

Detecção de *esforço reduzido* com os subtestes Memória de Dígitos e Vocabulário (WAIS-III) numa amostra com Traumatismo Crânio Encefálico examinada em contexto médico-legal

Sílvia Pedrosa¹, Mário R. Simões¹, Margarida B. Barreto², Mónica Fonseca¹,
António Mestre², Liliana B. Sousa¹, Isabel Cruz², Helena Gaspar²,
Graça Costa², Francisco Corte-Real²

⁽¹⁾ - Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra. Contactos relativos a este trabalho: Sílvia Pedrosa (smarinapedrosa@gmail.com); Mário R. Simões (simoesmr@fpce.uc.pt).

⁽²⁾ - Serviço de Clínica Forense, Delegação do Centro do Instituto Nacional de Medicina Legal.
Os autores agradecem a todas as pessoas que participaram neste estudo.

Detecção de esforço reduzido com os subtestes Memória de Dígitos e Vocabulário (WAIS-III) numa amostra com Traumatismo Crânio Encefálico examinada em contexto médico-legal

Resumo:

A utilidade dos indicadores de *esforço reduzido* provenientes de testes psicológicos comumente utilizados, como é por exemplo o caso dos instrumentos de exame da inteligência, tem sido reconhecida no âmbito da avaliação neuropsicológica forense. Neste contexto, e com o objectivo de analisar a relevância dos subtestes Memória de Dígitos e Vocabulário da WAIS-III, foi realizado um estudo comparativo dos desempenhos de um grupo de 40 litigantes com Traumatismo Crânio Encefálico examinados em contexto médico-legal (grupo experimental) e um grupo de 30 sujeitos da população geral (grupo controlo), equiparados quanto às variáveis sócio-demográficas.

Os resultados obtidos sugerem a necessidade de alguma prudência na utilização agregada dos indicadores de *esforço reduzido* derivados dos subtestes da WAIS-III, estando a sua validade e utilidade dependentes do recurso adicional a provas específicas de *enviesamento de resposta negativa* (*Test of Memory Malingering*, *Rey 15-Item Test* e escalas LI, AM e NI do *Structured Inventory of Malingered Symptomatology*). Em sujeitos com TCE, é indicada a potencial utilidade de um ponto de corte inferior a 7 na "Pontuação total bruta da Memória de Dígitos" como referência associada à presença de eventuais comportamentos de "esforço reduzido".

Palavras-chave: Esforço reduzido, Traumatismo Crânio Encefálico (TCE), Escala de Inteligência de Wechsler para Adultos – Terceira Edição (WAIS-III), Memória de Dígitos, Vocabulário, *Test of Memory Malingering* (TOMM), *Rey 15-Item Test*, *Structured Inventory of Malingered Symptomatology* (SIMS).

DETECTING INCOMPLETE EFFORT WITH DIGIT SPAN AND VOCABULARY SUBTESTES FROM WECHSLER ADULT INTELLIGENCE SCALE - THIRD EDITION (WAIS-III) WITH TRAUMATIC BRAIN INJURY SAMPLE IN MEDICO-LEGAL CONTEXT

Abstract:

The usefulness of reduced effort indicators derived from known and recurrently used psychological measures, like intelligence tests, has been emphasized in forensic neuropsychological assessment. With the purpose of analyze the WAIS-III Digit Span and Vocabulary subtests usefulness in this context, it was realized a comparative research analysis of the performance of 40 litigants with Traumatic Brain Injury examined in legal context (experimental group) and 30 subjects from general population (control group).

The results suggest the need of some caution in the aggregated use of the WAIS-III-based indicators of effort. It's validity and usefulness is essentially associated to additional measures specific of *negative response bias* (*Test of Memory Malingering*, *Rey 15-Item Test* and *LI*, *AM* and *NI Inventory of Malingered Symptomatology* scales) use. In TBI, the potential utility of a 7 cut-off score point to "Digit Span raw scores" is proposed as a reference to the presence of poor effort behaviours.

Keywords: Reduced effort, *Wechsler Adult Intelligence Scale-3rd edition* (WAIS-III), Digit Span, Vocabulary, *Test of Memory Malingering* (TOMM), *Rey 15-Item Test*, *Structured Inventory of Malingered Symptomatology* (SIMS)

Introdução

Em avaliação neuropsicológica forense, a validade dos desempenhos do sujeito constitui uma questão relevante, particularmente se atendermos a circunstâncias como a eventual presença de incentivos externos para a manifestação de incapacidade. Exemplo destas situações são os cada vez mais frequentes casos de queixas cognitivas em vítimas de Traumatismo Crânio Encefálico (TCE) decorrente de acidente de viação ou de trabalho. Neste contexto, os pacientes podem, de forma intencional, apresentar desempenhos baixos nos testes, ou seja, manifestar *esforço reduzido* de forma a exagerar os seus eventuais défices cognitivos (e/ou perturbações emocionais) e, assim, aumentar a recompensa pelos danos sofridos.

O termo *esforço reduzido*, intimamente associado a uma *performance* “incompleta” ou “insuficiente”, diz respeito aos casos ou circunstâncias em que um examinado não apresenta um desempenho correspondente ao melhor nível das suas capacidades. Ou seja, ao contrário do que é explicitamente solicitado na resposta aos testes, o sujeito não “faz o seu melhor”. O conceito de “*esforço reduzido*” é igualmente utilizado para caracterizar casos de sujeitos, que formulam pedidos de compensação, e que apresentam padrões de resultados substancialmente mais baixos nos testes do que os obtidos por indivíduos com perturbações neurológicas verdadeiras. O *esforço reduzido* engloba, assim, todos os desempenhos inferiores ao esperado no processo de avaliação psicológica, independentemente das motivações conscientes ou inconscientes (Strauss, Sherman, & Spreen, 2006; Rogers & Payne, 2006).

Em suma, sempre que estejam presentes incentivos de natureza externa (compensações de natureza financeira, possibilidade de desresponsabilização ou atenuação das consequências de comportamento criminal), o exame dos comportamentos de *enviesamento de resposta negativo e/ou esforço reduzido* deve constituir um ponto essencial da estratégia de avaliação psicológica (Bush *et al.*, 2005), pelo que o psicólogo deve considerar o eventual papel dos incentivos financeiros nos sintomas comunicados e nos desempenhos de incapacidade apresentados (Moss *et al.*, 2003).

