOMM na avaliação do esforço insuficiente em idosos, inclusive naqueles que manifestam algum grau de declínio cognitivo, corroborando quer no grupo de controlo, quer no grupo DCL, o valor do ponto de corte (≥ 45 acertos) proposto por Tombaugh (1996).

Palavras-chave: *Test of Memory Malingering (TOMM)*; Simulação; Esforço Insuficiente; Testes de Validade de Sintomas; Declínio Cognitivo Ligeiro; Validade; Neuropsicologia; Adultos Idosos.

Test of Memory Malingering (TOMM): Validation studies in old adults with Mild Cognitive Impairment

The issues of “malingering” and “insufficient effort” – the intentional act of pretending or exaggerating symptoms to obtain personal gain – are topics of growing importance in Neuropsychology by the increasingly frequent forensic cases involving neurological pathology. In medico-legal contexts, the neuropsychologists may be called to court to testify, as experts, on the authenticity of reported symptoms and, therefore, testify the validity of the neuropsychological tests results.

It's in this context that arises the interest in symptom validity tests (SVT), instruments specifically developed to detect insufficient and/or reduced effort. The Test of Memory Malingering (TOMM, Tombaugh, 1996) is a SVT with systematic studies in different countries and with various clinical groups.

In this study we analyze the profile of responses to the TOMM in 2 groups of older adults: control group (N=20); group of subjects with Mild Cognitive Impairment (MCI) (N=20).

TOMM results doesn't appear associated with variables like age, education, gender, employment status, intellectual verbal capacity, depression and functional capacity. The obtained results suggest the usefulness of the TOMM in assessment of insufficient effort in the elderly, including in those who express some degree of cognitive decline, supporting both in the control group, either in the DCL group, the value of the cutoff point (≥ 45 scores) proposed by Tombaugh (1996).

Key-Words: Test of Memory Malingering (TOMM); malingering; insufficient effort; symptom validity test; Mild Cognitive Impairment; validity; Neuropsychology; elderly.

Agradecimentos

Ao Professor Doutor Mário R. Simões, pelo exemplo de sabedoria que espero trilhar nos meus caminhos.

À Dr.^a Maria Teresa Gonçalves, estas linhas não são suficientes para expressar a minha gratidão. Obrigado por me ter ensinado a “sentir” os doentes e por me ensinar que podemos marcar a diferença na vida daqueles que nos pedem ajuda profissional. E porque a amizade pode surgir inesperadamente ...

À Mestre Manuela Vilar, obrigado por me ensinar que os psicólogos devem ser sempre melhores do que os testes que aplicam.

À Professora Doutora Margarida Lima, pela capacidade que tem de me fazer rir dos meus “maiores” problemas.

À minha madrinha, por tudo o que sou... pelo amor incondicional e indefinível, pela preocupação e incentivos constantes e, sobretudo, pelo sacrifício realizado para que eu pudesse prosseguir e ser mais...

Aos meus pais, obrigado por estarem sempre presentes e por esforçarem-se por entender e aceitar as minhas escolhas.

Aos colegas de Psicogerontologia Clínica, obrigado pela partilha de experiências e pela compreensão exibida nos momentos mais difíceis deste percurso tumultuoso. Principalmente à Catarina, à Maria Emília e à Patrícia, obrigado pelas palavras sábias nos momentos “exactos”, espero que continuem a partilhar comigo a vossa amizade.

À Ana Lúcia, uma amiga para toda a vida. Obrigado, por tudo aquilo que sabes ...

À Lílíana, obrigado pelo enorme empurrão inicial (e por todas as ajudas sempre que as pedi) ...

Às minhas amigas, por esperarem por mim todo este tempo ...

A ti, Emanuel, não vou agradecer porque o amor não se agradece, retribui-se. O melhor agradecimento que posso fazer é retribuir todo o carinho, dedicação, e compreensão. Obrigado por tornares os meus dias mais bonitos.

Um especial agradecimento a todos os sujeitos que aceitaram participar no presente estudo, pela disponibilidade que demonstraram... E, principalmente, ... por me terem permitido aprender com as suas experiências de vida.

Índice

Introdução	1
I – Enquadramento conceptual	2
1. Definição de simulação	2
2. Métodos na detecção da “simulação”	5
2.1. Testes de Validade de Sintomas (TVS)	6
2.1.1. <i>Test of Memory Malingering</i> (TOMM)	8
2.1.1.1. Estudos com adultos idosos	11
2.1.1.2. Estudos realizados em Portugal	13
II - Objectivos	15
III - Metodologia	16
1. Amostra	16
2. Instrumentos	18
2.1. <i>Test of Memory Malingering</i> (TOMM)	19
2.2. <i>Rey 15 Item Memory Test</i> (15-IMT)	19
2.3. Memória de Dígitos	20
2.4. Memória Lógica	20
2.5. Cenas de Família	21
2.6. <i>Montreal Cognitive Assessment</i> (MoCA)	21
2.7. Vocabulário	22
2.8. Instrumento de Avaliação Funcional de Adultos e Idosos (IAFAI)	23
2.9. <i>Geriatric Depression Scale</i> (GDS)	23
3. Tratamento Estatístico dos Dados	24
IV – Apresentação e discussão dos resultados	24
1. Efeito das variáveis género, idade, anos de escolaridade e situação laboral nos resultados no TOMM	24
2. Variáveis capacidade funcional, sintomatologia depressiva e capacidade intelectual nos desempenhos no TOMM	26
3. Intercorrelações e consistência Interna	28
4. Validade Discriminante	31
5. Validade referenciada a critérios externos	34
5.1. Exame da associação entre resultados no TOMM com respostas no 15-IMT	34
5.2. Relação entre os resultados obtidos no TOMM e desempenhos em testes de Memória	35
6. Análise crítica comparativa do valor dos pontos de corte propostos por Tombaugh (1996) para o TOMM (considerando os seus vários ensaios) e por Rey (1964) e Boone e colaboradores (2002) para o 15-IMT	37
7. Versão abreviada do TOMM	42
V - Conclusões	45
Bibliografia	48
Anexos	55
Anexo 1 – Critérios de diagnóstico de Slick et al. (1999)	56
Anexo 2 - Análises estatísticas relativas ao 15-IMT	57

Introdução

A determinação da presença e gravidade de défices de natureza cognitiva apoia-se, sobretudo, na colaboração, honestidade e objectividade do sujeito examinado no processo de avaliação (neuro)psicológica. Uma parte significativa das manifestações destes défices não é directamente observável, pelo que a estratégia de comunicação e o rigor das respostas ou a validade dos desempenhos de cada paciente constituem aspectos críticos do exame. Neste plano, as distorções do sujeito, intencionais ou não, introduzem uma maior complexidade nos procedimentos de exame (neuro)psicológico (Rogers, 2008). Em determinados casos, os dados da avaliação (neuro)psicológica podem ficar comprometidos em virtude da presença de motivações e comportamentos de “simulação” ou “esforço insuficiente” por parte do sujeito examinado.

A evidência da presença de “simulação” (“malingering”) ou “esforço insuficiente” (“insufficient effort”) na resposta aos testes tem-se revelado frequente em pessoas que se encontram em situação de litígio. Nestes casos, os ganhos secundários (incluindo incentivos de natureza financeira, redução de penas, alteração nas condições de trabalho, acesso a fármacos, reformas antecipadas) podem desempenhar um papel determinante.

Neste contexto, os (neuro)psicólogos são cada vez mais solicitados a avaliar se os défices cognitivos reportados pelo sujeito são reais ou simulados. O facto de, na maioria das situações, as técnicas tradicionais de avaliação (neuro)psicológica não serem suficientemente precisas na detecção do esforço insuficiente, conduziu ao desenvolvimento de testes específicos, designados de Testes de Validade de Sintomas (TVS), a serem utilizados na determinação da validade da informação obtida (Slick, Tan, Strauss & Hultsch, 2004). Destes, o *Test of Memory Malingering* (TOMM; Tombaugh, 1996) tem sido objecto de um considerável interesse do ponto de vista da prática e investigação clínicas. O TOMM é um TVS constituído por 50 itens (imagens de objectos comuns), concebido para discriminar “simuladores” de doentes com perturbações reais de memória. Sendo sensível à “simulação” de défices mnésicos, os resultados no TOMM não parecem ser influenciados pela presença de défices neurológicos que habitualmente produzem défices

mnésicos (Tombaugh, 1996, 1997).

Apesar do exame da “simulação” ou do “esforço insuficiente” constituir actualmente uma das áreas de maior crescimento no domínio da neuropsicologia clínica e forense, em Portugal é ainda muito rara e pouco sistemática a investigação com testes especificamente desenvolvidos para examinar a credibilidade das queixas e dos desempenhos cognitivos. Face a esta lacuna em termos de investigação e à crescente importância atribuída à validação do processo de avaliação (neuro)psicológica, especialmente em contexto clínico, forense e médico-legal, o presente trabalho pretende examinar a utilidade do TOMM em doentes com Defeito Cognitivo Ligeiro (grupo de particular interesse em que pode surgir a questão do défice real *versus* défice simulado), comparando os seus resultados aos obtidos por um grupo de sujeitos equivalente quanto às principais variáveis sócio-demográficas e sem problemas de natureza cognitiva.

I – Enquadramento conceptual

1. Definição de simulação

As referências históricas ao conceito de “simulação” remontam à época romana, especificamente a dois casos descritos pelo médico Galen: o de um paciente que simulou cólicas para evitar a comparecimento numa reunião; e o de um indivíduo que fingiu uma lesão no joelho para se escusar a acompanhar o seu mestre numa longa viagem. De acordo com Nies e Sweet (1994, *cit in* Tombaugh, 1996), o termo “simulação” surgiu, originalmente, no contexto militar, para descrever casos de soldados que fingiam doença como forma de evadir-se àquele serviço.

Presentemente, o exame dos comportamentos de “simulação” constitui um tópico de crescente importância e visibilidade, quer ao nível da investigação, quer nas práticas clínicas de avaliação e intervenção (neuro)psicológicas. Não é por isso surpreendente o aparecimento de várias definições e critérios para o diagnóstico de Simulação.

No DSM-IV-TR (APA, 2002, p.739), a Simulação é classificada como um comportamento merecedor de atenção clínica, não constituindo, no entanto, um distúrbio mental *per se*. Por esse motivo, não são indicados

critérios de diagnóstico formais, sendo apenas proporcionada uma definição geral, com diferentes índices clínicos suspeitos e uma breve revisão dos diagnósticos diferenciais. Assim, a definição de Simulação proposta pela DSM-IV-TR aponta para uma “produção intencional de sintomas físicos ou psicológicos, falsos ou exagerados, motivados por incentivos externos, tais como, evitar o serviço militar ou o trabalho, obter compensações económicas, evitar o cumprimento de penas ou obter drogas” (APA, 2002, p.739). Nesta definição estão patentes três aspectos importantes: a) invenção ou exagero de sintomas; b) motivação ou intenção consciente do sujeito para simular; e c) presença de incentivos externos.

Adicionalmente, alguns elementos devem ser considerados como passíveis de indicar maior probabilidade de presença de comportamentos de Simulação: a) exame realizado em contexto médico-legal; b) discrepância acentuada entre o sofrimento ou incapacidade referidos pelo sujeito e os dados objectivos da avaliação; c) falta de colaboração na avaliação diagnóstica e não adesão ao tratamento prescrito; e d) presença de uma Perturbação Anti-Social da Personalidade (DSM-IV-TR; APA, 2002, p.739; APA, 1952 versão original).

Esta definição de Simulação apontada no DSM tem sido objecto de diversas críticas, principalmente relacionadas com o facto dos critérios apresentados serem “vagos” e “minimalistas”. Slick, Sherman e Iverson (1999) acrescentam que esta definição comporta alguns problemas respeitantes à presença de julgamentos subjectivos sobre questões psicológicas complexas, tais como: 1) incentivos internos *versus* externos e 2) mecanismos inconscientes *versus* conscientes. Ademais, os critérios da DSM-IV-TR para a “simulação” proporcionam poucas linhas orientadoras para a avaliação do exagero ou “fabricação” de défices cognitivos (Berry & Schipper, 2008).

Face às críticas apontadas aos critérios de diagnóstico da DSM vários clínicos e investigadores têm indicado definições e critérios alternativos. Por exemplo, e com um outro enquadramento, Rogers (1990, *cit in* Slick et al, 1999) propôs critérios de diagnóstico específicos para a simulação de perturbações psiquiátricas. Usando um modelo similar ao da DSM, Rogers especificou um número mínimo de critérios em cada domínio, necessários ao diagnóstico de simulação, sugerindo três modelos explicativos deste

fenómeno: Modelo Patogénico, Modelo Criminológico e Modelo Adaptativo. Os dois primeiros modelos assumem que a simulação é determinada por “doença” (modelo Patogénico) ou por “maldade” (modelo Criminológico). Por outro lado, o modelo Adaptativo considera o comportamento de simulação como uma decisão tomada em circunstâncias adversas que, frequentemente, resulta de análises de custo-benefício das diferentes alternativas presentes.

Por outro lado, Ferguson (2004) aponta para uma outra forma de organizar a informação relativa a este problema sugerindo a existência de quatro padrões distintos de simulação: a) “Simulação pura” ou “comunicação de défices não existentes” – forma rara de simulação na qual o sujeito não apresenta sintomas ou défices resultantes de lesão mas, de forma fraudulenta, alega a sua presença no (seu) relatório comportamental; b) “Exagero de sintomas existentes” – o sujeito manifesta sintomas ou défices provocados por lesão, mas maximiza-os, constituindo esta a forma mais comum de “simulação”; c) “Manutenção ou agravamento dos sintomas” – na qual o sujeito, já recuperado dos sintomas de uma lesão, alega que estes persistem ou se agravaram; d) “Atribuição incorrecta dos sintomas” – quando o indivíduo atribui, intencionalmente, uma causa errónea aos sintomas que verdadeiramente experiencia.

Não podemos, de igual modo, negligenciar o trabalho de Slick, Sherman e Iverson (1999), que desenvolveram um conjunto de critérios formais de diagnóstico (consultar critérios de diagnóstico no anexo 1) de Simulação de Perturbação Neurocognitiva (*Malingered Neurocognitive Dysfunction* – MND). Nesta abordagem a Simulação é definida como “o exagero ou produção intencional de uma perturbação cognitiva, realizado com o objectivo de obter ganhos materiais substanciais, evitar ou escapar de responsabilidades formais ou de trabalho” (Slick et al, 1999, p.552). Houve da parte de Slick e colaboradores (1999) a preocupação em operacionalizar os termos utilizados, fornecendo ainda graus de certeza para acompanhar a identificação de MND: “definitivo”, “provável” e “possível”. Segundo Slick e colegas (1999), o clínico que pretenda avaliar a presença de MND, deve reger-se por quatro critérios base: A) Presença de incentivo(s) externo(s); B) Evidências na avaliação (neuro)psicológica; C) Evidências no auto-relato do sujeito; D) Comportamentos de “simulação” que não são melhor explicados

por doença psiquiátrica, neurológica ou desenvolvimental.

O termo esforço insuficiente, intimamente associado a um desempenho “incompleto” ou “insuficiente”, e mesmo ao conceito de simulação, corresponde ao “desempenho alcançado num teste que é significativamente mais reduzido do que os padrões de desempenho conhecidos e associados a perturbações neurológicas verdadeiras” (Simões, 2005, p.456; cf. Rogers, 2008). Este conceito adequa-se, de uma forma mais compreensiva, ao contexto do recurso a instrumentos de avaliação (neuro)psicológica, sendo utilizado para caracterizar casos de sujeitos, que formulam pedidos de compensação por dano neurológico, e que apresentam padrões de resultados substancialmente mais baixos nos testes do que os obtidos por indivíduos com perturbações neurológicas verdadeiras (Strauss, Sherman & Spreen, 2006). O facto de se estabelecer uma relação tão estreita entre “esforço insuficiente” e “simulação” pode ser sinónimo de cometimento de alguns erros, já que são múltiplos os factores (factores situacionais, educacionais e motivacionais, perturbações psiquiátricas) que podem “prejudicar” o esforço dispendido por um indivíduo no processo de avaliação.

Contudo, e à semelhança do que acontece com qualquer decisão diagnóstica, as conclusões relativas à presença de “simulação” (ou de “esforço insuficiente”) devem ser formuladas após a revisão cuidada de toda a informação disponível e a exclusão de explicações alternativas. É importante reconhecer que o diagnóstico de “simulação” corresponde a um dos juízos clínicos mais pejorativos, na medida em que acusa o sujeito de engano, fraude e perjúrio. Além disso, um diagnóstico falso tem consequências negativas substanciais para a vida pessoal, familiar e profissional do sujeito (Tombaugh, 1996). Aliás, segundo Tombaugh (2002), muitos sujeitos com traumatismo crânio-encefálico que simulam são “boas” pessoas surpreendidas em “más situações”.

2. Métodos na detecção da “simulação”

São várias as técnicas utilizadas na detecção das estratégias de simulação. De seguida, iremos abordar os testes de validade de sintomas.

2.1. Testes de validade de sintomas (TVS)

Segundo Iverson (1999) o exame do “esforço insuficiente” deve alicerçar-se no uso de vários instrumentos e recorrer a testes especificamente construídos para este efeito, os apelidados Testes de Validade de Sintomas, também conhecidos como testes de escolha forçada. Os TVS podem variar no conteúdo (estímulos verbais *vs* não verbais; evocação *vs* reconhecimento), formato (papel e lápis *vs* computador), grau de dificuldade, transparência (e.g. clareza do objectivo do teste) e duração do tempo de administração (Sweet, Condit & Nelson, 2008). Estes testes foram desenvolvidos com o objectivo de proporcionar medidas objectivas, baseadas em critérios psicométricos que permitissem discriminar sujeitos com perturbações genuínas daqueles que “simulam”. Os TVS são habitualmente constituídos por um número elevado de itens, apresentados num formato de escolha forçada, e sugestivos de avaliar uma função específica, de que é exemplo a memória.

