



**Universidade do Minho**  
Escola de Psicologia

Rui Miguel de Medeiros Paulo

## **Deteção da Simulação de Problemas de Memória em Contexto Forense**

Rui Miguel de Medeiros Paulo · Deteção da Simulação de Problemas de Memória em Contexto Forense

UMinho | 2012

Outubro de 2012



**Universidade do Minho**

Escola de Psicologia

Rui Miguel de Medeiros Paulo

## **Deteção da Simulação de Problemas de Memória em Contexto Forense**

Dissertação de Mestrado  
Mestrado Integrado em Psicologia  
Área de Especialização de Psicologia da Justiça

Trabalho realizado sob a orientação do  
**Professor Doutor Emanuel Pedro Viana Barbas de Albuquerque**

Outubro de 2012

## **DECLARAÇÃO**

**Nome**

Rui Miguel de Medeiros Paulo

**Endereço eletrónico**

a52941@alunos.uminho.pt

**Número do Bilhete de Identidade**

13626151

**Título dissertação**

Deteção da Simulação de Problemas de Memória em Contexto Forense

**Orientador**

Professor Doutor Emanuel Pedro Viana Barbas de Albuquerque

**Ano de conclusão**

2012

**Designação do Mestrado**

Mestrado Integrado em Psicologia – Área de Especialização em Psicologia da Justiça

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, 15 de outubro de 2012

Assinatura

## **Agradecimentos**

Gostaria de expressar os meus sentidos agradecimentos a todos aqueles que me apoiaram não só ao longo do último ano, mas também ao longo do meu percurso académico. Nomeadamente gostaria de agradecer:

Ao Professor Doutor Pedro Albuquerque não só pelo excelente apoio prestado ao longo do último ano mas também pela inspiração que transmitiu a mim e aos restantes alunos nas unidades curriculares que lecionou, mostrando ainda que é possível ser um ótimo investigador e, simultaneamente, um ótimo professor e orientador. Acima de tudo, agradeço por toda a disponibilidade demonstrada e por ter tornado a nossa colaboração, tanto presente como passada, tão gratificante.

Ao Grupo de Investigação em Memória Humana da Universidade do Minho, particularmente ao Doutor Karlos Luna, Beatriz Martín e Paulo Carvalho, pela troca de ideias, debate científico e conselhos que considero essenciais para o desenvolvimento de qualquer investigação. Agradeço em especial à Doutora Joana Arantes Silva pelo seu apoio, leitura e revisão desta dissertação.

A todos os participantes do meu estudo, à Doutora Marta Parreira, Sara Rego e Margarida Gonçalves pela disponibilidade demonstrada em colaborar nesta investigação.

À minha família, nomeadamente aos meus pais, ao meu irmão e à minha namorada por todas as razões demasiado extensas para incluir neste texto.

A todos os meus amigos, colegas de curso e colegas de estágio, que estiveram presentes e me apoiaram ao longo destes últimos anos, nomeadamente ao Alexandre Resende, Jorge Abreu, Daniel Vilaça, João Lima, João Petiz, Nelson Marques, Alexandre Cunha, Carlos Pinto, Filipa Oliveira e Inês Sousa.

Por fim, gostaria de agradecer a todos aqueles que deveriam de constar neste texto mas, pelas minhas habituais distrações, às quais provavelmente já estarão habituados, me esqueci de incluir nesta lista.

A todos, os meus sentidos agradecimentos!

## **Deteção da Simulação de Problemas de Memória em Contexto Forense**

### **Resumo**

A avaliação da simulação é um procedimento indispensável em muitas avaliações neuropsicológicas, clínicas e forenses. Uma vez que as provas de avaliação neuropsicológicas se mostram incapazes de detetar a simulação de sintomas, outras metodologias devem ser seguidas. A utilização de Testes de Validade de Sintomas (TVS) é uma das técnicas mais promissoras para este mesmo fim. Neste estudo, procurámos desenvolver um novo TVS baseado no TOMM - *Test Of Memory Malingering* (Tombaugh, 1996) contudo, no entanto, diversas alterações metodológicas, efetuadas com base nas conclusões de diversos estudos recentes, de forma a aumentar a eficácia e utilidade deste mesmo teste. Explorámos ainda qual o efeito de diversas variáveis na deteção da simulação de problemas de memória, nomeadamente, o tempo de resposta dos participantes, o treino dos avaliados e a ordem pela qual o TVS é aplicado na bateria de provas. Desenvolvemos um TVS (DETECTS – Teste de Deteção da Simulação de Problemas de Memória) completamente informatizado, com um tempo de aplicação reduzido e com pontos-de-corte eficazes, possuindo a prova um elevado índice de sensibilidade, especificidade e valor preditivo positivo e negativo. Em conjunto com a bateria de testes aplicada, a prova construída permitiu, para a amostra utilizada, que nenhum simulador fosse incorretamente classificado como detentor de uma perturbação mnésica legítima, e vice-versa.

## **Detection of Memory Malingering in the Forensic Context**

### **Abstract**

The evaluation of malingering is an essential procedure in most neuropsychological, clinical and forensic evaluations. Since neuropsychological tests are unable to detect malingering, other methods must be used for this purpose. One of the most promising procedures is the use of Symptom Validity Tests (SVT). In this study, we aimed to develop a new SVT, inspired on the TOMM - *Test of Memory Malingering* (Tombaugh, 1996), but with some modifications, based on recent findings from several studies, to increase the effectiveness and usefulness of this test. We also explored the effect of different variables on the detection of memory malingering, for instance, participants' response time, participants training, and the order in which the SVT is applied within the neuropsychological test battery. Therefore, we developed a fully computerized test (DETECTS – Memory Malingering Detection Test), with a reduced application time, and with effective cut-off points as well as an excellent sensitivity, specificity, positive and negative predictive power values. When applied among the test battery, this test allowed us not to wrongly classify any malingerer as a subject with legitimate memory deficits and vice versa.

## ÍNDICE

|   |    |
|---|----|
| I INTRODUÇÃO .....  | 8  |
| 1.1 Simulação de Sintomas.....  | 8  |
| 1.2 Detecção da Simulação .....   | 9  |
| 1.3 Testes de Validade de Sintomas .....  | 10 |
| 1.4 Metodologia de Estudo e Construção dos TVS.....                               | 12 |
| 1.5 Inovação nos TVS .....  | 14 |
| 1.6 TOMM – <i>Test Of Memory Malingering</i> .....                                | 15 |
| II O PRESENTE ESTUDO: OBJETIVOS, PROBLEMAS E HIPÓTESES DE ESTUDO .....            | 17 |
| III METODOLOGIA .....   | 20 |
| 3.1 Participantes .....   | 20 |
| 3.2 Equipamento e Materiais .....   | 22 |
| 3.2.1 DETECTS - Teste de Detecção da Simulação de Problemas de Memória.....       | 22 |
| 3.2.2 Bateria de Provas .....   | 24 |
| 3.2.3 Instruções de Simulação .....   | 25 |
| 3.2.4 Outros Materiais .....  | 25 |
| 3.3 Plano e Procedimento Experimental.....  | 26 |
| IV RESULTADOS.....  | 27 |
| 4.1 Ordem de Aplicação das Provas e Acertos no Teste de Simulação (DETECTS) ..... | 28 |
| 4.2 Tempos de Resposta.....   | 30 |
| 4.3 Determinação de Pontos-de-Corte no Teste de Simulação (DETECTS) .....         | 32 |
| 4.4 Outras Variáveis .....  | 35 |
| 4.4.1 Ordem de Aplicação dos Ensaios do Teste de Simulação (DETECTS).....         | 35 |
| 4.4.2 Provas de Memória (WMS-III) e Teste de Simulação (DETECTS) .....            | 36 |

|   |    |
|---|----|
| 4.4.3 Variáveis Sexo e Idade .....                              | 37 |
| 4.4.4 Provas de Memória (WMS-III) .....                         | 38 |
| V DISCUSSÃO DOS RESULTADOS .....                                | 38 |
| VI CONCLUSÃO .....  | 44 |
| 6.1 Limitações do Estudo e Sugestões para Estudos Futuros ..... | 44 |
| 6.2 Implicações para o Contexto Forense.....                    | 45 |
| VI REFERÊNCIAS .....  | 46 |

## **I – INTRODUÇÃO**

A simulação de sintomas apresenta-se como um fenómeno de extrema importância para a neuropsicologia e para a psicologia forense. A sua avaliação torna-se cada vez mais essencial tanto para a prática, como para a investigação nesta área (Boone, 2009; Greve & Bianchini, 2004; O’Bryant, Engel, Kleiner, Vasterling, & Black, 2007; O’Bryant & Lucas, 2006; Simões, 2006). Esta avaliação, bem como o consequente parecer do psicólogo, é muitas vezes um fator decisivo nos processos judiciais que envolvem pedidos de compensação por incapacidade. É também importante detetar os casos de simulação em processos judiciais onde os alegados défices se podem constituir como um fator atenuante da pena ou responsabilidade penal do arguido (Iverson, 2003; Oorsouw & Merckelbach, 2010; Simões, 2006). Na área dos processos cíveis, também os casos relacionados com acidentes de trabalho se apresentam como uma parcela importante da atividade de avaliação da simulação, uma vez que são, usualmente, altamente litigiosos (Simões & Sousa, 2011). Por fim, também na população militar a avaliação da simulação se revela particularmente importante devido ao grande número de pedidos de compensação ou dispensa de serviço militar existentes, após algum tipo de lesão sofrida pelo militar no decorrer do serviço (Lange, Pancholi, Bhagwat, Anderson-Barnes, & French, 2012).

Uma avaliação forense eficaz é habitualmente complementada com a aplicação de vários instrumentos de avaliação. No entanto, a maioria destes testes partilham o pressuposto básico de que o indivíduo procura responder fazendo uso de todas as suas competências e aplicando todo o seu esforço (Iverson, 2003; Simões, 2006; Tombaugh, 1996). Sabemos, no entanto, que, particularmente em contexto forense, o avaliado encontra-se muitas vezes motivado para simular problemas cognitivos ou para esconder e distorcer o seu funcionamento psicológico. Torna-se, assim, imprescindível medir e controlar até que ponto os resultados obtidos se devem às competências reais do avaliado, à sua motivação externa ou até ao esforço que este empenhou na realização dos testes (Bauer, O’Bryant, Lynch, McCaffrey, & Fisher, 2007; Bush et al., 2005; Green, Haley, & Allen, 2001; Iverson, 2003; Whiteside, Dumbar-Mayer, & Waters, 2009).

### **1.1 Simulação de Sintomas**

O fenómeno da simulação é definido no DSM-IV como a “produção intencional de sintomas, físicos ou psicológicos, falsos ou exagerados, motivados por incentivos externos, tais como evitar o trabalho, o cumprimento de penas, o serviço militar, obter compensações económicas ou ter acesso a drogas” (DSM-IV-TR, APA, 2000/2002, p.739). Esta definição realça, de forma adequada, o papel dos incentivos externos no fenómeno da simulação. Não fornece, no entanto, informação apropriada para distinguir a sintomatologia intencional da não intencional (Simões, 2006). Note-se também que a

simulação de sintomas pode variar desde a total criação de sintomas até ao exagero dos sintomas realmente sentidos. Assim, um diagnóstico positivo de simulação não significa que o avaliado não possua sintomatologia real (Haines & Norris, 1995). A simulação é ainda algo situacional, ou seja, a simulação de sintomas, por parte de determinado sujeito, é usualmente restrita a esse contexto ou situação, não devendo essa pessoa ser encarada como desonesta ou censurável (Tombaugh, 1996). Como refere Iverson (2003) quase todos nós, ao longo da nossa vida, simulámos sintomas, por exemplo, para faltar ao trabalho ou à escola.

Podemos simular vários tipos de sintomas como por exemplo: dor, desorientação, depressão, baixa concentração, mudanças na personalidade e problemas de memória (Iverson, 2003). Usualmente os avaliados simulam de forma seletiva, não procurando reproduzir diversas classes de sintomas durante o mesmo momento temporal (Binder, 1992). As queixas associadas à perda de memória são das mais predominantes em pacientes que sofreram algum tipo de lesão cerebral. Por consequência, a simulação destas é uma das estratégias mais utilizadas, por sujeitos instruídos, para obter ganhos financeiros após um acidente de viação (Oorsouw & Merckelbach, 2010; Simões, 2006). Este tipo de simulação é também particularmente proeminente em ofensores com um diagnóstico de psicopatia (Porter & Woodworth, 2007). Paradoxalmente, sabemos que pacientes com lesões cerebrais moderadas tendem a desproporcionar os seus sintomas chegando a apresentar maior sintomatologia do que pacientes com lesões cerebrais extensas (Green, 2011; Green et al., 2001). Percebemos assim a importância do efeito da variável esforço e motivação nos testes neuropsicológicos. A prevalência da simulação de sintomas foi alvo de várias estimativas sendo que estas variam muito consoante o estudo e o tipo de simulação considerado (Mittenberg, Patton, Canyock, & Condit, 2002). No estudo de Slick, Tan, Strauss e Hultsch (2004) foram entrevistados 39 profissionais desta área. Estes consideravam que, durante os 12 meses prévios à entrevista, pelo menos 10% dos avaliados estava, na sua opinião, a simular sintomas.

## **1.2 Deteção da Simulação**

Os instrumentos construídos para avaliar a simulação de sintomas procuram mostrar em que medida estes são realmente válidos, falsos ou exagerados (Haines & Norris, 1995; Simões, 2006). Existem diversos métodos para despistar a simulação de sintomas com diferentes graus de eficácia. Uma análise subjetiva do avaliador, como por exemplo a procura de resultados inconsistentes, discrepantes ou que não façam sentido de um ponto de vista clínico e/ou psicométrico pode ser o primeiro passo para analisar uma situação de simulação. No entanto o carácter subjetivo e impreciso desta avaliação torna evidentes diversas limitações, refletindo assim uma taxa de sucesso de identificação diminuta (Binder, 1992; Haines & Norris, 1995; Oorsouw & Merckelbach, 2010). Outra estratégia utilizada consiste em procurar que o avaliado responda incorretamente a questões muito simples, identificando assim os simuladores. No entanto, sendo as questões facilmente identificadas

pelo avaliado como muito simples, estes rapidamente detetam o seu propósito, invalidando assim a avaliação (Haines & Norris, 1995). Um desempenho anômalo em testes de memória, ou seja, a apresentação de resultados atípicos mesmo em pacientes com lesões cerebrais ou resultados que não obedecem às regras básicas da memória humana, pode também levantar a suspeita de simulação de sintomas. Por exemplo, se um avaliado apresenta um desempenho em tarefas de reconhecimento inferior ao que apresenta em tarefas de evocação, deverá ser examinada a possibilidade de estarmos perante uma simulação (Haines & Norris, 1995).

### **1.3 Testes de Validade de Sintomas**

Face à insatisfação existente em relação às metodologias previamente descritas, surgem os Testes de Validade de Sintomas (TVS) que se apresentam como os primeiros instrumentos especificamente elaborados para identificar a simulação, sendo também o método mais utilizado, e estudado, para este propósito (Bianchini, Mathias, & Greve, 2001; Gast & Hart, 2010; Simões, 2006). Atualmente, estes instrumentos são vastamente aceites pela comunidade científica, existindo inúmeras publicações acerca destes (Blaskewitz, Merten, & Kathmann, 2008; Gervais, Rohling, Green, & Ford, 2004; Haines & Norris, 1995), e há mesmo diversos autores que defendem que as avaliações neuropsicológicas e forenses devem incluir TVS (Simões, 2006; Slick et al., 2004).