Avaliação dos comportamentos de esforço reduzido

As estratégias para detecção de comportamentos de *simulação* e *esforço reduzido* têm sido objecto de consideráveis aperfeiçoamentos ao nível do seu desenvolvimento e validação. Entre os procedimentos de avaliação passíveis de ser utilizados com o objectivo de contribuir para a avaliação e diagnóstico de *esforço reduzido* encontram-se as entrevistas clínicas, a observação do comportamento, a recolha de dados de processos e, incontornavelmente, os testes psicológicos. Neste contexto, a utilização de testes psicológicos admite duas abordagens complementares. A primeira está associada à aplicação de testes construídos especificamente para a avaliação de *simulação e/ou esforço reduzido* (e.g. Testes de Validade de Sintomas); a segunda reporta-se à dupla utilidade dos resultados nos testes mais clássicos ou tradicionais como medidas reconhecidas de capacidade cognitiva e, simultaneamente, como potenciais indicadores de *esforço reduzido*. Nenhum método ou teste psicológico utilizado é suficiente, por si só, para o diagnóstico de *simulação* ou *esforço reduzido*, pelo que a convergência de

dados provenientes de múltiplos métodos e de diferentes testes é essencial para a sua determinação (Farkas *et al.*, 2006; Larrabee, 2007; Rogers, 1997). A dificuldade e complexidade na formulação deste tipo de diagnósticos e os constrangimentos temporais para a realização das avaliações, reforçam a importância da pesquisa orientada para o exame da utilidade de indicadores de *esforço reduzido* provenientes de provas neuropsicológicas conhecidas e frequentemente utilizadas neste âmbito, como é o caso da última versão da Escala de Inteligência de Wechsler para Adultos (WAIS-III, Wechsler, 1997, 2008), recentemente adaptada e aferida para a população portuguesa.

Os indicadores de validade interna oriundos de testes psicológicos clássicos (testes de inteligência, raciocínio, memória, atenção, etc.) têm suscitado investigação e interesse clínicos consideráveis, por vários motivos, incluindo: (a) o acréscimo de sensibilidade ao *esforço reduzido* da bateria de testes usada, sem necessidade de tempo adicional para a administração de mais testes; (b) a possibilidade de obter mais informação acerca da validade dos resultados do processo de avaliação quando surgem dúvidas, mesmo quando não são administrados testes especificamente construídos para detectar *esforço reduzido*; (c) comparativamente aos testes de validade de sintomas, os elementos de informação provenientes de testes clássicos são menos “evidentes” ou “perceptíveis” para o sujeito quanto ao objectivo da detecção de *simulação* ou *esforço insuficiente* tornando, desta forma, estes testes menos susceptíveis ao “treino” ou “dissimulação” (Etherton *et al.*, 2005; Mathias *et al.*, 2002; Meyers & Volbrecht, 1998); (d) o complemento dos indicadores provenientes de instrumentos especificamente construídos para avaliar *simulação* ou *esforço reduzido*; além de funcionarem como verificações adicionais da validade de toda a avaliação (Strauss, Sherman & Spreen, 2006), contribui ainda para a utilidade clínica desses mesmos testes psicológicos (Cercy *et al.*, 1997).

De um modo geral, é recomendável o recurso a uma abordagem multidimensional, que integre várias fontes de evidência para detectar dano cognitivo fictício ou exagerado (Nitch & Glassmire, 2007). Não existe nenhum padrão de referência (“*gold standard*”) que determine, de modo seguro e definitivo, se um indivíduo está a fingir ou a exagerar a presença de défice cognitivo (Farkas *et al.*, 2006; Ferguson, 2003; Larrabee, 2008; Rogers, 1997). Por isso, a utilização conjunta de indicadores de validade provenientes dos resultados quer de medidas específicas, quer de testes mais clássicos ou comumente usados (WAIS-III, p. ex.), parece ser o procedimento mais objectivo e fiável para detectar a presença de *esforço reduzido* (Bender, 2008; Nitch & Glassmire, 2007).

Escalas de Inteligência de Wechsler para Adultos (WAIS-R/WAIS III): Contributos para a detecção de esforço reduzido

A frequência da utilização das Escalas de Inteligência de Wechsler para Adultos (WAIS-R e WAIS-III) e o conhecimento geral disponível acerca destas provas, justificam o conjunto de investigações que têm sido recentemente realizadas com o intuito de determinar a sua utilidade na detecção de *esforço reduzido*. A pesquisa neste âmbito enfatiza o interesse particular do subteste Memória de Dígitos (MD) e da diferença entre pontuações no Vocabulário e Memória de Dígitos (V-MD). Os resultados no subteste Memória de Dígitos têm mostrado particular

sensibilidade ao *esforço reduzido* constituindo um potencial marcador de uma *performance* deliberadamente insuficiente ou incompleta numa avaliação neuropsicológica (Babikian & Boone, 2007; Iverson & Tulsky, 2003). A Memória de Dígitos é um subteste rápido e fácil de administrar, integrando várias baterias frequentemente utilizadas na avaliação de pacientes com TCE (cf. Etherton *et al.*, 2005). O subteste MD é percebido pelas pessoas em geral como uma medida de memória (Trueblood & Schmidt, 1993, *cit in* Babikian & Boone, 2007). No entanto, esta prova mede, sobretudo, a duração da atenção. Esta percepção errónea faz com que os indivíduos com intenções de fingir ou simular disfunção de memória escolham esta tarefa para evidenciar os seus “défices” de memória (Babikian & Boone, 2007). Aparentemente, a MD seria um subteste no qual pacientes com lesões cerebrais experimentariam dificuldades (Meyers & Volbrecht, 1998). No entanto, a investigação conclui que os resultados neste subteste não são afectados mesmo em casos de lesão cerebral severa (Capruso & Levin, 1992, *cit in* Babikian & Boone, 2007; Langeluddecke & Lucas, 2003) e que as pontuações *borderline* ou que reflectem incapacidade no subteste MD (i.e. pontuações padronizadas ≤ 5) ocorrem em menos de 5%, tanto da população normal como de grupos clínicos (Iverson & Tulsky, 2003). Estas são algumas das razões que tornam a MD numa tarefa particularmente apropriada para a detecção de *esforço reduzido*.