A grande maioria dos TVS aparenta um grau de dificuldade substancialmente maior do que os escassos recursos cognitivos que, na realidade, são necessários para a sua resolução (Tombaugh, 1996; Slick, Hopp, Strauss & Spellacy, 1996, *cit in* McCaffrey et al., 2003). A configuração de resposta de escolha forçada ao teste permite identificar os desempenhos deliberadamente reduzidos. De um modo geral, níveis de desempenho inferiores à probabilidade de resposta ao acaso ou abaixo dos pontos-de-corte determinados empiricamente sugerem a presença de respostas enviesadas, isto é, a possibilidade de esforço insuficiente para responder bem nas tarefas, a existência de motivação deliberada para apresentar desempenhos reduzidos e a persistência de um evitamento activo da resposta correcta.

Nos últimos anos verificou-se um acentuado desenvolvimento de novos instrumentos de exame da validade de sintomas, estudados do ponto

de vista clínico e psicométrico. Contudo, apenas alguns são regularmente utilizados na prática clínica (Slick, 2004). Entre os mais populares estão: o *Test of Memory Malingering* (TOMM; Tombaugh, 1996); o Teste de Memória de Palavras (*Word Memory Test* – WMT; Green, Astner & Allen, 1996); o Teste de Memória de Dígitos (*Digit Memory Test* – DMT; Hiscock & Hiscock, 1989); o Teste de Validade de Sintomas de Victoria (*Victoria Symptom Validity Test* – VSVT; Slick et al., 1997); o Teste Rey-15 Itens (*Rey-15 Item Memory Test* – 15-IMT; Rey, 1964; Boone, 2002); o Teste de Reconhecimento de Dígitos de Portland (*Portland Digit Recognition Test* - PDRT; Blinder & Willis, 1991) e o *Computerized Assessment of Response Bias* (CARB; Allen, Conder, Green & Cox, 1997).

Apesar da quantidade de TVS presentemente disponíveis, Slick e colegas (2004) verificaram que o TOMM é, actualmente, o teste mais frequentemente usado por neuropsicólogos para detectar o esforço insuficiente: mais especificamente, cerca 75% destes técnicos usa o TOMM como medida de esforço (Sweet, 2008). O *Rey 15-Item Memory Test* (15-IMT; Rey, 1964; Boone, 2002) sucede ao TOMM (Tombaugh, 1996) na lista dos testes mais utilizados (Slick, Tan, Strauss & Hultsch, 2004). Estas constatações, particularmente no que concerne ao 15-IMT são um pouco surpreendentes, dada a inúmera literatura que aponta para a falta de sensibilidade e especificidade¹ do teste (Strauss, Sherman & Spreng, 2006; Schretlen, Brandt, Krafft & Van Gorp, 1991 *cit in* McCaffrey, O`Bryant & Fisher, 2003; Arnett et al., 1995; Bernard & Fowler, 1990), principalmente em sujeitos idosos com algum grau de declínio cognitivo (Dean, Victor, Voone, Philpott & Hess, 2009). Perante estes dados, pode justificar a sua popularidade o baixo custo, a fácil administração e a simplicidade do material requerido.

Como já referido anteriormente, para um sujeito ser diagnosticado como “simulador” é necessário estarem presentes alguns elementos, a relembrar: a) sintomas falsos ou exagerados; b) intencionalmente

¹ A *sensibilidade* é um índice da probabilidade de um teste identificar correctamente um indivíduo com uma dada condição. Neste caso, a sensibilidade refere-se à proporção de sujeitos “simuladores” que são correctamente identificados por esse instrumento. Por sua vez, a *especificidade* refere-se à exactidão de um teste na identificação de sujeitos que não têm a condição, isto é, a proporção de examinados que não apresentam a condição de interesse e que, desta forma, são correctamente classificados. Neste contexto, a especificidade corresponde à classificação correcta de sujeitos como “não simuladores”.

produzidos; e c) motivados por incentivos externos (DSM-IV-TR, APA; 2002, p. 739). Os resultados nos TVS apenas contribuem para o diagnóstico de “simulação” ao demonstrar que os sintomas são falsos ou exagerados, sendo que informações sobre a intencionalidade e a motivação inerentes ao comportamento devem provir de outras fontes (Tombaugh, 1996).

O uso de pelo menos uma e, preferencialmente, duas ou mais medidas de validade de sintomas convenientemente validadas tem sido consensualmente recomendado (cf. Inman & Berry, 2002; Strauss, Sherman, & Sprenn, 2006; Slick et al., 2004). Estas medidas devem ser administradas, de modo sistemático, como elementos constituintes de qualquer processo de avaliação (neuro)psicológica, mas principalmente quando possam estar presentes potenciais incentivos para desempenhos reduzidos. Segundo Iverson e Binder (2000), uma avaliação que não inclua um meio objectivo de avaliação do esforço, da validade das respostas e da “simulação” é considerada como estando incompleta. No entanto, é importante reconhecer que os testes de validade de sintomas acabam por não identificar um número significativo de “simuladores”.

2.1.1. Test of Memory Malingering (TOMM)

Publicado em 1996 por Tombaugh, o *Test of Memory Malingering* (TOMM) é um teste concebido para discriminar “simuladores” de doentes com perturbações reais de memória. Sendo um teste com elevada validade facial (ao conter um elevado número de itens, os sujeitos desenvolvem a crença de que o teste é muito mais difícil do que é na realidade), baseia-se em três estratégias (de detecção do esforço insuficiente): (i) testes de validade de sintomas (baseia a classificação da simulação no facto de a proporção de itens falhados ser superior à que seria de esperar se o sujeito tivesse respondido ao acaso); (ii) metodologia de escolha forçada (baseia-se no facto de o desempenho ser considerado inferior ao esperado – tendo como referência dados normativos); e (iii) princípio do efeito tecto (identificação de itens cujo nível de dificuldade é reduzido, sendo alcançados com sucesso pela maioria das pessoas; uma minoria dos simuladores apresenta com frequência um padrão indiscriminado, podendo falhar nos itens mais simples) (Rogers, 2008).

O TOMM é um teste de reconhecimento visual originalmente

desenvolvido para adultos, que inclui dois Ensaios de Aprendizagem (Ensaio 1 e Ensaio 2) e um Ensaio de Retenção. Os dois Ensaios de Aprendizagem são constituídos por uma fase de estudo, em que são apresentados 50 desenhos de objectos comuns (objectos alvo) ao sujeito, um a um (durante 3 segundos e com um intervalo de 1 segundo entre cada uma das apresentações), e por uma fase de teste no qual são mostrados cada um dos 50 objectos-alvo emparelhados com um novo objecto distractor, sendo solicitado ao sujeito que indique, dos dois objectos, aquele que viu antes (tarefa de escolha forçada). Nos dois Ensaios de Aprendizagem são mostrados os mesmos 50 desenhos-alvo, ainda que numa ordem diferente. Já quanto aos objectos distractores, estes são apresentados uma única vez sendo que o examinador deve indicar ao sujeito se a resposta está certa ou errada (*feedback* explícito). O Ensaio de Retenção consiste apenas na aplicação da fase de teste, sendo administrado cerca de 15 minutos após o sujeito ter terminado de responder aos dois Ensaios de Aprendizagem (Tombaugh, 1996, 1997).

O TOMM possui quatro características que o tornam eficaz na detecção do esforço insuficiente: (1) a administração de um elevado número de estímulos torna a percepção do teste como mais difícil do que na realidade é, ou seja, é um teste apresentado como cognitivamente muito exigente para pessoas com problemas mnésicos, facilitando assim a manifestação de baixos desempenhos por parte de simuladores; (2) a apresentação de 50 desenhos-alvo torna o TOMM um instrumento de elevada validade facial enquanto teste de aprendizagem e memória, não sendo por isso habitualmente percebido como um teste de avaliação do esforço insuficiente; (3) o *feedback* explícito aumenta a diferença entre os resultados de pacientes com défices mnésicos e os desempenhos de possíveis “simuladores”, ou seja, o *feedback* afigura-se como uma oportunidade de aprendizagem para os sujeitos que estão verdadeiramente motivados e deve aumentar a precisão das suas respostas nos Ensaios subsequentes (Tombaugh, 1996); e (4) o teste é apenas sensível à simulação de défices mnésicos e não à presença de outros factores que habitualmente produzem um défice mnésico (Simões, 2006).

Os estudos de validade permitiram a Tombaugh (1996) concluir que um desempenho inferior a 45 acertos no Ensaio 2 ou no Ensaio de Retenção

poderá constituir um indicador de “esforço insuficiente”, independentemente da idade, condição neurológica ou distúrbios afectivos do sujeito.

Os estudos internacionais conduzidos por numerosos autores têm, de um modo geral, demonstrado que TOMM é um teste frequentemente administrado (Slick et al., 2004), que evidencia precisão no diagnóstico (O`Bryant & Lucas, 2006; Haber & Fichtenberg, 2006), sendo que os resultados parecem não ser influenciados pela idade (Horton, 2000; Constantinou & McCaffrey, 2003; Ashendorf, Constantinou & McCaffrey, 2004; Donders, 2005; Rees, Tombaugh, Gansler & Moczynski, 1998; Rees, Tombaugh & Boulay, 2001), anos de escolaridade (Tombaugh, 1996, 1997; Constantinou & McCaffrey, 2003; Teichner & Wagner, 2004; Rees et al., 1998; Horton, 2000), género e cultura (Tombaugh, 1996, 1997; Constantinou & McCaffrey, 2003; Rees, Tombaugh & Boulay, 2001), distúrbios afectivos (Tombaugh, 1996, 1997; Rees, Tombaugh & Boulay, 2001; Ashendorf, Constantinou & McCaffrey, 2004; Iverson et al., 2007; O`Bryant, Finlay & O`Jile, 2007), dor induzida em laboratório (Etherton et al., 2005); dor crónica (Iverson et al., 2007) ou défices cognitivos associados a diferentes formas de neuropatologia (Tombaugh, 1996, 1997; Haber & Fichtenberg, 2006; Rees, Tombaugh, Gansler & Moczynski, 1998).

Recentemente, a utilidade e validade clínica do Ensaio 1 do TOMM converteu-se também num tópico de investigação. Neste plano, Gavett, O`Bryant e McCaffrey (2004) examinaram a utilidade das pontuações obtidas no Ensaio 1 do TOMM numa amostra de sujeitos vítimas de traumatismo crânio-encefálico. Os resultados deste estudo sugeriram que o Ensaio 2 e o Ensaio de Retenção têm pouca utilidade quando os sujeitos obtêm uma pontuação ≥ 45 acertos no Ensaio 1. Resultados semelhantes foram encontrados por O`Bryant e colaboradores (2008). Horner, Bedwell e Duong (2006) foram mais longe, sugerindo um ponto de corte para o Ensaio 1, a partir do qual se deve ponderar o eventual recurso aos restantes dois Ensaios. Segundo estes autores, um resultado ≥ 36 acertos no Ensaio 1 indica, com 99% de probabilidade, que o sujeito passe no Ensaio 2 e/ou no Ensaio de Retenção do TOMM. Por sua vez, um resultado ≤ 27 acertos indica com 100% de probabilidade que o sujeito irá falhar no teste, ou seja, irá obter uma pontuação < 45 no Ensaio 2 e no Ensaio de Retenção.

Apesar do recente interesse numa versão abreviada do TOMM, no que diz respeito ao número de Ensaio administrados, a maior parte dos estudos continua a sugerir a utilização dos dois Ensaio de Aprendizagem para a avaliação do esforço insuficiente, sendo que alguns autores ainda defendem o recurso ao Ensaio de Retenção, uma vez que acrescenta poucos minutos ao teste, ajuda a corroborar os resultados e poderá detectar um grupo minoritário de simuladores que, de outro modo, não seria identificado como tal (Greve, Bianchini, Black, Heinly, Love & Swift, 2006).

Inicialmente construído para ser aplicados a adultos, o TOMM tem também, vindo a ser estudado em crianças a partir dos 5 anos de idade (Constantinou & McCaffrey, 2003; Donders, 2005; Blaskewitz, Mertenb & Kathmanna, 2008). De um modo geral, estes estudos parecem indicar que o teste constitui uma medida válida na avaliação do esforço cognitivo de crianças em idade escolar.

Recentemente, a validade clínica do TOMM na população idosa tem sido questionada, principalmente junto daqueles que já manifestam algum grau de declínio cognitivo (Dean et al., 2009). De seguida iremos então abordar os estudos que se debruçam sobre esta questão.

2.1.1.1. Estudos com adultos idosos

Nos estudos de validação do TOMM, Tombaugh (1996) concluiu que este instrumento é válido na detecção da “simulação” de défices mnésicos em adultos e em adultos idosos, incluindo aqueles com diagnóstico de demência. Da amostra de 42 adultos idosos com declínio cognitivo e de 40 adultos idosos diagnosticados com demência, respectivamente, 90.4% e 73% dos sujeitos avaliados, pontuaram acima do ponto de corte no Ensaio 2 do TOMM, suportando, segundo o referido autor, a validade deste instrumento em grupos de idosos com perturbação cognitiva.

Por sua vez, Ashendorf, Constantinou e McCaffrey (2004) investigaram o impacto da depressão (ligeira a moderada) e da ansiedade medidas, respectivamente, pelo BDI (*Beck Depression Inventory*) e STAI (*State-Trait Anxiety Inventory*) no desempenho ao TOMM, considerando para tal uma amostra de 197 idosos, com idades compreendidas entre os 55 e os 75 anos. A análise dos dados demonstrou que o TOMM não é sensível à presença de depressão e/ou de sintomatologia ansiosa. De acordo com estes

resultados, independentemente dos sujeitos obterem ou não uma pontuação elevada no BDI e/ou no STAI, todos os valores observados no Ensaio 2 do TOMM foram ≥ 48 acertos, sugerindo ausência de “esforço insuficiente”. Apesar da abundante literatura indicar um claro impacto da sintomatologia ansiosa e depressiva no desempenho em testes neuropsicológicos (Kizilbash, Vanderploeg & Curtiss, 2002; Strauss, Sherman & Spreen, 2006), os dados do estudo de Ashendorf e colaboradores (2004) sugerem que as pontuações no TOMM não são afectadas da mesma forma. Estes dados são relevantes, uma vez que a depressão tem associada uma forte componente motivacional. Os indivíduos deprimidos geralmente percebem as tarefas como sendo mais difíceis do que na realidade são, necessitando de um maior encorajamento para as desempenhar. Está aliás bem estabelecido na literatura que sujeitos deprimidos têm dificuldades em colocar o máximo esforço na tarefa, sendo este comportamento observado em testes neuropsicológicos que requerem um esforço contínuo, particularmente nos testes de memória (Rees, Tombaugh & Boulay, 2001). Uma vez que o TOMM é percebido como sendo um teste difícil (Tombaugh, 1996, 1997), poderia acontecer que os sujeitos deprimidos “fracassassem” ao não colocar o máximo esforço na tarefa e fossem assim incorrectamente identificados como possíveis “simuladores” (falsos positivos).

Em 2004, Teichner e Wagner conduziram um estudo com 78 idosos distribuídos por 3 grupos experimentais: a) demência; b) declínio cognitivo sem demência; e c) cognitivamente normais. Estes sujeitos responderam ao protocolo de avaliação constituído pelo TOMM, WAIS-III; teste de Orientação da WMS-III e MMSE. Os sujeitos dos grupos sem demência obtiveram resultados similares no TOMM, com 94.7% dos sujeitos a serem identificados como “honestos”. No grupo com demência os resultados foram mais baixos em todos os Ensaios, resultando num número elevado de falsos positivos (76% no Ensaio 2 e 71% no Ensaio de Retenção). Estes resultados reforçam a validade e utilidade clínicas do TOMM em pacientes idosos, sendo a especificidade do TOMM elevada quando são excluídos os casos de demência. Contudo, o TOMM parece ser sensível à disfunção cognitiva severa associada à demência, não sendo uma medida útil na avaliação da motivação de pessoas com esta gravidade de défice cognitivo.

Numa revisão da literatura, Dean e colaboradores (2009) constataram

a escassez dos estudos com pacientes diagnosticados com demência, uma vez que esta população clínica é frequentemente excluída das amostras de validação dos testes de “esforço reduzido”. Numa análise mais pormenorizada do número reduzido de estudos, estes autores verificaram que a especificidade do TOMM (Tombaugh, 1996, 1997) em amostras de sujeitos com demência oscila entre um máximo de 82% (Ensaio 2, Greve et al., 2006) e um mínimo de 24% (Ensaio 2, Teichner & Wagner, 2004). Após esta revisão da literatura, Dean e colaboradores (2009), analisaram o desempenho de sujeitos com demência em 12 testes frequentemente usados na avaliação do “esforço insuficiente” (Memória de Dígitos e Vocabulário da WAIS-III/WAIS-R, *Dot Counting Test*, *Warrington Recognition Memory Test-Words*, Memória Lógica da WMS-III, *Rey Word Recognition Memory Test*, *Finger Tapping*, *b-Test*, *Rey 15-Item*, TOMM, *Rey Auditory Verbal Learning Test* e Figura Complexa de Rey). De um modo geral, os dados sugerem que: (i) quanto mais baixos são os resultados obtidos no *Mini Mental State Examination* (MMSE), maior é a probabilidade dos sujeitos “não passarem” nos testes de esforço, e (ii) a maior parte dos testes de exame do esforço reduzido analisados apresentam uma taxa elevada (e inaceitável) de falsos positivos. Especificamente, os resultados obtidos no TOMM demonstraram que este instrumento tem uma especificidade de 45% em sujeitos com demência, sugerindo a necessidade de se modificar o ponto de corte (Ensaio 2 <28 acertos) com o objectivo de manter a especificidade do instrumento.

2.1.1.2. Estudos realizados em Portugal

Em Portugal começam a surgir alguns estudos empíricos com o TOMM, sendo que estes têm contemplado preferencialmente os contextos forenses. De um modo geral, estes trabalhos validam o ponto de corte sugerido por Tombaugh (1996) e corroborado pelos resultados de muitas outras investigações de natureza empírica.