Os TVS devem ser incluídos numa bateria de provas clínica, constituindo assim uma parte integrante da rotina de avaliação e, reduzindo a possibilidade do avaliado reconhecer ou identificar o seu verdadeiro propósito. No estudo de Slick e colaboradores (2004) aproximadamente 79% dos profissionais que constituíram a amostra do estudo afirmaram utilizar pelo menos um TVS nas suas avaliações. O despiste da simulação não deve ser baseado em apenas um teste, sendo aconselhável, sempre que possível, a aplicação de, no mínimo, dois testes (Boone, 2009; Farkas, Rosenfeld, Robbins, & Gorb, 2006; Gervais et al., 2004; Greve, Binder, & Bianchini, 2009; Iverson, 2003; Jelicic, Ceunen, Peters, & Merckelbach, 2011; Oorsouw & Merckelbach, 2010; Simões, 2006; Simões & Sousa, 2006; Slick et al., 2004). No entanto, devido ao escasso tempo disponível nas avaliações forenses, este critério é muitas vezes descurado e os TVS acabam por ser pouco utilizados (Bauer et al., 2007; Boone, 2009). O uso de apenas um TVS facilita também o treino dos avaliados por parte de terceiros (e.g., advogados), permitindo que estes identifiquem o TVS e simulem de uma forma mais sofisticada. Torna-se assim essencial o desenvolvimento, a validação e a aplicação de diversos TVS uma vez que, quanto maior o número de testes disponíveis, menor a probabilidade do avaliado possuir tempo e capacidade para estudar e reconhecer todos estes instrumentos (Bianchini et al., 2001; Chafetz, 2011; Haber & Fichtenberg, 2006; Oorsouw & Merckelbach, 2010). Atualmente, os investigadores procuram também desenvolver versões abreviadas destes testes com o intuito de permitir a aplicação de vários TVS em menos tempo (Boone, 2009; O'Bryant, et al., 2007). O efeito dos diferentes tipos de treino nas avaliações forenses e resposta aos TVS tem sido também alvo de

diversos estudos (Powell, Gfeller, Hendricks, & Sharland, 2004; Youngjohn, Lees-Haley, & Binder, 1999). Da mesma forma, a avaliação da eficácia de determinado TVS face a simuladores previamente treinados é uma etapa importante para a validação deste mesmo teste (Weinborn, Woods, Nulsen, & Leighton, 2012). Importa ainda realçar que pelo facto de um sujeito não ser identificado como simulador no TVS não significa, necessariamente, que não o está a fazer, podendo simplesmente ter identificado o teste como um TVS e estar, conseqüentemente, a simular de uma forma mais sofisticada dificultando a sua deteção. O avaliado pode também ter encarado o teste como não estando associado ao tipo de sintomas que procura simular e, conseqüentemente, estar a responder de forma esforçada. Assim, uma avaliação cuidada do clínico continua a mostrar-se de extrema importância (Gavett, O'Bryant, Fisher, & McCaffrey, 2005; Greve, Bianchini, & Doane, 2006; Greve et al., 2009; O'Bryant & Lucas, 2006; Tombaugh, 1996).

Os TVS são usualmente baseados em tarefas de escolha forçada, ou seja, integram tarefas de reconhecimento onde o avaliado tem de escolher, obrigatoriamente, uma das opções apresentadas. Tratando-se de tarefas de reconhecimento o avaliado deve, na presença de estímulos distrativos, identificar os estímulos que lhe foram previamente apresentados, podendo estes ser dígitos (e.g., *Portland Digit Recognition Test* e *Computerized Assessment of Response Bias*), palavras (e.g., *Word Memory Test*) ou desenhos, como é o caso do *Test Of Memory Malingering* (Simões, 2006). Sendo que, nestes testes, apenas existem duas opções de resposta a probabilidade de acertar ao acaso é de 50%. Por esta razão, numa fase inicial de desenvolvimento dos TVS, o avaliado era classificado como estando a simular caso obtivesse uma pontuação inferior ao acaso (Gervais et al., 2004). No entanto, rapidamente os investigadores se aperceberam que a utilização deste critério de decisão conduzia a taxas de deteção dos simuladores muito reduzidas, mais concretamente a uma baixa sensibilidade dos testes, critério este que iremos abordar de seguida (Haines & Norris, 1995). Assim, o recurso a pontos-de-corte mais rigorosos torna-se essencial nas avaliações de forma a aumentar a precisão do instrumento.

Existem pelo menos quatro indicadores que precisamos analisar para avaliar a utilidade de um TVS, bem como a eficácia dos pontos-de-corte estabelecidos: a especificidade<sup>1</sup>, a sensibilidade<sup>2</sup>, o valor preditivo positivo (VPP)<sup>3</sup> e o valor preditivo negativo (VPN)<sup>4</sup>. A especificidade refere-se à

---

<sup>1</sup> A sensibilidade é calculada através da divisão do número de simuladores corretamente classificados pelo número total de simuladores.

<sup>2</sup> A especificidade é calculada através da divisão do número de não simuladores corretamente classificados pelo número total de não simuladores.

<sup>3</sup> O VPP é calculado através da divisão do número de participantes corretamente classificados como simuladores pelo número total de participantes classificados, correta ou incorretamente, como simuladores.

<sup>4</sup> O VPN é calculado através da divisão do número de participantes corretamente classificados como não simuladores pelo número total de participantes classificados, correta ou incorretamente, como não simuladores.

proporção de avaliados que não apresentam a condição de interesse (não simuladores) corretamente classificados. Uma especificidade baixa está relacionada com um número elevado de falsos positivos. Trata-se de um aspeto fundamental a considerar, uma vez que um diagnóstico positivo pode ter consequências devastadoras para o indivíduo, independentemente de ser, ou não, o diagnóstico correto. Este critério está relacionado com o VPP, que se trata da probabilidade de um avaliado ser simulador quando é diagnosticado como tal. A sensibilidade refere-se à proporção de avaliados com a condição de interesse (simuladores) corretamente identificados (Simões, 2006). Este critério está relacionado com o VPN que se trata da probabilidade de um avaliado não ser simulador quando assim identificado. Os investigadores têm mostrado maior interesse em manter o índice de especificidade alto, por oposição ao índice de sensibilidade (Iverson, 2007). No entanto um número elevado de falsos negativos possui também as suas consequências, nomeadamente a atribuição inadequada de indemnizações monetárias a simuladores (Bianchini et al., 2001; Iverson, 2003).

Um TVS eficaz deve ainda cumprir alguns critérios básicos, como por exemplo, ser sensível apenas à simulação e não a outras variáveis como a inteligência, a memória ou até a lesões cerebrais. Deve ter aplicabilidade generalizada, isto é, não exigir necessidades de tradução ou adaptação a diferentes grupos linguísticos (Powell et al., 2004; Simões, 2006). Neste sentido, o uso de desenhos ou figuras, tal como ocorre no TOMM, torna-se uma vantagem, permitindo ao mesmo tempo uma maior homogeneidade entre os estímulos (MacAllister, Nakhutina, Bender, Karantzoulis, & Carlson, 2009; Tombaugh, 1996). Por fim, um TVS deve possuir uma dificuldade percebida maior do que a dificuldade real criando condições para que os simuladores exibam baixos desempenhos, por oposição aos restantes avaliados (Bianchini et al., 2001; Simões, 2006).

Outro assunto importante a explorar no estudo dos TVS é a pertinência de informar o avaliado, antes da realização dos testes neuropsicológicos, que um nível de esforço inadequado, bem como o exagero ou a produção de sintomas poderão ser detetados nos testes. Na prática, a maioria dos avaliadores não faz este tipo de advertência aos avaliados (Slick et al., 2004). Youngjohn e colaboradores (1999) referem mesmo que fazer este tipo de aviso não é uma boa estratégia porque os simuladores precavidos neste sentido são capazes de simular sintomas de uma forma ainda mais credível. Assim, os autores aconselham a apenas pedir aos avaliados que apliquem todo o seu esforço nas provas procurando também motivar o sujeito e perguntar, ao longo da avaliação, se algo o está a impedir de realizar a prova de forma adequada, e, caso esteja, ajudar no sentido de eliminar esse obstáculo (Iverson, 2007).

#### **1.4 Metodologias de Estudo e Construção dos TVS**

Existem duas principais metodologias utilizadas para estudar, desenvolver e/ou validar este tipo de testes. Ambas se mostraram essenciais para o estudo da simulação, possuindo, no entanto, as

suas próprias limitações (Bianchini et al., 2001). Atualmente, tal como no presente estudo, são utilizadas também metodologias experimentais que combinam estes dois métodos (Simões, 2006).

Na primeira metodologia são utilizados grupos clínicos e não clínicos. Os grupos não clínicos são constituídos por indivíduos que vivem na comunidade, recrutados com o objetivo de estimar a dificuldade percebida da prova. Os grupos clínicos são constituídos por pacientes com lesões cerebrais, recrutados com o intuito de determinar se o seu desempenho se assemelha, ou não, ao do grupo não clínico e aos suspeitos de simulação. O último grupo, habitualmente usado neste tipo de estudos, é constituído por suspeitos de simulação, ou seja, indivíduos que, no momento da administração dos TVS, se encontram em situações que potenciam o recurso à simulação, por exemplo, situações de litígio resultantes de acidentes de trabalho. É neste grupo que se encontra a principal limitação deste método pois é impossível ter a certeza que, neste tipo de grupos, se encontra apenas simuladores. Como vários autores referem, embora sejam suspeitos nada garante que sejam, de facto, simuladores (Bianchini et al., 2001; Greve & Bianchini, 2004; Haines & Norris, 1995).

O segundo método utilizado para o estudo e desenvolvimento dos TVS consiste na realização de experiências de simulação. Estas têm sido cada vez mais utilizadas porque, ao pedir diretamente a um grupo de participantes para desempenhar o papel de simuladores, podemos ter maior certeza que esse grupo está, de facto, a simular (Haines & Norris, 1995). Neste tipo de estudos são frequentemente utilizados alunos universitários que são divididos em dois grupos: um grupo de controlo, constituído por alunos que procuram responder ao teste aplicando todas as suas capacidades; e um grupo experimental, onde é pedido aos participantes para simularem sintomas tal como se estivessem numa situação litigiosa onde a simulação lhes traria benefícios financeiros. Para tornar este grupo mais real é, por vezes, dada informação acerca dos efeitos e sequelas cognitivas de algumas lesões cerebrais e/ou acerca da simulação de sintomas. Ainda que a influência deste tipo de informação no desempenho dos participantes seja discutível, Haines e Norris (1995) encontraram diferenças entre grupos de simuladores previamente informados acerca destes aspetos e grupos de simuladores não informados, tendo o primeiro grupo sido mais bem sucedido na simulação credível de sintomas.

Este segundo método apresenta também algumas limitações. O facto de os simuladores utilizados neste tipo de experiências não possuírem incentivos externos, ao contrário do que acontece com os verdadeiros simuladores, é uma das principais críticas a esta metodologia de investigação (Tan, Slick, Strauss, & Hultsch, 2002). Embora possa ser dada alguma recompensa aos participantes, por exemplo, um prémio simbólico atribuído ao participante que simular com maior sucesso os sintomas, esta recompensa é incomparável com as que resultam de um processo de indemnização. Assim, é provável que os participantes não se envolvam tanto na tarefa e não se esforcem o suficiente para simular de uma forma credível, o que é corroborado no estudo de Haines, Norris e Montaldi (1995) onde os resultados obtidos mostram que o uso de incentivos simbólicos não tem qualquer tipo

de efeito no desempenho dos sujeitos. Atualmente, alguns autores sugerem que, tendo em conta o nível educacional acima da média dos participantes tipicamente utilizados nestes estudos, os alunos universitários, é de esperar que estes consigam desenvolver estratégias de simulação mais elaboradas do que a população geral. Desta forma, embora estes participantes possam estar menos motivados na simulação credível dos sintomas, em comparação aos simuladores em casos de compensação reais, o seu nível educacional permite que estes simulem de forma mais elaborada podendo assim compensar o seu esforço decrescido (Bianchini et al., 2001). Outras limitações, como a inexistência de um evento traumático, vivido por este tipo de participantes, bem como a ausência do risco de ser detetado como simulador, existente nas avaliações forenses, podem também enviesar os resultados neste tipo de estudos. Não obstante, os investigadores que têm utilizado esta metodologia encontram frequentemente diferenças consistentes entre os grupos de simuladores e não simuladores, demonstrando assim a eficácia da mesma (Haines & Norris, 1995).

### **1.5 Inovações nos TVS**

Atualmente existem várias inovações na aplicação dos TVS. Uma destas inovações, a aplicação informatizada, permite manter o teste fiel à sua estrutura original garantindo assim a sua legitimidade (Vanderslice-Barr, Miele, Jardin, & McCaffrey, 2011), possuindo ainda inúmeras vantagens como, por exemplo, permitir registar de forma precisa o tempo de resposta dos participantes (Bianchini et al., 2001; Haines & Norris, 1995). Este registo é essencial visto que uma das estratégias mais utilizadas pelos avaliados para simular sintomas é fornecer a resposta de forma mais lenta do que habitual (Tan et al., 2002; Willison & Tombaugh, 2006). Da mesma forma, ao fornecer uma resposta errada, sabendo, no entanto, qual é a resposta correta, estes indivíduos necessitarão de mais tempo para efetuar a resposta do que os restantes sujeitos (Haines & Norris, 1995). Assim, a medição do tempo de resposta do participante aumenta o poder discriminativo da prova. A aplicação informatizada dos testes permite também a apresentação aleatória dos estímulos, aumentando a homogeneidade dos itens no que concerne ao seu grau de dificuldade. Segundo Haines e Norris (1995) este tipo de aplicação permite ainda poupar tempo pessoal ao avaliador e reduzir a transparência da prova, ou seja, diminuir a facilidade com que os participantes identificam a dimensão que o teste pretende avaliar (a simulação), tornando-o mais semelhante com um teste de memória.

As versões informatizadas potenciam também um maior rigor na apresentação, bem como na medição e na análise dos dados. No estudo de Tan e colaboradores (2002) os participantes afirmaram ainda ter tido dificuldades em simular sintomas, tal como lhes tinha sido pedido, na prova TOMM uma vez que foi a única aplicada pessoalmente pelos investigadores. Desta forma, a aplicação informatizada dos testes parece também aumentar o à-vontade dos avaliados, incentivando assim os simuladores a apresentar pontuações mais extremas e facilitando a deteção. Como limitação, a familiaridade reduzida que os avaliados podem ter com o computador pode condicionar as suas

respostas, sobretudo os tempos de resposta. Atualmente existem diversas versões informatizadas de distintos TVS como por exemplo o PDRT-C – *Portland Digit Recognition Test - Computerized*, o MMDT – *Malingered Memory Deficit Test* e o TOMM – *Test Of Memory Malingering* (Haines & Norris, 1995).

### **1.6 TOMM – *Test Of Memory Malingering***

O TOMM é um instrumento robusto na detecção da simulação de problemas de memória, sendo um dos testes mais extensivamente estudados pela comunidade científica e, em simultâneo, mais utilizado pelos peritos (Bauer et al., 2007; Bianchini et al., 2001; Etherton, Bianchini, Greve, & Ciota, 2005; Greiffenstein, Greve, Bianchini, & Baker, 2008; O’Bryant, Finlay, & O’Jile, 2006; Slick et al., 2004). Apresenta uma taxa de identificação correta dos sujeitos não simuladores de 91% a 95% conforme se inclua, ou não, na análise os pacientes com demência, devendo o uso do TOMM ser feito com particular prudência perante este tipo de pacientes (Greve & Bianchini, 2004; Rogers, 2008; Tombaugh, 1996). Num estudo de Powell e colaboradores (2004) o TOMM apresentou uma taxa geral de reconhecimento correto (simuladores e não simuladores) de 96% bem como uma taxa de sensibilidade e especificidade elevada (94% e 100%, respetivamente). O teste mantém a sua elevada sensibilidade em participantes que foram instruídos acerca das consequências das lesões cerebrais (93%), bem como em participantes instruídos sobre as diversas estratégias de simulação (96%), garantindo assim a eficiência da prova em indivíduos previamente treinados (Jelicic, Ceunen, Peters, & Merckelbach, 2011; Tan et al., 2002). O TOMM não avalia como simuladores indivíduos que possuem algum tipo de défices cognitivos nem discrimina os participantes em função de variáveis como a inteligência, idade, capacidade de concentração ou o nível de escolaridade, podendo ser aplicado a crianças a partir dos 5 anos (Batt, Shores, & Chekaluk, 2008; Blaskewitz, et al., 2008; Gast & Hart, 2010; Greve et al., 2006; Simon, 2007). Estas duas últimas variáveis (idade e nível de escolaridade) explicam menos de 2% da variância encontrada, podendo assim o avaliador ter um elevado grau de certeza que quando um avaliado falha no TOMM, tal acontece devido ao seu intuito de simular sintomas ou devido ao seu reduzido envolvimento na tarefa (Donders, 2005; Greve et al., 2006; Haber & Fichtenberg, 2006; Iverson, 2007, Kirk et al., 2011; Tombaugh, 1996). Este teste pode ainda ser aplicado em indivíduos com perturbações afetivas (e.g., sintomas de ansiedade e depressão), perturbações psicóticas, dor crónica e epilepsia, uma vez que tais condições não influenciam os resultados do teste (Duncan, 2006; Etherton et al., 2005; Iverson, LePage, Koehler, Shojania, & Badii, 2007; MacAllister et al., 2009; O’Bryant, Finlay, & O’Jile, 2006; Yanez, Fremouw, Tennant, Strunk, & Coker, 2006)

A administração do TOMM envolve vários ensaios. No primeiro é pedido ao avaliado que memorize 50 desenhos de objetos comuns, sendo, no final da apresentação, pedido ao sujeito que identifique as imagens previamente apresentadas através de um teste de reconhecimento imediato com

escolha forçada composto por 50 pares de imagens. No segundo ensaio o procedimento é semelhante, são apresentados novamente os mesmos 50 desenhos, numa ordem diferente, e é mais uma vez aplicado um teste de reconhecimento imediato com novos pares de imagens. Por fim, após um intervalo de retenção de 15 minutos, pode ser aplicado um novo teste de reconhecimento, desta vez sem que as 50 imagens sejam novamente apresentadas ao avaliado (Tombaugh, 1996). A opção pela aplicação do último teste de reconhecimento do TOMM tem sido alvo de debate entre os investigadores, parecendo ser útil apenas quando o TOMM é o único TVS aplicado. Numa avaliação mais extensa, com a aplicação de vários TVS, este último ensaio pode ser excluído (Greve & Bianchini, 2006). No estudo de Tan e colaboradores (2002) verificou-se que, em qualquer dos testes de reconhecimento que compõem o TOMM, os indivíduos simuladores apresentaram menores pontuações do que os restantes. Assim, versões abreviadas do TOMM, sem o último ensaio, têm sido desenvolvidas com o objetivo de reduzir o tempo de aplicação deste instrumento, podendo até ser suficiente a aplicação isolada do primeiro ensaio desta prova quando os resultados obtidos são suficientemente esclarecedores (Bauer et al., 2007; Gavett et al., 2005; Horner, Bedwell, & Duong, 2006; O'Bryant et al., 2007; O'Bryant et al., 2008). A título de exemplo, Hilsabeck, Gordon, Hietpas-Wilson e Zartman (2011), consideram que, caso o sujeito obtenha uma pontuação superior a 41 acertos ou inferior a 25 acertos no primeiro ensaio do TOMM, não existe a necessidade de aplicar os restantes ensaios, sendo mais útil para o avaliador, utilizar o tempo poupado para aplicar outros TVS.