Os diversos estudos que utilizam o subteste MD para detecção de *esforço reduzido* permitiram identificar vários indicadores de validade, nomeadamente: (i) *pontos de corte “simples”*, relativos ao desempenho do sujeito no teste (e.g., pontos de corte para a “Pontuação total bruta da MD”, “Pontuação total padronizada da MD”, “Pontuação bruta em sentido directo”, “Pontuação bruta em sentido inverso”, “Maior número de dígitos em sentido directo” e “Maior número de dígitos em sentido inverso”) e (ii) *novos indicadores*, resultantes da combinação de índices (e.g., *Reliable Digit Span*, RDS). Apesar da sua atractividade e simplicidade, os indicadores baseados em *pontos de corte simples* tendem a demonstrar sensibilidade baixa e especificidade moderada. Do mesmo modo, a sensibilidade dos indicadores construídos através da *combinação de índices* é tipicamente baixa; porém, a especificidade destas medidas tende a ser superior à encontrada com base no recurso a pontos de corte simples (Strauss, Sherman & Spreen, 2006).

O subteste Vocabulário tem sido igualmente utilizado na detecção de *esforço reduzido*. É uma prova relativamente resistente à disfunção do sistema nervoso central e proporciona resultados bastante estáveis (Miller *et al.*, 2004; Groth-Marnat, 2003). A pertinência deste subteste para a determinação de *esforço reduzido* surge, essencialmente, quando ele é associado ao subteste MD, constituindo-se a diferença entre Vocabulário e MD (V-MD) como um novo indicador.

O presente estudo pretende: 1. examinar a validade e utilidade dos indicadores de *esforço reduzido* obtidos a partir dos subtestes Memória de Dígitos e Vocabulário da WAIS-III; 2. analisar as relações entre pontuações em provas especificamente construídas para detectar *enviesamento de resposta negativo de natureza cognitiva* (*Rey 15-Item Test*; TOMM e SIMS) e indicadores de validade interna derivados de instrumentos clínicos tradicionais (subtestes Memória de Dígitos e Vocabulário da WAIS-III); 3. examinar a sensibilidade e especificidade dos indicadores de *esforço reduzido* derivados dos subtestes Memória de Dígitos e Vocabulário da WAIS-III.

Metodologia

Participantes

Amostra 1: Litigantes com Traumatismo Crânio Encefálico

A amostra de litigantes com TCE (grupo experimental; cf. Quadro 1) é constituída por 40 indivíduos adultos envolvidos, no momento da avaliação, em processos civis e num contexto médico-legal) potencialmente compensatório. Esta amostra integra 28 sujeitos do sexo masculino (70%) e 12 do sexo feminino (30%), com idades compreendidas entre os 19 e os 78 anos ($M=42.32$; $DP=15.37$) e anos de escolaridade entre 2 e 17 ($M=7.50$; $DP=3.91$).

Os sujeitos desta amostra foram avaliados, em média, 67.34 meses após a lesão ter ocorrido ($DP=49.67$). Na maioria dos casos, as lesões decorreram de acidentes de viação (62.5%), tendo resultado, em menor número, de acidentes de trabalho (15%) e de acidentes de viação em trabalho (22.5%). Os TCE são predominantemente fechados (26 sujeitos; 65%) e a maioria dos sujeitos teve TCE com perda de conhecimento durante mais de 24 horas (19 sujeitos; 47.5%). No que diz respeito ao tempo de coma, os dados recolhidos em 38 dos 40 sujeitos da amostra indicam que, em média, os sujeitos desta amostra permaneceram 13.45 dias em coma ($DP=41.22$).

Amostra 2: Grupo de Controlo (Comunidade)

A amostra controlo é constituída por 30 sujeitos da população geral, sem qualquer envolvimento com o sistema legal no momento da avaliação. Os sujeitos deste grupo são, maioritariamente, do sexo masculino (66.7%) e apresentam idades compreendidas entre os 19 e os 77 anos, sendo a média de idades de 47 anos ($DP=15.68$). Quanto à escolaridade, os sujeitos da amostra controlo estudaram, em média, 6.43 anos ($DP=3.51$). As amostras 1 e 2 não diferem significativamente em termos de idade ($t(68)=-1.302$; $p=.197$), anos de escolaridade ($t(68)=1.179$; $p=.242$) e género (qui-quadrado(1)=.088; $p=.766$).

Instrumentos

Escala de Inteligência de Wechsler para Adultos - Terceira edição (WAIS-III; Wechsler, 1997, 2008). Esta escala proporciona uma medida geral das capacidades intelectuais e pode ser aplicada a sujeitos dos 16 aos 89 anos de idade. Nas análises referentes aos indicadores de esforço da WAIS-III serão utilizados os seguintes pontos de corte: <7 para a "Pontuação total bruta na MD" (MD/PB); ≤ 5 para a "Pontuação total padronizada na MD" (MD/PP); ≤ 5 para a "Pontuação bruta na MD em sentido directo" (MD/PB em SD); ≤ 4 para o "Maior número de dígitos em sentido directo" (MD/Maior nº em SD); <3 para a "Pontuação bruta na MD em sentido inverso" (MD/PB em SI); <3 para o "Maior número de dígitos em sentido inverso" (MD/Maior nº em SI); ≤ 6 para o *Reliable Digit Span* (RDS); <7 para a "Pontuação total padronizada no Vocabulário" (V/PP); e <2 para a "Diferença entre a Memória de Dígitos e o Vocabulário" (V-MD) (cf. p. ex. Axelrod *et al.*, 2006; Babikian *et al.*, 2006; Demakis, 2004; Greve *et al.*, 2003; Mathias *et al.*, 2002; Millis *et al.*, 1998; Mittenberg *et al.*, 2001; Schwarz *et al.*, 2006).