Assim, no estudo de Maior, Silva, Pedrosa, Dias, Simões e Pinho (2006) o TOMM foi aplicado a dois grupos de sujeitos: indivíduos reclusos ($n=41$) e indivíduos não reclusos ($n=40$), entre os quais não se evidenciaram diferenças com significado estatístico. Os resultados obtidos mostram que

ambas as amostras obtiveram um desempenho superior ao ponto de corte proposto por Tombaugh (1996) quer no Ensaio 2, quer no Ensaio de Retenção do TOMM.

Mais recentemente Maior (2008), aplicou o TOMM e a SIMS a 116 sujeitos. Os resultados obtidos permitiram concluir que o TOMM apresenta resultados satisfatórios (E2: $M=49.46$ $dp=1.28$; ER: $M=49.46$ $dp=1.28$) e que pode ser aplicado à população portuguesa. Os dados obtidos também permitiram concluir que o TOMM pode ser administrado a sujeitos com diagnóstico psiquiátrico ou neurológico. Contrariamente aos estudos internacionais, os mesmos dados sugerem que os resultados no TOMM podem ser influenciados, em casos pontuais, pela escolaridade e pela idade.

Num outro estudo realizado por Mota, Simões, Amaral, Dias, Luís, Pedrosa e colaboradores (2008) com uma amostra constituída por 103 reclusos, foi analisado o perfil de resposta no TOMM, na SIRS (*Structured Interview of Reported Symptoms*), na SIMS (*Structured Inventory of Malingered Symptomatology*) e numa medida de inteligência (subteste de Cubos da WAIS-III). Os resultados obtidos sugerem que o ponto de corte proposto por Tombaugh (1996) é eficaz para a distinção entre “participantes honestos” e possíveis “simuladores” (E2: $M=49.29$ $dp=1.72$; ER: $M=49.46$ $dp=1.40$). Os dados obtidos neste estudo, também sugerem que os resultados no TOMM podem ser influenciados, em casos pontuais, pela escolaridade e pela capacidade intelectual.

Em 2008, Cruz analisou o perfil de resposta de 4 grupos de sujeitos no TOMM: normativo ($n=30$), “simuladores” ($n=30$), “*coaching*²” ($n=30$) e um grupo de adultos idosos ($n=50$), alguns dos quais com defeito cognitivo ligeiro e depressão. Os resultados comprovam o valor do ponto de corte inicialmente proposto por Tombaugh (1996), quer na amostra normativa (E2: $M=49.57$ $dp=0.90$), quer na amostra de adultos idosos [adultos idosos com defeito cognitivo (E2: $M=45.88$ $dp=5.88$); adultos idosos sem defeito cognitivo (E2: $M=48.88$ $dp=2.73$); adultos idosos com depressão (E2: $M=45.67$ $dp=5.68$); adultos idosos sem depressão (E2: $M=49.55$ $dp=0.99$)].

Faustino (2009) aplicou o TOMM e o 15-IMT a reclusos violentos, reclusos não violentos e a não reclusos não violentos. A autora sugere que os

² Conceito que designa o conhecimento por parte do sujeito avaliado, das consequências médicas resultantes de lesões cerebrais, bem como das estratégias que devem utilizar para “passar” nos testes.

pontos de corte previamente sugeridos na literatura (45 no Ensaio 2 e/ou Ensaio de Retenção do TOMM; 9 na Evocação Imediata e 20 no Reconhecimento³ do 15-IMT) parecem ser adequados à população portuguesa.

Pedrosa (2009) aplicou o TOMM e o 15-IMT a dois grupos de sujeitos: 40 litigantes com Traumatismo Crânio-Encefálico (TCE) e 30 sujeitos da população geral. No grupo de controlo os resultados revelam-se satisfatórios no TOMM (E2: $M=46.86$ $dp=2.54$; ER: $M=49.43$ $dp=0.98$) e no 15-IMT (Evoc: $M=12.13$ $dp=3.06$; Reconh: $M=23.87$ $dp=6.17$), acentuando a pertinência e eficácia destas provas na identificação do esforço reduzido. Os resultados obtidos pelo grupo de litigantes com TCE foram inferiores e sugestivos de esforço insuficiente, quer no TOMM (E2: $M=36.74$ $dp=6.60$; ER: $M=37.11$ $dp=9.40$) quer no Ensaio de Reconhecimento do 15-IMT ($M=16.85$ $dp=7.90$)

II - Objectivos

O principal objectivo do presente estudo consistiu na aplicação do TOMM em doentes com Defeito Cognitivo Ligeiro (DCL) comparando os seus resultados aos obtidos por sujeitos cognitivamente normais. O DCL é um grupo clínico de particular interesse para examinar a presença de défice real vs défice simulado ou exagerado (ou esforço reduzido). Mais especificamente, pretendeu-se avaliar:

- i. Efeito das variáveis género, idade, escolaridade e situação laboral nos resultados no TOMM;
- ii. Relação entre variáveis relativas à capacidade funcional, sintomatologia depressiva e capacidade intelectual geral com os desempenhos no TOMM;
- iii. Consistência interna e inter-correlações entre os Ensaio do TOMM;
- iv. Validade discriminante, ou seja, examinar a capacidade dos resultados no TOMM na caracterização e diferenciação dos desempenhos de

³ Pontuação combinada (resultado combinado do reconhecimento = número itens correctamente evocados + (número de itens correctamente reconhecidos – número de falsos positivos).

- idosos normais e de idosos com Declínio Cognitivo Ligeiro;
- v. Validade referenciada a critérios externos:
 - Examinar a associação entre os resultados no TOMM e o *Rey 15-Item Memory Test*, um outro instrumento específico conhecido na avaliação do “esforço insuficiente” (validade concorrente);
 - Explorar a relação entre medidas de memória (nas suas componentes visual, verbal e auditiva) e desempenhos no TOMM (teste de avaliação da simulação de défices mnésicos)
 - vi. Análise crítica comparativa do valor dos pontos de corte propostos por Tombaugh (1996) para o TOMM (considerando os seus vários Ensaio), e por Rey (1964) e Boone e colaboradores (2002) para o *Rey 15-Item Memory Test* (considerando os dois grupos em estudo);
 - vii. Perspectivar a possibilidade de uma versão abreviada do TOMM, em termos de número de Ensaio administrados, que passe pelo estabelecimento de um ponto de corte para o Ensaio 1.

III - Metodologia

1. Amostra

As amostras do presente estudo (cf. Quadro2) são constituídas por um grupo controlo de 20 sujeitos, com idades compreendidas entre os 61 e os 75 anos e por um grupo clínico de 20 sujeitos com Declínio Cognitivo Ligeiro (DCL) com idades compreendidas entre os 60-75 anos.

Mais especificamente, o grupo de controlo é constituído por 10 (50%) sujeitos do género masculino e 10 (50%) do género feminino, sendo que apenas 13 (65%) estão reformados. Este grupo tem uma média de idades de 67 anos ($dp=4.36$) e uma frequência escolar média de 5 anos ($dp=3.28$). O grupo com DCL é constituído por 8 (40%) sujeitos do género masculino e por 12 (60%) do género feminino. A média de idades e anos de escolaridade é de 67.8 ($dp=4.92$) e 4.9 ($dp=3.24$) anos, respectivamente. Destes, 16 (80%) estão reformados.

Os dois grupos não diferem estatisticamente entre si na idade ($U=178.500$; $p=.565$), nos anos de escolaridade ($U=178.000$; $p=.565$), no estado civil ($\chi^2(2)=.667$; $p=.717$), no género ($\chi^2(1)=.404$; $p=.525$), na situação

laboral ($\chi^2_{(1)}=1.129$; $p=.288$), na inteligência verbal/cristalizada ($U=158.500$; $p=.265$), na sintomatologia depressiva ($U=132.500$; $p=.068$) e na capacidade funcional ($U=147.000$; $p=.152$). De acordo com o esperado, os dois grupos diferem na capacidade cognitiva ($U=102.000$; $p=.007$) medida pelo MoCA, sendo que o grupo de controlo obteve resultados mais elevados ($M=21.85$; $dp=3.13$) comparativamente ao grupo DCL ($M=18.7$; $dp=3.7$).

Quadro 2. Caracterização da Amostra

N=40	Controlo (N=20)	DCL (N=20)	U/ χ^2	P
Idade	M=67; DP=4,36 (Min=61; Máx=75)	M=67,8; DP=4,92 (Min=60; Máx=75)	178.500	.565
Anos de Escolaridade	M=5; DP=3,28 (Min=3; Máx=17)	M=4,9; DP=3,24 (Min=3; Máx=15)	178.000	.565
Nível de Escolaridade	1º ciclo: 16(80%) 2º ciclo: 3(15%) Licenciatura: 1(5%)	1º ciclo: 17(85%) 2º ciclo: 1(5%) Secundário: 1(5%) Licenciatura: 1 (5%)	2.030	.566
Estado Civil	Solteiro: 5% Casado: 85% Viúvo: 10%	Solteiro: 10% Casado: 85% Viúvo: 5%	.667	.717
Género	10 M; 10 F (50%M; 50%F)	8 M; 12 F (40%M; 60%F)	.404	.525
Situação Laboral	Reformado: 13 (65%) Activo: 7 (35%)	Reformado: 16 (80%) Activo: 4 (20%)	1.129	.288
Vocabulário	M=29.55; DP=8,59 (Min=12; Máx=53)	M=27,45; DP=10,57 (Min=13; Máx=54)	158.500	.265
MoCA	M=21.85; DP=3,13 (Min=17; Máx=26)	M=18.7; DP=3,7 (Min=14; Máx=26)	102.000	.007**
GDS	M=10,9; DP=5,65 (Min=1; Máx=20)	M=14.25; DP=5,19 (Min=5; Máx=20)	132.500	.068
IAFAI	M=.43; DP=.33 (Min=.0; Máx=1,21)	M=.87; DP=1.25 (Min=.04; Máx=5.78)	147.000	.152

* $p<.05$; ** $p<.01$

Legenda: MoCA – *Montreal Cognitive Assessment*; GDS – *Geriatric Depression Scale*; IAFAI – Instrumento de Avaliação Funcional de Adultos e Idosos

Os participantes foram recrutados no Serviço de Neurologia do Hospital Geral do Centro Hospitalar de Coimbra. A sua selecção ocorreu no contexto das consultas externas de Neurologia Geral e das avaliações neuropsicológicas realizadas no Laboratório de Neuropsicologia.

Na selecção dos sujeitos do **grupo DCL** foram considerados os seguintes **critérios de inclusão**: (1) idades compreendidas entre os 60-75

anos; (2) diagnóstico clínico de DCL e/ou diagnóstico de DCL, de acordo com critérios diagnósticos actualmente considerados, nomeadamente queixas de problemas mnésicos, défice cognitivo⁴ (valores no MoCA >1.5 *dp* inferior às normas ajustadas para a idade) e actividades de vida diária intactas (Tabert et al., 2006). Foram ainda contemplados os seguintes **critérios de exclusão**: (1) presença de perturbações psicopatológicas com significado clínico (e.g. depressão grave, psicose); (2) escolaridade <3^a classe; (3) presença de défices visuais e/ou auditivos não corrigidos.

Os participantes do grupo de controlo foram abordados em contexto hospitalar, tendo, tal como os anteriores, concordado em participar voluntariamente no presente estudo. Os doentes, e respectivos familiares, foram informados acerca da importância da investigação e do relevo da sua participação, quer esta fosse inserida na sua rotina habitual de acompanhamento ou não. O consentimento informado oral foi fornecido pelo doente, por familiares ou por acompanhantes. A motivação e disponibilidade física e psicológica do doente para participar na investigação foram sempre consideradas.

2. Instrumentos

Para a recolha dos dados foram utilizados instrumentos específicos ao exame do esforço insuficiente (TOMM e Rey 15-Item Memory Test), de avaliação do funcionamento cognitivo (MoCA), de avaliação da memória (Memória Lógica, Memória de Dígitos e Cenas de Família) e de avaliação da capacidade intelectual verbal (Vocabulário). Do protocolo de avaliação consta, também, uma escala de avaliação da sintomatologia depressiva (GDS) e um instrumento de avaliação funcional (IAFAI). Passamos a

⁴ Para preencher este critério foram utilizados os dados actualmente disponíveis relativos aos estudos realizados no âmbito do projecto “Validação de Provas de Memória e de Inventários de Avaliação Funcional e da Qualidade de Vida” (financiado pela Fundação Calouste Gulbenkian; Proc. 74569), a decorrer na Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra. Foram calculados os valores médios no MoCA para sujeitos cognitivamente normais (controlo). Destes, apenas foram seleccionados sujeitos com idades compreendidas entre os 60-75 anos e com escolaridade $\geq 3^a$ classe. Para este grupo (N=73), com idades e escolaridades médias de 68.05 (*dp*=4,54) e 5.59 (*dp*=3,80) respectivamente, obtiveram-se valores médios para o MoCA de 22.75 (*dp*=3.91). Atendendo ao critério acima citado de >1.5 *dp* inferior às normas ajustadas para a idade (neste caso, de 5.85), verificamos que valores <17 neste teste são sugestivos de Declínio Cognitivo Ligeiro.

descrever estas provas

2.1. *Test of Memory Malingering (TOMM)*

O *Test of Memory Malingering* (TOMM) é um TVS, de escolha forçada, desenvolvido por Tombaugh (1996), constituído por 50 itens (objectos comuns). Originalmente desenvolvido para adultos, permite avaliar a “simulação” de défices mnésicos. É composto por dois Ensaio de Aprendizagem e por um Ensaio de Retenção, de aplicação opcional. É atribuído 1 ponto por cada resposta correcta. Assim, em cada um dos ensaios, as pontuações podem variar entre o valor mínimo de 0 (nenhuma resposta correcta) e o valor máximo de 50 (todas as respostas correctas). Segundo Tombaugh (1996), um resultado inferior a 45 acertos no Ensaio 2 ou no Ensaio de Retenção do TOMM é sugestivo de “simulação” (ou “esforço insuficiente”).

2.2. *Rey 15-Item Memory Test (15-IMT)*

O *Rey 15-Item Memory Test* (Rey, 1964) é, depois do TOMM, um dos instrumentos mais usados (Slick, Tan, Strauss & Hultsch, 2004) para detectar a “simulação” de défices mnésicos (Reznek, 2005) ou o “esforço insuficiente” (Slick et al., 1999). Consiste num cartão com 15 itens (5 linhas, com 3 caracteres cada), apresentado ao sujeito durante 10 segundos, após os quais o cartão é removido e lhe é pedido para reproduzir o máximo dos 15 itens que conseguir recordar (Ensaio de Evocação Imediata). Boone e colegas (2002) alargaram o protocolo do teste propondo um Ensaio de Reconhecimento, que consiste na apresentação dos 15 itens do cartão estímulo original e 15 distractores, semelhantes aos itens do cartão inicial. Para além da consideração do número total de itens correctamente reproduzidos (independentemente da sua localização espacial) no Ensaio de Evocação Imediata, os indicadores relativos ao Ensaio de Reconhecimento contemplam: (1) o número de itens correctamente identificados; (2) o número de falsos positivos (itens incorrectamente assinalados); e (3) a pontuação combinada (resultado combinado do reconhecimento = número itens correctamente evocados + (número de itens correctamente

reconhecidos – número de falsos positivos).

2.3. Memória de Dígitos

Este subteste da *Wechsler Memory Scale – Third Edition* (WMS-III; Wechsler, 1997, 2008) permite examinar as capacidades de evocação imediata, reversibilidade (dos números em sentido directo para os números em sentido inverso), concentração e atenção, sequenciação auditiva e aprendizagem, sendo ainda conceptualizado como uma medida de avaliação da memória a curto prazo. Especificamente, este subteste exige que o examinado repita várias séries de números, de diferentes extensões, ditos verbalmente pelo examinador pela ordem directa (associada à eficiência dos processos atencionais) e inversa (requer o armazenamento de alguns elementos durante um curto período de tempo, enquanto são trabalhados mentalmente para que sejam evocados em sentido contrário).

A capacidade para repetir dígitos na ordem de apresentação verbal (sentido directo) é considerada uma medida de concentração e, em média, um adulto consegue evocar 7 dígitos com uma variabilidade de 2 dígitos acima ou abaixo deste valor (Miller, 1956; Spitz, 1972, *cit in* Wechsler, 2008). Esta tarefa tende a manter-se relativamente estável ao longo da vida. Contudo, a tarefa em sentido inverso é mais afectada pela idade e pela incapacidade, verificando-se um decréscimo acentuado a partir dos 70 anos (Lezak, 1995, *cit in* Wechsler, 2008).

2.4. Memória Lógica

Este subteste da WMS-III (Wechsler, 1997, 2008) é constituído por três tarefas administradas em dois momentos: Memória Lógica I, na qual são apresentadas 2 histórias curtas ao sujeito, que devem ser reproduzidas imediatamente (a segunda história é lida 2 vezes e evocada também 2 vezes); e, Memória Lógica II (apresentado entre 25 a 30 min. após o subteste Memória Lógica I), que requer a evocação diferida das duas histórias antes apresentadas, sendo imediatamente seguida de um ensaio de reconhecimento.

Este subteste avalia a memória através da apresentação verbal de

material conceptual, sendo que o desempenho do indivíduo vai depender de factores como a capacidade para ouvir e processar sons falados, capacidade para compreender estruturas gramaticais lógicas e concentração.

As cotações são atribuídas de acordo com a precisão com que o sujeito é capaz de repetir cada uma das histórias (Unidades de História), bem como a sua capacidade para recordar os temas da história (Unidades Temáticas).