Devido à quantidade de estímulos apresentados, bem como à sua natureza visual, o TOMM aparenta ser bastante mais difícil do que é na realidade, permitindo que os simuladores sejam o único grupo de sujeitos que apresenta baixos desempenhos neste teste. Assim, até um resultado aparentemente alto (e.g., um resultado inferior a 45 acertos no 2º ensaio) levanta a suspeita de estarmos perante um caso de simulação. Também o *feedback* proporcionado imediatamente após cada resposta permite ao simulador controlar o número de respostas que erra, potenciando assim um baixo desempenho por parte destes (Haines & Norris, 1995). A estrutura do teste, composta por sucessivos ensaios de apresentação e reconhecimento de imagens, bem como a sua dificuldade aparente diminui também a transparência do teste. Uma vez que mesmo os indivíduos com danos cerebrais possuem uma capacidade elevada para reconhecer figuras do dia a dia, tal como as apresentadas nesta prova, esta não é sensível a este tipo de pacientes, pois também estes apresentam elevados desempenhos na prova de simulação, tal como os avaliados sem qualquer tipo de lesão que respondem de forma esforçada à prova. Esta é a razão pela qual a maioria dos testes de simulação utiliza algum tipo de prova de reconhecimento e não de evocação (Tombaugh, 1996).

Como referimos anteriormente os TVS devem ser integrados numa bateria de provas neuropsicológicas. Assim, o efeito da ordem pela qual o TOMM é aplicado na bateria de provas, no desempenho dos avaliados tem sido alvo de diversos estudos. No estudo de Ryan, Glass, Hinds e Brown (2010) os autores afirmam que a ordem pela qual este teste é aplicado não possui um efeito

significativo nos resultados da prova, exceto quando é aplicado o BNT (*Boston Naming Test*) imediatamente antes do TOMM, pelo que se torna desaconselhável fazê-lo. Gast e Hart (2010) alertam para a necessidade de verificar em estudos futuros a possibilidade do TOMM ser mais eficaz quando aplicado no início da bateria de provas pois a sua aplicação tardia poderá aumentar a sua transparência devido ao contacto que o avaliado poderá ter tido com provas de memória mais difíceis o que o poderá levar, conseqüentemente, a estranhar a reduzida dificuldade desta prova.

Atualmente existem inúmeros testes que procuram despistar a simulação de sintomas como, por exemplo, o *Word Memory Test* (WMT), o *Amsterdam Short Term Memory Test* (ASTMT), o *Computerized Assessment of Response Bias* (CARB), entre outros. Problemas de alguns testes, como a falta de um ponto-de-corte universalmente aceite, a classificação incorreta de pacientes com danos cerebrais como simuladores, a duração de administração da prova ou a sua fraca semelhança com um teste de memória levaram Tombaugh a desenvolver o TOMM (Tombaugh, 1996). Não obstante, o TOMM, ou qualquer outro TVS, não deve ser a única fonte de informação para realizar um diagnóstico de simulação (Armistead-Jehle & Gervais, 2011). É essencial uma avaliação qualitativa do técnico, baseada nas entrevistas realizadas, bem como a utilização de outros instrumentos de avaliação. No estudo de Gervais e colaboradores (2004), a aplicação do TOMM é identificada como tendo maior utilidade em casos extremos de simulação e de esforço reduzido, sendo outros testes, como o CARB ou o WMT mais fiáveis na identificação de casos de simulação mais discretos.

## **II - O PRESENTE ESTUDO: OBJETIVOS, PROBLEMAS E HIPÓTESES DE ESTUDO**

Como referido anteriormente, embora exista já um grande número de TVS, é fundamental o desenvolvimento de novas provas para que o avaliador possua à sua disposição diferentes testes, podendo escolher qual melhor se aplica a cada avaliação e conseguindo assim ultrapassar obstáculos como o treino dos avaliados por terceiros. É igualmente importante desenvolver novos testes que possuam um tempo de aplicação mais adequado às avaliações forenses e neuropsicológicas. Por fim, é essencial aplicar os conhecimentos e técnicas provenientes dos recentes estudos acerca da simulação na construção de novos TVS.

Conseqüentemente, no presente estudo procurámos: (1) construir e desenvolver um novo TVS baseado na estrutura do TOMM, mas utilizando diversos procedimentos inovadores de forma a otimizar a utilidade e o poder discriminativo desta prova, nomeadamente, construindo uma prova completamente informatizada com um tempo de aplicação reduzido; (2) desenvolver pontos-de-corte eficazes de forma a maximizar a sensibilidade, a especificidade e o valor preditivo da prova e (3) explorar o efeito de diversas variáveis na deteção da simulação, como por exemplo, o tempo de

resposta dos participantes, o treino dos avaliados ou a ordem pela qual o TVS é aplicado na bateria de testes neuropsicológicos.

Com base nestes objetivos, formulámos os seguintes problemas de investigação e respetivas hipóteses de estudo fundamentadas na bibliografia existente, as quais procurámos testar ao longo deste estudo.

### **Problema 1**

A ordem de aplicação do teste de simulação na bateria de provas que o inclui, bem como a ordem de aplicação dos dois ensaios desta prova altera o desempenho dos participantes?

### **Hipótese 1**

Com base na bibliografia (ex.: Ryan et al., 2010) e na natureza dos estímulos que englobam cada ensaio da prova, prevemos que a ordem de aplicação da prova e dos ensaios que a constituem é irrelevante no que diz respeito ao desempenho dos avaliados

### **Problema 2**

Diferentes grupos de participantes (grupo normativo, grupo de simuladores e grupo clínico) obtêm desempenhos (número de acertos) diferentes na prova de simulação?

### **Hipótese 2**

O grupo normativo e o grupo clínico irão apresentar resultados elevados e semelhantes na prova de simulação, ao contrário do grupo de simuladores que irá apresentar um desempenho consideravelmente mais baixo.

### **Problema 3**

Diferentes grupos de participantes apresentam diferentes tempos de resposta no teste de simulação?

### **Hipótese 3**

O grupo de simuladores demorará mais tempo a responder aos itens do teste, comparativamente ao grupo normativo pois, o primeiro grupo, para além de identificar a resposta correta, terá de decidir se pretende acertar ou errar a resposta. O grupo clínico demorará mais tempo a responder, em comparação ao grupo normativo, pois, as dificuldades cognitivas que estes apresentam poderão afetar a sua rapidez de raciocínio e tomada de decisão. É imprevisível, com base na literatura, se será o grupo de simuladores ou o grupo clínico a apresentar tempos de resposta mais longos no teste de simulação.

#### **Problema 4**

Os tempos de resposta dos participantes variam em função do tipo de resposta (acerto ou erro)?

#### **Hipótese 4**

Com base na literatura, prevemos que, para todos os grupos de participantes, as respostas corretas serão mais rápidas do que as respostas erradas.

#### **Problema 5**

Será o teste de simulação capaz de identificar eficazmente os simuladores, mesmo quando estes foram previamente informados acerca da intenção do investigador em detetar a simulação, acerca das perturbações que pretendem simular, bem como acerca da simulação de sintomas e dos diversos métodos para a sua deteção?

#### **Hipótese 5**

O teste é capaz de identificar eficazmente este tipo de simuladores pois estes são incapazes de discriminar qual dos testes aplicados na bateria de provas é o teste de simulação.

#### **Problema 6**

O número de acertos dos participantes, bem como os seus tempos de resposta, variam em função de se tratar do primeiro ensaio ou do segundo ensaio a que estes respondem?

#### **Hipótese 6**

É previsível que o número de acertos aumente, e que o tempo de resposta diminua, do primeiro para o segundo ensaio a que os participantes respondem, pois, no segundo ensaio respondido, os participantes não só terão tido a oportunidade de treinar a execução das respostas, no primeiro ensaio, mas terão também tido acesso pela segunda vez aos estímulos alvo (uma vez em cada ensaio da prova) e recebido feedback das suas respostas certas e erradas. Este aumento no desempenho dos participantes não irá ocorrer para o grupo de simuladores, pois este grupo irá procurar obter baixos desempenhos independentemente se tratar do primeiro ou do segundo ensaio a que respondem.

#### **Problema 7**

Variáveis como o sexo, a idade ou a memória dos participantes estão associadas ao desempenho dos mesmos na prova de simulação?

## **Hipótese 7**

Uma vez que o teste de simulação não procura avaliar nem discriminar nenhuma destas variáveis, prevemos que estas não estão associadas ao desempenho dos participantes no teste de simulação.

De forma a atingir os objetivos anteriormente descritos, testar as hipóteses formuladas e testar a eficácia da prova de simulação construída, aplicámos este teste a três grupos de participantes: um grupo normativo, um grupo de simuladores e um grupo clínico. Para os dois primeiros grupos referidos, inserimos a prova de simulação numa bateria de provas de memória que nós próprios construímos, utilizando uma versão adaptada para o computador de três testes da *Wechsler Memory Scale – Third Edition* (WMS-III): Teste de Reconhecimento de Faces (“*Faces*”- no original), Teste de Reprodução Visual e Teste de Localização Espacial (Wechsler, 1997). De seguida, iremos abordar mais pormenorizadamente cada um destes testes, bem como a metodologia por nós adotada.

## **III – METODOLOGIA**

### **3.1 Participantes**

A amostra total foi constituída por 69 participantes, sendo 59 (85.5%) do sexo feminino e 10 (14.5%) do sexo masculino. A idade dos participantes variou entre os 18 e os 46 anos, com uma média de 22.22 e com um desvio padrão de 5.90. Esta amostra foi constituída por três grupos de participantes: grupo normativo, grupo de simuladores e grupo clínico.

O grupo normativo foi constituído por 41 estudantes universitários que frequentavam o Mestrado Integrado em Psicologia da Universidade do Minho, sendo 37 (90.2%) do sexo feminino e quatro (9.8%) do sexo masculino. Aos participantes deste grupo foram atribuídos créditos pela participação no presente estudo, os quais poderiam mais tarde ser utilizados para subir parcialmente a sua classificação em determinadas Unidades Curriculares, tal como proposto no atual programa de creditação das experiências da Escola de Psicologia da Universidade do Minho. A idade dos participantes variou entre os 18 e os 28 anos, com uma média de 19.71 e com um desvio padrão de 1.93. A este grupo de participantes foi pedido que respondessem à bateria de provas de forma a obterem o melhor resultado possível.

O grupo de simuladores foi constituído por 24 estudantes universitários que frequentavam o quarto ano do Mestrado Integrado em Psicologia da Universidade do Minho, nas áreas de especialização de Psicologia Clínica ou Psicologia da Justiça, sendo 21 (87.5%) do sexo feminino e três (12.5%) do sexo masculino. A razão de apenas termos incluído neste grupo alunos do quarto ano do curso e áreas de especialização previamente mencionadas, prende-se com o facto de estes mesmos alunos terem frequentado unidades curriculares relacionadas com a memória humana, avaliação psicológica e/ou avaliação forense, potenciando assim a utilização de estratégias de simulação mais elaboradas da parte destes, o que, por sua vez, permitiu uma melhor testagem da eficácia da prova de simulação construída. Embora uma minoria dos participantes deste grupo tenha participado no estudo sem receber qualquer tipo de recompensa pela sua participação, à maioria dos participantes foram atribuídos créditos por participar na experiência. A idade dos participantes variou entre os 21 e os 35 anos, com uma média de 23.33 e com um desvio padrão de 3.66. Embora o grupo fosse inicialmente composto por 25 participantes, um deles foi excluído pois mostrou não ter compreendido as instruções das provas. Para este grupo foi pedido aos participantes que respondessem à bateria de provas imaginando que estariam a simular uma perturbação mnésica com o objetivo de obter uma avultada compensação monetária da sua companhia de seguros.

O grupo clínico foi constituído por quatro participantes que, no momento da avaliação, estavam a ser acompanhados no Serviço de Psicologia da Universidade do Minho, realizando sessões semanais de reabilitação cognitiva. Este grupo foi composto por quatro participantes, um (25%) do sexo feminino e três (75%) do sexo masculino. Os participantes deste grupo cooperaram voluntariamente nesta investigação. A idade dos participantes variou entre os 30 e os 46 anos, com uma média de 41.25 e com um desvio padrão de 7.54. Uma vez que estes participantes tinham sido, ou estavam a ser, avaliados no âmbito das sessões de reabilitação cognitiva, foi pedido à terapeuta responsável, com o consentimento dos participantes, que fornecesse o diagnóstico clínico de cada indivíduo. Assim, verificámos que dois participantes possuíam um diagnóstico de Doença de Alzheimer com visíveis défices mnésicos e cognitivos associados, outro participante exibia défices de memória e linguagem associados a um acidente vascular cerebral (AVC) e, por fim, o último participante exibia diversos défices cognitivos e mnésicos, bem como sintomas de comprometimento da função executiva, resultantes de uma lesão cerebral após um acidente de trabalho. Embora a amostra final utilizada, para este grupo, seja composta por apenas quatro participantes, a amostra inicial era composta por mais três sujeitos que foram excluídos devido a possuírem um diagnóstico de Doença de Alzheimer em estado avançado, tendo demonstrado ser incapazes de compreender e reter a informação relativa às instruções da prova. Ao grupo clínico, tal como ao grupo normativo, foi pedido a todos os participantes que respondessem à bateria de provas de forma a obterem o melhor resultado possível.

## 3.2 Equipamento e Materiais

### 3.2.1 DETECTS - Teste de Detecção da Simulação de Problemas de Memória

Tal como referido anteriormente, um dos principais objetivos deste estudo foi o de construir uma prova de simulação, baseada no TOMM (Tombaugh, 1996) e na bibliografia existente acerca dos TVS. O teste desenvolvido (DETECTS – Teste de Detecção da Simulação de Problemas de Memória) é completamente informatizado e possui um tempo de aplicação reduzido (cerca de 15 minutos). Este teste é composto por dois ensaios. Em cada ensaio, o avaliado observa 50 imagens, seguindo-se uma prova de reconhecimento imediata onde são apresentados 50 pares de imagens, sendo que o participante deve, obrigatoriamente, escolher uma das imagens, mais concretamente, a imagem que pensa ter sido apresentada anteriormente. Embora as figuras apresentadas sejam iguais em ambos os ensaios e, conseqüentemente, as opções corretas nas duas provas de reconhecimento também o sejam, os estímulos distrativos são diferentes para cada uma das provas de reconhecimento. Em ambos os ensaios, os estímulos são apresentados aleatoriamente, tanto na etapa de processamento, como na etapa de reconhecimento. A ordem pela qual os participantes visualizam e respondem a cada ensaio (ensaio um e ensaio dois) foi controlada (metade dos participantes realizou o ensaio um, seguido pelo ensaio dois, e a outra metade realizou o procedimento contrário). Assim, por primeira fase do teste referimo-nos ao ensaio que dado participante visualizou e respondeu em primeiro lugar (independentemente de se tratar do ensaio um ou do ensaio dois) e, por segunda fase, referimo-nos ao segundo ensaio a que cada participante respondeu, mais uma vez, independentemente de se tratar do ensaio um ou dois.