Quadro 1. Caracterização da amostra de litigantes com TCE (n=40)

		N	Frequências	Média	Desvio-padrão	Amplitude
Género	Masculino	28	70%			
	Feminino	12	30%			
	Total	40	100%			
Idade		40		42.32	15.37	19-78
Anos de escolaridade		40		7.50	3.91	2-17
Intervalo entre a avaliação e a lesão (meses)		38		67.34	49.67	8-240
Tipo de acidente	Viação	25	62.5%			
	Trabalho	6	15%			
	Viação em trabalho	9	22.5%			
	Total	40	100%			
Tempo de hospitalização (dias)		38		58.60	85.13	1-369
Tipo de TCE	Aberto	14	35%			
	Fechado	26	65%			
	Total	40	100%			
Perda de conhecimento	Sem perda de conhecimento	5	12.5%			
	Menos de 30 minutos (ligeiro)	3	7.5%			
	Entre 1 a 24 horas (moderado)	7	17.5%			
	Mais de 24 horas (severo)	19	47.5%			
	Com perda de conhecimento (tempo não especificado)	6	15%			
	Total	40	100%			
Tempo de coma (dias)		38		13.45	41.22	0-240
Escala de Coma de Glasgow	Total	14	35%	10.71	3.73	7-15
	Inferior a 8 pontos (severo)	6	15%			
	9 a 12 pontos (moderado)	3	7.5%			
	13 a 15 pontos (ligeiro)	5	12.5%			
	Total	14	35%			
Exames do SNC (TAC e EEG)	Com alterações	18	45%			
	Sem alterações	3	7.5%			
	Sem informação	19	47.5%			
Área afectada	Frontal	2	5%			
	Frontal bilateral e occipital direito	1	2.5%			
	Frontal direito	2	5%			
	Frontal temporal	1	2.5%			
	Frontal temporal direito	2	5%			
	Lesão difusa	2	5%			
	Occipital esquerdo	1	2.5%			
	Parietal esquerdo	1	2.5%			
	Temporal direito	1	2.5%			
	Temporal esquerdo	1	2.5%			
	Temporo parietal	1	2.5%			
	Total	15	37.5%			
Disability Rating Scale	Sem incapacidade (0)	40	100%	4.00	2.82	0-14
	Ligeira (1)	1	2.5%			
	Parcial (2-3)	3	7.5%			
	Moderada (4-6)	19	47.5%			
	Moderadamente grave (7-11)	11	27.5%			
	Grave (12-16)	5	12.5%			
	Extremamente grave (17-21)	1	2.5%			
	Estado vegetativo (22-24)	0	0%			
	Estado vegetativo extremo (25-29)	0	0%			
	Morte (30)	0	0%			
	Total	40	100%			

Test of Memory Malingering (TOMM; Tombaugh, 1996). O TOMM é um teste de escolha forçada, desenvolvido para avaliar *esforço reduzido*, constituído por 50 itens (desenhos de objectos comuns) para cada ensaio (dois ensaios de aprendizagem e um ensaio de retenção opcional). Apesar de originariamente desenvolvido para adultos, o TOMM pode ser administrado a partir dos 5 anos de idade (Constantinou & McCaffrey, 2003).

Rey 15-Item Test (Rey, 1964, *cit in* Strauss, Sherman & Spreen, 2006; Boone, Salazar *et al.*, 2002). O *Rey 15-Item Test* destina-se à detecção de *esforço reduzido*. Inclui 15 itens simples e redundantes, apresentados em 5 linhas com 3 caracteres cada. Foi utilizado o formato standardizado da prova, recorrendo às instruções de Goldberg e Miller (1986, *cit in* Strauss, Sherman & Spreen, 2006) e, imediatamente a seguir, o ensaio de reconhecimento proposto por Boone, Salazar e colaboradores (2002).

Structured Inventory of Malingered Symptomatology (SIMS; Widows & Smith, 2005). O SIMS é um inventário em formato de resposta verdadeiro/falso, desenvolvido para o rastreio de *enviesamento de resposta negativo* de sintomas psicopatológicos e/ou neuropsicológicos. Inclui 75 itens agrupados em cinco escalas: Psicose (P), Perturbações Afectivas (AF), Incapacidade Neurológica (NI), Amnésia (AM) e Baixa Inteligência (LI). Pode ser administrado a partir dos 18 anos de idade.

Procedimento

Num primeiro momento foi obtida autorização por parte da Delegação do Centro do Instituto Nacional de Medicina Legal para a realização da investigação e pedida a colaboração dos peritos do Serviço de Clínica Forense para a sinalização de casos com TCE. Para a constituição da Amostra 1 (grupo experimental) foram seleccionados indivíduos adultos, com indicação de TCE no processo, referenciados para avaliação no Serviço de Clínica Forense da Delegação do Centro do Instituto Nacional de Medicina Legal. A recolha de dados relativa a esta amostra teve lugar no decurso da avaliação psicológica solicitada pelo Tribunal ou como complemento da avaliação realizada por peritos de outras áreas do Serviço de Clínica Forense (Psiquiatria, Neurologia, Avaliação de dano). O processo de recolha de dados decorreu ao longo de aproximadamente 2 anos e incluiu, posteriormente, a consulta de dados dos processos.

Relativamente à Amostra 2 (grupo controlo), foram seleccionados sujeitos adultos da população geral (comunidade), de forma a aproximar as características destes da amostra de litigantes com TCE relativamente às variáveis género, idade e anos de escolaridade, salvaguardando assim a equivalência entre os grupos quanto a estas variáveis. Constituiu critério de exclusão a presença de perturbações sensório-perceptivas, história de TCE e história de tratamento motivado por perturbação psiquiátrica.

As análises estatísticas foram efectuadas com o auxílio do programa *Statistical Package for the Social Sciences* (SPSS - versão 16.0). Nas análises efectuadas, recorreremos a métodos paramétricos e não paramétricos para as variáveis que não cumpriam os pressupostos de utilização de provas paramétricas, nomeadamente: (i) aderência à curva normal (teste de Shapiro-Wilk);

(ii) homogeneidade das variâncias (teste de Levene); e (iii) N da amostra inferior a 30 (Pestana & Gageiro, 2003). Assim, quando comparámos os resultados entre grupos utilizámos o teste t de Student (paramétrico) e o teste U de Mann-Whitney (não paramétrico). Para as variáveis nominais utilizámos o teste do Qui-Quadrado, e nas análises da correlação, utilizámos a versão não paramétrica (Correlação de Spearman).