2.5. Cenas de Família

O subteste Cenas de Família foi concebido para avaliar a capacidade de recordar personagens em determinadas cenas, bem como as suas actividades e localizações. Este subteste é composto por duas tarefas administradas em dois momentos: Cenas de Família I, em que inicialmente é apresentada ao sujeito uma gravura com uma família (7 personagens), seguindo-se uma série com 4 cenas em que aparecem algumas das 7 personagens. Após a apresentação de cada uma das cenas, segue-se uma tarefa de evocação imediata (quem estava presente na cena, em que local e que actividade estava a desempenhar); Cenas de Família II: subteste realizado 25 a 30 min. após o anterior, no qual o sujeito deverá responder às perguntas colocadas precedentemente sem voltar a visualizar as imagens. A interpretação dos resultados neste subteste considera o peso relativo dos três elementos incluídos no procedimento de cotação: memória da personagem, actividade e localização. As pontuações baseiam-se especialmente nas personagens, ou seja, a atribuição de pontos pela actividade e localização correctas é apenas considerada quando existe uma correcta identificação da personagem (Wechsler, 1999, 2008).

2.6. Montreal Cognitive Assessment (MoCA)

O MoCA é um instrumento de rastreio cognitivo criado por Nasreddine e colaboradores (2005) para a detecção de casos de DCL, situação onde o MMSE se tem revelado pouco sensível. Avalia oito dimensões através de diversas tarefas: Memória a Curto-Prazo (evocação diferida, 5 pontos); Capacidades Visuoespaciais (desenho do relógio, 3 pontos; cópia de um cubo, 1 ponto); Funções Executivas (*trail-making test*

B, 1 ponto; fluência verbal fonémica, 1 ponto; abstracção verbal, 2 pontos); Atenção, Concentração e Memória de Trabalho (detecção do alvo, 1 ponto; subtracção em série, 3 pontos; sequência numérica, 2 pontos); Linguagem (nomeação, 3 pontos; repetição de frases, 2 pontos); e Orientação (dia, mês, ano, dia da semana, lugar e localidade, 6 pontos) (Nasreddine et al., 2005; Martins, 2007; Duro, 2008; Freitas, 2008).

No estudo original, que incluiu sujeitos com DCL, Doença de Alzheimer (DA) ligeira e um grupo de controlo (indivíduos sem deteriorização cognitiva), os autores recorreram a um ponto de corte de 26 pontos (em 30 possíveis), comum às versões inglesa e francesa. Com base neste ponto de corte, foi possível observar uma sensibilidade excelente do MoCA na identificação do DCL e da DA (90% e 100%, respectivamente), bem como uma especificidade muito boa (87%).

No estudo original do MoCA verificou-se que os sujeitos com 12 ou menos anos de escolaridade tendiam a obter piores desempenhos. Para corrigir os efeitos da escolaridade, é atribuído um ponto adicional aos sujeitos com escolaridade ≤ 12 anos (Nasreddine et al., 2005).

2.7. Vocabulário

O subteste de Vocabulário pertence à escala verbal da WAIS-III (Wechsler, 1997, 2008) e é uma tarefa que envolve a capacidade para expressar ideias por palavras, avaliando a capacidade intelectual verbal. A dimensão avaliada é a inteligência (verbal, cristalizada, pré-morbida), que se acredita manter estável ao longo da vida. Neste teste, a tarefa do sujeito consiste em definir oralmente um conjunto de palavras (por ordem de dificuldade crescente) que o examinador lê em voz alta, ao mesmo tempo que as apresenta impressas num cartão. As respostas são cotadas por comparação, quer com os critérios de correcção, quer com os exemplos de respostas apresentados no manual. As pontuações variam entre 2, 1 ou 0 pontos, sendo que o resultado total oscila entre 0 e 66 pontos.

2.8. Instrumento de Avaliação Funcional de Adultos e Idosos (IAFAI, versão experimental)

O IAFAI é um instrumento de avaliação funcional, constituído por 5 módulos: Avaliação Funcional Geral; Actividades Básicas de Vida Diária (domínios: alimentação, vestir, higiene, controlo dos esfíncteres, mobilidade e transferências); Actividades Instrumentais Familiares de Vida Diária (domínios: conversação e uso do telefone, preparar as refeições/fazer compras, cuidar da casa, segurança em casa); Actividades Instrumentais Avançadas de Vida Diária (domínios: compreensão e comunicação, tomada de decisão relativa à saúde, finanças e correspondência, deslocação e uso de transportes e lazer e relações interpessoais).

Este instrumento pode ser respondido pelo doente ou pelo cuidador. O sujeito respondente deve reportar-se ao último mês e considerar um gradiente de nível de dificuldade. Os níveis de análise contemplados no IAFAI foram conceptualizados à luz da definição de funcionalidade e incapacidade da Organização Mundial de Saúde, enquadrados na Classificação Internacional de Funcionalidade (CIF; OMS, 2001). Ou seja, radicando no modelo biopsicosocial de funcionalidade da OMS em que esta constitui um resultado da interacção entre o estado de saúde do indivíduo e factores contextuais, o IAFAI parte da consideração de vários níveis para a determinação da capacidade funcional, nomeadamente do nível de aptidão funcional do indivíduo e a sua condição de saúde, bem como o contexto em que se insere, incluindo a assistência disponível/barreiras e os factores de natureza pessoal.

2.9. Geriatric Depression Scale (GDS)

A Escala de Depressão Geriátrica (*Geriatric Depression Scale* - GDS; Yesavage et al., 1983; tradução portuguesa de Barreto, Leuschner, Santos & Sobral, 2008) é uma escala breve de auto-resposta, especificamente concebida para proceder ao rastreio de sintomas depressivos na população geriátrica. O tempo de administração é de aproximadamente 5-10 minutos. Nesta escala foram contemplados os sintomas afectivos e comportamentais desta perturbação e excluídos sintomas somáticos e vegetativos (Strauss, Sherman & Spreen, 2006).

A GDS é constituída por 30 questões que requerem respostas do tipo sim/não, considerando o que o sujeito tem sentido durante a última semana. Cada resposta é cotada com 1 ou 0 pontos, consoante esta indique a presença ou ausência, respectivamente, de determinado sintoma. A pontuação total pode variar entre 0 e 30. As respostas “sim” nas questões 2-4, 6,8,10-14, 16-18, 20, 22-26, 28 são cotadas com 1 ponto. As restantes perguntas são cotadas com 1 ponto se a sua resposta for “não” (sentido inverso). Resultados mais elevados na GDS são indicativos de maior gravidade de depressão. Quando o indivíduo tem uma pontuação entre 0-10 considera-se “ausência de depressão”; nos casos em que a pontuação se situa no intervalo 11-20, existirão sintomas de “depressão ligeira” e, no intervalo 21-30, é considerado ter “depressão grave” (Barreto, Leuschner, Santos & Sobral, 2008).

3. Tratamento Estatístico dos Dados

Todas as análises estatísticas foram efectuadas através do *Statistical Package for the Social Sciences* (SPSS), versão 16.0 para o Windows. Um $p < .05$ foi considerado estatisticamente significativo. Quanto aos métodos estatísticos empregues, para a comparação de grupos recorreremos ao teste não paramétrico *U* de *Mann-Whitney* (2 grupos). No estudo das variáveis nominais, foi utilizado o teste de Qui-Quadrado. Finalmente, para o estudo da associação entre as variáveis foram efectuadas correlações de Spearman.

IV – Apresentação e discussão dos resultados

1. Efeito das variáveis género, idade, anos de escolaridade e situação laboral nos resultados no TOMM

Para a análise das variáveis idade e anos de escolaridade, foram estudadas as correlações destas com os desempenhos no TOMM. Em ambos os grupos, os coeficientes de correlação⁵ entre a idade e os resultados no

⁵ Os coeficientes de correlação variam entre -1 e 1, sendo que quanto mais próximo o valor do coeficiente estiver destes extremos, maior será a associação linear entre as variáveis. O sinal negativo da correlação significa que as variáveis se relacionam

TOMM, considerando os 3 Ensaio, não se demonstraram estatisticamente significativos ($p < .05$), como apontado no Quadro 3. Estes resultados estão de acordo com o referido originalmente por Tombaugh (1996) e corroborado posteriormente, noutros trabalhos (Tombaugh, 1997; Horton, 2000; Constantinou & McCaffrey, 2003; Ashendorf, Constantinou & McCaffrey, 2004; Donders, 2005; Rees, Tombaugh, Gansler & Moczynski, 1998; Rees et al., 1998), inclusive em estudos portugueses (cf. Mota, 2008). No entanto, estes resultados são opostos aos obtidos por Teichner e Wagner (2004) que obtiveram correlações significativas entre a idade e o Ensaio 2 ($r = -.40$) e a idade e o Ensaio de Retenção ($-.50$), na sua amostra de adultos idosos.

No que concerne à variável anos de escolaridade, verificamos que, no geral, esta não influencia os resultados no TOMM. A única exceção foi encontrada no grupo DCL onde se verificou uma associação significativa, moderada e negativa, entre os anos de escolaridade e o Ensaio de Retenção ($r_s = -.478$), sugerindo que quanto mais escolarizados são os sujeitos com DCL, menores são as pontuações no Ensaio de Retenção do TOMM. Esta associação completamente inesperada poderá dever-se ao facto de durante o processo de avaliação ter eventualmente ocorrido um desinteresse e desatenção progressivos dos sujeitos na tarefa de Retenção, em consequência de uma concomitante maior confiança na sua resposta ao teste. Com esta exceção, os resultados corroboram estudos que convergentemente assinalam que as pontuações no TOMM não são influenciadas pela escolaridade (cf. Horton, 2000; Tombaugh, 1996, 1997; Teichner & Wagner, 2004; Rees et al., 1998; Faustino, 2009). Contrariamente, os dados obtidos por Mota (2008) e Cruz (2008), parecem sugerir uma possível influência da escolaridade nos resultados do TOMM. As diferenças encontradas entre os estudos de Mota (2008) e Cruz (2008) e o presente estudo poderão dever-se ao facto de os primeiros usaram nas suas amostras sujeitos menos escolarizados (respectivamente 1º e 2º ano de escolaridade).

em sentido contrário, isto é, as categorias mais elevadas de uma variável estão associadas a categorias mais baixas da outra variável. Por convenção, sugere-se que R menor que 0.2 indica uma associação muito baixa; entre 0.2 e 0.39 baixa; entre 0.4 e 0.69 moderada; entre 0.7 e 0.89 alta; e por fim entre 0.9 e 1 uma associação muito alta (Pestana & Gageiro, 2003, p. 189).

Quadro 3. TOMM: Efeito das variáveis Género, Idade, Escolaridade e Situação Laboral

	Variáveis	TOMM_E1	TOMM_E2	TOMM_ER
Controlo (N=20)	Género	$U=39,50$	$U=37,00$	$U=44,00$
	Idade	$r_s=-,108$	$r_s=,231$	$r_s=,341$
	Anos Escolaridade	$r_s=-,221$	$r_s=-,242$	$r_s=-2,10$
	Sit. Laboral	$U=29,00$	$U=29,50$	$U=25,50$
DCL (N=20)	Género	$U=37,00$	$U=45,00$	$U=41,500$
	Idade	$r_s=,000$	$r_s=,045$	$r_s=-,035$
	Anos Escolaridade	$r_s=,104$	$r_s=-,109$	$r_s=-,478^*$
	Sit. Laboral	$U=21,50$	$U=22,00$	$U=29,00$

* $p < .05$; ** $p < .01$

No sentido de avaliar as diferenças de género e da situação laboral nos resultados no TOMM, os desempenhos em ambos os grupos foram comparados nestas variáveis. Como reportado no Quadro 3, não foram observadas diferenças de género nos resultados obtidos em ambos os grupos ($p > .05$), sendo que esta variável parece não influenciar os desempenhos na prova. Estes dados vão ao encontro dos obtidos por Tombaugh (1996, 1997) e por Rees e colaboradores (2001).

Para a variável situação laboral também não foram observadas diferenças significativas nos desempenhos dos sujeitos reformados ou no activo ($p > .05$). Contudo, esta é uma variável que não surge nos estudos disponíveis pelo que os nossos resultados não podem ser comparados.

2. Variáveis capacidade funcional, sintomatologia depressiva e capacidade intelectual nos desempenhos no TOMM

Para avaliar a relação entre o TOMM e as variáveis sintomatologia depressiva, inteligência cristalizada/verbal e capacidade funcional (cf. Quadro 4) foram efectuadas análises de correlação, para cada um dos grupos em análise.

De um modo geral, não se observaram associações estatisticamente significativas ($p > .05$) entre a capacidade intelectual verbal (medida pelo subteste do Vocabulário), a sintomatologia depressiva (medida pela GDS), a capacidade funcional (medida pelo IAFAI) e os desempenhos em cada um dos Ensaio do TOMM. Apenas se verificou, no grupo de controlo, uma correlação negativa e moderada, estatisticamente significativa, entre a sintomatologia depressiva e o Ensaio 2 do TOMM ($r_s = -.475$). Este resultado

sugere que para os sujeitos cognitivamente normais, a sintomatologia depressiva influencia os resultados obtidos no Ensaio 2, mas não ao ponto do TOMM identificar incorrectamente estes sujeitos como possíveis “simuladores”. Aliás, os sujeitos do grupo de controlo obtiveram resultados médios bastante superiores ao ponto de corte sugerido para o Ensaio 2 ($M=49.30$; $dp=1.22$; $Min=46$; $Máx=50$).

Quadro 4. TOMM: Efeito das variáveis capacidade funcional, sintomatologia depressiva e capacidade intelectual geral

	Medidas	TOMM E1	TOMM E2	TOMM ER
Controlo (N=20)	Vocabulário	$r_s=.088$	$r_s=-.139$	$r_s=.140$
	GDS	$r_s=-.159$	$r_s=-.475^*$	$r_s=-.273$
	IAFAI	$r_s=-.155$	$r_s=.005$	$r_s=.135$
DCL (N=20)	Vocabulário	$r_s=.363$	$r_s=.012$	$r_s=-.114$
	GDS	$r_s=-.224$	$r_s=-.189$	$r_s=.170$
	IAFAI	$r_s=.006$	$r_s=-.124$	$r_s=-.094$

* $p<.05$; ** $p<.01$

Tentando associar a sintomatologia depressiva aos desempenhos no TOMM, verificamos que, como refere Tombaugh (1996) e testado noutros estudos (cf. Iverson et al., 2007; Rees, Tombaugh & Boulay, 2001; O`Bryant, Finlay & O`Jile, 2007; Yanez et al., 2006), este instrumento não é, de um modo geral, influenciado por esta variável. Estas conclusões estão, igualmente em consonância com as conclusões de Tombaugh (1996, 1997) e de Ashendorf, Constantinou e McCaffrey (2004) ao apontarem que o desempenho no TOMM não é afectado pelos efeitos da depressão em adultos idosos. Controversamente, os resultados obtidos no presente estudo são opostos aos encontrados por Cruz (2008). No seu estudo, Cruz (2008) sugere que os resultados no TOMM podem ser influenciados pela sintomatologia depressiva em idosos, sendo que a presença de um estado depressivo poderá contribuir para que o TOMM aponte, erradamente, alguns idosos deprimidos como possíveis “simuladores”. É, no entanto, necessário enfatizar que o estudo de Cruz (2008) abrange sujeitos com depressão grave (que foram excluídos do presente estudo) e desconhece-se o tipo e a severidade dos défices apresentados por estes indivíduos.

Para a variável capacidade intelectual verbal, também não foram observados coeficientes de correlação estatisticamente significativos com as pontuações no TOMM. Estes resultados revelam-se contrários aos obtidos

recentemente em Portugal. Mota (2008a, 2008b) sugere que os resultados no TOMM podem ser influenciados pela inteligência, tendo registado correlações significativas e moderadas entre o subteste de Cubos da WAIS-III e os três ensaios do TOMM (respectivamente, .352; .325; .237). Pedrosa (2009) concluiu que um QIEC (WAIS-III) baixo (<80) se associa a resultados igualmente baixos em provas de esforço reduzido (TOMM e 15-IMT). Faustino (2009) também comprovou a existência de relações entre os Ensaios do TOMM e medidas cognitivas (TFVS⁶ Animais; TFVS Alimentos; TFVS Animais e Alimentos; Semelhanças, Semelhanças e Cubos; TMT A tempo e erros; TMT B tempo). É, no entanto, necessário enfatizar que estes estudos não abrangem sujeitos idosos.

Quanto à variável capacidade funcional, também não foram observadas associações significativas com as pontuações no TOMM (coeficientes variam entre -.155 (muito baixo) e .135 (muito baixo)), não sendo possível a comparação destes dados com outras investigações, uma vez que esta variável não é considerada neste tipo de estudos. A inclusão de um instrumento de avaliação funcional teve como objectivo inicial caracterizar as amostras em estudo. A selecção do IAFAI, inseriu-se num estudo mais alargado de validação do instrumento, onde se considerou importante explorar a sua aplicabilidade junto a idosos, cognitivamente normais e com declínio cognitivo.

Perante estes resultados, e dadas as discrepâncias encontradas com os estudos realizados em Portugal, consideramos essencial replicar estas análises com outros grupos de adultos idosos para melhor compreender se as variáveis depressão grave (cf. Cruz, 2008) e inteligência (cf. Mota, 2008a, 2008b; Faustino, 2009; Pedrosa, 2009) podem ou não enviesar os resultados no TOMM.

3. Intercorrelações e Consistência interna

Pela análise do Quadro 5, constata-se que nos dois grupos considerados é evidente a tendência para associações com significado estatístico ($p < .05$) entre os resultados obtidos nos vários Ensaios do TOMM, verificando-se que a intensidade das correlações varia entre .470 (moderada)

⁶ Teste de Fluência Verbal Semântica

e .797 (alta).