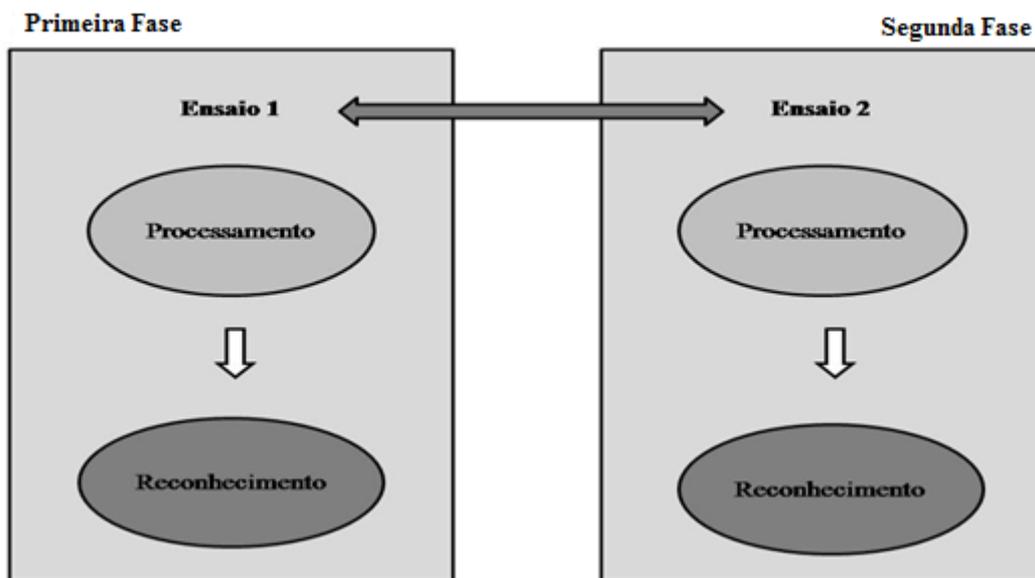


Figura 1 – Esquema da Estrutura da Prova de Simulação (DETECTS).

O DETECTS foi elaborado em formato computadorizado através do programa informático Superlab 4.5 (Cedrus Corporation, San Pedro, CA, USA). Os participantes responderam às provas de reconhecimento utilizando o teclado do computador, mais pormenorizadamente, carregando na tecla correspondente à letra adjacente às imagens que consideravam ter visto anteriormente. Após efetuarem a sua resposta, os participantes receberam feedback imediato acerca da exatidão da mesma (Certo ou Errado). Para além de registar a resposta dos participantes a cada um dos diapositivos, o programa utilizado registou também o tempo que cada participante demorou a efetuar cada uma das suas respostas. A prova inclui ainda as instruções de resposta no próprio procedimento, bem como uma fase inicial de treino.

No que diz respeito aos estímulos apresentados nesta prova, optámos por utilizar estímulos provenientes da base de dados de imagens de Snodgrass e Vanderwart (Snodgrass & Vanderwart, 1980). Escolhemos esta base de dados não só por ter uma política de utilização livre, mas também por conter um elevado número de imagens (260 imagens), permitindo assim a categorização e consequente escolha das figuras mais adequadas. Assim, procedemos à categorização de todas as imagens, tendo sido elaboradas 100 categorias diferentes. Estas categorias foram construídas de forma a serem bastante específicas e restritas (ex.: veículos terrestres, veículos aéreos, veículos náuticos, etc.). Todas as imagens escolhidas foram meticulosamente tratadas para que os estímulos fossem o mais uniformes possível (ex.: todas as imagens têm o mesmo tamanho e são apresentadas na parte central do ecrã). O elevado número de categorias criado permitiu-nos escolher 50 imagens, todas elas de categorias diferentes, para a lista de imagens alvo a apresentar individualmente em ambos os ensaios. Criámos também duas listas de 50 estímulos distrativos, uma para cada ensaio. Mais uma vez, dentro de cada lista não há repetição de categorias, ou seja, cada imagem pertence a uma categoria diferente. Da mesma forma, os estímulos distrativos, emparelhados com os estímulos alvo nas provas de reconhecimento, nunca são da mesma categoria que estes. Os dois estímulos distrativos utilizados para o mesmo estímulo alvo, um para cada ensaio, são sempre da mesma categoria (entre si).

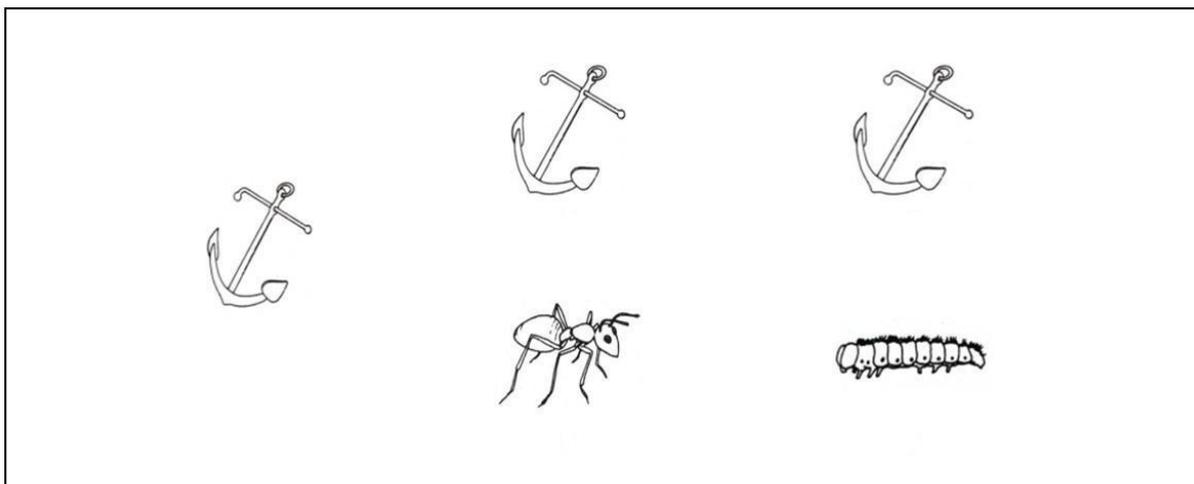


Figura 2 – Exemplo da sequência de estímulos apresentados, para o estímulo alvo “âncora”, nas diversas etapas da prova de simulação (DETECTS).

De forma a clarificar o parágrafo anterior, observemos a figura 2. A imagem “âncora”, apresentada como estímulo alvo em ambos os ensaios, é apresentada também em ambas as provas de reconhecimento, adjacente a dois estímulos distrativos diferentes. O estímulo alvo (“âncora”) é de uma categoria diferente dos estímulos distrativos (“formiga” e “lagarta”). No entanto, uma vez que ambos os estímulos distrativos estão adjacentes ao mesmo estímulo alvo (“âncora”) estes são da mesma categoria (“insetos”). Esta metodologia foi adotada de forma a diminuir a dificuldade da prova, permitindo assim a não discriminação dos participantes em função das suas capacidades mnésicas.

### **3.2.2 Bateria de Provas**

Como referimos anteriormente, quer para o grupo normativo, quer para o grupo de simuladores o DETECTS foi aplicado juntamente com três provas da WMS-III adaptadas para computador: a prova de reprodução visual, localização espacial e reconhecimento de faces (Wechsler, 1997). A prova de reconhecimento de faces (“*Faces*” – no original) é um teste de memória visual, não verbal, que se consiste na apresentação de um conjunto de faces e execução de uma consequente tarefa de reconhecimento onde o avaliado deve afirmar se já visualizou, ou não, as diversas faces apresentadas. A prova de reprodução visual é também uma prova de memória visual, não verbal, que inclui uma tarefa de reconstrução visuográfica dos estímulos pois, neste teste, é pedido aos avaliados para, após visualizar cada uma das figuras geométricas apresentadas, representá-las graficamente no papel. Embora as duas provas descritas sejam ambas provas de memória visual, não verbal, utilizam estímulos bastante diferentes (faces e figuras geométricas), sendo assim expectável que o mesmo participante possa obter desempenhos bastantes distintos nas duas provas, pois, a título de exemplo, o facto de determinado participante possuir uma elevada capacidade de memorização de faces não significa que este exiba também uma elevada capacidade para memorizar figuras geométricas, bem como para as reconstruir graficamente. Por fim, a prova de localização espacial pretende avaliar a memória de trabalho. Nesta prova é pedido aos avaliados para repetir por ordem direta e, numa segunda fase da prova, por ordem inversa diferentes sequências que o próprio investigador executa tocando num conjunto de blocos com a sua mão (este teste é também conhecido por prova dos *Blocos de Corsi*).

Aplicámos a prova de simulação em conjunto com estes três testes, não só de forma a evitar que os participantes identificassem o DETECTS como uma prova de simulação, diminuindo assim a transparência deste teste, mas também de forma a avaliar se a prova por nós construída estaria, ao contrário do pretendido, a avaliar os participantes em função da sua capacidade mnésica. O facto de as provas escolhidas avaliarem diferentes tipos de memória e conterem estímulos bastante distintos bem como regras de aplicação e cotação bem definidas, as quais podem ser consultadas no manual da WMS-III, torna-as indicadas para este fim. Todas as provas foram adaptadas para formato computadorizado de forma a manter a uniformidade da bateria de provas. Não obstante, todos os

parâmetros de aplicação e cotação dos testes foram idênticos aos das provas originais. Para a prova de reprodução visual os participantes utilizaram, para além do computador, uma folha e um lápis para desenhar as imagens apresentadas, tal como utilizado na prova original.

### ***3.2.3 Instruções de Simulação***

Para o grupo de simuladores foi ainda criada uma apresentação de diapositivos, utilizando o mesmo programa informático (Superlab 4.5), com o objetivo de informar os participantes acerca das diversas perturbações mnésicas e das suas consequências, bem como acerca da simulação destas perturbações e das várias técnicas existentes para a deteção da simulação (ex.: TVS). Esta apresentação procurou recriar um contexto real onde o avaliado, de forma a simular mais eficazmente, pesquisou informação, previamente à avaliação, de forma a saber como proceder perante esta. Estes diapositivos continham também algumas indicações para ajudar o participante a recriar mentalmente o contexto de uma avaliação neuropsicológica e/ou forense. Assim, nestes diapositivos, foi pedido a cada participante para imaginar que, após um acidente de viação onde sofreu uma pequena lesão sem qualquer tipo de consequências para o seu funcionamento cognitivo e mnésico, decidiu apresentar um pedido de indemnização à sua companhia de seguros afirmando que o acidente tinha afetado a sua memória. Como consequência, a seguradora marcou uma avaliação neuropsicológica onde este pretende enganar o avaliador de forma a fazê-lo crer que possui algum tipo de perturbação mnésica. Foi ainda pedido a cada participante para imaginar que a presente investigação, em que iriam participar de seguida, se tratava, de fato, da avaliação neuropsicológica previamente mencionada, sendo o seu objetivo iludir o investigador, fazendo-o crer que possuíam uma verdadeira perturbação de memória.

### ***3.2.4 Outros Materiais***

Os participantes do grupo normativo e do grupo de simuladores foram avaliados individualmente numa cabine insonorizada do laboratório de Cognição Humana da Escola de Psicologia da Universidade do Minho, de forma a eliminar qualquer estímulo exterior que pudesse influenciar os resultados. Para estes grupos de participantes, todos os testes foram aplicados utilizando o mesmo computador (*Fujitsu Siemens Computers - Esprimo P3500*), o mesmo monitor LCD (*Fujitsu Siemens Computers – Scenicview A17-2*), o mesmo teclado (*Fujitsu Siemens Computers – KB SC USB P*) e o mesmo rato (*Fujitsu Siemens Computers - Logitech M-UAE96*). Os participantes do grupo clínico responderam apenas à prova de simulação utilizando um computador *ASUS FS8*, na sala onde habitualmente realizavam as sessões de reabilitação. Por questões de disponibilidade dos próprios participantes, e de forma a interromper o menos possível as sessões terapêuticas, foi impossível deslocar tais participantes para as cabines insonorizadas onde os restantes dados foram recolhidos.

Não obstante, foram sempre garantidas as condições necessárias para que a prova fosse aplicada de forma adequada a estes participantes (ex.: silêncio, conforto, ausência de terceiros na sala, etc.). Verificámos também, em todos os computadores utilizados, que os estímulos apresentados eram facilmente perceptíveis e que os restantes materiais, utilizados pelos participantes para fornecer a sua resposta, funcionavam adequadamente

### 3.3 Plano e Procedimento Experimental

Todos os participantes foram avaliados individualmente. Foi explicado a cada participante que iria participar num estudo relacionado com a memória humana e que os seus dados seriam tratados com confidencialidade, ou seja, que apenas seriam observados e analisados pelos investigadores do presente estudo. Todos os participantes deram o seu consentimento oral para a participação.

No que diz respeito ao grupo normativo, foi indicado aos participantes que iriam realizar um conjunto de provas de avaliação psicológica para avaliar diversos tipos de memória (ex.: memória visual, memória de trabalho, etc.). Foi pedido aos participantes para aplicarem o seu máximo esforço de forma a apresentarem o melhor desempenho possível. De seguida, os participantes foram encaminhados para a cabine insonorizada onde realizaram todas as provas. A ordem de aplicação dos diversos testes foi controlada. Assim, criámos quatro ordens de aplicação da bateria de provas e distribuímos os participantes equitativamente pelas quatro condições experimentais. A razão de termos escolhido criar quatro ordens de aplicação dos testes prende-se com o facto de este ser o número mínimo para que a prova de simulação fosse imediatamente precedida e sucedida por todas as restantes provas. Para cada uma das ordens de aplicação dos testes estabelecidas, 50% dos sujeitos realizaram primeiro o ensaio um da prova de simulação (E1), seguido pelo ensaio dois (E2), e os restantes 50% realizaram o procedimento contrário. A figura seguinte ilustra as diferentes condições experimentais que acabámos de explorar.

| <b>1ª Ordem de Aplicação da Bateria de Provas</b>                            | <b>3ª Ordem de Aplicação da Bateria de Provas</b> |
|--|---|
| D <sup>1</sup> (E1-E2) → LE <sup>2</sup> → RV <sup>3</sup> → RF <sup>4</sup> | LE → RF → D (E1-E2) → RV                          |
| D (E2-E1) → LE → RV → RF   | LE → RF → D (E2-E1) → RV                          |
| <b>2ª Ordem de Aplicação da Bateria de Provas</b>                            | <b>4ª Ordem de Aplicação da Bateria de Provas</b> |
| LE → D (E1-E2) → RF → RV   | RF → LE → RV → D (E1-E2)                          |
| LE → D (E2-E1) → RF → RV   | RF → LE → RV → D (E2-E1)                          |

**Legenda:** <sup>1</sup> Prova de Simulação (DETECTS); <sup>2</sup> Prova de Localização Espacial; <sup>3</sup> Prova de Reprodução Visual; <sup>4</sup> Prova de Reconhecimento de Faces

Figura 3 – Esquema das diversas ordens de aplicação dos testes na bateria de provas construída.

Para o grupo de simuladores foi adotado um procedimento muito semelhante ao acima descrito, com apenas algumas diferenças metodológicas. Assim, para este grupo de participantes, foi indicado, imediatamente antes da realização dos testes, que os participantes deveriam de imaginar o seguinte contexto: *“Imagine ter sido vítima de um acidente de viação no qual sofreu uma pequena lesão na cabeça. Esta lesão não provocou qualquer consequência ao nível do seu funcionamento cognitivo e mnésico. No entanto, para obter uma avultada compensação monetária da sua companhia de seguros, deverá procurar convencer o avaliador, nomeado pela sua companhia de seguros, que o acidente teve consequências ao nível do seu funcionamento mnésico. Assim, imagine que o procedimento que irá realizar de seguida não faz parte de uma investigação científica, mas sim de uma avaliação neuropsicológica, sendo eu o técnico nomeado pela sua seguradora para o avaliar. O seu objetivo é fazer com que eu, através da análise dos testes que irá realizar, seja incapaz de o identificar como simulador, fazendo-me crer que possui legítimos problemas de memória”*. Foi ainda apresentado a todos os participantes um conjunto de diapositivos que esclareciam, novamente, estas instruções de forma a eliminar possíveis equívocos. Tal como referido anteriormente, e pelas razões anteriormente descritas, estes diapositivos continham também informação acerca das perturbações mnésicas, da simulação destas perturbações e dos tipos de testes utilizados para a deteção da simulação.