Apresentação, análise e discussão dos resultados

Os resultados a seguir apresentados (cf. Quadro 2) reportam-se aos diferentes indicadores de esforço derivados dos subtestes Memória de Dígitos e Vocabulário da WAIS-III.

Quadro 2. Análises descritivas e testes estatísticos relativos aos indicadores de esforço derivados dos subtestes Memória de Dígitos e Vocabulário da WAIS-III

	Ponto de corte	Litigantes com TCE (n= 40)			Controlo (n=30)			(t)/U
		M	DP	Amp.	M	DP	Amp.	
MD/PB	<7	9.75	2.83	4-16	12.10	3.81	1-22	380.00**
MD/PP	≤5	6.85	2.22	3-11	9.10	2.48	5-15	298.00**
MD/PB em SD	≤5	6.02	1.59	2-9	7.07	1.60	4-10	(-2.704)**
MD/Maior nº em SD	≤4	4.38	0.95	2-6	4.80	1.09	2-7	479.00
MD/PB em SI	<3	3.72	1.71	1-7	4.67	2.01	2-10	437.50*
MD/Maior nº em SI	<3	3.12	0.99	2-5	3.70	1.12	2-7	427.50*
RDS	≤6	5.98	1.46	3-9	6.90	1.67	5-10	401.50**
V/PP	<7	7.82	3.33	3-15	10.53	2.43	7-16	299.00**
V-MD	>2	0.98	2.76	-5-6	1.43	2.74	-3-9	560.50

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

Os dados indicam a presença de diferenças estatisticamente significativas entre os grupos para a maior parte dos indicadores de esforço da WAIS-III em estudo ($p < .05$), sendo as pontuações médias obtidas pelo grupo de litigantes com TCE inferiores às pontuações do grupo de controlo. Apenas não foram encontradas diferenças significativas entre os grupos para os indicadores MD/Maior nº em SD e o V-MD ($U=479.00$ e $U=560.50$; $p > 0.05$, respectivamente).

Se atendermos aos pontos de corte propostos para estes indicadores (MD e Vocabulário) verificamos que, no que diz respeito ao grupo de controlo, foram obtidos desempenhos médios que não indiciam a presença de *esforço reduzido*. A análise das amplitudes de resultados indica, no entanto, que poderão surgir situações de identificação incorrecta de *esforço reduzido* neste grupo, em todos os indicadores de esforço considerados. O único indicador para o qual não foram encontradas pontuações abaixo do ponto de corte e sugestivas de *esforço reduzido* foi a "Pontuação padronizada no Vocabulário" (V/PP <7; amplitude: 7-16).

Analisando os resultados médios do grupo de litigantes com TCE para os indicadores de *esforço reduzido* derivados dos subtestes Memória de Dígitos e Vocabulário, verificamos que as pontuações obtidas em MD/PB ($M=9.75$; $DP=2.83$), MD/PP ($M=6.85$; $DP=2.22$), MD/PB em SD ($M=6.02$; $DP=1.59$), MD/PB em SI ($M=3.72$; $DP=1.71$), MD/Maior nº em SD ($M=4.38$; $DP=0.95$), MD/Maior nº em SI ($M=3.12$; $DP=0.99$), e V/PP ($M=7.82$; $DP=3.33$) encontram-se acima dos

pontos de corte propostos, não apontando, portanto, para a presença de *esforço reduzido*, tal como ocorre com o indicador V-MD ($M=0.98$ e $DP=2.76$) que possui um valor inferior ao ponto de corte. Apenas os desempenhos médios no RDS ($M=5.98$ e $DP=1.46$) foram inferiores ao ponto de corte ≤ 6 , considerado como potencialmente indicativo de *esforço reduzido*. No entanto, devem considerar-se as amplitudes de resultados obtidas nestes indicadores, que sugerem a existência de casos de possível *esforço reduzido*.

Na interpretação destes resultados, deverá ter-se em consideração que as amostras consideradas nos estudos de validação destes indicadores diferem da amostra utilizada no presente trabalho, uma vez que esta não é constituída por indivíduos suspeitos de *enviesamento de resposta negativo* e/ou *esforço reduzido* (*known-groups design*) ou por sujeitos treinados para esse objectivo (*simulation design*) mas, meramente, por indivíduos que se encontram num contexto potencialmente favorável ao aparecimento de tais comportamentos (contexto médico-legal). Deste modo, os resultados relativos ao desempenho do grupo de litigantes com TCE indicam que a diminuição da *performance* nos indicadores referidos não foi tão acentuada como aquela que parece surgir em grupos de simuladores suspeitos ou com indicações de *enviesamento de resposta negativo*. Apenas o indicador RDS teve resultados médios (inferiores a 6) sugestivos de *esforço reduzido*.

Relações entre diferentes indicadores de “esforço reduzido” derivados dos subtestes Memória de Dígitos e Vocabulário

Analisando a matriz de correlações respeitante ao grupo litigante (Quadro 3), observamos que as correlações alcançadas estatisticamente significativas, oscilam entre .408 (RDS e V/PP) e .943 (MD/PB em SI e MD/Maior nº em SI), evidenciando a existência de uma relação, moderada a forte, entre os diferentes indicadores. Os resultados obtidos foram superiores aos observados por Larrabee (2008), que comparou os desempenhos de um grupo de litigantes com “*definite malingered neurocognitive dysfunction*” ($n=26$) em cinco indicadores de esforço reduzido derivados de provas neuropsicológicas (*Visual Form Discrimination Test*; *Finger Tapping Test*; *Reliable Digit Span*; *Wisconsin Card Sorting Test*; *Fake Bad Scale* do MMPI-2) observando correlações que oscilaram entre a .008 e -.517. Estes dados apontam para o interesse de uma utilização conjunta destas medidas na determinação do *esforço reduzido*, ao contrário do presente trabalho, onde as associações mais elevadas entre os vários indicadores indicam que estes não são independentes. Deste modo, a utilização de todos estes indicadores na determinação de *esforço reduzido* é susceptível de proporcionar informação convergente mas ao mesmo tempo redundante, pelo que podem não existir benefícios da sua utilização conjunta (cf. Larrabee, 2008; Nelson *et. al.*, 2003). De acordo com estes investigadores, os benefícios da utilização de múltiplas medidas para a detecção de *esforço reduzido* apenas se verificam caso exista independência ou baixas inter-correlações entre os testes ou indicadores de validade utilizados. Isto é, se os resultados nas diferentes provas se encontrarem fortemente correlacionados, eles não proporcionam individualmente informação autónoma exclusiva.