No grupo de controlo observam-se correlações moderadas e significativas entre as pontuações obtidas no Ensaio 1 e as obtidas nos dois últimos Ensaios do TOMM (respectivamente, $r=.470$ e $r=.495$). Contudo, a associação entre o Ensaio 2 e o Ensaio de Retenção não se revelou estatisticamente significativa ($p>.05$). No que concerne às correlações observadas para o grupo DCL, verificaram-se correlações satisfatórias, especialmente entre o Ensaio 2 e o Ensaio de Retenção ($r=.797$). Assim, enquanto no grupo de controlo, sem défices cognitivos, a pontuação do Ensaio de Retenção se revela associada de forma significativa ao desempenho no Ensaio 1 (onde poderá haver um maior investimento e motivação para a tarefa), no grupo DCL, eventualmente devido aos problemas cognitivos, é a pontuação no Ensaio 2 que se revela associada de uma forma mais evidente ao desempenho no Ensaio de Retenção.

Estes resultados sugerem que um resultado elevado num dos Ensaios está associado a resultados elevados nos restantes e que um resultado baixo num dos Ensaios está associado a resultados baixos nos Ensaios restantes. Resultados semelhantes foram obtidos por McCaffrey, O`Bryant e Fisher (2003), por O`Bryant, Finlay e O`Jile (2007), Maior (2008) e Mota (2008).

Quadro 5. TOMM: Intercorrelações entre resultados, para os grupos (Controlo, DCL).

	Controlo (n=20)			DCL (n=20)		
	Ens_1	Ens_2	Ens_Ret	Ens_1	Ens_2	Ens_Ret
Ens_1	—	.470*	.495*	—	.571**	.479*
Ens_2	—	—	.377	—	—	.797**
Ens_Ret	—	—	—	—	—	—

* $p<.05$; ** $p<.01$

Especificamente, McCaffrey e colaboradores (2003) encontraram correlações significativas e muito altas entre pontuações do Ensaio 1 com os resultados dos Ensaios 2 e de Retenção (respectivamente .928 e .904), bem como entre os resultados do Ensaio 2 e do Ensaio de Retenção (.952). Por sua vez, O`Bryant e colegas (2007) encontraram associações significativas e altas entre os desempenhos do Ensaio 1 e os Ensaios 2 e de Retenção (respectivamente, .84 e .86) e muito altas entre o Ensaio 2 e o Ensaio de Retenção (.94). As associações mais fortes, comparativamente às encontradas no presente estudo, obtidas por McCaffrey e colaboradores

(2003) poder-se-ão explicar pelo facto destes autores terem utilizado uma amostra de sujeitos com traumatismo crânio-encefalico, mais jovem ($M=41.1$; $dp=2.7$) e com mais anos de escolaridade ($M=13.1$; $dp=2.7$). No estudo de O`Bryant e colegas (2007) também se verificou que os sujeitos eram relativamente jovens ($M=43.94$; $dp= 16.57$), com mais anos de escolaridade ($M=12.91$; $dp=3.03$) e estavam diagnosticados com uma grande diversidade de perturbações (e.g. depressão major, demência vascular, PTSD, esforço insuficiente/simulação, distúrbio bipolar, dependência alcoólica, demência fronto-temporal, défice de atenção/hiperactividade...).

Maior (2008) encontrou correlações significativas, de valor moderado (.395 entre o Ensaio 1 e o Ensaio de Retenção; .440 entre o Ensaio 1 e o Ensaio 2) a satisfatório (.656 entre o Ensaio 2 e o Ensaio de Retenção). No caso de Mota (2008a), foram encontradas correlações satisfatórias entre as pontuações no Ensaio 1 e os dois últimos Ensaios do TOMM (respectivamente, .678 e .637), mas especialmente entre os desempenhos do Ensaio 2 e o Ensaio de Retenção (.814) (todos com $p<.01$).

A **consistência interna**⁷ foi calculada, num primeiro momento para cada um dos ensaios, considerando também cada um dos grupos em estudo. O Ensaio 1 revela bons valores de consistência interna ($\alpha=.80$ no grupo de controlo e $\alpha=.83$ no grupo DCL), sendo os resultados nos restantes dois Ensaios inferiores: inadequados no Ensaio 2 e no Ensaio de Retenção do grupo de controlo, bem como no Ensaio 2 do grupo DCL; adequado com algumas carências no Ensaio de Retenção do grupo DCL. Contudo, saliente-se que nestes casos (Ensaio 2 e Ensaio de Retenção) a ausência de variância em alguns itens compromete uma análise mais exacta dos valores da consistência interna.

Num segundo momento, foram então calculados os valores de consistência para cada um dos grupos, relativamente aos 3 Ensaios do TOMM. Os valores obtidos revelam-se adequados com algumas carências no grupo de controlo ($\alpha=.63$), sendo inadequados no grupo DCL ($\alpha=.56$).

⁷ Por convenção sugere-se que $\alpha<.60$ é inadequado; $.60\leq\alpha<.70$ é adequado mas com algumas carências; $.70\leq\alpha<.80$ é adequado; $.80\leq\alpha<.85$ é bom; $\alpha\geq.85$ é excelente (Prieto & Muñiz, 2000)

Quadro 6. TOMM: Valores relativos à consistência interna para os grupos (Controlo, DCL).

	Controlo			DCL		
	Ens_1	Ens_2	Ens_Ret	Ens_1	Ens_2	Ens_Ret
Nº Itens	50	50	50	50	50	50
α	.63			.56		
α	.80	.56	.55	.83	.59	.65

Os valores de consistência interna obtidos no Ensaio 1 ($\alpha=.80$ no grupo de controlo e $\alpha=.83$ no grupo DCL) aproximam-se dos obtidos por Mota (2008a) ($\alpha=.85$ no Ensaio 1). O mesmo não se verifica nos restantes Ensaio, onde se verificou que os valores de consistência interna do presente estudo são inadequados (Grupo de controlo: $\alpha=.56$ no Ensaio 2 e $\alpha=.55$ no Ensaio de Retenção; Grupo DCL: $\alpha=.59$ no Ensaio 2) ou mesmo adequados com algumas carências ($\alpha=.65$ no Ensaio de Retenção do grupo DCL) comparativamente aos valores adequados obtidos por Mota (2008a) (E2 $\alpha=.78$; ER $\alpha=.74$). As diferenças observadas ao nível dos valores de consistência interna, nestes dois estudos, poderão dever-se ao facto da amostra utilizada por Mota (2008 a) ser significativamente maior ($n=103$) do que a utilizada no presente estudo (n controlo=20 e n DCL=20).

Desconhecemos outros resultados relativos à consistência interna do TOMM que possam funcionar como termo de comparação para os aqui discutidos.

4. Validade discriminante: Examinar a capacidade do TOMM na discriminação dos desempenhos de idosos normais e de idosos com Declínio Cognitivo Ligeiro

Um tópico controverso no que concerne ao TOMM diz respeito à sua aplicabilidade e fiabilidade em populações com défices cognitivos (cf. Dean et al., 2009). Assim, procuramos avaliar a existência de diferenças estatisticamente significativas entre os dois grupos quanto aos desempenhos em cada um dos ensaios do TOMM.

O Quadro 7 permite-nos atender mais pormenorizadamente aos valores obtidos no TOMM em cada um dos ensaios, pelos adultos idosos com e sem défice cognitivo. Com base nos resultados obtidos, verificamos que não existem diferenças estatisticamente significativas entre os dois grupos, quer no Ensaio 1 ($U=175.50$; $p=.512$), quer no Ensaio 2 ($U=184.00$;

$p=.678$), quer ainda no Ensaio de Retenção ($U=199.00$; $p=.989$), o que significa que os idosos do grupo DCL e do grupo de controlo apresentam desempenhos semelhantes no TOMM. Estes resultados vão ao encontro dos obtidos por Teichner e Wagner (2004), no qual os idosos sem e com defeito cognitivo (não dementes) apresentaram um padrão de resposta semelhante em todos os ensaios do TOMM. Controversamente, Cruz (2008) encontrou diferenças significativas ($p<.05$) ao nível da relação entre os resultados do TOMM nos idosos e a presença de defeito cognitivo, em todos os ensaios, sugerindo que o desempenho no TOMM poderá ser afectado pela presença de defeito cognitivo. No que diz respeito às discrepâncias verificadas neste estudo relativamente ao estudo de Cruz (2008), não se pode deixar de constatar alguns aspectos relevantes e que podem explicar os resultados encontrados. Em primeiro lugar, se compararmos as médias de idades deste estudo com o acima mencionado, verificamos que os sujeitos da amostra de Cruz (2008) representam uma faixa etária superior. Em segundo lugar, no estudo de Cruz (2008) desconhece-se o tipo e severidade dos défices dos sujeitos, cujo único despiste efectuado foi feito com recurso ao *Mini-Mental State Examination* (MMSE; Guerreiro, 1998; Guerreiro et al., 2008). Deste modo, podem existir vários défices na sua amostra de idosos com defeito cognitivo, incluindo estados demenciais ou pré-demenciais que poderão justificar os resultados obtidos por Cruz (2008). Convém referir uma vez mais que a literatura começa a apontar uma influência negativa da demência nos resultados do TOMM (cf. Dean et al., 2009; Teichner & Wagner, 2004). Em terceiro lugar, Cruz (2008) utilizou na sua amostra sujeitos com depressão severa (excluídos da amostra do presente trabalho), o que também poderá ter contribuído para os resultados obtidos.

Deste modo, consideramos essencial replicar estas análises, abrangendo outras amostras, principalmente idosos em estados iniciais de demência, para verificar se uma maior deterioração cognitiva influi nos resultados do TOMM.

Quadro 7. TOMM: Resultados em função do grupo (Controlo e DCL)

	Controlo (N=20)	DCL (N=20)	U	P
TOMM_E1	M=44,90; DP=4,77 (Min=33; Máx=50)	M=44,35; DP=4,77 (Min=33; Máx=50)	175,50	.512
TOMM_E2	M=49,30; DP=1,22 (Min=46; Máx=50)	M=49,18; DP=1,34 (Min=45; Máx=50)	184,00	.678
TOMM_ER	M=49,35; DP=1,14 (Min=46; Máx=50)	M=49,25; DP=1,33 (Min=45; Máx=50)	199,00	.989

Observando-se as correlações entre as pontuações do MoCA e os três Ensaio do TOMM (Quadro 8), verificam-se apenas associações significativas e moderadas para o Ensaio 1 e para o Ensaio 2 do TOMM (.538 e .572, respectivamente) no grupo de controlo. Contudo, esta associação significativa deixa de existir no Ensaio de Retenção. No grupo DCL não se verificam associações entre os resultados no MoCA e cada um dos Ensaio do TOMM.

Quadro 8. TOMM: Correlações com o MoCA, para o grupo de controlo e para o grupo DCL

	Medida	TOMM_E1	TOMM_E2	TOMM_ER
Controlo (N=20)	MoCA	$r=.538^*$	$r=.572^{**}$	$r=.360$
DCL (N=20)	MoCA	$r=.207$	$r=.064$	$r=-.027$

* $p<.05$; ** $p<.01$

Os resultados alcançados sugerem que, para o grupo de controlo, a capacidade cognitiva, avaliada pelo MoCA, demonstra estar associada aos resultados nos dois primeiros Ensaio de Aprendizagem (Ensaio 1 e Ensaio 2). Para o grupo DCL, os resultados são contrários, uma vez que o desempenho no MoCA não se revela associado de forma significativa às pontuações em nenhum dos Ensaio do TOMM (E1: $r=.207$; E2: $r=.064$; ER: $r=-.027$). Estes dados sugerem a presença de uma dimensão “cognitiva” no TOMM, mas apenas em sujeitos sem problemas cognitivos. Quando os défices cognitivos estão presentes, esta dimensão de natureza cognitiva deixa de ser relevante e não se revela associada aos resultados em cada um dos Ensaio do TOMM.

5. Validade referenciada a critérios externos:

5.1. Exame da associação entre os resultados no TOMM com as respostas no 15-IMT

O Quadro 9 reporta as correlações entre as pontuações em cada um dos Ensaio do TOMM e as pontuações nos Ensaio de Evocação Imediata e Reconhecimento⁸ do 15-IMT⁹, um outro instrumento específico na avaliação do esforço insuficiente. Apenas se verificaram associações estatisticamente significativas no grupo de controlo ($p < .05$). Especificamente, neste grupo observam-se correlações altas entre o Ensaio 1 do TOMM e o Ensaio de Evocação Imediata do 15-IMT ($r = .73$) e moderadas entre o Ensaio de Reconhecimento do 15-IMT e o Ensaio 1 e 2 do TOMM (respectivamente, $r_s = .579$ e $r_s = .478$). Contrariamente ao grupo de controlo, no grupo DCL não se verificaram associações com significado estatístico entre os desempenhos nos Ensaio do TOMM e do 15-IMT ($p > .05$).

Quadro 9. TOMM: Coeficientes de correlação com pontuações no Rey 15-Item Memory Test

	Controlo			DCL		
	Ens_1	Ens_2	Ens_Ret	Ens_1	Ens_2	Ens_Ret
REY_EVOC	.726**	.432	.215	.246	-.104	-.291
REY_REC	.579**	.478*	.225	.007	-.190	-.330

* $p < .05$; ** $p < .01$

McCaffrey, O`Bryant e Fisher (2003) encontraram associações significativas e moderadas entre a pontuação no Ensaio de Evocação Imediata do 15-IMT e os Ensaio 1 e 2 do TOMM (respectivamente, $r = .689$ e $r = .687$) e altas entre a Evocação Imediata do 15-IMT e o Ensaio de Retenção do TOMM ($r = .776$). Contudo, estes autores não calcularam as associações entre o Ensaio de Reconhecimento do 15-IMT e os três ensaios do TOMM. As diferenças verificadas entre os dois estudos devem ser cuidadosamente analisadas, uma vez que o estudo de McCaffrey e colaboradores (2003) foi realizado com uma amostra de sujeitos com traumatismo crânio-encefálico relativamente jovem ($M = 41.1$; $dp = 12.3$) e com uma maior escolaridade

⁸ Pontuação combinada

⁹ No Anexo 2 poderão ser consultados os resultados mais detalhados desta prova para ambos os grupos em análise.

($M=13.1$; $dp=2.7$), comparativamente à amostra utilizada no presente estudo. Recentemente Faustino (2009) observou que o TOMM e o 15-IMT, de modo geral, apresentam correlações baixas entre si, ainda que por vezes atinjam valores moderados (entre $r=-.205$ e $r=.486$). Estas diferenças de resultados podem ser compreendidas pelo facto da literatura sugerir uma influência de variáveis como a idade, inteligência e défice cognitivo nos resultados no 15-IMT (Frazen et al., 1990; Goldberg & Miller, 1986; Lee, Loring & Martin, 1992, *cit in* McCaffrey, O'Bryant & Fisher, 2003; Constantinou & McCaffrey, 2003), variáveis estas que distinguem as amostras de ambos os estudos.

5.2. Relação entre os resultados obtidos no TOMM e desempenhos em testes de Memória

O TOMM é um teste concebido para discriminar “simuladores” de doentes com perturbações reais de memória. Sendo um teste que avalia a “simulação” de défices mnésicos considerou-se relevante explorar a relação entre medidas de memória (nas suas componentes visual, verbal e auditiva) e desempenhos no TOMM.

São apresentados no Quadro 10, para ambos os grupos, os coeficientes de correlação entre as pontuações em cada um dos Ensaio do TOMM e os resultados (padronizados) obtidos nos testes de memória utilizados.

No grupo de controlo verificam-se associações estatisticamente significativas ($p<.05$) entre os resultados na Memória Lógica (Evocação Imediata e Diferida) e o Ensaio 1 do TOMM. Especificamente, as pontuações na Evocação Imediata e Diferida deste subteste da WMS-III correlacionam-se de forma positiva e moderada com o desempenho no Ensaio 1 do TOMM (respectivamente, $r_s=.553$ e $r_s=.478$). Também se verificam associações moderadas entre a pontuação na Evocação das Unidades Temáticas Diferidas (subteste Memória Lógica) e os resultados no Ensaio 2 ($r_s=.472$) e no Ensaio de Retenção ($r_s=.462$). No subteste Cenas de Família da WMS-III observam-se correlações significativas satisfatórias entre a pontuação no Ensaio 2 do TOMM e a Evocação Imediata ($r_s=.447$) e Diferida ($r_s=.514$) deste subteste.

No grupo de DCL apenas se observam associações moderadas entre o

desempenho no Ensaio Imediato da Memória Lógica e os resultados no Ensaio 1 do TOMM ($r_s=.464$), bem como entre as pontuações do Ensaio Diferido da Memória Lógica e as obtidas no Ensaio de Retenção do TOMM ($r_s=.502$).

Quadro 10. TOMM: Coeficientes de correlação com pontuações em testes de Memória

	Controlo			DCL		
	Ens_1	Ens_2	Ens_Ret	Ens_1	Ens_2	Ens_Ret
Mem_Dígitos_Total	.209	-.079	.201	.196	.121	-.041
Mem_Lógica_Imed	.553*	.400	.263	.464*	.355	.333
Mem_Lógica_Difer	.478*	.408	.375	.442	.345	.502*
M_Lógica_Temáticas_Imed	.293	.170	.299	.281	-.119	-.077
M_Lógica_Temáticas_Difer	.363	.472*	.462*	.321	.009	.221
Mem_Lógica_%Reten	.032	.181	.172	.401	.149	.353
Cenas_Família_I	.098	.447*	.363	.169	.328	.292
Cenas_Família_II	.239	.514*	.276	.221	.307	.186
Cenas_Família_Reten	.135	-.044	-.182	.238	-.026	-.253

* $p < .05$; ** $p < .01$

Podemos verificar que, apesar do TOMM se conceptualizar como um teste simples de reconhecimento visual de objectos, os coeficientes de correlação encontrados apontam para a existência de uma componente mnésica verbal na realização da prova. Os dados relativos à observação do comportamento durante as sessões de avaliação corroboram este facto, uma vez que, apesar da instrução contrária, os sujeitos insistiam em nomear os objectos apresentados, retendo não só a apresentação visual do objecto como guardando na sua memória o nome do objecto.