Relativamente ao grupo clínico, também pelas razões anteriormente descritas, foi apenas aplicado o DETECTS. Metade dos participantes deste grupo responderam ao primeiro ensaio da prova, seguido pelo segundo ensaio e outra metade realizou o procedimento contrário. A este grupo foram dadas instruções idênticas ao grupo normativo, ou seja, foi-lhes indicado que iriam realizar uma prova de memória e que deveriam de responder a esta de forma a obterem o melhor desempenho possível.

O teste de simulação foi cotado através do Superlab 4.5, tendo, para tal, sido registados pelo próprio programa informático todos os acertos e erros de cada participante em cada um dos ensaios da prova, bem como os seus tempos de resposta, tanto para acertos, como para erros. As provas de memória aplicadas foram cotadas tal como sugerido no manual da WMS-III.

#### **IV – RESULTADOS**

Previamente à realização de testes paramétricos foi efetuada a análise exploratória de dados de forma a avaliar se os pressupostos necessários para a realização destes mesmos testes (normalidade da distribuição e homogeneidade das variâncias) estavam cumpridos. Quando tais pressupostos mostraram estar cumpridos, utilizaram-se testes paramétricos. Quando tais pressupostos mostraram

não estar cumpridos, optou-se por realizar ambos os testes (paramétricos e não paramétricos). Neste caso, tal como sugerido por Field (2009), optámos por relatar apenas o teste não paramétrico, quando a sua significância aponta num sentido diferente da significância do teste paramétrico (um dos testes é significativo e o outro não o é) e optámos por relatar apenas o teste paramétrico quando a significância de ambos os testes aponta no mesmo sentido, por exemplo, quando ambos os testes se revelam significativos.

Sempre que necessário recorrer a testes post-hoc, optámos por utilizar o Teste de Gabriel pois trata-se de um teste adequado para análises que incluam grupos com diferentes números de participantes (Martins, 2011). Quando utilizámos múltiplos testes t em simultâneo, foi aplicada a correção de Bonferroni. Para todas as análises foi adotado um nível de significância de 95%.

#### 4.1 Ordem de Aplicação das Provas e Acertos no Teste de Simulação (DETECTS)

Um dos objetivos do nosso estudo era compreender o efeito da ordem de aplicação do DETECTS face às restantes provas aplicadas, nos resultados obtidos pelos participantes nesta mesma prova. Após realizarmos diversas análises (ANOVAs Unifatoriais), verificámos que, relativamente ao eventual efeito da ordem pela qual a prova de simulação foi aplicada, no número de acertos obtidos nesta mesma prova, não há diferenças significativas ao nível do número de acertos em qualquer uma das fases da prova simulação. Esta afirmação aplica-se para todos os grupos de participantes. Consequentemente, nas análises seguintes, optámos por não examinar separadamente os dados dos participantes em função da ordem pela qual estes realizaram as diversas provas, uma vez que, tal como referido, a ordem de aplicação das provas não tem qualquer efeito no desempenho dos participantes no teste de simulação.

Tabela 1 – Média, Desvio Padrão e análises estatísticas realizadas para o número de acertos dos diversos grupos de participantes nas duas fases da prova de simulação (DETECTS), em função da ordem pela qual os diversos testes da bateria de provas foram aplicados.

| <b>Grupo de Participantes</b> | <b>Fase da Prova</b> | <b>Acertos (1ª Ordem)</b> | <b>Acertos (2ª Ordem)</b> | <b>Acertos (3ª Ordem)</b> | <b>Acertos (4ª Ordem)</b> | <b>ANOVAs Unifatoriais</b>       |
|-------------------------------|----------------------|---------------------------|---------------------------|---------------------------|---------------------------|----------------------------------|
| Normativo                     | 1ª                   | 48.64<br>(1.57)           | 48.10<br>(1.37)           | 49<br>(1.63)              | 49.40<br>(1.08)           | $F(3, 37) = 1.49,$<br>$p = .230$ |
| Simuladores                   | 1ª                   | 25.50<br>(14.60)          | 33<br>(10.97)             | 34.83<br>(4.88)           | 29.10<br>(3.82)           | $F(3, 20) = 1.11,$<br>$p = .367$ |
| Normativo                     | 2ª                   | 49.73<br>(.65)            | 49.60<br>(.70)            | 49.50<br>(.53)            | 49.90<br>(.32)            | $F(3, 37) = .92,$<br>$p = .439$  |
| Simuladores                   | 2ª                   | 23.17<br>(15.69)          | 35.17<br>(11.75)          | 35.17<br>(7.36)           | 27.83<br>(4.07)           | $F(3, 20) = 1.83,$<br>$p = .100$ |

Um dos pontos essenciais do nosso estudo foi a análise do número de acertos dos diversos grupos de participantes nas diversas fases do DETECTS, de forma a compreender que papel a contagem e análise das respostas corretas nesta prova pode ter na detecção da simulação. Assim, efetuámos diversas análises estatísticas (ANOVA Bifatorial Mista e Testes t para amostras emparelhadas) para verificar qual o efeito das variáveis: (1) grupo de participantes (três níveis: normativo, simuladores e clínico) e (2) fase da prova (dois níveis: primeira e segunda fase) no número de acertos na prova de simulação.

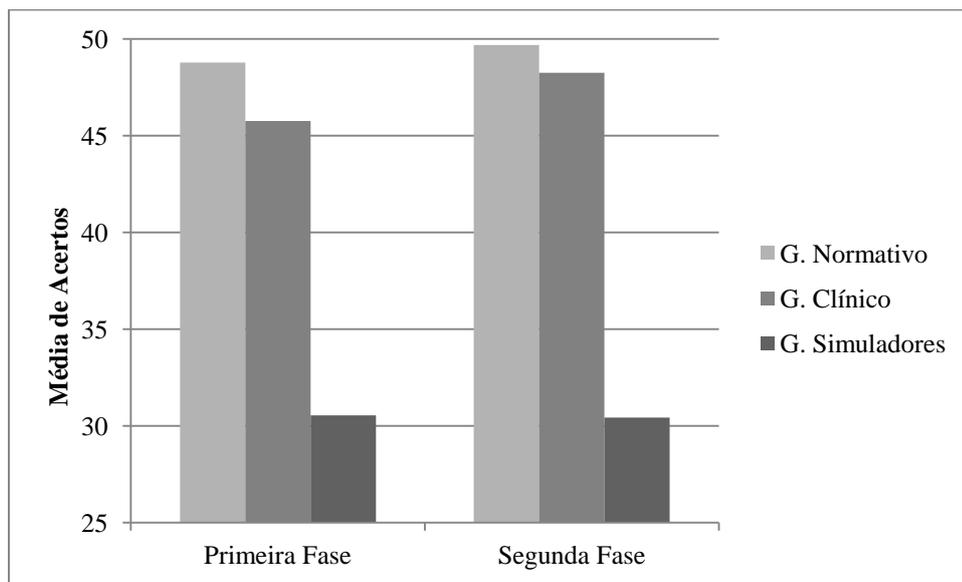


Figura 4 – Média do número de acertos dos diversos grupos de participantes nas duas fases do teste de simulação (DETECTS).

No que concerne à amostra total, ou seja, quando incluídos todos os grupos de participantes, existe um efeito principal significativo da fase do teste de simulação no número de acertos nesta mesma prova ( $Wilks \Delta = .94$ ,  $F(1, 66) = 4.29$ ,  $p = .042$ ). Tal como esperado, na segunda fase do DETECTS a média do número de acertos dos participantes é mais elevada ( $\bar{X}_{F1} = 41.72$ ,  $\bar{X}_{F2} = 42.76$ ). De forma mais pormenorizada, verificámos que, para o grupo normativo, há diferenças significativas ao nível do número de acertos, em função da fase da prova de simulação ( $t(40) = -4.45$ ,  $p < .001$ ), sendo que o número de acertos na segunda fase desta prova é mais elevado do que na primeira. Não obstante, não há diferenças significativas entre o número de acertos na primeira fase do DETECTS e o número de acertos na segunda fase desta prova quer para o grupo de simuladores ( $t(23) = .35$ ,  $p = .728$ ) quer para o grupo clínico ( $t(3) = -1.73$ ,  $p = .182$ ).

Existe um efeito principal significativo do grupo de participantes no número de acertos no DETECTS, independentemente da fase da prova em questão ( $F(2,66) = 71.70, p < .001$ ). O teste post-hoc aplicado a esta análise revela que existem diferenças significativas entre o grupo de simuladores e o grupo normativo ao nível do número de acertos nesta prova ( $p < .001$ ) ( $\bar{X}_{G. Normativo} = 49.23, \bar{X}_{G. Simuladores} = 30.48$ ). Há também diferenças significativas entre o grupo de simuladores e o grupo clínico ao nível do número de acertos no DETECTS ( $p < .001$ ) ( $\bar{X}_{G. Simuladores} = 30.48, \bar{X}_{G. Clínico} = 47$ ). Como esperado, não encontramos diferenças significativas entre o grupo normativo e o grupo clínico ao nível do número de acertos no DETECTS ( $p = 1$ ) ( $\bar{X}_{G. Normativo} = 49.23, \bar{X}_{G. Clínico} = 47$ ). Por fim, não há um efeito de interação significativo entre a fase da prova de simulação e o grupo de participantes no número de acertos nesta prova ( $Wilks \Delta = .93, F(2, 66) = 2.61, p = .081$ ).

Tabela 2 - Média e Desvio Padrão do número de acertos dos diversos grupos de participantes nas duas fases da prova de simulação (DETECTS).

|                   | Fase 1       | Fase 2        |
|-------------------|--------------|---------------|
| Grupo Normativo   | 48.78 (1.46) | 49.68 (.57)   |
| Grupo Simuladores | 30.63 (9.71) | 30.33 (11.23) |
| Grupo Clínico     | 45.75 (2.99) | 48.25 (2.36)  |

Em suma, o grupo normativo e o grupo clínico apresentaram um maior número de acertos na prova de simulação, em relação ao grupo de simuladores, facilitando assim a deteção dos mesmos.

## 4.2 Tempos de Resposta

Outro procedimento bastante promissor na deteção da simulação é a medição e análise dos tempos de resposta de cada participante (ex.: Haines & Norris, 1995). Assim, efetuámos diversas análises estatísticas (ANOVA Multifatorial Mista e Testes t para amostras emparelhadas) de forma a verificar qual o efeito das variáveis: (1) grupo de participantes (três níveis: normativo, simuladores e clínico), (2) fase da prova (dois níveis: primeira e segunda fase) e (3) tipo de resposta (dois níveis: acerto ou erro) no tempo de resposta dos participantes.

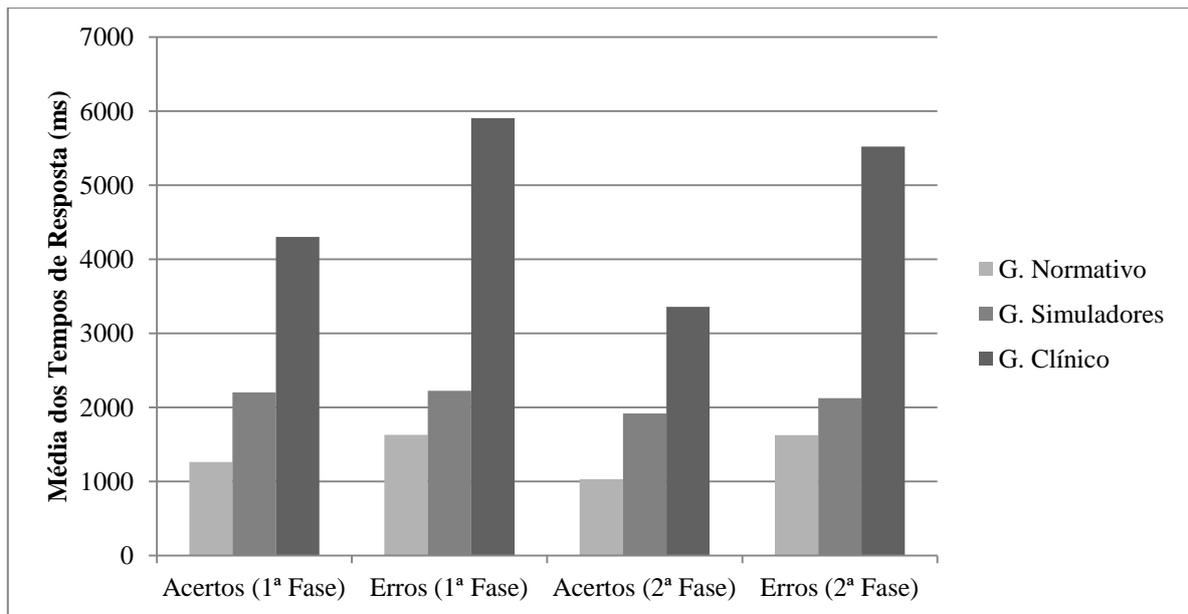


Figura 5 - Tempos médios de resposta dos diversos grupos de participantes, para acertos e erros, nas duas fases da prova de simulação (DETECTS).

Verificámos que há um efeito principal significativo da fase do DETECTS nos tempos de resposta dos participantes ( $Wilks \Delta = .78, F(1, 29) = 8.29, p = .007$ ) sendo que, os participantes respondem de forma mais rápida na segunda fase da prova do que na primeira fase deste mesmo teste, independentemente do grupo a que pertencem ou do tipo de resposta que fornecem ( $\bar{X}_{TR F1} = 2918, \bar{X}_{TR F2} = 2594$ ).

Existe também um efeito principal significativo do tipo de resposta (erros ou acertos) no tempo de resposta dos participantes ( $Wilks \Delta = .31, F(1, 29) = 66.15, p < .001$ ). Independentemente do grupo de participantes e da fase da prova em questão, os sujeitos fornecem, de forma mais rápida, respostas corretas (acertos) do que as respostas incorretas ou erros ( $\bar{X}_{TR acertos} = 2342, \bar{X}_{TR erros} = 3169$ ). Mais pormenorizadamente, para o grupo normativo, há diferenças significativas entre o tempo de resposta para os acertos e o tempo de resposta para os erros, apenas no que diz respeito à primeira fase do DETECTS ( $t(18) = -4.45, p < .001$ ). O tempo de resposta para os erros é mais longo do que o tempo de resposta para os acertos. Para os restantes grupos, bem como para a segunda fase do teste de simulação não encontramos diferenças significativas entre o tempo de resposta para acertos e o tempo de resposta para erros, pois, embora algumas análises tenham revelado um valor de significância inferior a .05, após aplicar a correção de Bonferroni, tais diferenças revelaram-se não significativas.

Encontrámos um efeito principal significativo do grupo de participantes nos tempos de resposta dos mesmos ( $F(2, 29) = 22.67, p < .001$ ). De forma mais pormenorizada, verificámos que existem diferenças significativas entre participantes do grupo normativo e participantes do grupo de

simuladores ao nível dos tempos de resposta ( $p = .028$ ) ( $\bar{X}_{G. Norm.} = 1384, D.P. = 225; \bar{X}_{G. Sim.} = 2114, D.P. = 136$ ). Existem também diferenças significativas entre participantes do grupo clínico e participantes do grupo de simuladores ao nível dos tempos de resposta ( $p < .001$ ) ( $\bar{X}_{G. Norm.} = 1384, D.P. = 225; \bar{X}_{G. Sim.} = 2114, D.P. = 136$ ). Por fim, existem diferenças significativas entre participantes do grupo normativo e participantes do grupo clínico ao nível desta mesma variável ( $p < .001$ ) ( $\bar{X}_{G. Norm.} = 1384, D.P. = 225; \bar{X}_{G. Clin.} = 4769, D.P. = 450$ ). Os participantes do grupo normativo são os que respondem mais rapidamente ao DETECTS, seguidos pelos participantes do grupo de simuladores e, por último, pelos participantes do grupo clínico, mostrando estes ser o grupo que mais tempo demora a fornecer uma resposta, independentemente de esta ser correta ou errada.

Não encontramos um efeito de interação significativo entre o tipo de resposta (acerto ou erro) e o grupo de participantes ao nível dos tempos de resposta ( $Wilks \Delta = .90, F(1, 29) = 3.18, p = .085$ ). Da mesma forma, não observamos um efeito de interação significativo entre a fase do DETECTS e o grupo de participantes ao nível dos tempos de resposta ( $Wilks \Delta = .91, F(2, 29) = 1.42, p = .258$ ). Por fim, não existe um efeito de interação significativo entre a fase do DETECTS, o grupo de participantes e o tipo de resposta ao nível do tempo de resposta dos participantes ( $Wilks \Delta = .98, F(2, 29) = .29, p = .75$ ).