Quadro 3. Correlações entre os indicadores de esforço derivados dos subtestes Memória de Dígitos e Vocabulário para o grupo de litigantes com TCE

	MD/PP	MD/ PB em SD	MD/ Maior nº em SD	MD/ PB em SI	MD/ Maior nº em SI	RDS	V/PP	V-MD
MD/PB	.915**	.849**	.765**	.900**	.867**	.834**	.486**	-.161
MD/PP	-	.791**	.738**	.837**	.796**	.726**	.578**	-.108
MD/PB em SD		-	.838**	.576**	.567**	.666**	.442**	-.157
MD/Maior nº em SD			-	.578**	.553**	.508**	.447**	-.091
MD/PB em SI				-	.943**	.784**	.468**	-.090
MD/Maior nº em SI					-	.772**	.455**	-.081
RDS						-	.408**	-.085
V/PP							-	.723**

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

No que diz respeito ao grupo controlo, a análise dos dados com significância estatística da matriz de correlações (Quadro 4) permite observar associações baixas entre V/PP e RDS (.373) e entre V-MD e RDS (-.369). Os coeficientes de correlação entre os restantes indicadores são de magnitude moderada a elevada, oscilando entre .411 (V/PP e MD/PB em SI) e .969 (MD/PB em SI e MD/Maior nº em SI).

Quadro 4. Correlações entre os indicadores de *esforço reduzido* derivados dos subtestes Memória de Dígitos e Vocabulário para o grupo controlo

	MD/PP	MD/ PB em SD	MD/ Maior nº em SD	MD/ PB em SI	MD/ Maior nº em SI	RDS	V/PP	V-MD
MD/PB	.943**	.893**	.776**	.882**	.887**	.807**	.426**	-.501**
MD/PP	-	.823**	.709**	.795**	.818**	.739**	.448**	-.537**
MD/PB em SD		-	.887**	.740**	.745**	.768**	.308	-.496**
MD/Maior nº em SD			-	.663**	.649**	.754**	.209	-.495**
MD/PB em SI				-	.969**	.795**	.328	-.411*
MD/Maior nº em SI					-	.711**	.421*	-.340
RDS						-	.373*	-.369*
V/PP							-	.463*

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

Os dados obtidos no grupo controlo são igualmente contraditórios com os resultados encontrados por Larrabee (2008) que observou baixas correlações (entre .009 e .398) diferentes indicadores de *esforço reduzido* derivados de provas neuropsicológicas (*Visual Form Discrimination Test*; *Finger Tapping Test*; *Reliable Digit Span*; *Wisconsin Card Sorting Test*; *Fake Bad Scale* do MMPI-2) numa amostra de TBI sem indicações de *esforço reduzido*. De facto, no grupo controlo do presente estudo apenas se observaram dois coeficientes de correlação baixos (V/PP e RDS; V-MD e RDS), sendo as restantes correlações moderadas a elevadas, o que indica, uma vez mais, uma convergência da informação obtida pelos diferentes indicadores. Acrescentamos ainda que, uma vez que o nosso grupo de controlo não apresenta uma natureza clínica como

a amostra de Larrabee (2008), incluindo apenas sujeitos da população geral sem qualquer incentivo para apresentar *enviesamento de resposta negativo*, torna-se imperativa a replicação desta análise com uma população clínica com TCE, fora do âmbito litigioso.

Relações entre diferentes provas específicas de enviesamento de resposta negativo e indicadores de esforço reduzido derivados dos subtestes Memória de Dígitos e Vocabulário

Recorrendo aos dados do Quadro 5, verifica-se em termos gerais que no grupo de litigantes com TCE são observadas correlações positivas, baixas a moderadas, entre os indicadores derivados da WAIS-III e provas específicas de *enviesamento de resposta negativo* (.346 a .650 no Rey 15-Item Test; .385 a .493 no TOMM; -.313 a -.517 nas escalas neuropsicológicas do SIMS).

Concluimos que estas provas (TOMM, Rey 15-Item Test e SIMS) constituem indicadores de *enviesamento de resposta negativo* e/ou *esforço reduzido* em termos cognitivos relativamente independentes, podendo ser utilizados conjuntamente com indicadores de esforço da WAIS-III uma vez que, atendendo à magnitude dos coeficientes de correlação obtidos, estes vários indicadores e as provas específicas de *enviesamento de resposta negativo* não fornecem informação redundante (cf. Nelson *et. al.*, 2003).

Quadro 5. Correlações entre indicadores de esforço derivados dos subtestes Memória de Dígitos e Vocabulário e indicadores de esforço de provas específicas no grupo de litigantes com TCE (n=40)

		Indicadores derivados dos subtestes MD e Vocabulário								
		MD/PB	MD/P P	MD/PB em SD	MD/ Maior nº em SD	MD/ PB em SI	MD/ Maior nº em SI	RDS	V/PP	V-MD
Rey 15- Item Test	Evocação	.532**	.358*	.356*	.346*	.551**	.650**	.474**	.210	-.038
	Pont. combinada	.540**	.373*	.377*	.414**	.517**	.599**	.493**	.209	-.074
TOMM	Ensaio 2	.385*	.255	.305	.272	.386*	.440**	.443**	.095	-.083
	Ensaio de Retenção ¹	.356	.340	.293	.102	.355	.327	.466*	.493*	.479*
SIMS	LI	-.428**	-.314*	-.512**	-.517**	-.290	-.233	-.365*	-.108	.141
	AM	-.313*	-.235	-.504**	-.419**	-.106	-.205	-.267	-.109	.064
	NI	-.373*	-.299	-.393*	-.443**	-.273	-.306	-.322*	-.275	-.101

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

¹ Apenas 19 sujeitos deste grupo responderam ao Ensaio de retenção do TOMM, pelo que os coeficientes de correlação aqui apresentados se restringem a este grupo. Este ensaio é opcional: apenas os sujeitos com pontuações inferiores a 45 no Ensaio 2 responderam igualmente a este Ensaio de retenção.