Estes resultados apontam a necessidade de mais estudos, com o objectivo de se averiguar se o desempenho mnésico deve ser considerado aquando da avaliação do “esforço insuficiente” em idosos com défices mnésicos. Convém lembrar que Hill e colegas (2003) encontraram relações moderadas e significativas entre medidas de memória declarativa e o desempenho no TOMM, em sujeitos com e sem patologia neurológica documentada. No seu estudo, as medidas de aprendizagem verbal e de memória verbal (medidas pelo CVLT e pela Memória Lógica da WMS-R) surgiram como preditores do desempenho no Ensaio 2 e/ou Ensaio de Retenção do TOMM. No entanto, Hill e colaboradores (2003) referem que apesar dos défices em medidas de memória declarativa poderem influenciar

negativamente o desempenho no TOMM, isto não se verifica ao ponto do teste identificar incorrectamente sujeitos “honestos” como “simuladores” (resultados médios acima do ponto de corte sugerido por Tombaugh, 1996).

6. Análise crítica comparativa do valor dos pontos de corte propostos por Tombaugh (1996) para o TOMM (considerando os seus vários Ensaios), e por Rey (1964) e Boone e colaboradores (2002) para o 15-IMT *Memory Test* considerando os dois grupos (DCL, controlo)

A identificação dos pontos de corte sugestivos de “esforço insuficiente” constitui uma tarefa essencial para a interpretação dos resultados neste tipo de instrumentos.

Como já referido anteriormente, Tombaugh (1996) indicou o ponto de corte <45 acertos no Ensaio 2 ou no Ensaio de Retenção para a identificação de sujeitos com “esforço insuficiente”, ponto de corte este que tem sido sistematicamente validado em muitos outros estudos. Quanto ao *Rey 15-Item Memory Test*, os dados relativos a esta questão não são convergentes, uma vez que têm sido propostos diferentes pontos de corte para o Ensaio de Evocação Imediata: 14 (DiCarlo, Gfeller & Drury citados em Frederick, 2002); 10 (Greiffenstein et al., 1996); 9 (p. ex., Goldberg & Miller, 1986); 8 (Bernard & Fowler, 1990); 7 (p.ex., Lee et al., 1992) e até 6 (Arnett, Hammeke & Schwartz, 1995). Recentemente, Boone e colaboradores (2000), com o objectivo de aumentar os valores de sensibilidade da prova sem sacrificar os parâmetros de especificidade, desenvolveram um Ensaio de Reconhecimento¹⁰, sugerindo para este ensaio um ponto de corte <20.

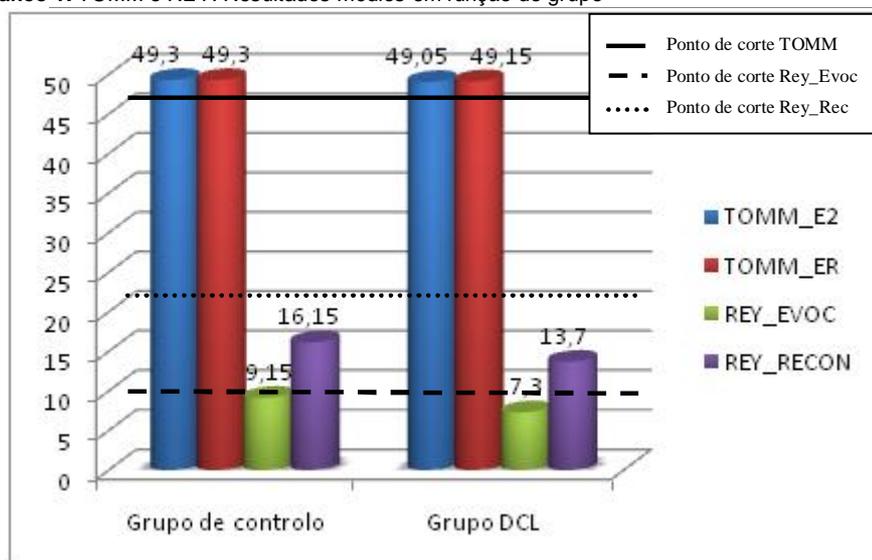
O Gráfico 1 expõe os resultados médios obtidos no TOMM e no 15-IMT, em cada um dos ensaios, para cada um dos grupos.

Se atendermos aos respectivos valores no TOMM (Quadro 12), verificamos que a média de acertos no Ensaio 2 e no Ensaio de Retenção no grupo de controlo (E2: $M=49.30$; $dp=1.22$; ER: $M=49.35$; $dp=1.12$) é superior ao ponto de corte proposto por Tombaugh (1996). O mesmo se

¹⁰ Resultado combinado do reconhecimento = número itens correctamente evocados + (número de itens correctamente reconhecidos – número de falsos positivos).

verifica para o grupo DCL, que obteve valores médios de 49.05 ($dp=1.47$) no Ensaio 2 e de 49.15 ($dp=1.53$) no Ensaio de Retenção. Importa salientar que em ambos os grupos, nenhum indivíduo pontuou abaixo dos pontos de corte considerados quer no Ensaio2, quer no Ensaio de Retenção. A média dos resultados para o Ensaio 1 também se revela elevada, quer no grupo de Controlo (E1: $M=44.90$; $dp=4.75$), quer ainda no grupo DCL (E1: $M=43.80$; $dp=5.29$).

Gráfico 1. TOMM e REY: Resultados médios em função do grupo



Estes resultados são convergentes com os encontrados em outros estudos (Horton, 2000; Tombaugh, 1996, 1997; Iverson et al., 2007; O`Bryant et al., 2007; Rees, Tombaugh, Gansler & Moczynski, 1998; Weinborn et al., 2003; Yanez et al., 2006; Rees, Tombaugh & Boulay, 2001; Haber & Fichtenberg, 2006; Greve et al., 2006; Etherton, Bianchini, Greve & Ciota, 2005), incluindo estudos portugueses (cf. Maior, Silva, Pedrosa, Dias, Simões & Pinho, 2006; Mota, 2008; Pedrosa, 2009; Maior, 2008; Faustino, 2009). Particularmente, os dados obtidos convergem com estudos que apontam para a utilização do ponto de corte ≥ 45 acertos no Ensaio 2 e/ou no Ensaio de Retenção, proposto por Tombaugh (1996), em idosos (Ashendorf, Constantinou & McCaffrey, 2004; Cruz, 2008), inclusive naqueles que manifestam algum grau de declínio cognitivo (Teichner & Wagner, 2004; Cruz, 2008).

Quadro 12. TOMM e REY: Resultados em função do grupo estudado (Controlo, DCL)

	Controlo (N=20)	DCL (N=20)	Pontos de corte
TOMM_E1	M=44.90; <i>dp</i> =4.75 (Min=33; Máx=50)	M=43.80; <i>dp</i> =5.29 (Min=33; Máx=50)	—
TOMM_E2	M=49.30; <i>dp</i> =1.22 (Min=46; Máx=50)	M=49.05; <i>dp</i> =1.47 (Min=45; Máx=50)	≥45
TOMM_ER	M=49.35; <i>dp</i> =1.12 (Min=46; Máx=50)	M=49.15; <i>dp</i> =1.53 (Min=45; Máx=50)	≥45
REY_EVoc	M=9.15; <i>dp</i> =3.22 (Min=3; Máx=14)	M=7.30; <i>dp</i> =2.92 (Min=3; Máx=14)	<9
REY_REC	M=16.15; <i>dp</i> =6.69 (Min=5; Máx=28)	M=13.70; <i>dp</i> =6.03 (Min=4; Máx=28)	<20

No que concerne ao 15-IMT, o grupo de controlo obteve valores médios de 9.15 (*dp*=3.22) no Ensaio de Evocação Imediata e de 16.15 (*dp*=6.69) no Ensaio de Reconhecimento. Por sua vez, o grupo DCL obteve valores médios de 7.30 (*dp*=2.92) no Ensaio de Evocação Imediata e de 13.70 (*dp*=6.03) no Ensaio de Reconhecimento. Tanto no grupo de controlo como no grupo de DCL, os resultados médios revelam-se inferiores aos pontos de corte sugeridos, exceptuando o Ensaio de Evocação Imediata no grupo de controlo (*M*=9.15; *dp*=3.22) que se encontra ligeiramente acima do valor mais comumente proposto (9).

Os dados obtidos pelo grupo DCL são convergentes com estudos que apontam para a falta de especificidade do ponto de corte habitualmente utilizado (9) para o Ensaio de Evocação Imediata, em populações com défice cognitivo (Dean et al., 2009; D'Arcy & McGlone, 2000). No mesmo sentido encontramos o estudo de Simões e colaboradores (submetido para publicação), que sugere que em grupos de adultos idosos, é necessário ajustar os pontos de corte para valores inferiores a 9 no Ensaio de Evocação Imediata e inferiores a 20 no Ensaio de Reconhecimento.

No quadro 13 são apresentadas as percentagens de sujeitos que pontuam acima e abaixo dos pontos de corte apontados para ambos os testes.

No grupo de controlo e no grupo DCL, 100% dos sujeitos pontuam acima do ponto de corte para o Ensaio 2 e para o Ensaio de Retenção, ou seja, o TOMM identificou todos os sujeitos como “não simuladores” ou como casos não sugestivos de desempenhos característicos de “esforço insuficiente”.

Quadro 13. TOMM e REY: Comparação na identificação de sujeitos com resultados sugestivos da presença de esforço insuficiente

	Ponto de corte	Controlo (N=20)	DCL (N=20)
TOMM_E2	<45	—	—
	≥45	20 (100%)	20 (100%)
TOMM_ER	<45	—	—
	≥45	20 (200%)	20 (100%)
REY_EVoc	<9	8 (40%)	14 (70%)
	≥9	12 (60%)	6 (30%)
REY_REC	<20	14 (70%)	15 (75%)
	≥20	6 (30%)	5 (25%)

No que respeita ao 15-IMT, no grupo de controlo 40% dos sujeitos pontuaram abaixo do ponto de corte para o Ensaio de Evocação Imediata e 70% no Ensaio de Reconhecimento. No Grupo DCL, estes valores percentuais aumentaram para 70% no Ensaio de Evocação Imediata e 75% no Ensaio de Reconhecimento. Novamente, estes resultados são convergentes com outros estudos que constataam que muitos sujeitos “não simuladores” (défice cognitivo, demência) pontuam abaixo do ponto de corte (9) sugerido para o Ensaio de Evocação Imediata (Dean et al., 2009; Simões et al., submetido para publicação). É importante salientar que em ambos os grupos, o 15-IMT identifica sujeitos com “esforço insuficiente”, sendo esta percentagem superior no grupo de sujeitos com DCL. Contrariamente ao esperado, no Ensaio de Reconhecimento houve uma maior percentagem de sujeitos identificados com “esforço insuficiente” (70% do grupo de controlo e 75% do grupo DCL), comparativamente ao Ensaio de Evocação Imediata.

Com vista a uma melhor compreensão dos pontos de corte relativos ao Ensaio 2 e ao Ensaio de Retenção do TOMM mas também ao Ensaio de Evocação Imediata e ao resultado combinado do Ensaio de Reconhecimento do 15-IMT, foram calculados os percentis, para cada um dos grupos. Os dados relativos aos Ensaio do TOMM são apresentados no Quadro 14 e os dados relativos aos Ensaio do 15-IMT são apresentados no Quadro 15.

No Ensaio 2, atendendo ao Percentil 5 dos resultados, o ponto de corte ≥45 acertos parece apresentar-se como adequado para o grupo de controlo e para o grupo DCL. Considerando o Ensaio de Retenção e atendendo também ao Percentil 5, o ponto de corte sugerido apresenta-se, do mesmo modo, adequado quer no grupo de controlo, quer no grupo DCL.

Quadro 14. TOMM: Percentis em função do grupo

Percentis	Controlo			DCL		
	Ens_1	Ens_2	Ens_Ret	Ens_1	Ens_2	Ens_Ret
5	33.15	46.00	46.05	33.05	45.05	45.05
10	36.30	46.30	47.10	34.30	46.10	46.10
25	42.00	49.00	49.00	39.50	49.00	48.25
50	46.00	50.00	50.00	45.00	50.00	50.00
75	48.75	50.00	50.00	48.75	50.00	50.00
90	50.00	50.00	50.00	50.00	50.00	50.00
95	50.00	50.00	50.00	50.00	50.00	50.00

Estes dados estão de acordo com o apontado em outros estudos, considerando grupos clínicos com lesões cerebrais de diferentes etiologias (Tombaugh, 1996, 1997; Haber & Fichtenberg, 2006; Rees, Tombaugh, Gansler & Moczynski, 1998), distúrbios afectivos (Tombaugh, 1996, 1997; Rees, Tombaugh & Boulay, 2001; Iverson et al., 2007; O`Bryant, Finlay & O`Jile, 2007; Etherton, Bianchini, Greve & Ciota, 2005) e mesmo idosos (Tombaugh, 1996; Ashendorf, Constantinou & McCaffrey, 2004; Cruz, 2008), inclusive aqueles com declínio cognitivo que não demência (Teichner & Wagner, 2004; Cruz, 2008).

Quadro 15. Rey 15-Item Memory Test. Percentis em função do grupo (Controlo, DCL)

Percentis	Controlo		DCL	
	REY_EVOC	REY_RECON	REY_EVOC	REY_RECON
5	3.05	5.05	3.00	4.10
10	4.20	6.10	3.10	6.20
25	6.00	10.25	6.00	9.25
50	9.00	17.50	6.50	12.50
75	12.00	21.00	9.75	19.00
90	13.00	24.90	11.90	21.90
95	13.95	27.85	13.90	27.70

No Ensaio de Evocação Imediata do 15-IMT, e com base no percentil 5, o ponto de corte sugerido (9) revela-se inadequado para os sujeitos deste estudo, quer estes façam parte do grupo de controlo (3.05), quer do grupo DCL (3.00). Uma pontuação média de 9 apenas se observa no percentil 50 para o grupo de controlo, situando-se perto do percentil 75 para o grupo DCL. Apesar de mais aceitável, o ponto de corte de ≤ 6 , apontado por Darcy e McGlone (2000) também se revela inadequado para ambos os grupos deste

estudo correspondendo a um percentil 25 de resultados. Estes resultados são convergentes com os obtidos por Simões e colaboradores (submetido para publicação) que anotam, para um ponto de corte de 6, a sua eventual inadequação para sujeitos idosos com défice cognitivo.

Ao desenvolver o Ensaio de Reconhecimento para o 15-IMT, Boone e colaboradores (2002) procuraram melhorar a sensibilidade do teste mantendo um elevado nível de especificidade. De facto, estes autores concluíram que o uso da Evocação livre combinada com as pontuações obtidas no Ensaio de Reconhecimento (utilizando o ponto de corte <20) aumenta a sensibilidade da prova de 47% para 71%. Contudo, neste estudo e considerando o percentil 5, foram observados resultados muito inferiores ao ponto de corte de 20 em ambos os grupos: 5.05 para o grupo de controlo e 4.10 para o grupo DCL. Resultados próximos ao ponto de corte sugerido foram apenas obtidos no percentil 75 no grupo de controlo, e no percentil 90 no grupo DCL. Convergentes com estes dados são os resultados de Simões e colaboradores (submetido para publicação), que no seu estudo obtiveram resultados substancialmente inferiores ao ponto de corte <20 no grupo dos adultos predominantemente idosos: 2.00 para sujeitos idosos com défices cognitivos e 11.70 para o grupo de controlo constituído por sujeitos idosos (percentil 5 dos resultados).

Com base nos resultados obtidos pelos dois grupos, o presente estudo sugere a necessidade de ajustar os pontos de corte para valores inferiores aos comumente utilizados. Neste âmbito, e embora sejam necessários mais estudos, os resultados encontrados sugerem que os pontos de corte devem ser ajustados para valores próximos a 3 (corresponde ao Percentil 5) no Ensaio de Evocação Imediata e para valores próximos a 5 (corresponde ao Percentil 5) no Ensaio de Reconhecimento, considerando os resultados no grupo de controlo e no grupo DCL.

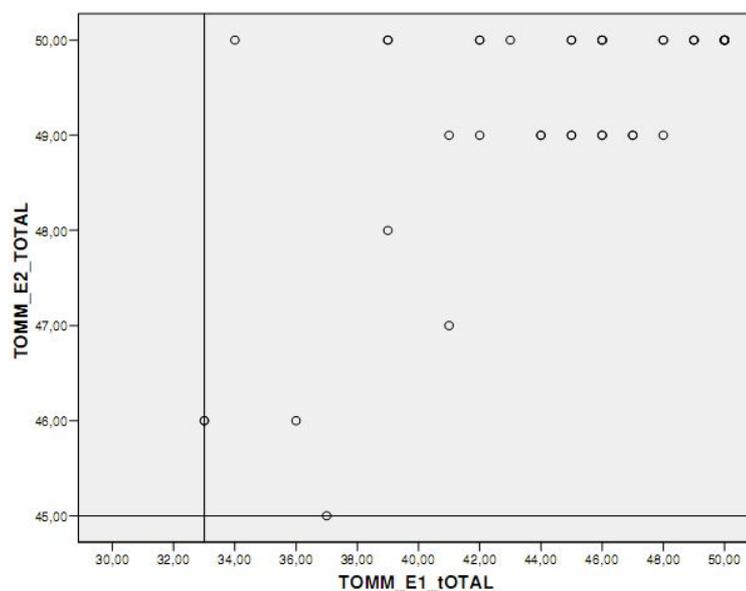
7. Versão abreviada do TOMM

Como anteriormente referido, o TOMM é um teste de validade de sintomas no qual um valor inferior a 45 acertos no Ensaio 2 ou no Ensaio de Retenção sugere “esforço insuficiente” na avaliação, não sendo os valores obtidos no Ensaio 1 habitualmente utilizados. Este estudo tem também como

objectivo obter um valor de ponto de corte no Ensaio 1 acima do qual os ensaios seguintes não necessitem de ser administrados, retomando e explorando sugestões de trabalhos recentes (cf. O'Bryant et al., 2007; Horner et al., 2006). Para tal, devido à não existência de diferenças estatisticamente significativas nos resultados no TOMM nos dois grupos e ao facto dos resultados se terem revelado superiores ao ponto de corte, foram utilizados, na prossecução deste objectivo, os sujeitos de ambas as amostras ($n=40$).