Tabela 3 - Média e Desvio Padrão dos tempos de resposta (ms) dos diversos grupos de participantes nas duas fases da prova de simulação (DETECTS).

|                   | Acertos    |            | Erros      |            |
|-------------------|------------|------------|------------|------------|
|                   | Fase 1     | Fase 2     | Fase 1     | Fase 2     |
| Grupo Normativo   | 1261 (215) | 1027 (243) | 1627 (220) | 1624 (318) |
| Grupo Simuladores | 2199 (130) | 1916 (147) | 2221 (133) | 2122 (192) |
| Grupo Clínico     | 4299 (430) | 3355 (486) | 5902 (440) | 5521 (637) |

#### 4.3 Determinação de Pontos-de-corte no Teste de Simulação (DETECTS)

Tal como referido anteriormente, para que uma prova de simulação seja eficaz na deteção da mesma, é necessário o recurso a pontos-de-corte rigorosos que permitam aumentar a precisão do instrumento. Para tal é fundamental calcular os índices de Sensibilidade, Especificidade, Valor Preditivo Positivo (VPP) e Valor Preditivo Negativo (VPN) da prova, quando adotados diversos possíveis pontos-de-corte. Relembra-mos novamente que a especificidade se refere à proporção de avaliados que não apresentam a condição de interesse, ou seja, não simuladores, corretamente

classificados, estando este critério relacionado com o VPP, que se trata da probabilidade de um avaliado ser simulador quando é diagnosticado como tal. A sensibilidade refere-se à proporção de avaliados com a condição de interesse, ou seja, simuladores, corretamente identificados, estando este critério relacionado com o VPN que se trata da probabilidade de um avaliado não ser simulador quando assim identificado.

Nas duas tabelas seguintes apresentamos os índices de sensibilidade, especificidade, VPP e VPN para diversos possíveis pontos-de-corte a adotar para ambas as fases do DETECTS. Embora, para o estabelecimento dos pontos-de-corte desta prova nos tenhamos baseado no cálculo destes índices com a inclusão de todos os grupos constituintes da nossa amostra, calculámos também estes índices sem incluir, nessa mesma análise, o grupo clínico pois tais valores poderão ser interessantes para uma futura comparação dos índices por nós obtidos, com os de outros estudos que não incluam na sua amostra um grupo clínico (ex.: experiências de simulação). Esta análise complementar não foi efetuada para o índice de sensibilidade pois este mantém o mesmo valor independentemente de ser incluído, ou não, o grupo clínico na análise.

Tabela 4 – Índices de sensibilidade, especificidade, VPP e VPN para diversos possíveis pontos-de-corte a adotar para a primeira fase do DETECTS.

|                    | Acertos | Sensibilidade | Especificidade * | Especificidade | VPP * | VPP  | VPN * | VPN |
|--------------------|---------|---------------|------------------|----------------|-------|------|-------|-----|
| <b>Simuladores</b> | 37      | 79%           | 100%             | 100%           | 100%  | 100% | 89%   | 90% |
|                    | 38      | 79%           | 100%             | 100%           | 100%  | 100% | 89%   | 90% |
|                    | 39      | 83%           | 100%             | 100%           | 100%  | 100% | 91%   | 92% |
|                    | 40      | 83%           | 100%             | 100%           | 100%  | 100% | 91%   | 92% |
|                    | 41      | 88%           | 100%             | 100%           | 100%  | 100% | 93%   | 94% |
|                    | 42      | 96%           | 100%             | 100%           | 100%  | 100% | 100%  | 98% |
| <b>Não Sim.</b>    | 43      | 96%           | 100%             | 98%            | 100%  | 96%  | 98%   | 98% |
|                    | 44      | 96%           | 100%             | 98%            | 100%  | 96%  | 98%   | 98% |
|                    | 45      | 96%           | 98%              | 91%            | 96%   | 85%  | 98%   | 98% |
|                    | 46      | 96%           | 90%              | 84%            | 85%   | 77%  | 97%   | 97% |

\*Análises efetuadas sem incluir o grupo clínico.

Tabela 5 – Índices de sensibilidade, especificidade, VPP e VPN para diversos possíveis pontos-de-corte a adotar para a segunda fase do DETECTS.

|                    | Acertos | Sensibilidade | Especificidade * | Especificidade | VPP * | VPP  | VPN * | VPN |
|--------------------|---------|---------------|------------------|----------------|-------|------|-------|-----|
| <b>Simuladores</b> | 40      | 83%           | 100%             | 100%           | 100%  | 100% | 91%   | 92% |
|                    | 41      | 83%           | 100%             | 100%           | 100%  | 100% | 91%   | 92% |
|                    | 42      | 88%           | 100%             | 100%           | 100%  | 100% | 93%   | 94% |
|                    | 43      | 88%           | 100%             | 100%           | 100%  | 100% | 93%   | 94% |
|                    | 44      | 88%           | 100%             | 100%           | 100%  | 100% | 93%   | 94% |
| <b>Não Sim.</b>    | 45      | 92%           | 100%             | 98%            | 100%  | 96%  | 95%   | 96% |
|                    | 46      | 92%           | 100%             | 98%            | 100%  | 96%  | 95%   | 96% |
|                    | 47      | 92%           | 100%             | 98%            | 100%  | 96%  | 98%   | 98% |
|                    | 48      | 92%           | 95%              | 98%            | 92%   | 85%  | 98%   | 98% |

\*Análises efetuadas sem incluir o grupo clínico.

Com base nestes valores, e através da análise das tabelas 4 e 5, optámos por estabelecer o ponto-de-corte em 42 acertos para a primeira fase da prova, e 44 acertos para a segunda fase da prova. Assim, quem apresentar um resultado igual ou inferior a estes critérios em ambos os ensaios será classificado como potencial simulador. A razão de termos escolhido tais critérios, prende-se com o facto de que, tal como referido na bibliografia (ex.: Greve & Bianchini, 2004), um teste de simulação deve manter o índice de especificidade e VPP o mais elevados possível. Assim, optámos por manter estes índices com o valor máximo (100%), evitando, desta forma, falsos diagnósticos positivos que poderão acarretar severas consequências para o avaliado. Não obstante, os pontos-de-corte estabelecidos permitem também manter índices de sensibilidade e VPN elevados. Os critérios adotados para estabelecer os pontos-de-corte, bem como a forma mais adequada de os utilizar em contexto de avaliação serão abordados mais à frente, na discussão dos resultados.

Após a definição destes pontos-de-corte, e a partir de uma base de dados onde o grupo de cada participante foi omissivo, aplicámos os pontos-de-corte estabelecidos a toda a amostra recolhida, reclassificando todos os participantes como simuladores ou não simuladores. Note-se que este procedimento foi extremamente importante para o presente estudo pois permitiu avaliar novamente a eficácia do DETECTS no que diz respeito à deteção da simulação, através da análise do número de diagnósticos corretos e errados que esta prova, e os respetivos critérios estabelecidos, produzem na classificação dos participantes da presente amostra. Através de uma segunda análise qualitativa e individual dos diagnósticos incorretos originados pela prova de simulação, foi também possível avaliar qual a implicação destas classificações erradas na futura utilização da prova em contexto forense e neuropsicológico.

Nesta análise apenas três sujeitos foram classificados incorretamente, tendo os restantes 66 participantes sido classificados corretamente. Um dos participantes foi incorretamente classificado pois, embora pertencesse ao grupo de simuladores, apresentou uma pontuação de 50 acertos (pontuação máxima) em ambas as fases do DETECTS. Uma vez que este participante também apresentou resultados elevados e condizentes com o grupo normativo nas provas de memória aplicadas, considerámos que este pertenceria ao grupo normativo. Dois participantes, também do grupo de simuladores, apresentaram um valor igual ou inferior ao ponto-de-corte anteriormente estabelecido para apenas uma das fases do DETECTS. Mais uma vez, também estes apresentaram resultados elevados e compatíveis com o grupo normativo nas provas de memória aplicadas, pelo que pensámos também que pertenceriam a este grupo. Assim, todos os participantes incorretamente classificados foram sujeitos que, embora pertencessem ao grupo de simuladores, apresentaram resultados elevados em todas as provas aplicadas, ou seja, apresentaram resultados condizentes com os apresentados pelos participantes do grupo normativo em todas as provas aplicadas. Assim, nenhum destes participantes apresentou pontuações semelhantes ao grupo clínico. Consequentemente, concluímos que estes três participantes (simuladores) em contexto forense e/ou neuropsicológico não seriam classificados como detentores de algum tipo de limitação mnésica, mas sim como possuindo um funcionamento mnésico normal, pelo que a sua tentativa de simulação seria igualmente frustrada. Na discussão dos resultados iremos abordar mais pormenorizadamente as implicações práticas destes resultados.

#### **4.4 Outras Variáveis**

Por fim, realizámos diversas análises de forma a compreender qual o efeito de variáveis como a ordem de aplicação dos ensaios do teste de simulação, bem como o sexo e a idade dos participantes no desempenho dos mesmos no teste de simulação (DETECTS). Analisámos ainda a presença de possíveis correlações entre os diversos testes aplicados de forma a avaliar se o DETECTS estaria a discriminar os participantes em função de variáveis como o seu desempenho mnésico, bem como de forma a compreender os padrões de resposta dos diversos grupos de participantes à bateria de provas aplicada.

##### ***4.4.1 Ordem de Aplicação dos Ensaios do Teste de Simulação (DETECTS)***

Através da realização de Testes t para amostras independentes, averiguámos qual o efeito da ordem de aplicação dos ensaios do DETECTS, nos resultados obtidos em ambas as fases desta mesma prova. Para o grupo normativo, há diferenças significativas ao nível do número de acertos na primeira fase do teste de simulação em função da ordem de aplicação dos ensaios ( $t(39) = 2.41, p = .021$ ). Assim, participantes que respondem primeiro ao ensaio um e depois ao ensaio dois obtêm uma maior pontuação na primeira fase da prova de simulação, ou seja, no primeiro ensaio a que respondem, em relação aos participantes que realizam o procedimento contrário ( $\bar{X}_{E1-E2} = 49.29, \bar{X}_{E2-E1} = 48.25$ ).

Para os restantes grupos (grupo clínico e grupo de simuladores), bem como para a segunda fase da prova, não encontramos diferenças ao nível do número de acertos no DETECTS em função da ordem de aplicação dos ensaios. Assim, concluímos que a aplicação do ensaio um, seguido pelo ensaio dois, permite um melhor desempenho dos participantes do grupo normativo no ensaio um.

#### **4.4.2 Provas de Memória (WMS-III) e Teste de Simulação (DETECTS)**

Analisámos também a existência de possíveis associações entre o desempenho dos participantes nas provas de memória e o seu desempenho no DETECTS. Tal como esperado, para o grupo normativo, não há uma associação significativa entre o desempenho dos participantes no teste de localização espacial e o seu desempenho quer na primeira fase do DETECTS ( $r = -.02, p = .910$ ) quer na segunda fase deste mesmo teste ( $r = .001, p = .993$ ). No entanto, existe uma associação significativa entre estas variáveis para o grupo de simuladores, quer para a primeira fase da prova ( $r = .59, p = .002$ ), quer para a segunda fase ( $r = .59, p = .002$ ). Assim, para o grupo de simuladores, um valor mais elevado no teste de localização espacial está associado a um valor mais elevado no DETECTS.

Relativamente ao teste de reconhecimento de faces, para o grupo normativo, não há uma associação significativa entre o desempenho dos participantes no teste de reconhecimento de faces e o seu desempenho, quer na primeira fase da prova de simulação ( $r = .09, p = .558$ ), quer na segunda fase da mesma ( $r = .01, p = .931$ ). Para o grupo de simuladores, há uma associação entre o seu desempenho no teste de reconhecimento de faces e o seu desempenho, quer na primeira fase do teste de simulação ( $r = .60, p = .002$ ), quer na segunda fase desta mesma prova ( $r = .62, p = .001$ ). Assim, para o grupo de simuladores, um valor mais elevado no teste de reconhecimento de faces está também associado a um valor mais elevado em ambas as fases do DETECTS.

Não obstante, no que diz respeito ao teste de reprodução visual, não há associação significativa entre o desempenho dos participantes neste mesmo teste e o seu desempenho na primeira fase do teste de simulação quer para o grupo normativo ( $r = -.08, p = .604$ ), quer para o grupo de simuladores ( $r = .26, p = .226$ ). O mesmo se verificou na segunda fase da prova de simulação, quer para o grupo normativo ( $r = .20, p = .215$ ), quer para o grupo de simuladores ( $r = .32, p = .124$ ).

Para compreender melhor as correlações estatísticas encontradas e conseqüente relação existente entre algumas das provas anteriormente referidas realizámos testes de diferenças (Testes t para amostras independentes) de forma a explorar se existem diferenças ao nível do número de acertos em ambas as fases do DETECTS em função do resultado obtido nas diversas provas da WMS-III. Para tal, utilizando como critério de divisão o percentil 50, dividimos o grupo de simuladores em dois novos grupos: (1) *BLE* - grupo de simuladores com pontuação baixa no teste de localização espacial (50% dos participantes) e (2) *ELE* - grupo de simuladores com pontuação elevada no teste de

localização espacial (50% dos participantes). Verificámos que, para o grupo de simuladores, existem diferenças significativas ao nível do número de acertos no teste de simulação em função do resultado obtido no teste de localização espacial, quer para a primeira fase do teste de simulação ( $t(22) = -30, p = .007$ ), quer para a segunda fase deste mesmo teste ( $t(22) = -2.66, p = .014$ ). Em suma, participantes com um valor elevado no teste de localização espacial, obtêm mais acertos em ambas as fases do DETECTS (*Primeira Fase:  $\bar{X}_{BLE} = 25.5, \bar{X}_{ELE} = 35.8$ ; Segunda Fase:  $\bar{X}_{BLE} = 24.9, \bar{X}_{ELE} = 35.8$* ).

Utilizando um procedimento semelhante ao descrito no parágrafo anterior, dividimos novamente a amostra em dois grupos (BRF – grupo de participantes com pontuação baixa no teste de reprodução visual e ARF – grupo de participantes com pontuação elevada neste mesmo teste). Verificámos que existem também diferenças significativas ao nível do número de acertos no teste de simulação, em função do resultado obtido no teste de reconhecimento de faces, quer para a primeira fase do teste de simulação ( $t(22) = -2.87, p = .009$ ) quer para a segunda fase deste mesmo teste ( $t(22) = -3.31, p = .003$ ). Participantes com um valor elevado no teste de reconhecimento de faces, obtêm mais acertos em ambas as fases do DETECTS (*Primeira Fase:  $\bar{X}_{BRF} = 25.67, \bar{X}_{ARF} = 35.58$ ; Segunda Fase:  $\bar{X}_{BRF} = 24, \bar{X}_{ARF} = 36.67$* ).

Em suma, verificámos que, apenas para o grupo de simuladores, os resultados obtidos pelos participantes nas provas de localização espacial e reconhecimento de faces estão associados aos resultados obtidos no DETECTS (um valor mais elevado nestas provas está associado a um valor mais elevado no teste de simulação), existindo ainda diferenças no número de acertos na prova de simulação em função do resultado obtido nestes dois testes.

#### **4.4.3 Variáveis Sexo e Idade**

No presente estudo, verificámos que não há diferenças entre participantes do sexo feminino e participantes do sexo masculino no que concerne ao seu desempenho em ambas as fases do teste de simulação, independentemente do grupo a que pertencem.

Também no que diz respeito à idade dos participantes, não encontramos associações significativas entre a idade e o desempenho dos participantes em ambas as fases do teste de simulação, com a exceção da associação negativa significativa encontrada entre a idade dos participantes do grupo de simuladores e o seu desempenho na primeira fase do teste de simulação ( $r_{sp} = -.47, p < .001$ ). Assim, para este grupo, maior idade dos participantes está associada a menos acertos na primeira fase do teste. Não obstante, e em suma, observámos que esta variável não está genericamente associada ao desempenho dos participantes no DETECTS.

#### **4.4.4 Provas de Memória (WMS-III)**

Por fim, analisámos também a possível existência de associações entre o desempenho dos participantes nas diversas provas de memória aplicadas, de forma a compreender os padrões de resposta dos diversos grupos de participantes a estes mesmos testes. Para o grupo normativo, não encontrámos nenhum tipo de associação entre os resultados obtidos nos diversos testes. Para o grupo de simuladores, apenas encontrámos uma associação significativa entre o desempenho destes participantes no teste de localização espacial e o seu desempenho no teste de reconhecimento de faces ( $r = .66, p < .001$ ). Assim, um valor mais elevado no teste de localização espacial está associado a um valor mais elevado no teste de reconhecimento de faces. Em suma, encontrámos uma associação entre duas das provas aplicadas, apenas para o grupo de simuladores.