Análise da sensibilidade e especificidade dos indicadores de “esforço reduzido” derivados dos subtestes Memória de Dígitos e Vocabulário

A impossibilidade de analisar as questões associadas à sensibilidade dos indicadores de esforço derivados dos subtestes Memória de Dígitos e Vocabulário, devido à natureza da amostra de litigantes, constituída apenas por indivíduos que se encontram num contexto potencialmente favorável ao aparecimento de comportamentos de *enviesamento de resposta negativo* e/ou *esforço reduzido*, conduziu-nos à formação de quatro grupos distintos a partir das duas amostras iniciais. Deste modo, a sensibilidade e especificidade destes indicadores serão analisados a partir dos desempenhos de sujeitos litigantes que apresentam resultados inferiores ao ponto de corte no Ensaio 2 do TOMM (<45 acertos) e, simultaneamente, na Evocação do Rey 15-Item Test (<9 evocações correctas; *Grupo de TCE com esforço reduzido*), sujeitos litigantes com resultados sugestivos de *esforço reduzido* em apenas uma destas provas (*Grupo de TCE com validade questionável*), sujeitos litigantes com TCE que pontuaram acima dos pontos de corte nas duas provas específicas de *esforço reduzido* (*Grupo de TCE sem esforço reduzido*) e, ainda, no Grupo controlo inicialmente constituído.

Observando os valores de sensibilidade e especificidade alcançados para todos os indicadores de esforço derivados dos subtestes Memória de Dígitos e Vocabulário (Quadro 6), constatamos que os valores de sensibilidade oscilam entre 18.2% e 63.6%. Por outro lado, os valores de especificidade variam entre 52.6% e 94.7% para os sujeitos com TCE e entre 53.3% e 100% para o grupo controlo. Estes dados indicam que a proporção de sujeitos com *esforço reduzido* que são correctamente identificados pelo indicador (sensibilidade) é baixa a moderada para os indicadores derivados da WAIS-III. Por outro lado, a proporção de examinados que não apresentam *esforço reduzido* e que são correctamente identificados como tal (especificidade) é moderada a elevada.

Os resultados obtidos estão de acordo com investigações relativas à validade dos referidos pontos de corte para indicadores da WAIS-III com *grupos conhecidos*, nomeadamente, sujeitos com TCE (Axelrod *et al.*, 2006; Babikian *et al.*, 2006; Greve *et al.*, 2003; Heinly *et al.*, 2005; Mittenberg *et al.*, 2001), que apontam para valores baixos de sensibilidade (entre 21 e 39%) e valores elevados de especificidade ($\geq 90\%$). A única excepção é encontrada na especificidade para a diferença entre o Vocabulário e a Memória de Dígitos que varia entre 63% e 86% (Greve *et al.*, 2003; Millis *et al.*, 1998; Mittenberg *et al.*, 2001). No entanto, note-se que não existem estudos com grupos de referência (*known groups design*) relativos ao ponto de corte <3 para o indicador MD/PB em SI.

A pesquisa em torno da sensibilidade e especificidade destes indicadores salienta a utilidade de um ponto de corte <7 pontos na “Pontuação total bruta da MD” (especificidade de 94.7-100% e sensibilidade de 18.2-40%), bem como de um ponto de corte <3 pontos na “Pontuação bruta na MD em sentido inverso” (especificidade: 84%-87%; sensibilidade: 45%-50%) e no “Maior número de dígitos em sentido inverso” (especificidade: 84%-87%; sensibilidade: 36%-60%), atendendo aos respectivos valores de especificidade. Estes dois últimos indicadores devem, no entanto, ser utilizados com prudência, pois, apesar de se encontrarem perto do que seria aceitável para a sua utilização (90%; Greve & Bianchini, 2004), possuem uma taxa de erro superior a 10%, o que poderá dever-se ao reduzido tamanho dos grupos.

Quadro 6. Sensibilidade e especificidade dos indicadores de *esforço reduzido* derivados dos subtestes Memória de Dígitos e Vocabulário

	Ponto de corte	Sensibilidade		Especificidade	
		TCE com esforço reduzido (n=10)	TCE com validade questionável (n=11)	TCE sem esforço reduzido (n=19)	Controlo (n=30)
MD/PB	<7	40% (4)	18.2% (2)	94.7% (18)	100% (30)
MD/PP	≤5	40% (4)	45.5% (5)	63.2% (12)	93.3% (28)
MD/PB em SD	≤5	60% (6)	36.4% (4)	78.9% (15)	80% (24)
MD/Maior nº em SD	≤4	60% (6)	45.5% (5)	52.6% (10)	60% (18)
MD/PB em SI	<3	50% (5)	45.5% (5)	84.2% (16)	86.7% (26)
MD/Maior nº em SI	<3	60% (6)	36.4% (4)	84.2% (16)	86.7% (26)
RDS	≤6	60% (6)	63.6% (7)	57.9% (11)	53.3% (16)
V/PP	<7	50% (5)	45.5% (5)	63.2% (12)	100% (30)
V-MD	>2	20% (2)	54.5% (6)	73.7% (14)	63.3% (19)

Existe algum consenso relativamente à necessidade de uma elevada especificidade em provas de esforço (ver recomendações de Greve & Bianchini, 2004), a fim de evitar as graves consequências da classificação de sujeitos honestos como apresentando *esforço reduzido* e, em última instância, como *simuladores*. A reduzida sensibilidade observada não inviabiliza a utilização de provas (e indicadores) de esforço uma vez que, na determinação de *esforço reduzido* não deve ser considerado apenas um único indicador, mas antes informação convergente do cruzamento de dados provenientes de diferentes provas (gerais e específicas), aspecto que contribui para o aumento da sensibilidade sem sacrificar a especificidade (cf. Larrabee, 2003, 2008; Farkas *et al.*, 2006).