A distribuição das pontuações obtidas no Ensaio 2 e no Ensaio de Retenção, em função dos valores no Ensaio 1 são mostrados nas Figuras 1 e 2.

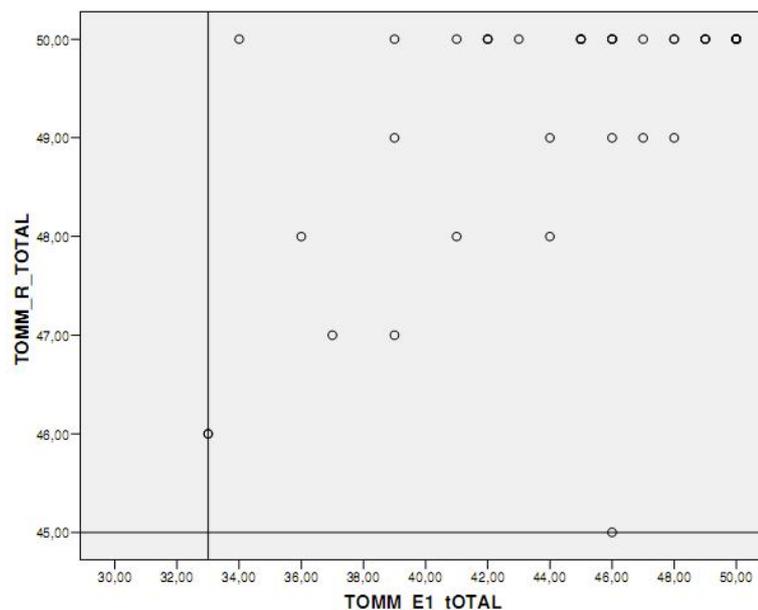
Figura 1. Gráfico de dispersão dos valores no Ensaio 2 do TOMM (N=40) em função dos valores no Ensaio 1



Legenda: Cada ponto representa um valor obtido pelos sujeitos. A linha horizontal indica o ponto de corte recomendado, abaixo do qual o teste sugere esforço insuficiente; a linha vertical indica o ponto de corte sugerido acima do qual o teste pode ser interrompido, sugerindo esforço suficiente na tarefa.

Um ponto de corte ≥ 33 no Ensaio 1 revela, com 100% de probabilidade, que o sujeito irá passar no Ensaio 2 e no Ensaio de Retenção. Isto é, os sujeitos que obtiveram valores iguais ou superiores a 33 no Ensaio 1, obtiveram valores iguais ou superiores a 45 no Ensaio 2 e no Ensaio de Retenção, sendo classificados como “não simuladores”.

Figura 2. Gráfico de dispersão dos valores no Ensaio de Retenção do TOMM (N=40) em função dos valores no Ensaio 1



Legenda: Cada ponto representa um valor obtido pelos sujeitos. A linha horizontal indica o ponto de corte recomendado, abaixo do qual o teste sugere esforço insuficiente; a linha vertical indica o ponto de corte sugerido acima do qual o teste pode ser interrompido após o Ensaio 1.

Estes resultados vão ao encontro dos obtidos por O`Bryant e colaboradores (2007) ao sugerirem que o Ensaio 1 do TOMM demonstra elevada precisão diagnóstica ao utilizar um ponto de corte > 32 acertos, parecendo ser uma opção viável para o exame do “esforço insuficiente”. Apesar de ligeiramente inferior, o ponto de corte aqui obtido aproxima-se do sugerido por Horner e colaboradores (2006) que indicam um ponto de corte igual ou superior a 36 no Ensaio 1 como traduzindo 99% de probabilidade dos sujeitos serem identificados no TOMM como “não simuladores”. Estamos de acordo com O`Bryant e colaboradores (2008) quando estes propõem que os participantes que obtêm pontuações ≥ 45 acertos no Ensaio 1 do TOMM irão continuar a obter resultados ≥ 45 acertos nos Ensaio seguintes, sendo que estes dois últimos ensaios se tornam dispensáveis.

Os presentes resultados apontam que, para amostras similares às utilizadas neste estudo, um valor igual ou superior a 33 no Ensaio 1 indica que os Ensaio seguintes não necessitam ser administrados. Porém, dado o impacto negativo da ansiedade e da depressão nos desempenhos no Ensaio 1 do TOMM (cf. O`Bryant, Finlay, & O`Jile, 2007), aconselha-se a aplicação

dos Ensaio 2 e de Retenção sempre que os sujeitos obtenham pontuações <33 acertos no Ensaio 1.

V – Conclusões

Actualmente, a pesquisa e a prática em avaliação (neuro)psicológica clínica e forense enfatizam a indispensabilidade do exame da “simulação” de alterações ou défices cognitivos. Neste plano, é cada vez mais consensual que a inexistência de testes de validade de sintomas (de que o TOMM é o exemplo mais conhecido) nos protocolos de avaliação torna o exame (neuro)psicológico incompleto e deve, por isso, ser justificado (Bush et al., 2005).

Em Portugal os estudos empíricos com este tipo de instrumentos estão ainda numa fase muito inicial e têm sido implementados especialmente em contexto forense. Pretende-se com este trabalho contribuir para a pesquisa desta problemática no nosso país, certos, porém, de que há ainda muito por fazer.

O estudo desenvolvido teve como principal objectivo verificar a aplicabilidade do TOMM em sujeitos com Declínio Cognitivo Ligeiro, por comparação a uma amostra de adultos idosos saudáveis (equiparáveis quanto às características demográficas). Os dados obtidos no presente estudo sugerem que os resultados no TOMM não parecem ser influenciados pela idade, escolaridade, género e situação laboral. Também não foram observadas associações estatisticamente significativas entre inteligência verbal, sintomatologia depressiva, capacidade funcional e desempenhos em cada um dos Ensaio do TOMM. Perante estes resultados, e dadas as discrepâncias encontradas com outros estudos realizados em Portugal (estudos que têm vindo a sugerir uma possível influência da idade, escolaridade e sintomatologia depressiva nos resultados no TOMM), consideramos essencial replicar estas análises com outros grupos de adultos idosos (idosos analfabetos e com depressão grave) para melhor compreender se estas variáveis podem determinar os resultados no TOMM.

Um tópico controverso no que concerne ao TOMM (e a outros testes de exame da simulação ou esforço reduzido) diz respeito à sua aplicabilidade e validade em populações com défices cognitivos (cf. Dean et al., 2009). No

presente estudo verificou-se que os idosos com DCL apresentam um padrão de resposta no TOMM semelhante ao apresentado pelo grupo de controlo. Especificamente, todos os sujeitos do grupo DCL pontuaram acima do ponto de corte proposto por Tombaugh (1996), sugerindo que as pontuações no TOMM não são influenciadas pela presença de Declínio Cognitivo Ligeiro.

Na análise das associações entre o TOMM e os testes de memória utilizados neste estudo, os resultados encontrados (de um modo geral correlações positivas mas de magnitude reduzida) apontam para a possibilidade de algum envolvimento mnésico na realização das tarefas, sugerindo a necessidade de considerar a variável memória aquando da avaliação do “esforço insuficiente” em pacientes neurológicos.

A identificação dos pontos de corte sugestivos de “esforço insuficiente” constitui uma tarefa essencial para a interpretação dos desempenhos neste tipo de instrumentos. Os resultados deste estudo, comprovam o valor do ponto de corte inicialmente proposto para o TOMM, quer no grupo controlo, quer mais significativamente, no grupo DCL. Comparativamente ao TOMM, o 15-IMT surge como um teste menos específico, ou mais problemático, na avaliação do esforço insuficiente, com resultados aquém dos esperados, quer para o grupo normativo, quer para o grupo DCL. Por sua vez, e embora sejam necessários mais estudos, os presentes resultados sugerem que os pontos de corte propostos na investigação publicada para o 15-IMT < 9 (Ensaio de Evocação Imediata) e <20 (Ensaio de Reconhecimento¹¹) não parecem ser adequados para o exame de comportamentos de “esforço insuficiente” por parte de adultos idosos, independentemente da presença ou não de problemas cognitivos. Os baixos resultados obtidos nos dois ensaios deste teste poderão dever-se a outras variáveis que não (apenas) o “esforço insuficiente” ou a intenção de “simular” défices de natureza cognitiva (memória, atenção). Consequentemente, consideramos importante ajustar os pontos-de-corte para valores inferiores aos propostos. Neste sentido, e com base nos dados do presente estudo os valores de 3, no Ensaio de Evocação Imediata, e de 5 no Ensaio de Reconhecimento poderão ser considerados, muito provisoriamente, como resultados sugestivos de esforço insuficiente ou reduzido. Estes valores são congruentes com os sugeridos por Simões e

¹¹ Resultado Combinado

colaboradores (2009, submetido para publicação) que sugerem a necessidade de ajustar os pontos de corte para valores inferiores a 6 (Ensaio de Evocação Imediata) e perto de 2 (Ensaio de Reconhecimento) aquando da avaliação de sujeitos idosos com défices cognitivos. Assim, e à semelhança do sugerido na literatura, o 15-IMT parece não ser insensível a formas de declínio e mesmo à idade.

Este trabalho procurou ainda indicar um valor de ponto de corte no Ensaio 1, que uma vez atingido torna dispensável o recurso ao Ensaio 2 e ao Ensaio de Retenção na determinação do “esforço insuficiente” do sujeito na tarefa. Os resultados apontam um ponto de corte igual ou superior a 33, no Ensaio 1, como traduzindo 100% de probabilidade dos sujeitos serem identificados no TOMM como “não simuladores”. É necessário sublinhar que os presentes resultados são preliminares, ou seja, de natureza ou valor exploratório. É necessário replicar este estudo com outros tipos de amostras de modo a aumentar a confiança no ponto de corte sugerido. Contudo, estes dados apontam um caminho prático para diminuir consideravelmente o tempo de administração deste instrumento. Isto poderá ser particularmente vantajoso em contexto clínico, no qual apenas uma bateria breve de testes (neuro)psicológicos é exequível devido ao tempo limitado para as avaliações e ao maior cansaço deste grupo de adultos idosos.

Em suma, no que concerne ao TOMM, e tendo em conta os dados do presente trabalho, bem como as conclusões de outras investigações portuguesas (Mota, 2008a; Cruz, 2008; Pedrosa, 2009; Maior, 2008; Faustino, 2009), este parece vislumbrar-se como um instrumento com elevado potencial de uso no nosso país, e deve por isso ser incluído nos protocolos de avaliação (neuro)psicológica.

O presente estudo apresenta alguns **limites**. Neste plano, é necessário referir que as amostras usadas não apresentavam qualquer evidência de suspeita de “simulação”, incentivos externos ou ganhos secundários. Futuros estudos deverão incluir grupos de idosos suspeitos de “esforço insuficiente” ou “simulação” (com motivação para exagerar sintomas), uma vez que o recurso a este tipo de testes supõe investigação empírica que evidencie que apenas os “simuladores” falham em tais medidas. Uma outra limitação remete para a natureza e o tamanho reduzido das amostras clínicas de

adultos idosos com DCL. A amostra de natureza mais “normativa” (grupo de controlo) deverá ser consideravelmente alargada de modo a incluir um maior número de adultos idosos, e sujeitos menos escolarizados, incluindo analfabetos. Será também interessante observar o desempenho de adultos idosos com depressão severa e mesmo com demência, uma vez que Cruz (2008) sugere que a depressão e os défices cognitivos poderão influenciar os resultados no TOMM em adultos idosos.

Por fim, é importante salientar que o diagnóstico de “simulação” deve ser estabelecido apenas quando o (neuro)psicólogo possui informação convergente, corroborada com recurso a vários instrumentos e registos, tendo em conta as consequências que tal diagnóstico pode desencadear na vida de outrem.

Bibliografia

- American Psychiatric Association (APA, 2002). *DSM-IV-TR: Manual de diagnóstico e estatística das perturbações mentais* (4ªed., texto revisto). Lisboa: Climepsi Editores.
- Andrade, S. A. O. (2007). *Geriatric Depression Scale (GDS-30): Estudos de validação em grupos com Declínio Cognitivo Ligeiro e Demência*. Tese de Mestrado Integrado. Coimbra: Faculdade de Psicologia e de Ciências da Educação.
- Arnett, P. A., Hammeke, T. A., & Schwartz, L. (1995). Quantitative and qualitative performance on Rey's 15-Item Test in neurological patients and dissimulators. *The Clinical Neuropsychologist*, 9, 17-26.
- Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology*, 19, 125-130.
- Barreto, J., Leuschner, A., Santos, F. & Sobral, M. (2008). Escala de Depressão Geriátrica. In Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* (2ª ed., pp.71-72). Lisboa: Novartis.
- Bender, S. D. (2008). Malingered Traumatic Brain Injury. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 69-86). New York: The Guilford Press.
- Bernard, L. C., & Fowler, W. (1990). Assessing the validity of memory complaints: Performance of brain-damaged and normal individuals on Rey's task to detect malingering. *Journal of Clinical Psychology*, 46, 434-436.
- Berry, D. T. R. & Schipper, L. J. (2008). Assessment of feigned cognitive impairment using standard neuropsychological tests. In R. Rogers (Ed), *Clinical assessment of malingering and deception* (pp. 237-255). New York: The Guilford Press.

- Blaskewitz, N., Merten, T., & Kathmann, N. (2008). Performance of child on symptom validity tests: TOMM, MSVT and FIT. *Archives of Clinical Neuropsychology*, 23, 379-391.
- Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., Shamieh, E., & Warner-Chacon, K. (2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology*, 17, 625-642.
- Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., et al. (2005). Symptom validity assessment: Practice issues and medical necessity. *Archives of Clinical Neuropsychology*, 20, 419-426.
- Chafetz, M. D. (2008). Malingering on the Social Security Disability Consultative Exam: Predictors and base rates. *The Clinical Neuropsychologist*, 22, 529-546.
- Constantinou, M., & McCaffrey, R. J. (2003). Using the TOMM for evaluating children`s effort to perform optimally on neuropsychological measures. *Child Neuropsychology*, 9 (2), 81-90.
- Cruz, A. P. S. (2008). *Quatro estudos exploratórios com o Test of Memory Malingering (TOMM)*. Tese de Mestrado em Psicologia (área de especialização em Avaliação Psicológica). Coimbra: Faculdade de Psicologia e de Ciências da Educação.
- D`Arcy, R. C., & McGlone, J. (2000). Profound amnesia does not impair performance on 36-item Digit Memory Test: A test of malingered memory. *Brain and Cognition*, 44, 54-58.
- Dean, A. C., Victor, T. L., Boone, K. B., Philpott, L. M. & Hess, R. A. (2009). Dementia and effort test performance. *The Clinical Neuropsychologist*, 23, 133-152.
- Donders, J. (2005). Performance on the Test of Memory Malingering in a mixed pediatric sample. *Child Neuropsychology*, 11, 221-227.
- Duro, D. F. D. (2008). *Montreal Cognitive Assessment (MoCA): Características psicométricas e capacidade diagnóstica num contexto de Consulta de Memória*. Tese de Mestrado Integrado em Psicologia (área de especialização em Psicologia Clínica e da Saúde, sub-especialização em Psicogerontologia Clínica). Coimbra: Faculdade de Psicologia e de Ciências da Educação.
- Etherton, J. L., Bianchini, K. J., Greve, K. W., & Ciota, M. A. (2005). Test of Memory Malingering performance is unaffected by laboratory-induced pain: Implications for clinical use. *Archives of Clinical Neuropsychology*, 20, 375-384.
- Faustino, J. D. P. R. (2009). *Violência, funcionamento cognitivo e “esforço reduzido”*: Estudos com grupos de reclusos. Tese de Mestrado Integrado em Psicologia (área de especialização em Psicologia Clínica e da Saúde, sub-especialização em Psicologia Forense). Coimbra: Faculdade de Psicologia e Ciências da Educação.
- Ferguson, K. E. (2004). Detecting malingering in forensic neuropsychological evaluations in litigants with mild traumatic brain injury. In William O`Donohue & Levensky (Eds.), *Handbook of forensic psychology* (pp. 301-314). San Diego, CA: Elsevier Academic Press.
- Frederick, R. I. (2002). A review of Rey`s strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology*, 2, 1-25.

- Freitas, S., Simões, M. R., & Santana, I. (2008). Dados normativos preliminares da versão portuguesa do Montreal Cognitive Assessment (MoCA). Comunicação pessoal.
- Gavett, B. E., O'Bryant, S. E., & McCaffrey, R. J. (2004). The clinical utility of TOMM Trial 1 in detecting suboptimal performance: A cross-validation study. *Archives of Clinical Neuropsychology*, *19*, 860.
- Goldberg, T. O. & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology*, *42*, 792-795.
- Greiffenstein, M. F., Baker, W. J., & Gola, T. (1996). Comparison of multiple scoring methods for Rey's malingered amnesia measures. *Archives of Clinical Neuropsychology*, *11*, 283-293.
- Greve, K. W., Bianchini, K. J., Black, F. W., Heinly, M. T., Love, J. M., Swift, D. A., et al (2006). Classification accuracy of the Test of Memory Malingering in industrial toxins: Results of know-groups analysis. *Archives of Clinical Neuropsychology*, *19*, 439-448.
- Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2008). Detecting malingering in Traumatic Brain Injury and Chronic Pain: A comparison of three forced-choice symptom validity tests. *The Clinical Neuropsychologist*, *22*, 896-918.
- Grote, C. L. & Hook, J. N. (2007). Forced-choice effort measures. In G. J. Larrabee (Ed), *Assessment of malingered neuropsychological deficits* (pp. 44-80). New York: Oxford University Press.
- Guerreiro, M. (1998). Contributo da Neuropsicologia para o estudo das Demências. Dissertação de Doutoramento não publicada. Faculdade de Medicina de Lisboa.
- Guerreiro, M., Silva, A. P., Botelho, M. A., Leitão, O., Castro-Caldas, A., & Garcia, C. (2008). Avaliação Breve do Estado Mental. In A. Mendonça & M. Guerreiro (Coords.). *Escalas e Testes na Demência* (2.^a ed.; pp.33-39). Lisboa: Grupo de Estudos de Envelhecimento Cerebral e Demência
- Haber, A. H. & Fichtenberg, N. L. (2006). Replication of the Test of Memory Malingering (TOMM) in a traumatic brain injury and head trauma sample. *The Clinical Neuropsychologist*, *20*, 524-532.
- Heubrock, D., & Petermann, F. (1998). Neuropsychological assessment of suspected malingering: Research results, evaluation techniques, and further direction of research and application. *European Journal of Psychological Assessment*, *14*, 211-225.
- Hill, S. K., Ryan, L. M., Kennedy, C. H., & Malamut, B. L. (2003). The Relationship between measures of declarative memory and the Test of Memory Malingering in patients with and without temporal lobe dysfunction. *Journal of Forensic Neuropsychology*, *3*(3), 1-16.
- Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology*, *11*, 967-974.
- Horner, M. D., Bedwell, J. S. & Duong, A. (2006). Abbreviate form of the Test of Memory Malingering. *International Journal of Neuroscience*, *116*, 1181-1186.
- Horton, A. M. (2000). Book and test review. *Archives of Clinical Neuropsychology*, *15*(5), 649-651.