## **V - DISCUSSÃO DOS RESULTADOS**

A simulação de problemas de memória é um fenómeno frequente nas avaliações forenses e neuropsicológicas, podendo-se verificar por diversas razões, como por exemplo, obter uma compensação monetária ou evitar responsabilidade penal. Consequentemente, a avaliação da simulação é um procedimento fundamental neste tipo de avaliações. No entanto, este procedimento é poucas vezes utilizado, ou efetuado de forma pouco rigorosa. O presente estudo procura não só alertar os técnicos para a necessidade de avaliar adequadamente a simulação, mas também explorar o efeito de diversas variáveis na deteção da simulação (ex.: tempo de resposta dos participantes) e, por fim, construir e testar um TVS eficaz, completamente informatizado e com um tempo de aplicação adequado às avaliações forenses e neuropsicológicas.

Tal como esperado, no que concerne à ordem de aplicação do DETECTS na bateria de provas construída, verificámos que esta não possui qualquer tipo de efeito no desempenho dos participantes neste mesmo teste de simulação. Assim, não verificámos a hipótese postulada por Gast e Brown (2010), uma vez que não encontrámos nenhuma vantagem em aplicar o TVS no início da bateria de provas. Torna-se assim indiferente a ordem pela qual os vários testes são aplicados. Estes dados são condizentes com os resultados de Ryan, Glass, Hinds e Brown (2010).

Uma das medidas mais importantes do DETECTS para a deteção da simulação é o número de acertos de cada participante. Independentemente da fase da prova em questão, encontrámos sólidas diferenças entre o grupo de simuladores e os restantes grupos no que diz respeito ao número de acertos no DETECTS, sendo que os simuladores apresentam resultados significativamente mais baixos. Entre o grupo clínico e o grupo normativo, tal como esperado, diferenças ao nível dos acertos no DETECTS

não foram encontradas. Mais uma vez, tais resultados sugerem que a prova construída não discrimina sujeitos com perturbações cognitivas e mnésicas de sujeitos sem este tipo de perturbações. Assim, tal como pretendido, pelas características inerentes à prova construída (ex.: dificuldade aparente elevada e dificuldade real reduzida) ambos os grupos (simuladores e grupo clínico) apresentam altos desempenhos, sendo que apenas o grupo de simuladores apresenta baixos desempenhos. Em suma, este teste diferencia apenas os simuladores dos restantes grupos.

Para o grupo normativo, existe um aumento significativo do número de acertos da primeira, para a segunda fase do DETECTS. Este efeito era já expectável, sendo também vastamente referido na bibliografia (ex.: Tombaugh, 1996). O facto de, na segunda prova de reconhecimento aplicada, os participantes terem visualizado pela segunda vez os estímulos alvo (uma vez em cada fase da prova), bem como terem tido oportunidade de receber feedback acerca das suas respostas, também na primeira fase da prova, permite aos mesmos corrigir os erros que cometeram na primeira fase e, consequentemente, apresentar melhores desempenhos. Não obstante, não encontramos diferenças entre o número de acertos obtidos na primeira e segunda fase do DETECTS, quer para o grupo de simuladores, quer para o grupo clínico. Estes dados eram já previstos para o grupo de simuladores, pois, uma vez que estes participantes procuram obter baixos desempenhos na prova de simulação, tentam-no fazer independentemente do ensaio. O feedback recebido durante as provas de reconhecimento permite ainda a estes participantes controlar o número de respostas erradas em ambos os ensaios, de forma a exibir sempre desempenhos baixos e semelhantes. Assim, este aumento de acertos, da primeira para a segunda fase da prova, acontece apenas para o grupo normativo, mas não para o grupo de simuladores, permitindo uma mais fácil diferenciação entre os participantes destes dois grupos. Para o grupo clínico, seria expectável um aumento no número de acertos da primeira para a segunda fase da prova tal como se verificou para o grupo normativo. O reduzido número da nossa amostra clínica poderá justificar tais resultados, constituindo assim uma limitação do presente estudo.

Os dados recolhidos em relação aos tempos de resposta dos participantes são também bastante encorajadores, sugerindo que a análise desta variável poderá ajudar na deteção da simulação. Todos os grupos de participantes diferem significativamente entre si no que concerne a esta variável, sendo os participantes do grupo normativo os que demoram menos tempo a responder a cada tarefa de reconhecimento da prova, seguidos pelos participantes do grupo de simuladores e, por último, pelos participantes do grupo clínico. Estas diferenças verificam-se independentemente da resposta prestada estar certa ou errada, e podem ser explicadas por diversas razões apontadas na bibliografia. Nomeadamente, vários autores (ex.: Willison & Tombaugh, 2006) apontam que uma das habituais estratégias dos simuladores é responder de forma mais lenta do que habitual, explicando assim as diferenças encontradas entre os tempos de resposta do grupo normativo e os tempos de resposta dos grupos de simuladores. Outros autores (ex.: Haines & Norris, 1995) referem que os simuladores, ao fornecer uma resposta errada, sabendo, no entanto, qual é a resposta correta, necessitarão de mais

tempo para efetuar a resposta do que os restantes sujeitos. Ambos os argumentos poderão explicar estes resultados.

Era igualmente esperado que as consequências de algumas perturbações cognitivas, apresentadas pelos indivíduos da nossa amostra clínica, dificultassem consideravelmente a execução de diversas operações mentais, aumentando assim o tempo que os participantes do grupo clínico necessitariam para efetuar as suas respostas, comparativamente aos participantes do grupo normativo. No entanto, como referimos anteriormente, era imprevisível qual seria o grupo de participantes a exibir maiores tempos de resposta: o grupo de simuladores ou o grupo clínico. No presente estudo, o grupo clínico apresentou maiores tempos de resposta, em comparação aos restantes grupos, sugerindo assim que, mesmo após a tentativa deliberada do grupo de simuladores em responder de forma mais lenta aos testes de reconhecimento, os pacientes com verdadeiras perturbações ao nível do funcionamento cognitivo e mnésico acabam por demorar ainda mais tempo a efetuar as suas respostas devido ao diagnóstico que possuem.

Averiguámos ainda que todos os grupos de participantes respondem mais rapidamente à segunda fase do DETECTS, em comparação à primeira. Este resultado era expectável pois, na segunda fase da prova, os participantes tiveram a oportunidade de visualizar duas vezes as imagens alvo (uma vez em cada fase da prova), mantendo assim um maior grau de confiança nas escolhas efetuadas e precisando, conseqüentemente, de menos tempo para excluir a resposta incorreta e executar a resposta correta ou a resposta desejada, no caso dos simuladores. Também devido ao treino no que concerne à execução das próprias respostas, adquirido na primeira fase da prova, os diferentes grupos de participantes terão uma maior facilidade em executar mais rapidamente as suas respostas na segunda fase da prova (ex.: os participantes precisarão de menos tempo para localizar as teclas correspondentes a cada imagem).

Em suma, os resultados recolhidos relativos aos tempos de respostas dos diversos participantes, sugerem que, no contexto clínico, neuropsicológico e forense, particularmente caso permaneçam dúvidas quanto ao diagnóstico de determinado avaliado após a observação e contagem do número de acertos deste na prova de simulação, o avaliador poderá recorrer à média e ao desvio padrão dos tempos de resposta de cada grupo de participantes, previamente apresentados, como forma de esclarecer as suas dúvidas e situar o avaliado num dos grupos anteriormente mencionados. Assim, e uma vez que todos os grupos de participantes diferem entre si no que concerne a esta variável, será possível ao avaliador classificar determinado avaliado em função do seu tempo de resposta. Não obstante, aconselhamos que este procedimento seja utilizado como complementar à análise dos acertos de cada participante, e não como elemento único ou principal de classificação.

Em relação aos pontos-de-corte definidos para a prova de simulação construída (DETECTS), optámos por estabelecer o ponto-de-corte em 42 acertos para a primeira fase da prova, e 44 acertos

para a segunda fase da prova pois, tal como apontado na bibliografia (ex.: Greve & Bianchini, 2004) o ponto-de-corte de um teste de simulação deve manter o índice de especificidade (proporção de não simuladores corretamente classificados) e VPP (Valor Preditivo Positivo - probabilidade de um avaliado ser simulador quando é diagnosticado como tal) o mais alto possível. Assim, optámos por manter tais índices com um valor de 100% (valor máximo), uma vez que um valor inferior poderia originar falsos diagnósticos positivos, os quais, tal como abordado anteriormente, acarretam severas consequências para o avaliado. Para além de certificarmos-nos que os pontos-de-corte estabelecidos permitiam manter um índice de especificidade e VPP a 100%, evitando assim falsos diagnósticos positivos, procurámos ainda que estes possibilitassem manter o índice de sensibilidade (proporção de simuladores corretamente identificados) e VPN (Valor Preditivo Negativo - probabilidade de um avaliado não ser simulador quando assim identificado) o mais elevados possível de forma a minimizar também o número de diagnósticos negativos incorretos.

Não obstante, o avaliador poderá optar por utilizar outros pontos-de-corte. Com base na tabela 4 e 5, o próprio técnico poderá analisar a especificidade, sensibilidade, VPP e VPN do DETECTS, consoante os diferentes pontos-de-corte adotados, realizando assim uma avaliação semelhante à que por nós foi efetuada e considerando quais as vantagens e desvantagens inerentes a cada ponto-de-corte, podendo assim selecionar os critérios que considera mais adequados para a sua avaliação. Da mesma forma, embora no presente estudo aconselhemos a que o avaliador apenas considere potencial simulador o avaliado que apresenta um valor igual ou inferior ao respetivo ponto-de-corte em ambas as fases do DETECTS, o avaliador poderá optar, com base na informação que recolheu durante a sua avaliação, considerar como potencial simulador o avaliado que apresentou um resultado igual ou inferior ao respetivo ponto-de-corte em apenas um dos ensaios da prova. Mais uma vez, optámos por estabelecer tal critério de forma a evitar diagnósticos positivos errados. Em caso de dúvida, no que concerne ao diagnóstico de determinado avaliado, a avaliação multimodal torna-se particularmente importante. Também nestas situações, a análise dos tempos de resposta dos avaliados poderá ser particularmente relevante. Mais uma vez, realçamos que os pontos-de-corte por nós estabelecidos não tornam dispensável uma avaliação qualitativa e multimodal do avaliador, como por exemplo, a cotação e interpretação das restantes provas aplicadas, cuja importância iremos abordar de seguida.

Como já referimos, após o estabelecimento e definição dos pontos-de-corte, reclassificámos todos os participantes como simuladores ou não simuladores em função dos seus resultados na bateria de provas. Este procedimento levou a que apenas três participantes, do grupo de simuladores, fossem classificados “incorretamente” como não simuladores. No entanto, através da cotação e análise dos testes de memória aplicados, verificámos que estes três participantes apresentaram resultados condizentes com a população normativa também nestas provas. Assim, em contexto clínico, neuropsicológico e/ou forense estes sujeitos seriam classificados como não tendo qualquer tipo de perturbação mnésica, pelo que a sua tentativa de simulação seria, de qualquer forma, infrutífera. Uma

vez que, no presente estudo, nenhum participante do grupo clínico ou normativo foi classificado como simulador, e nenhum dos participantes do grupo de simuladores foi classificado como pertencendo ao grupo clínico, concluímos que, no que diz respeito ao objetivo mais prático desta prova de impedir que os simuladores sejam classificados como detentores de verdadeiros défices e os não simuladores de ser incorretamente identificados como simuladores, a eficácia do DETECTS, quando aplicado em conjunto com a bateria de testes proposta e no que diz respeito à amostra utilizada, foi de 100%. Concluímos ainda que os três participantes incorretamente reclassificados na análise anteriormente descrita, terão apresentado tais desempenhos atípicos e semelhantes à população normativa, pois, sabendo que uma pontuação demasiado baixa poderia chamar a atenção do investigador para a sua tentativa de simulação (esta informação foi incluída nos diapositivos apresentados a todos os simuladores) e sendo incapazes de distinguir a prova de simulação das restantes provas, optaram por utilizar um estilo de simulação demasiado discreto em todas as provas, de tal forma que acabaram por exibir resultados elevados, semelhantes aos apresentados pelo grupo normativo.

Relativamente à ordem de aplicação dos ensaios do DETECTS, verificámos que, para o grupo normativo, os participantes que respondem primeiro ao ensaio um e depois ao ensaio dois obtêm uma maior pontuação na primeira fase da prova, ou seja, no primeiro ensaio a que respondem, comparativamente aos participantes que realizam o procedimento contrário. Uma vez que os estímulos alvo apresentados em ambos os ensaios são exatamente iguais, este efeito poderá apenas dever-se aos estímulos distrativos que, embora sejam de categorias idênticas, diferem entre cada ensaio. Assim, concluímos que aplicar o ensaio um, seguido pelo ensaio dois, será o procedimento mais vantajoso pois permite um melhor desempenho apenas dos participantes do grupo normativo, permitindo assim uma melhor diferenciação entre simuladores e não simuladores

No que diz respeito à relação entre os desempenhos obtidos pelos participantes nos testes de memória aplicados e os desempenhos obtidos pelos mesmos no DETECTS, verificámos que, para o grupo normativo, o desempenho destes participantes nas diversas provas de memória aplicadas nunca está associado ao seu desempenho no teste de simulação. Estes resultados significam que o DETECTS não discrimina os avaliados em função de nenhum dos diferentes tipos de memória avaliados pelas restantes provas. Para o grupo de simuladores, encontramos uma associação entre o desempenho dos participantes na prova de localização espacial e reconhecimento de faces, com o seu desempenho no DETECTS. Estes resultados são também encorajadores pois sugerem que os simuladores não identificam a prova de simulação como sendo diferente das restantes provas aplicadas, utilizando um estilo de simulação semelhante (mais extremo ou mais reservado) para todas as provas, e garantindo assim a eficácia do TVS. Estes resultados foram ainda corroborados pela utilização de testes de diferenças, tal como foi abordado anteriormente.

Como prevíamos, o DETECTS não discrimina os avaliados em função do seu sexo, não existindo associação entre o sexo dos participantes e os resultados obtidos nesta mesma prova. De igual forma, a prova de simulação não discrimina genericamente os participantes em função da sua idade, não tendo sido encontrada uma associação entre a idade dos participantes e o seu desempenho nesta prova, exceto para o grupo de simuladores e em apenas uma das fases da prova. Concluimos assim que estas variáveis não produzem um efeito considerável nos resultados apresentados pelos diversos grupos de participantes no DETECTS, podendo esta prova ser utilizada independentemente da idade ou do sexo dos avaliados. Não obstante, uma vez que a amostra do presente estudo é relativamente homogénea no que concerne à idade dos participantes, incluindo apenas sujeitos dos 18 aos 46 anos ( $\bar{X} = 22.22$ , D.P. = 5.90), estudos futuros devem comprovar a eficácia desta prova em participantes com idades mais extremadas. Por fim, a associação entre a idade e o desempenho dos participantes encontrada apenas para o grupo de simuladores numa das fases do teste de simulação, não parece ser explicável pela bibliografia existente. Caso tal associação se verificasse para o grupo normativo, poderíamos estar perante uma situação problemática pois levantar-se-ia a hipótese de que, possivelmente, os participantes mais novos poderiam estar a obter um melhor desempenho devido à sua maior capacidade mnésica, o que, conseqüentemente, significaria que a nossa prova, ao contrário do esperado, estaria a discriminar participantes em função da sua capacidade mnésica. No entanto, uma vez que esta associação apenas se verifica para o grupo de simuladores, é do nosso parecer que se deve a características internas da nossa amostra, sendo improvável que tal associação se verificasse numa replicação do presente estudo. Note-se ainda que, independentemente da idade, os simuladores obtiveram desempenhos significativamente mais baixos do que os restantes participantes, garantindo assim a eficácia desta prova e tornando a associação previamente mencionada de pouco relevo para o presente estudo.

Relativamente às provas de memória aplicadas, para o grupo normativo, não encontramos nenhuma associação entre os resultados obtidos pelos participantes nas diversas provas. O facto de estes testes avaliarem diferentes tipos de memória (ex.: memória visual e memória de trabalho) e utilizarem estímulos diferentes (ex.: faces e figuras geométricas), poderá explicar a ausência de associações. Para o grupo de simuladores, verificámos que o desempenho dos participantes no teste de localização espacial e o seu desempenho no teste de reconhecimento de faces está associado. Este resultado é compreensível pois os simuladores, sendo incapazes de detetar qual o teste de simulação, tendem a adotar estilos de simulação semelhantes (mais extremos ou mais reservados) para todas as provas. Assim, indivíduos que procuraram exibir um baixo desempenho na prova de localização espacial, procuram fazê-lo também na prova de reconhecimento de faces. No entanto, para este mesmo grupo, não foram encontradas associações entre os resultados obtidos no teste de reprodução visual e os resultados obtidos nas restantes provas aplicadas. Uma vez que a prova de reprodução visual requer um tipo de resposta diferente do que é habitualmente pedido aos avaliados em outras tarefas de

memória (tarefa de reconstrução visuográfica dos estímulos), os simuladores poderão estar a exibir dificuldade em controlar o seu desempenho nessa mesma prova, explicando assim a ausência de correlações entre este e os restantes testes aplicados.