Conclusão

Ao longo deste trabalho foi salientada a pertinência da avaliação dos comportamentos de *simulação* e de *esforço reduzido*. A utilização e validação de diferentes instrumentos e técnicas que possibilitam a avaliação de tais comportamentos constituíram o fio condutor deste estudo, com particular ênfase na utilidade de indicadores de *esforço reduzido* derivados de instrumentos clínicos tradicionais.

Este estudo pretendeu colmatar a escassez de investigação em torno deste problema no nosso país, constituindo, tanto quanto é do nosso conhecimento, a primeira investigação em Portugal que avalia *enviesamento de resposta negativo* em litigantes com Traumatismo Crânio Encefálico (TCE) examinados em contexto médico-legal. A análise da utilidade da WAIS-III, um instrumento de avaliação neuropsicológica muito utilizado com este grupo clínico, na detecção de comportamentos de *esforço reduzido* proporciona dados relevantes para avaliações de litigantes com TCE em contexto forense.

Os resultados da presente investigação são sugestivos da validade de alguns dos pontos de corte propostos para os indicadores de esforço da WAIS-III, nomeadamente, o ponto de corte <7 pontos na "Pontuação total bruta da MD" (especificidade de 94.7-100% e sensibilidade de 18.2-40%), bem como de um ponto de corte <3 pontos na "Pontuação bruta na MD em sentido inverso" (especificidade: 84%-87%; sensibilidade: 45%-50%) e no "Maior número de dígitos

em sentido inverso” (especificidade: 84%-87%; sensibilidade: 36%-60%), atendendo aos respectivos valores de especificidade. As correlações baixas e moderadas entre estes indicadores e provas específicas de *enviesamento de resposta negativa* (*Test of Memory Malingering*, *Rey 15-Item Test* e escalas LI, AM e NI do *Structured Inventory of Malingered Symptomatology*), remetem para a necessidade da sua utilização conjunta e complementar. A análise da utilidade de diferentes pontos de corte de provas específicas de *enviesamento de resposta negativa* e as respectivas correlações com diferentes pontos de corte para os indicadores derivados da WAIS-III não foi considerada no presente estudo, uma vez que não constituía o objectivo deste trabalho. Este tópico será objecto de um estudo futuro.

O reduzido tamanho das amostras e a ausência de um grupo constituído por indivíduos com TCE recente (bem documentado do ponto de vista clínico, incluindo dados objectivos acerca do grau de gravidade da lesão) e sem qualquer envolvimento com o sistema legal constituem limitações da presente investigação.

BIBLIOGRAFIA

- Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer. (2006). Detecting incomplete effort with digit span from Wechsler Adult Intelligence Scale - Third edition. *The Clinical Neuropsychologist*, 20, 513-523.
- Babikian, T., & Boone, K. B. (2007). Intelligence tests as measures of effort. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 103-127). New York: Guilford Press.
- Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspected effort. *The Clinical Neuropsychologist*, 20, 145-159.
- Bender, S. D. (2008). Malingered Traumatic Brain Injury. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.; pp. 69-86). New York: Guilford Press.
- Blackburn, R. (2006). Relações entre psicologia e direito. In A. C. Fonseca, M. R. Simões, M. C. Simões, M. S. Pinho (Eds.), *Psicologia forense* (pp. 3-24). Coimbra: Almedina.
- Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item recognition trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology*, 24(5), 561-573.
- Bush, S. S., Ruff, R. M., Troster, 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. *Archives of Clinical Neuropsychology*, 20, 419-426.
- Cercy, S., Schretlen, D. J., & Brandt, J. (1997). Simulated amnesia and the pseudo-memory phenomena. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed.; pp.108-129). New York: Guilford Press.
- Constantinou, M., & McCaffrey, R. (2003). Using the TOMM for evaluating children's effort to perform optimally on neuropsychological measures. *Child Neuropsychology*, 9, 81-90.
- Demakis, G. J. (2004). Application of clinically-derived malingering cutoffs on the California Verbal Learning Test and the Wechsler Adult Intelligence Test-Revised to an analog malingering study. *Applied Neuropsychology*, 11(4), 220-226.
- Denney, R. L. (2008). Negative response bias and malingering during neuropsychological assessment in criminal forensic settings. In R. L. Denney, J. P. Sullivan (Eds.), *Clinical neuropsychology in the criminal forensic setting* (pp. 91-134). New York: Guilford Press.
- Etherton, J., Bianchini, K., Greve, K., & Heinly, M. (2005). Sensitivity and specificity of Reliable Digit Span in malingered pain-related disability. *Assessment*, 12, 130-136.
- Farkas, M. R., Rosenfeld, B., Robbins, R. & van Gorp (2006). Do tests of malingering concur? Concordance among malingering measures. *Behavioral Sciences & the Law*, 24, 659-671.
- Ferguson, K. (2003). Detecting malingering in forensic neuropsychological evaluations in litigants with mild traumatic brain injury. In W. T. O'Donohue, E. R., Levensky (Eds.), *Handbook of forensic psychology* (pp.301-314). San Diego, CA: Elsevier Academic Press.
- Fonseca, A. C. (2006). Psicologia forense: Uma breve introdução. In A. C. Fonseca, M. R. Simões, M. C. Simões, M. S. Pinho (Eds.), *Psicologia forense* (pp. 3-24). Coimbra: Almedina.
- Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2003). Detecting malingered performance on the Wechsler Intelligence Scale: Validation of Mittenberg's approach in traumatic brain injury. *Archives of Clinical Neuropsychology*, 18, 245-260.
- Greve, K., & Bianchini, K. (2004). Setting empirical cut-offs on psychometric indicators of negative response bias: A methodological commentary with recommendations. *Archives of Clinical Neuropsychology*, 19, 533-541.
- Groth-Marnat, G. (2003). Wechsler intelligence scales. In G. Groth-Marnat (Ed.), *Handbook of psychological assessment* (4th ed.; pp.129-196). New York: John Wiley and Sons, Inc.
- Heinly, M. T., Greve, K. W., Bianchini, K. J., Love, J. M., & Brennan, A. (2005). WAIS digit span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. *Assessment*, 12(4), 429-444.
- Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS-III: Unusual digit span performances patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology*, 18, 1-9.