- Inman, T. H., & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and malingering observations. *Archives of Clinical Neuropsychology, 17*, 1-23.
- Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology, 13*(2), 77-90.
- Iverson, G. L., Page, J., Koehler, B. E., Shojania, K., & Badii, M. (2007). Test of Memory Malingering (TOMM) scores are not affected by chronic pain or depression in patients with fibromyalgia. *The Clinical Neuropsychologist, 21*, 532-546.
- Iverson, G. L. & Binder, L. M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. *Journal of Head Trauma Rehabilitation, 15*, 829-858.
- Kitchen, R. (2003). Investigating benefit fraud and illness deception in the United Kingdom. In P. W. Halligan, C. Bass, D. A. Oakley (Eds.), *Malingering and illness deception* (pp.313-323). New York: Oxford University Press.
- Kizilbash, A. H., Vanderploeg, R. D. & Curtiss, G. (2002). The effects of depression and anxiety on memory performance. *Archives of Clinical Neuropsychology, 17*, 57-67.
- Lee, G. P., Loring, D. W., & Martin, R. C. (1992). Rey's 15-item visual memory test for the detection of malingering: Normative observations on patients with neurological disorders. *Psychological Assessment, 4*, 43-46.
- Lezak, M. D., Howieson, D. B., & Loring, D. M. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.
- Maior, F. S., Silva, I., Pedrosa, C., Dias, I., Simões, M. R., & Pinho, S. (2006). *Test of Memory Malingering (TOMM): Estudo exploratório com uma amostra de reclusos*. Poster apresentado no IX Congresso Internacional e Multidisciplinar "A maldade humana: Fatalidade ou educação?". Coimbra.
- Maior, F. S. (2008). *Structured Inventory of Malingered Symptomatology (SIMS) e Test of Memory Malingering (TOMM): Estudos de validação numa amostra normativa*. Tese de Mestrado Integrado. Coimbra: Faculdade de Psicologia e Ciências da Educação.
- Maior, F. S., Simões, M. R., Silva, I., Pedrosa, C., Dias, I., & Pinho, M. S. (2009, aceite para publicação). Test of Memory Malingering (TOMM): Estudos de validação para a população portuguesa em amostras de reclusos e estudantes universitários. *Revista de Psicologia Militar*.
- Martins, M. S. S. (2007). *Montreal Cognitive Assessment (MoCA): Validade e utilidade clínica no diagnóstico do Défice Cognitivo Ligeiro e da Demência*. Tese de Mestrado Integrado em Psicologia (área de especialização em Psicologia Clínica e da Saúde, sub-especialização em Psicogerontologia Clínica). Coimbra: Faculdade de Psicologia e de Ciências da Educação.
- McCaffrey, R. J., O'Bryant, S. E., & Fisher, J. M. (2003). Correlations among the TOMM, Rey-15, and MMPI-2 validity scales in a sample of TBI litigants. *Journal of Forensic Neuropsychology, 3*(3), 45-51.
- Mota, M. (2008a). *Comportamentos de "simulação" e "esforço insuficiente": Estudos com a Structured Interview of Reported Symptoms (SIRS), o Structured Inventory of Malingered Symptomatology (SIMS) e o Test of Memory Malingering (TOMM) numa amostra de reclusos*. Tese de Mestrado

- Integrado em Psicologia (área de especialização em Psicologia Clínica e da Saúde, sub-especialização em Psicologia Forense). Coimbra: Faculdade de Psicologia e Ciências da Educação.
- Mota, M., Simões, M. R., Amaral, L., Dias, I., Luís, D., Pedrosa, C., Soto Maior, F., & Silva, I. (2008b). *Test of Memory Malingering (TOMM): Estudos de validação numa amostra de reclusos. Psiquiatria, Psicologia & Justiça, 2*, 23-41.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699.
- O`Bryant, S. E., Engle, L., Kleiner, J. S., Vasterling, J. J., & Black, F. W. (2007). Test of Memory Malingering (TOMM) Trial 1 as a screening measure for insufficient effort. *The Clinical Neuropsychologist, 21*, 511-521
- O`Bryant, S. E., Finlay, C., & O`Jile, J. R. (2007). TOMM performance and self-reported symptoms of depression and anxiety. *Journal of Psychopathology and Behavioral Assessment, 29*, 111-114.
- O`Bryant, S. E., Gavett, B. E., McCaffrey, R. J., O`Jile, J. R., Huerkamp, J. K., Smitherman, T. A., & Humphreys, J. D. (2008). Clinical utility of Trial 1 of the Test of Memory Malingering (TOMM). *Applied Neuropsychology, 15*, 113-116.
- O`Bryant, S. E., & Lucas, J. A. (2006). Estimating the predictive value of the Test of Memory Malingering: An illustrative example for clinicians. *The Clinical Neuropsychologist, 20*, 533-540.
- Pedrosa, M. (2009). *Detecção de esforço reduzido: Estudos com os subtestes Memória de Dígitos e Vocabulário (WAIS-III) numa amostra com Traumatismo Crânio-Encefálico*. Tese de Mestrado Integrado em Psicologia (área de especialização em Psicologia Clínica e da Saúde, sub-especialização em Psicologia Forense). Coimbra: Faculdade de Psicologia e Ciências da Educação.
- Pedrosa, S., Simões, M.R., Barreto, M., Fonseca, M., Mestre, A., Sousa, L.B., Cruz, I., Gaspar, H., Costa, G., & Corte-Real, F. (2009, no prelo). Detecção de esforço reduzido com os subtestes Memória de Dígitos e Vocabulário (WAIS-III) numa amostra com Traumatismo Crânio Encefálico examinada em contexto médico-legal. *Psiquiatria, Psicologia & Justiça, 3*.
- Pestana, M. H., & Gageiro, J. N. (2003). *Análise de dados para Ciências Sociais: A complementaridade do SPSS (3ª ed.)*. Lisboa: Edições Sílabo.
- Prieto, G., & Muñoz, J. (2000). *Un modelo para evaluar la calidad de los tests utilizados en España*. Retirado em Dezembro 3, 2007, do Consejo General Colegios Oficiales de Psicólogos Web Site: <http://www.cop.es/vernumero.asp>.
- Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology, 16*, 501-506.
- Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment, 10*, 10-20.
- Rogers, R. (2008). Detection strategies for malingering and defensiveness. In R. Rogers (Ed.). *Clinical assessment of malingering and deception* (pp. 14-38).

New York: The Guilford Press.

- Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Report Symptoms: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Rogers, R., Sewell, K. W. & Goldstein, A. M. (1994). Explanatory models of malingering: A prototypical analysis. *Law and Human Behavior*, 18, 543-552.
- Rogers, R., & Schuman, D. (2000). *Conducting insanity evaluations*. New York: Guilford.
- Rogers, R., & Schuman, D. (2005). *Fundamentals of forensic practice, mental health and criminal law*. New York: Springer Science.
- Simões, M. R. (2005). O exame dos comportamentos de simulação em avaliação (neuro)psicológica. In C. Vieira, A. M. Seixas, A. Matos, M. P. Lima, M. Vilar e M. R. Pinheiro (Eds.), *Ensaio sobre o comportamento humano: Do diagnóstico à intervenção. Contributos nacionais e internacionais* (pp. 453-481). Coimbra: Almedina.
- Simões, M. R. (2006). Testes de validade de sintomas na avaliação de comportamentos de simulação. In A. C. Fonseca, M. R. Simões, M. C. T. Simões e M. S. Pinho (Eds.), *Psicologia forense* (pp. 279-309). Coimbra: Almedina
- Simões, M. R., Sousa, L. B., Duarte, P., Firmino, H., Pinho, S., Gaspar, N., Pires, L., Batista, P., Silva, A. R., Silva, S., Ferreira, A. R., & França. S. (submetido para publicação). Avaliação da simulação ou esforço insuficiente com o *Rey 15 Item Memory Test* (15-IMT): Estudos de validação para a população portuguesa.
- Slick, D. J., Sherman, E. M. S. & Iverson, G. L. (1999). Diagnostic criteria for Malingered Neurocognitive Dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, 4, 545-561.
- Slick, D. J., Tan, J. E., Strauss, E. H. & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, 19, 465-473
- Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms and commentary* (3rd ed.). New York: Oxford University Press.
- Sweet, J. J., Condit, D. C., & Nelson, N. W. (2008). Feigned amnesia and memory loss. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 218-236). New York: The Guilford Press.
- Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., Zamora, D., Goodkind, M., Bell, K., Stern, Y., & Devanand, D.P. (2006). Neuropsychological prediction of conversion to Alzheimer Disease in patients with Mild Cognitive Impairment. *Archives of General Psychiatry*, 63, 916-924.
- Tan, J. E., Slick, D. J., Strauss, E. & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on Symptom Validity Tests. *The Clinical Neuropsychologist*, 16, No. 4, 495-505.
- Teichner, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired and elderly patients with dementia. *Archives of Clinical Neuropsychology*, 19, 455-464.

- Tombaugh, T. N. (1996). *Test of Memory Malingering*. Canada: Multi-Health Systems.
- Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*, 260-268.
- Tombaugh, T. N. (2002). The Test of Memory Malingering (TOMM) in forensic psychology. In J. Hom, & R.L. Denney (Eds.) *Detection of response bias in forensic neuropsychology* (pp. 69-96). New York: The Haworth Medical Press.
- Trueblood, W., & Schmidt, M. (1993). Malingering and other validity considerations in the neuropsychological evaluation of mild head injury. *Journal of Clinical and Experimental Neuropsychology, 15*, 578-590.
- Wechsler, D. (1997). Wechsler Adult Intelligence Scale-III. San Antonio TX: The Psychological Corporation.
- Wechsler, D. (2008). *Escala de Inteligência de Wechsler para Adultos – III (WAIS-III) Manual*. Lisboa: Cegoc.
- Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A validation of the Test of Memory Malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology, 25*(7), 979-990.
- Yanez, Y. T., Fremouw, W., Tennant, J., Strunk, J., & Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology, 21*, 161-165.
- Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatry Research, 17* (1), 37-49.

Anexos

Anexo 1. Sumário dos critérios de MND de Slick, Sherman e Iverson (1999)

<p>A. Presença de incentivo(s) externo(s).</p> <p>B. Evidências na avaliação neuropsicológica</p> <ol style="list-style-type: none">1. Enviesamento definitivo da resposta2. Enviesamento provável3. Discrepância entre os dados dos testes e os padrões conhecidos de funcionamento cognitivo4. Discrepância entre os dados dos testes e o comportamento observado5. Discrepância entre os dados dos testes e dados de informadores fiáveis6. Discrepância entre os dados dos testes e a história documental do sujeito <p>C. Evidências no auto-relato do sujeito</p> <ol style="list-style-type: none">1. Discrepância entre o auto-relato e a história documental2. Discrepâncias entre o auto-relato e os padrões conhecidos de funcionamento cognitivo3. Discrepâncias entre o auto-relato e comportamentos observados4. Discrepância entre os sintomas auto-relatados e as informações obtidas de informadores fiáveis.5. Evidência de disfunção psicológica fabricada ou exagerada <p>D. Comportamentos que preenchem os critérios dos grupos B e C mas não podem ser totalmente explicados por factores de ordem psiquiátrica, neurológica ou desenvolvimental</p>

Adaptado de Slick, Sherman e Iverson (1999, pp.553-555)

Anexo 2. Análises estatísticas relativas ao 15-IMT

Efeito das variáveis Género, Idade, Escolaridade e Situação Laboral nos resultados no 15-IMT

N=40	Variáveis	REY-EVOC	REY-RECON
Controlo (N=20)	Género	$U=42.00$	$U=49.00$
	Idade	$r_s=-.145$	$r_s=-.103$
	Anos Escolaridade	$r_s=-.076$	$r_s=.064$
	Sit. Laboral	$U=32.00$	$U=42.00$
DCL (N=20)	Género	$U=47.00$	$U=38.00$
	Idade	$r_s=-.094$	$r_s=-.102$
	Anos Escolaridade	$r_s=-.555^*$	$r_s=.453^*$
	Sit. Laboral	$U=27.00$	$U=27.00$

* $p<.05$; ** $p<.01$

Relação entre os Ensaio do 15-IMT e outras medidas.

	Controlo (N=20)		DCL (N=20)	
	REY-EVOC	REY-REC	REY-EVOC	REY-REC
Vocabulário	.198	.335	.382	.343
GDS	-.240	-.223	-.200	-.055
IAFAI	-.369	-.309	-.338	-.306
MoCA	.626**	.678**	.710**	.779**
Mem_Dígitos_Total	.601**	.457*	-.012	-.052
Mem_Lógica_Imed	.620**	.658**	.524*	.471*
Mem_Lógica_Difer	.586**	.609**	.258	.380
M_Lógica_Temáticas_Imed	.378	.545*	.623**	.560*
M_Lógica_Temáticas_Difer	.347	.477**	.258	.380
Mem_Lógica_%Reten	.073	-.013	.176	.261
Cenas_Família_I	.107	.161	.225	.304
Cenas_Família_II	.314	.478*	.398	.438
Cenas_Família_Reten	.196	.456*	.658**	.497*

* $p<.05$; ** $p<.01$

Intercorrelações entre os indicadores do 15-IMT para cada um dos grupos em estudo

		Evocação Imediata	Nº itens reconhecidos	Nº falsos positivos
Grupo de Controlo	Evocação Imediata	—	—	—
	Nº Itens reconhecidos	.584**	—	—
	Nº falsos positivos	-.009	-.144	—
	R_Res. Combinado	.805**	.912**	-.255
Grupo DCL	Evocação Imediata	—	—	—
	Nº Itens reconhecidos	.712**	—	—
	Nº falsos positivos	.000	.142	—
	R_Res. Combinado	.889**	.761**	-.279

* $p<.05$; ** $p<.01$

Valores relativos à consistência interna do 15-IMT para os dois grupos.

	Controlo		DCL	
	Evocação	Reconhecimento	Evocação	Reconhecimento
Nº Itens	15	30	15	30
α	.738	.441	.750	.604

Resultados obtidos no 15-IMT em função do grupo (Controlo e DCL)

	Controlo (N=20)	DCL (N=20)	U	P
REY_EVoc	M=9.15; $dp=3.22$ (Min=3; Máx=14)	M=7.30; $dp=2.92$ (Min=3; Máx=14)	133.000	.067
REY_REC	M=16.15; $dp=6.69$ (Min=5; Máx=28)	M=13.70; $dp=6.03$ (Min=4; Máx=28)	157.000	.244

Evocação: Análise dos itens (% itens evocados)

15-IMT	Controlo (N=20)	DCL (N=20)
A	18 (90%)	19 (95%)
B	16 (80%)	16 (80%)
C	18 (90%)	15 (75%)
1	16 (80%)	16 (80%)
2	16 (80%)	16 (80%)
3	16 (80%)	17 (85%)
A	10 (50%)	5 (25%)
B	11 (55%)	6 (30%)
C	10 (50%)	6 (30%)
○	6 (30%)	7 (35%)
□	8 (40%)	9 (45%)
△	4 (20%)	1 (5%)
I	11 (55%)	4 (20%)
II	10 (50%)	4 (20%)
III	10 (50%)	4 (20%)

Reconhecimento: Análise dos itens (% itens assinalados)

	15-IMT	Controlo (N=20)	DCL (N=20)
Itens Correctamente Reconhecidos	A	19 (95%)	19 (95%)
	B	18 (90%)	17 (85%)
	C	14 (70%)	12 (60%)
	1	14 (70%)	15 (75%)
	2	14 (70%)	16 (80%)
	3	14 (70%)	15 (75%)
	A	10 (50%)	10 (50%)
	B	12 (60%)	7 (35%)
	C	14 (70%)	17 (85%)
	○	7 (35%)	8 (40%)
	□	6 (30%)	8 (40%)
	△	7 (35%)	3 (15%)
	I	9 (45%)	4 (20%)
	II	9 (45%)	7 (35%)
	III	9 (45%)	9 (45%)
Falsos Positivos	—	0 (0%)	0 (0%)
	≡	4 (20%)	3 (15%)
	≡≡	2 (10%)	2 (10%)
	◇	2 (10%)	1 (5%)
	◡	2 (10%)	1 (5%)
	▱	4 (20%)	3 (15%)
	D	6 (30%)	4 (20%)
	E	1 (5%)	3 (15%)
	F	0 (0%)	3 (15%)
	d	6 (30%)	2 (10%)
	e	4 (20%)	2 (10%)
	f	1 (5%)	2 (10%)
	4	1 (5%)	5 (25%)
	5	1 (5%)	3 (15%)
	6	1 (5%)	3 (15%)