Em suma, a análise dos resultados obtidos ao longo deste estudo permitiu-nos compreender que o teste de simulação construído (DETECTS) possui uma elevada eficácia na deteção da simulação, particularmente quando aplicado em conjunto com a bateria de provas sugerida. Através da análise do número de acertos dos participantes no DETECTS, dos tempos de resposta que estes exibem, e da cotação dos testes de memória aplicados, é possível classificar determinado avaliado como potencial simulador, ou não simulador, com um elevado grau de confiança. Por fim, o presente estudo mostrou também que variáveis como o sexo, a idade, ou a capacidade mnésica dos avaliados não influenciam o seu desempenho no DETECTS, garantindo assim a eficácia do teste.

## **VI - CONCLUSÃO**

### **6.1 Limitações do Estudo e Sugestões para Estudos Futuros**

A principal limitação do presente estudo prende-se com o número reduzido da amostra clínica, constituída por apenas quatro participantes. Este número reduzido da amostra clínica impossibilitou-nos de explorar qual o desempenho que avaliados com outro tipo de perturbações, não incluídos neste mesmo grupo, poderão exibir na prova construída. Desta forma, a capacidade de generalização dos resultados referentes ao grupo clínico, face à população clínica geral, é reduzida. Adicionalmente, note-se que três participantes deste grupo foram excluídos do estudo devido a possuírem um diagnóstico de Doença de Alzheimer em estado avançado e, conseqüentemente, serem incapazes de compreender e reter as instruções do DETECTS. Assim, em contexto clínico, neuropsicológico e/ou forense, este tipo de pacientes terão de ser avaliados através de outras metodologias uma vez que esta prova de simulação é inadequada para estes casos. Note-se ainda que o grupo clínico, ao contrário dos restantes grupos, não era composto por estudantes universitários. Assim, os participantes integrantes deste grupo possuíam uma idade consideravelmente superior, bem como um nível de escolaridade e familiaridade com o computador menor, comparativamente com o grupo normativo e o grupo de simuladores. Estas variáveis poderão ter contribuído para as diferenças encontradas entre o grupo clínico e os restantes grupos, nomeadamente no que diz respeito aos tempos de resposta dos participantes. No entanto, a bibliografia existente (ex.: Gast & Hart, 2010; Greve et al., 2006; Simon, 2007), bem como os resultados do presente estudo, sugerem que tais variáveis não têm um efeito significativo no desempenho dos participantes no teste de simulação. Adicionalmente, sendo o nosso estudo maioritariamente baseado numa metodologia tipicamente utilizada nas experiências de

simulação, a qual já explorámos anteriormente, não era o nosso objetivo incluir um grupo clínico numeroso. Contudo, a realização de estudos futuros, baseados noutra tipo de metodologias que incluam a aplicação da bateria de provas construída a um grupo clínico amplo, com uma idade e nível de escolaridade semelhante aos restantes grupos (grupo de simuladores e grupo normativo), é essencial para melhor testar a eficácia da prova de simulação.

Neste estudo, o grupo de simuladores foi constituído por estudantes universitários, e não por suspeitos de simulação. Esta metodologia confere ao estudo diversas vantagens, como por exemplo, permite-nos ter uma maior certeza de que o grupo de simuladores estava, de facto, a simular. Não obstante, possui também algumas limitações, nomeadamente a diminuta motivação que os estudantes universitários poderão apresentar face aos verdadeiros simuladores, uma vez que os incentivos atribuídos são qualitativamente diferentes. Não obstante, tal como referido no capítulo da introdução, as possíveis estratégias de simulação mais complexas utilizadas pelos estudantes universitários, podem, de certa forma, minimizar este problema. Fatores como a inexistência de um evento traumático vivido pelo tipo de participantes utilizado neste estudo, bem como a ausência do risco dos participantes em serem detetados como simuladores e sofrerem as respetivas consequências, vividas em contexto real, poderão também enviesar os resultados obtidos neste estudo, em comparação com os resultados que seriam obtidos numa investigação em contexto real. Assim, futuros estudos com verdadeiros suspeitos de simulação constituirão também um passo importante para melhor testar a eficácia da prova construída (Bianchini et al., 2001).

## **6.2 Implicações para o Contexto Forense**

Não desconsiderando as limitações anteriormente apontadas, o presente estudo permitiu construir e testar uma nova e promissora prova de simulação (DETECTS – Teste de Detecção da Simulação de Problemas de Memória) eficaz na deteção da simulação de problemas de memória. Esta prova demonstrou também trazer inúmeras vantagens para a prática clínica, neuropsicológica e, particularmente, para a prática forense. Assim, a prova construída revela ser extremamente útil na deteção da simulação de problemas mnésicos uma vez que: (1) quando inserida na bateria de provas sugerida, permitiu distinguir eficazmente todos os simuladores do presente estudo dos pacientes com perturbações cognitivas e mnésicas sem gerar falsos diagnósticos positivos; (2) não foi identificada pelos avaliados como uma prova de simulação; (3) possui um tempo de aplicação reduzido e adequado às avaliações forenses; e (4) é aplicada através do computador permitindo uma utilização fácil e rigorosa, bem como o registo exato do tempo de resposta dos participantes, variável esta que demonstrou ser também extraordinariamente importante na deteção da simulação de problemas de memória.

Em suma, considerando a vulnerabilidade dos testes neuropsicológicos face à simulação de sintomas, a elevada prevalência da simulação de problemas de memória em contexto forense e ainda

as vantagens que a prova de simulação construída demonstrou apresentar face a outras metodologias e outros TVS existentes, consideramos aconselhável a utilização do DETECTS sempre que existir a suspeita que determinado avaliado possa estar a simular problemas de memória.

## VII – REFERÊNCIAS

- American Psychiatric Association (2002). *DSM-IV-TR: Manual de diagnóstico e estatística das perturbações mentais* (4ª ed., texto revisto). Lisboa: Climepsi Editores.
- Armistead-Jehle, P., & Gervais, R. O. (2011). Sensitivity of the test of memory malingering and the nonverbal medical symptom validity test: a replication study. *Applied Neuropsychology*, *18*(4), 284-290.
- Batt, K., Shores, E. A., & Chekaluk, E. (2008). The effect of distraction on the word memory test and test of memory malingering performance in patients with a severe brain injury. *Journal of the International Neuropsychological Society: JINS*, *14*(6), 1074-1080.
- Bauer, L., O'Bryant, S. E., Lynch, J. K., McCaffrey, R. J., & Fisher, J. M. (2007). Examining the test of memory malingering trial 1 and word memory test immediate recognition as screening tools for insufficient effort. *Assessment*, *14*(3), 215-222.
- Bianchini, K., Mathias, C., & Greve, K. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist*, *15*, 19-45.
- Binder, L. M. (1992). Deception and Malingering. In A. Puente & R. McCaffrey (Eds.), *Handbook of neuropsychological assessment: A biopsychosocial approach* (pp. 353-374). New York: Plenum.
- Blaskewitz, N., Merten, T., & Kathmann, N. (2008). Performance of children on symptom validity tests: TOMM, MSVT, and FIT. *Archives of Clinical Neuropsychology*, *23*(4), 379-391.
- Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *The Clinical Neuropsychologist*, *23*(4), 729-741.
- Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., Reynolds, C. R., & Silver, C. H. (2005). Symptom validity assessment: practice issues and medical necessity NAN policy & planning committee. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *20*(4), 419-426.
- Chafetz, M. (2011). Reducing the probability of false positives in malingering detection of social security disability claimants. *The Clinical Neuropsychologist*, *25*(7), 1239-1252.
- Donders, J. (2005). Performance on the test of memory malingering in a mixed pediatric sample. *Child Neuropsychology*, *11*(2), 221-227.

- Duncan, A. (2005). The impact of cognitive and psychiatric impairment of psychotic disorders on the test of memory malingering (TOMM). *Assessment, 12*(2), 123-129.
- 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: The Official Journal of the National Academy of Neuropsychologists, 20*(3), 375-384.
- Farkas, M. R., Rosenfeld, B., Robbins, R., & Gorp, W. (2006). Do tests of malingering concur? Concordance among malingering measures. *Behavioral Sciences & the Law, 24*(5), 659-671.
- Field, A. P. (2009). *Discovering statistics using SPSS* (3<sup>a</sup> ed). SAGE Publications Ltd.
- Gast, J., & Hart, K. (2010). The performance of juvenile offenders on the test of memory malingering. *Journal of Forensic Psychology Practice, 10*(1), 53-68.
- Gavett, B. E., O'Bryant, S. E., Fisher, J. M., & McCaffrey, R. J. (2005). Hit rates of adequate performance based on the test of memory malingering (TOMM) trial 1. *Applied Neuropsychology, 12*(1), 1-4.
- Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head-injury disability claimants. *Archives of Clinical Neuropsychology, 19*, 475-487.
- Green, P. (2011). Comparison between the test of memory malingering (TOMM) and the nonverbal medical symptom validity test (NV-MSVT) in adults with disability claims. *Applied Neuropsychology, 18*(1), 18-26.
- Green, P., Rohling, M., Lees-Haley, P., & Allen, L. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15*(12), 1045-1060.
- Greiffenstein, M. F., Greve, K. W., Bianchini, K. J., & Baker, W. J. (2008). Test of memory malingering and word memory test: a new comparison of failure concordance rates. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 23*(7-8), 801-807.
- Greve, K. W., & Bianchini, K. J. (2004). Setting empirical cut-offs on psychometric indicators of negative response bias: a methodological commentary with recommendations. *Archives of Clinical Neuropsychology, 19*, 533-541.
- Greve, K. W., & Bianchini, K. J. (2006). Should the retention trial of the test of memory malingering be optional? *Archives of Clinical Neuropsychology, 21*, 117-119.
- Greve, K. W., Bianchini, K. J., Black, F. W., Heinly, M. T., Love, J. M., Swift, D. A., & Ciota, M. (2006). Classification accuracy of the test of memory malingering in persons reporting exposure to environmental and industrial toxins: Results of a known-groups analysis. *Archives of Clinical Neuropsychology, 21*(5), 439-448.

- Greve, K. W., Bianchini, K. J., & Doane, B. M. (2006). Classification accuracy of the test of memory malingering in traumatic brain injury: Results of a known-groups analysis. *Journal of Clinical and Experimental Neuropsychology*, 28(7), 1176-1190.
- Greve, K. W., Binder, L. M., & Bianchini, K. J. (2009). Rates of below-chance performance in forced-choice symptom validity tests. *The Clinical Neuropsychologist*, 23(3), 534-544.
- Greve, K. W., Etherton, J. L., Ord, J., Bianchini, K. J., & Curtis, K. L. (2009). Detecting malingered pain-related disability: Classification accuracy of the test of memory malingering. *The Clinical Neuropsychologist*, 23(7), 1250-1271.
- Gunn, D., Batchelor, J., & Jones, M. (2010). Detection of simulated memory impairment in 6- to 11-year-old children. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 16(2), 105-118.
- 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(3), 524-532.
- Haines, M. E., & Norris, M. P. (1995). Detecting the malingering of cognitive deficits: An update. *Neuropsychology Review*, 5, 125-148.
- Hilsabeck, R. C., Gordon, S. N., Hietpas-Wilson, T., & Zartman, A. L. (2011). Use of trial 1 of the test of memory malingering (TOMM) as a screening measure of effort: Suggested discontinuation rules. *The Clinical Neuropsychologist*, 25(7), 1228-1238.
- Horner, M. D., Bedwell, J. S., & Duong, A. (2006). Abbreviated form of the test of memory malingering. *The International Journal of Neuroscience*, 116(10), 1181-1186.
- Iverson, G. L. (2003). Detecting malingering in civil forensic evaluations. In A. M. Horton Jr. & L. C. Hartlage (Eds.), *Handbook of Forensic Neuropsychology* (pp. 137-177). New York: Springer.
- Iverson, G. L. (2007). Identifying exaggeration and malingering. *Pain Practice*, 7(2), 94-102.
- Iverson, G. L., Le 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(3), 532-546.
- Jelicic, M., Ceunen, E., Peters, M. J. V., & Merckelbach, H. (2011). Detecting coached feigning using the test of memory malingering (TOMM) and the structured inventory of malingered symptomatology (SIMS). *Journal of Clinical Psychology*, 67(9), 850-855.
- Kirk, J. W., Harris, B., Hutaff-Lee, C. F., Koelemay, S. W., Dinkins, J. P., & Kirkwood, M. W. (2011). Performance on the test of memory malingering (TOMM) among a large clinic-referred pediatric sample. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 17(3), 242-254.

- Lange, R. T., Pancholi, S., Bhagwat, A., Anderson-Barnes, V., & French, L. M. (2012). Influence of poor effort on neuropsychological test performance in U.S. military personnel following mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *34*(5), 453–466.
- MacAllister, W. S., Nakhutina, L., Bender, H. A., Karantzoulis, S., & Carlson, C. (2009). Assessing effort during neuropsychological evaluation with the TOMM in children and adolescents with epilepsy. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *15*(6), 521-531.
- Martins, C. (2011). *Manual de análise de dados quantitativos com recurso ao IBM SPSS: Saber decidir, fazer, interpretar e redigir* (1ª ed., Vol. 1). Braga: Psiquilímbros Edições.
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *18*, 265-275.
- O’Bryant, S. E., Engel, L. R., 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*(3), 511-521.
- O’Bryant, S. E., Finlay, C. G., & O’Jile, J. R. (2006). TOMM performances and self-reported symptoms of depression and anxiety. *Journal of Psychopathology and Behavioral Assessment*, *29*(2), 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*(3), 533-540.
- Oorsouw, K., & Merckelbach, H. (2010). Detecting malingered memory problems in the civil and criminal arena. *Legal and Criminological Psychology*, *15*(1), 97-114.
- Porter, S., & Woodworth, M. (2006). «I’m sorry I did it ... but he started it»: A comparison of the official and self-reported homicide descriptions of psychopaths and non-psychopaths. *Law and Human Behavior*, *31*, 91-107.
- Powell, M. R., Gfeller, J. D., Hendricks, B. L., & Sharland, M. (2004). Detecting symptom- and test-coached simulators with the test of memory malingering. *Archives of Clinical Neuropsychology*, *19*(5), 693-702.
- Rogers, R. (2008). *Clinical assessment of malingering and deception* (pp. 14-34). New York: Guilford Press.
- Ryan, J. J., Glass, L. A., Hinds, R. M., & Brown, C. N. (2010). Administration order effects on the test of memory malingering. *Applied Neuropsychology*, *17*(4), 246-250.

- 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 & M. S. Pinho, (Eds.), *Psicologia Forense* (pp. 280-309). Coimbra: Almedina.
- Simões, M. R., & Sousa, L. B. (2011). Traumatismos crânio-encefálicos no âmbito do direito de trabalho: Avaliação (neuro)psicológica e elaboração de relatório para tribunal. In M. Matos, R. Abrunhosa & C. Machado, (Eds.), *Manual de psicologia forense: Contextos práticas e desafios* (pp. 351-374). Braga: Psiquilibrios Edições.
- Simon, M. J. (2007). Performance of mentally retarded forensic patients on the test of memory malingering. *Journal of Clinical Psychology*, 63(4), 339-344.
- Slick, D. J., Tan, J. E., Strauss, E., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, 19, 465-473.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, 6, 174-215.
- Superlab (4.5) [programa informático]. San Pedro, Cedrus Corporation.
- 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(4), 495-505.
- Tombaugh, T. M. (1996). *Test of Memory Malingering (TOMM)*. Canada: Multi-Health Systems.
- Vanderslice-Barr, J. L., Miele, A. S., Jardin, B., & McCaffrey, R. J. (2011). Comparison of computerized versus booklet versions of the TOMM™. *Applied Neuropsychology*, 18(1), 34-36.
- Wechsler, D. (1997). *WMS-III: Wechsler memory scale administration and scoring manual* (3rd. ed.). Psychological Corp.
- Weinborn, M., Woods, S. P., Nulsen, C., & Leighton, A. (2012). The effects of coaching on the verbal and nonverbal medical symptom validity tests. *The Clinical Neuropsychologist*, 26(5), 832-849.
- Whiteside, D. M., Dunbar-Mayer, P., & Waters, D. P. (2009). Relationship between Tomm performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist*, 23(3), 523-533.
- Willison, J., & Tombaugh, T. N. (2006). Detecting simulation of attention deficits using reaction time tests. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 21(1), 41-52.
- 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: The Official Journal of the National Academy of Neuropsychologists*, 21(2), 161-165.
- Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology*, 14(6), 511 -515.